

# Modeling Seasonal Sudden Stratospheric Warming Climatology Based on Polar Vortex Statistics

MATTHEW F. HORAN AND THOMAS REICHLER

*Department of Atmospheric Sciences, University of Utah, Salt Lake City, Utah*

(Manuscript received 19 April 2017, in final form 3 August 2017)

## ABSTRACT

This study investigates the climatological frequency distribution of sudden stratospheric warmings (SSWs). General circulation models (GCMs) tend to produce SSW maxima later in winter than observations, which has been considered as a model deficiency. However, the observed record is short, calling into question the representativeness of the observational record. To study the seasonality of SSWs and the factors behind it, the authors use observations, a long control simulation with a stratosphere resolving GCM, and also a simple statistical model that is based on the climatological seasonal cycle of the polar vortex winds. From the combined analysis, the authors conclude that the late-winter SSW maximum seen in most climate models is realistic and that observations would also have a late-winter SSW maximum if more data were available. The authors identify the seasonally varying strengths of the polar vortex and stratospheric wave driving as the two main factors behind the seasonal SSW distribution. The statistical model also indicates that there exists a continuum of weak polar vortex states and that SSWs simply form the tail of normally distributed stratospheric winds.

## 1. Introduction

It is now widely accepted that conditions in the stratosphere matter for the troposphere and that the representation of the stratosphere in numerical models is important for simulating the atmosphere on a range of time scales. Of particular importance are sudden stratospheric warmings (SSWs), which are extreme wintertime circulation anomalies of the Northern Hemisphere that occur at irregular intervals in about every second year. At the surface, SSWs tend to be associated with the negative phase of the North Atlantic Oscillation, a southward-shifted North Atlantic storm track, and weather extremes over the North Atlantic sector (Thompson et al. 2002; Kidston et al. 2015). The simulation of SSWs in models contributes significantly to tropospheric seasonal forecast skill (Sigmond et al. 2013; Domeisen et al. 2015; Jia et al. 2017).

Despite the significance of SSWs for surface weather and climate, considerable uncertainty remains as to when in winter SSWs are most likely and what the reasons for their temporal distribution are. Previous studies (Charlton et al. 2007;

Charlton-Perez et al. 2008; Butchart et al. 2011) note that in observations SSWs occur mostly during mid-winter, while several modeling studies have shown that general circulation models (GCMs) tend to favor late-winter SSWs. One potential reason for such model biases could be differences in the number of simulated split or displacement events, since each event type has different seasonal cycles (Seviour et al. 2016). However, we believe that it is more important that SSWs are rare events, leading to a small observed monthly sample size. In the 67 winters of the NCEP–NCAR reanalysis (NNR; Kalnay et al. 1996), there are less than 10 SSWs in each month, creating large uncertainty in the shape of the seasonal distribution of SSWs. The first goal of the present study is therefore to decide how significant the differences in the seasonal timing of SSWs between observations and models are.

The seasonal distribution of SSWs and how it is simulated by models leads us also to ask what factors determine the climatological frequency of the number of SSWs during any given winter. SSWs rely on the upward propagation of planetary waves from the troposphere into the stratosphere and on the absorption of the waves in the stratosphere (Matsuno 1971). Previous studies attribute the onset of SSWs to an

---

*Corresponding author:* Thomas Reichler, thomas.reichler@utah.edu

increasing amount of planetary waves (Holton and Mass 1976; Schoeberl and Strobel 1980; Yoden 1987; Polvani and Waugh 2004; Limpasuvan et al. 2005), leaving the impression that pulses of planetary waves are important prerequisites for the creation of SSWs. Jucker (2016), however, claims that increased wave propagation is not the primary trigger for SSWs and that during most of Northern Hemisphere winter there is always enough upward-directed wave energy for SSWs. Similarly, Albers and Birner (2014) argue that increased planetary wave propagation is only responsible for the preconditioning of the polar vortex to favor SSWs but that the wave convergence is not responsible for the SSW itself. Irrespective of the question of exactly which aspect of the upward-directed planetary wave energy is most important for SSWs, the seasonally varying strength of the polar vortex is another potential factor that must be considered, as shown by Jucker et al. (2014) in the context of an idealized climate model. Since a reversal of the stratospheric winds is an important prerequisite for SSWs, it is reasonable to assume that to first order a stronger polar vortex requires more stratospheric wave driving to produce a wind reversal and an SSW.

In the climatological mean, both the stratospheric wave driving and also the strength of the polar vortex maximize during January. Hence, going from winter to spring, the weakening polar vortex strength (creating more SSWs) and reducing planetary wave driving (creating fewer SSWs) may have opposing effects on the seasonal timing of SSWs. This idea leads to the second goal of this research, namely, to understand the roles of seasonally varying stratospheric wave driving and polar vortex strength for the creation of SSWs and how the interplay between the two influences the timing of maximum SSWs in winter. We accomplish our goal by statistically modeling the daily evolution of SSW probability, using only information about the seasonally varying statistics of the zonal mean zonal winds at 10 hPa and 60°N (U1060) and assuming that on any given day U1060 is normally distributed around its mean. In the statistical model, the strength of the polar vortex is represented by the mean of U1060, and the wave driving and how effective the waves propagate and break in the polar vortex region are indirectly represented by the interannual standard deviation of daily U1060. The idea behind the latter assumption is that the eddy driving leads to intermittent weakening of the polar vortex and hence to interannual variability of daily U1060. Other potential influences on this variability, like from ENSO or the QBO, are assumed to be small, cyclic in nature, and irrelevant in the context of

climatological means. We then use the statistical model to analyze the influence of the individual SSW definition requirements (Charlton and Polvani 2007, hereafter CP07) and also the importance of polar vortex strength and stratospheric wave driving in determining the seasonality of SSWs. As we will show, the seasonally decreasing strength of the polar vortex is the primary factor, and the planetary wave driving is of secondary importance.

This work is structured as follows. Section 2 describes our data, methods, and assumptions. Section 3 discusses the statistical model and the results thereof in terms of the input parameters and the actual occurrence of SSWs. We then analyze how sensitive the statistical model is to changes in input parameters. Finally, section 4 presents a summary of our findings and our conclusions.

## 2. Methodology

Most of our analysis is based on daily output from a nearly 10 000-yr-long control run conducted with a stratosphere-resolving version of the Geophysical Fluid Dynamics Laboratory's (GFDL) Climate Model, version 2.1 (CM2.1) (Staten and Reichler 2014). The original version of the model (Delworth et al. 2006) has 24 vertical levels, while our version uses 48 vertical levels, with the additional levels being mostly concentrated in the stratosphere. Greenhouse gases, ozone concentrations, and other external forcings are prescribed to the model at 1990 levels and held constant through time. The well-resolved stratosphere is essential for the simulation of a more accurate stratospheric circulation (Charlton-Perez et al. 2013). We also use daily data from the 1948–2015 NNR (referred to as “observations”) and from select models of phase 5 of the Coupled Model Intercomparison Project (CMIP5). From each data source, we extract continuous multiyearlong daily time series of U1060 to create two SSW event distributions: an empirical distribution, based solely on the actual number of events that occur in the dataset itself, and a statistical distribution, based on the statistical model described below.

The daily climatologies of the mean, interannual standard deviation, and skewness of U1060 are empirically derived from the investigated dataset. The values are denoted as  $\mu(t)$ ,  $\sigma(t)$ , and  $\gamma(t)$ , respectively, where  $t$  is the day of the year. All values are smoothed in time using a Gaussian smoother with a kernel of three standard deviations. Additionally, we denote  $r(t, \tau)$  as the climatological lagged autocorrelation of U1060 for each day  $t$  and lag  $\tau$ . One of the basic

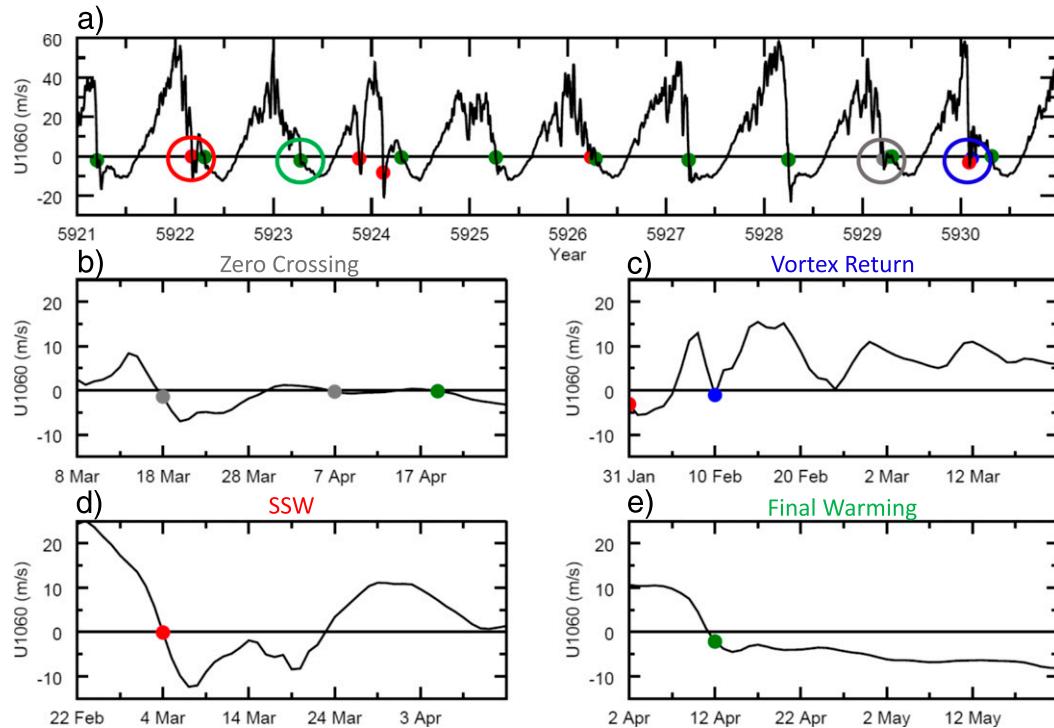


FIG. 1. U1060 event definition. (a) An arbitrary 10-yr-long time series of daily U1060. Color indicates zero crossings (gray), vortex returns (blue), SSWs (red), and final warmings (green). (b)–(e) Close-ups of events encircled in (a).

assumptions of this work is that on any given day, U1060 is approximately normally distributed. In other words, we assume that the most likely value of U1060 is at its daily mean and the actual values are evenly distributed around that mean following a Gaussian distribution. As we will show, this assumption is reasonable for most of our work.

Although some debate exists on how to best define SSWs (Butler et al. 2015), CP07 provide a widely used definition for major SSWs, which is similar to what we use in this work. The CP07 definition is based on a reversal of the polar vortex winds (as represented by U1060), while ignoring minor warmings. In addition, CP07 define some extra conditions that must be fulfilled for an SSW, leading us to describe the probabilities for the following four events:

- 1) Zero crossings are days when U1060 shifts from westerly to easterly. Zero crossings can occur on any day of the year, but empirically do so only during fall, winter, and spring.
- 2) Vortex returns are zero crossings that are additionally conditioned on 10 subsequent and consecutive days of positive values of U1060, which must occur before 31 May (rather than 30 April as in CP07). The date of a vortex return is the date of the

preceding zero crossing, and it must occur between 1 November and 31 March.

- 3) SSWs are isolated vortex returns that are separated from each other by at least 20 days. Should two vortex returns occur within 20 days of each other, then only the first is considered as an SSW.
- 4) Final warmings are the final zero crossings of the winter. Each winter has exactly one final warming.

Figure 1 shows examples of the above four events in terms of an arbitrary segment of a U1060 time series. Of note is the vortex return shown by the blue circle in Fig. 1c, which does not qualify as an SSW since it is preceded by another SSW within less than 20 days (red circle). Our definition of final warmings differs somewhat from that of Black et al. (2006): for consistency reasons, we base all our event definitions on U1060, whereas Black et al. (2006) define final warmings based on the zonal mean zonal wind at 70°N and 50 hPa and the condition that the wind does not exceed  $5 \text{ m s}^{-1}$  after a zero crossing.

Zero crossings are rare events that form the basis for the three other events described above. In Fig. 2 we present an example of model-derived U1060 distributions on two adjacent January days. During this time of the year, the relative frequency of easterly

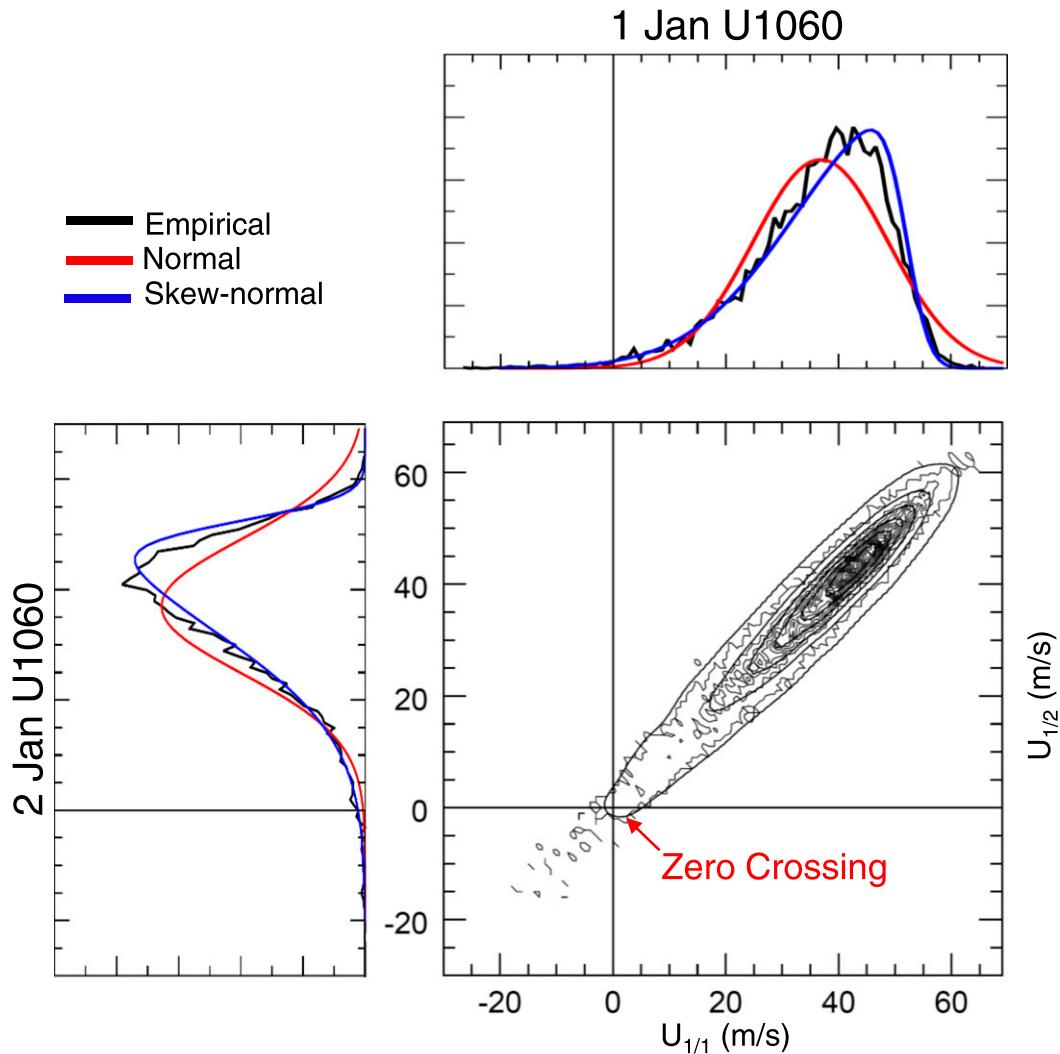


FIG. 2. U1060 distribution on 1 and 2 Jan in CM2.1. Shown are outcomes from the 10 000 samples of the empirical data (black), along with normal (red) and skew-normal (blue) (Azzalini 1985) distributions using parameters calculated from the empirical data. Bin size is  $1 \text{ m s}^{-1}$ . (bottom right) The corresponding two-dimensional distribution, with first contour representing 1 event per 10 000 and additional contours at intervals of 5. The arrow in (bottom right) highlights what we define as zero crossings.

U1060 is small, and the probability of a directional shift in the zonal winds, or zero crossing, is even smaller. In our statistical model, we aim to portray the

probability of such zero crossings  $P(\text{ZC})_n$  on any given day  $n$ . To this end, we integrate a multivariate normal distribution in two dimensions following

$$P(\text{ZC})_n = \int_{z'_2 = -\frac{\mu_n}{\sigma_n}}^{z'_2 = \frac{\mu_n}{\sigma_n}} \int_{z'_1 = -\frac{\mu_{n-1}}{\sigma_{n-1}}}^{z'_1 = \infty} \frac{1}{(2\pi)^{k/2} \sqrt{\det[\mathbf{R}]}} \exp\left[-\frac{\mathbf{z}^T [\mathbf{R}]^{-1} \mathbf{z}}{2}\right] dz_1 dz_2 = P(U_n^- \cup U_{n-1}^+), \quad (1)$$

(adapted from Wilks 2006), where  $\mathbf{R}$  is a  $2 \times 2$  lagged (for days  $n - 1$  and day  $n$ ) correlation matrix consisting of the corresponding  $r$  values,  $\mathbf{z}$  is a vector  $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$  of normalized U1060 for days  $n - 1$  ( $z_1$ ) and  $n$  ( $z_2$ ),  $k = 2$  is

the dimensionality of the problem (i.e., the number of days considered), and the  $\cup$  symbol represents a union of multiple events. The  $-\mu/\sigma$  in terms of normalized  $\mathbf{z}$  constitutes zero in the actual U1060. Thus, the inner integral in (1) represents the probability that U1060 is

positive on day  $n - 1$ , denoted by  $P(U_{n-1}^+)$ . The outer integral represents the probability of U1060 being negative on day  $n$ , denoted by  $P(U_n^-)$ . To calculate multivariate integrals like (1) we use an efficient numerical algorithm (Genz et al. 2004), but this algorithm does not allow incorporating the skewness of the distribution. This means that our statistical model is only controlled by the mean, standard deviation, and autocorrelations. If we were able to take into account the skewness then it would lead to a more accurate representation of the negative tail of U1060 (Fig. 2). However, as we will show below, skewness is small during late winter when most SSWs occur, and the neglect of skewness does not impact our results in major ways.

Finding the probability of a vortex return  $P(\text{VR})_n$  on day  $n$  is based on a similar technique, but we additionally incorporate the requirement of 10 consecutive days of westerly U1060. This can be written as

$$P(\text{VR}^*)_n = [P(\text{ZC})_n \cup U_{n+1}^+ \cup U_{n+2}^+ \dots \cup U_{n+10}^+] \\ + [P(\text{ZC})_n \cup U_{n+1}^- \cup U_{n+2}^+ \dots \cup U_{n+11}^+] \dots \\ + [P(\text{ZC})_n \cup U_{21\text{May}}^- \cup U_{22\text{May}}^+ \dots \cup U_{31\text{May}}^+].$$

Mathematically, this union is resolved by multiple nested integrals similar to (1). However, as defined above,  $P(\text{VR}^*)_n$  also includes the chance that the 10-day criterion is met multiple times after the first zero crossing. As detailed in the appendix, we correct for this chance to arrive at the final probability of a vortex return  $P(\text{VR})_n$ .

We next determine the probability of SSWs, using the chance of a vortex return on day  $n$  and requiring at least 20 days of separation between individual SSW events. To this end we multiply the chance of a vortex return by the sum of chances of an SSW in the previous 20 days under the simplifying assumption that SSWs are independent events following

$$P(\text{SSW})_n = P(\text{VR})_n \left[ 1 - \sum_{i=n-20}^{i=n-1} P(\text{SSW})_i \right].$$

We use a similar methodology to determine the probability of final warmings  $P(\text{FW})$  and integrate (1) requiring that day  $n - 1$  has a positive U1060 while day  $n$  and all subsequent days through 31 May have negative values. This can be written as

$$P(\text{FW}) = P(U_{n-1}^+ \cup U_n^- \cup U_{n+1}^- \dots \cup U_{31\text{May}}^-).$$

As the  $U1060 > 0$  requirement for day  $n - 1$  creates mutually exclusive events, there is no need to adjust these results for multiple events.

We employ Eliassen–Palm (EP) flux analysis in pressure coordinates (Kushner and Polvani 2004) to diagnose the impact of waves on SSWs. The stratospheric wave driving is given by the vertical component of the EP flux vector at 100 hPa, averaged between 20 and 90°N. The negative of this quantity, denoted by  $-F2$ , is similar to the meridional eddy heat flux  $\overline{v^*T^*}$  (Polvani and Waugh 2004). We also calculate the EP flux convergence within a box bounded by 40 and 90°N and 100 and 1 hPa to capture the wave activity in the vicinity of the polar vortex.

### 3. Results

We first examine the relationship between stratospheric wave driving and variability in U1060. Next, we determine how robust the differences in SSW seasonality between our long GCM control integration and the observations are. We then use our statistical model to replicate the finescale temporal evolution of SSWs seen in the GCM. By varying some of the statistical model's input parameters, we investigate the factors most responsible for setting the seasonal structure of SSWs.

#### a. Wave driving and polar vortex variability

Our statistical modeling work is partly based on the variability of the polar vortex, measured by the interannual standard deviation of daily U1060. The black curve in Fig. 3a shows the climatological annual cycle of this quantity. We understand this quantity as a proxy for the intensity of the eddy driving of the polar vortex by upward-propagating Rossby waves from the troposphere. To investigate this conjecture, we first focus on the amount of upward-propagating waves at the 100-hPa level, which is also known as the stratospheric wave driving. As shown in Fig. 3a, wave driving and U1060 standard deviation have fairly similar annual cycles, but there are also differences. These differences may be related to the fact that not all waves entering the stratosphere from below get absorbed in the polar vortex region and contribute to the eddy driving. Some waves escape into the mesosphere, or get refracted away from the vortex region into the tropics. Therefore, a better measure of the eddy driving of the vortex should be the actual convergence of the EP flux in the polar vortex region. This convergence is not only a function of the wave driving from below, but also of the seasonally varying stratospheric background flow and how it refracts and breaks the waves. The EP flux convergence in the vortex region, shown by the blue curve in Fig. 3a, agrees even better with the U1060 standard deviation than the stratospheric wave driving. This confirms our earlier assumption that the variability of the polar vortex is closely

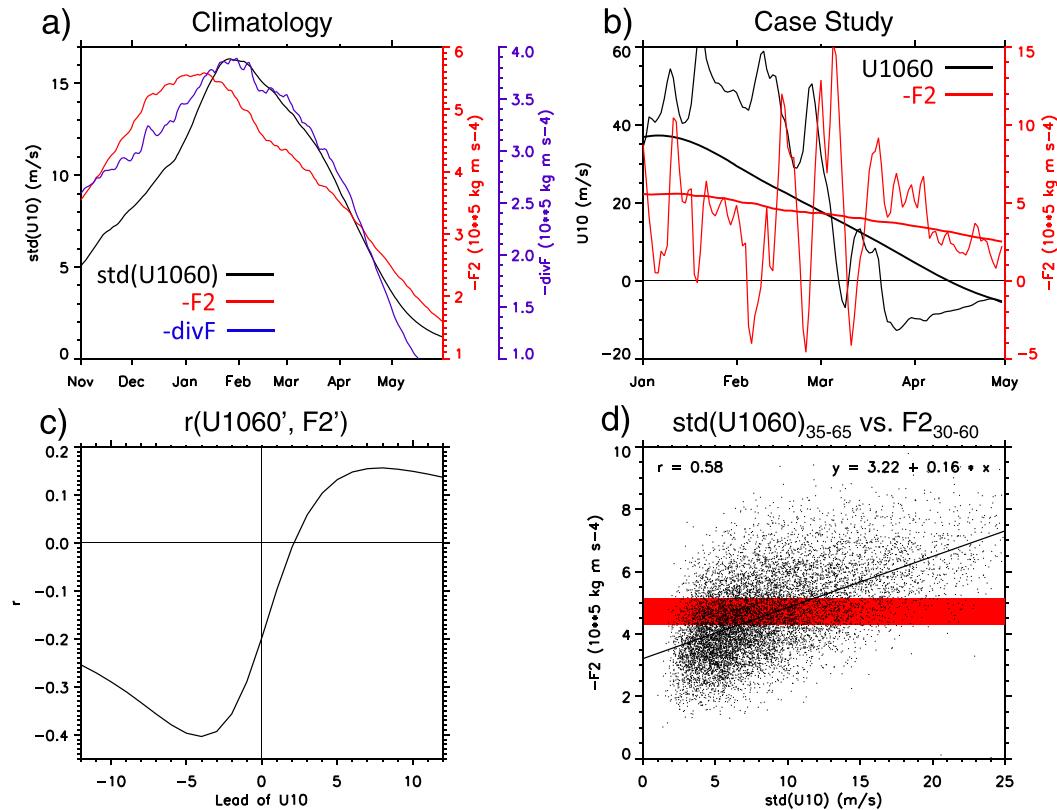


FIG. 3. U1060, wave driving, and EP flux convergence from the CM2.1 model. (a) Seasonal cycle climatology of EP flux convergence (blue;  $10^5 \text{ kg m s}^{-4}$ ), wave driving (red;  $10^5 \text{ kg m s}^{-4}$ ), and U1060 standard deviation (black;  $\text{m s}^{-1}$ ). EP flux convergence is integrated from  $40^\circ$  to  $90^\circ\text{N}$  and 100 to 1 hPa; positive values indicate a net convergence of EP wave activity. Wave driving is the vertical component of the EP flux vectors, measured at 100 hPa and averaged from  $20^\circ$  to  $90^\circ\text{N}$ ; positive values indicate upward propagation. (b) Evolution of wave driving (red) and U1060 (black) for an arbitrary winter; thick lines show climatological mean. (c) Cross correlation between daily anomalies in wave driving and U1060; lag is positive if U1060 leads; the underlying data are very large so that correlations even at the 10% level are highly significant; and positive correlations at positive lags are presumably due to the oscillatory nature of the two quantities. (d) Scatterplot between yearly values of wave driving, averaged from day 30 to 60, and daily standard deviation of U1060 between day 35 and 65 of the same year; each dot represents outcome from one year; outcomes are very similar for other day ranges; and the horizontal red band indicates climatological seasonal range of wave driving during selected period.

related to the eddy driving and that the two are roughly proportional to each other. There are some discrepancies during early winter (November–December). During this time, the standard deviation of U1060 somewhat underestimates the eddy convergence.

We next turn our attention again to the lower-stratospheric wave driving, as this quantity is also related to vortex variability and more commonly used in the literature. We investigate how the wave driving and the polar vortex are related on shorter, intra-seasonal time scales. Figure 3b shows the evolution of both during an arbitrary winter. The chosen winter contains one SSW and is associated with reduced (compared to the climatology) wave driving during January and early February and increased wave driving thereafter. The two curves are anticorrelated

at a few days lag. This becomes clearer when all years and all days of the year are included to calculate the cross correlation between U1060 and wave driving (Fig. 3c). The anticorrelation maximizes when the wave driving leads U1060 by about four days. This is apparently the time for the waves to propagate from the lower to the middle stratosphere and undergo wave–mean flow interaction.

The scatterplot in Fig. 3d compares the February-averaged (days 30–60) wave driving with the daily U1060 standard deviation during the same period at a 5-day lag. Each dot represents the outcome from one year. As before for the climatological seasonal cycle (Fig. 3a), stronger wave driving is associated with increased U1060 variability. The results are similar when different periods are selected (e.g., January or March). The

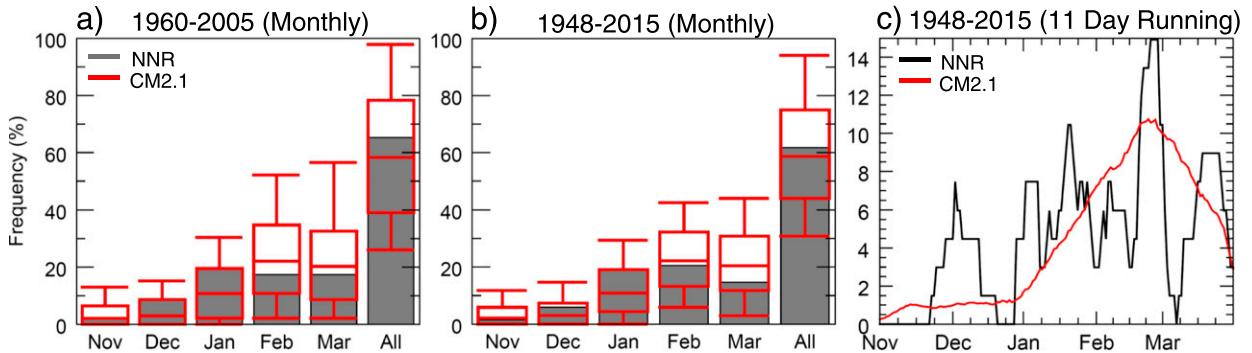


FIG. 4. SSW seasonality. Shown is the SSW frequency (events per period in %) from reanalysis (gray or black) and the CM2.1 GCM (red). (a),(b) Red symbols indicate mean, minimum, maximum and 95% confidence interval, derived from bootstrapping. (c) SSW distribution based on an 11-day-long sliding interval.

relatively good relationship between wave driving and vortex variability has a simple explanation: absence of substantial wave driving creates a cold and stable vortex and hence reduced variability; periods of increased wave driving, however, tend to be associated with bursts of wave energy (Fig. 3b), leading to strong perturbations of U1060. In summary, the analysis of Fig. 3 demonstrates a fairly good relationship between lower-stratospheric wave driving, eddy driving of the polar vortex, and interannual variability of daily U1060.

#### b. SSW seasonality

In evaluating the month-to-month SSW distribution, we note that the reanalysis period is relatively short, resulting in a very modest number of events (10 or fewer) during individual months. This introduces large uncertainty in the seasonal distribution of SSWs, which in the following we attempt to quantify. We begin by comparing the month-to-month distribution of SSWs seen in the observations and the GCM by focusing on the same  $n = 45$  winter-long period (1960–2005) (Fig. 4a) as in Charlton-Perez et al. (2008). As indicated by the gray bars, January has the most number of observed SSWs. In contrast, the GCM (red symbols) has most SSWs in February, a situation representative for many other climate models. However, when considering the longer  $n = 67$  winter period (1948–2015) (Fig. 4b), the observations also exhibit a February SSW maximum, just like the GCM. This nicely demonstrates the sensitivity of the observed distribution to the period of the underlying data. To better quantify the corresponding uncertainty, we use bootstrapping (without replacement) of the GCM data to generate a large number (10 000) of synthetic seasonal SSW distributions, each based on the same number of winters ( $n$ ) as the observations but using randomly selected winters [November–March (NDJFM)] from the long GCM

run. The outcome from the bootstrapping in terms of mean, extreme values and 95% confidence interval is shown by the red symbols in Figs. 4a and 4b. The monthly number of SSWs seen in the observations is in each case within the 95% confidence interval of the GCM produced SSWs, indicating that one cannot reject the null hypothesis that GCM and observed SSWs stem from the same distribution. In other words, the two distributions cannot be distinguished from each other, unless the error rate is increased above 5%. We note, however, that the difference between GCM and observations during January cannot be fully ignored either; the observed number of SSWs is at the upper end of the GCM generated 95% limit, hinting at the small chance of a certain inconsistency between the two datasets.

We further investigate the SSW seasonality and reduce our bin size from the monthly time period to a running 11-day average (Fig. 4c). Interestingly, now there is a relatively good visual agreement between the two distributions, with a maximum at the end of February in both datasets. This suggests that the January maximum seen in the NNR (Fig. 4a) is partially related to the fixed dates of each month in combination with large SSW variability within each month. It should also be noted that January contains more days than February, leading to more January SSWs for artificial reasons. The difference between the two datasets stems from both the GCM showing fewer SSWs in December and January and the observations having a relative minimum in early February.

Getting back to our question about the relative importance of decreasing polar vortex strength and decreasing stratospheric wave driving in controlling the seasonality of SSWs, we note that the late-February SSW maximum in the model is long after the early-January maximum in vortex strength and wave driving. It suggests

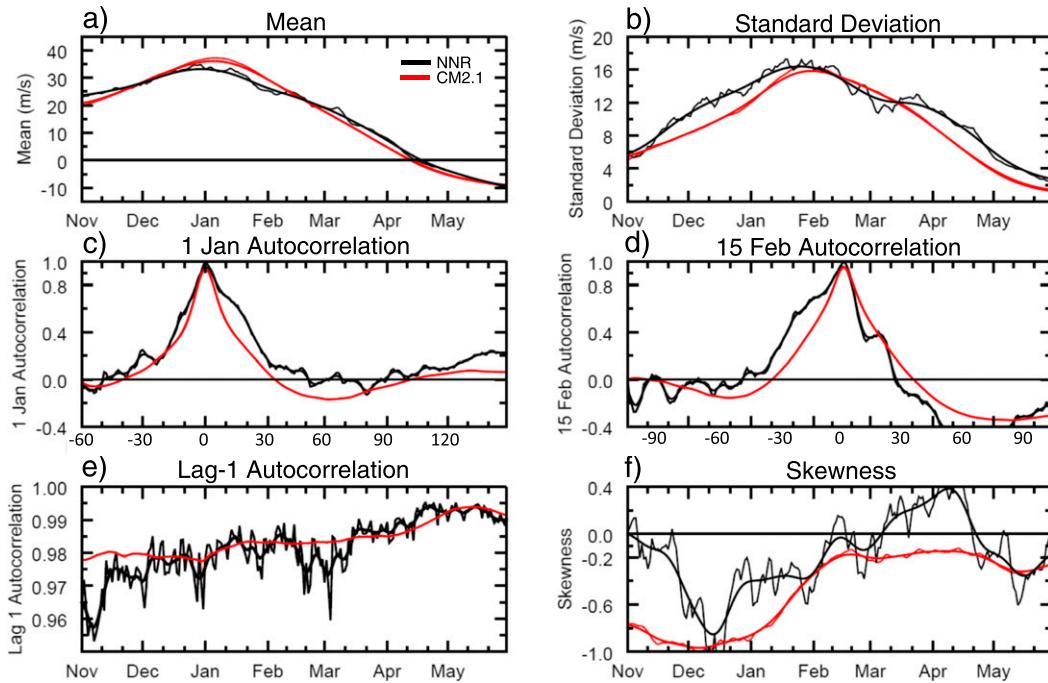


FIG. 5. U1060 input parameters. Tick lines are smoothed (using a Gaussian with a kernel of three standard deviations) and represent the actual input data for the statistical model.

that the decreasing polar vortex strength is the dominant factor for the occurrence of SSWs. This late-February SSW maximum may be during the time when the climatological vortex is weak enough to allow for a large number of zero crossings but also during the time when the vortex is still strong enough to create a decent chance for a vortex return, thus generating an SSW. Better understanding this idea motivated us to construct our statistical model, which determines the daily climatological probability of SSWs using only the statistical properties of the polar vortex strength.

### c. Input parameters to the statistical model

We now discuss the temporal evolution of the moments and correlations of U1060 (Fig. 5), which were empirically derived from GCM and observations. These values drive our statistical model, and we refer to them simply as input parameters. The data from the reanalysis (black) and the GCM (red) are similar. The mean U1060 (Fig. 5a) maximizes at the beginning of January and reaches zero at about mid-April. The U1060 standard deviation (Fig. 5b) maximizes at the end of January and then tends to decrease to small values in summer. The autocorrelation of U1060 is shown exemplarily for two different dates and various lags (Figs. 5c and 5d). Autocorrelation is important because it directly enters our statistical model through

(1), and physically it can be interpreted as the stiffness or persistence of the polar vortex. In other words, higher autocorrelation means it is increasingly difficult to disturb the vortex. There are some minor differences between autocorrelation of GCM and observations at lags of 10–20 days, and both datasets show an interesting weak anticorrelation at lags of 60–75 days. This anticorrelation reflects the low-frequency variability of the stratospheric circulation (Christiansen 2000), which includes warm (SSW like) and also cold vortex states. For example, we find a systematic delay of the final warming date in winters with a prior SSW. The temporal evolution of the persistence of U1060 is indicated by the lag-1 autocorrelation (Fig. 5e). This quantity is key for our calculation of zero crossings, and it increases more or less monotonically from winter to spring. We note that the temporal evolution of correlations at higher lags (not shown) is similar to that of lag 1.

Although not included in our statistical model, the skewness of the U1060 distribution (Fig. 5f) provides insight into potential errors caused by our assumption of unskewed normality. The skewness in the GCM is generally negative, implying that the U1060 distributions have a long negative tail and that our assumption of a standard normal distribution underestimates the number of negative U1060 values. However, during late winter and early spring, the skewness is only moderately

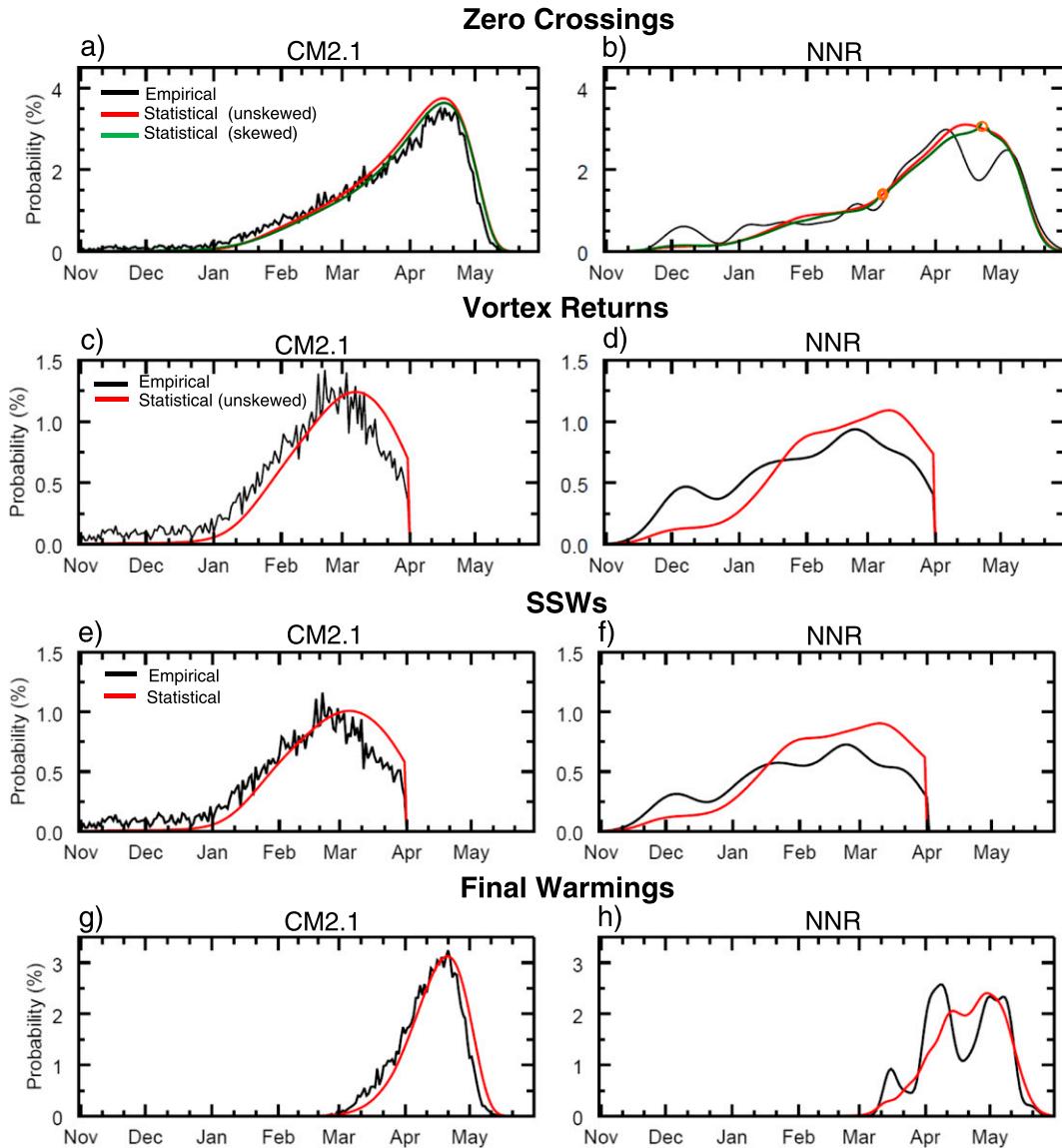


FIG. 6. Seasonal evolution of event probabilities ( $\% \text{ day}^{-1}$ ). Shown are empirical data (black), those based on the statistical model (red; nonskewness assumption), and those from the statistical model using a bivariate skew-normal distribution (green). The empirical NNR results (black lines) are smoothed using a Gaussian smoother with a kernel of three standard deviations.

negative, and this is the time when most of the SSWs occur.

#### d. Zero crossings

Zero crossings of U1060 are the most basic of the four events described above. The distributions of the empirically derived zero crossings (black) from GCM and reanalysis are very similar (Figs. 6a and 6b), with low probabilities in early winter, a gradual increase starting after the polar vortex strength has reached its maximum strength, and a steep decrease in May. Zero crossings maximize in mid-April, the time when the mean wind

changes from westerlies to easterlies and the probability of positive or negative U1060 is near equal. In May, zero crossings rapidly decrease, due primarily to the large persistence of U1060 at this time.

Our statistical model calculates zero crossings by integrating a bivariate normal distribution, giving the probability that U1060 on day  $n - 1$  is positive and on day  $n$  it is negative. The model (red) replicates the above-described empirically derived results well, including the more modest increase of NNR probabilities between January and March as compared to the GCM. There are also some minor but systematic flaws: for both

datasets, the statistical approach underpredicts zero crossing in early to midwinter, while there is a slight overestimate in April.

The underestimate is only marginally related to not incorporating skewness in our statistical model. This is demonstrated by the green curve in Fig. 6a, which was derived from calculations that consider skewness using a bivariate skew-normal distribution (Azzalini and Dalla Valle 1996) and which only led to minor improvements. A better explanation for the underestimate during early to midwinter comes from Fig. 3a. There, it is shown that during November and December the standard deviation of U1060, which is an important input parameter to our model, does not describe very well the amount of lower-stratospheric wave driving and wave convergence in the polar vortex region. Hence, U1060 somewhat underestimates the true variability of the polar vortex during this time of the year.

The overestimate of the red curve at the April maximum, though small, could potentially be due to the fact that the correlation is calculated from many years and that it is therefore influenced by a combination of two different atmospheric states: one in which the polar vortex still exists (leading to lower correlations) and one in which the polar vortex is already broken down (leading to higher correlations). Nonlinearities in calculating correlations, which become more important as correlation approaches unity, may lead to the seen discrepancies.

#### e. Vortex returns

The black curves in Figs. 6c and 6d show the probability of empirically derived vortex returns, which are zero crossings with the additional constraint that U1060 becomes positive for 10 consecutive days anytime afterward. From November through mid-February, the number of zero crossings and vortex returns is almost identical. This is because the stratosphere strongly favors westerlies at that time, and 10 consecutive days of westerlies are almost guaranteed to happen. In mid-February, however, this is no longer true, and the probability of vortex returns increases more gradually than that of zero crossings. The vortex return probability maximizes in late February, as the requirement for 10 consecutive days of positive U1060 is increasingly more difficult to fulfill and begins to outweigh the still increasing chance of a zero crossing. The drop on 1 April is caused by the somewhat arbitrary definition (CP07).

Our statistical model of vortex returns is based on a multivariate normal distribution, which conditions the likelihood of a zero crossing on a subsequent 10-day-long period of westerlies (see methods). The statistical model mirrors the empirical probabilities of zero

crossings in early winter well, decreasing in slope once the probability of 10 consecutive positive days becomes less than one and reaching a maximum while the probability of a zero crossing is still increasing. For NNR, the statistical model also finds larger probabilities for a vortex return than for the GCM, which agrees well with the empirical results. Thus, subtle differences in the input parameters for our model are important and lead to meaningful differences between the two datasets.

#### f. SSWs

An SSW is a vortex return separated from a previous SSW by at least 20 days. The outcome of the statistical modeling of the evolution of SSW probabilities is presented in Figs. 6e and 6f. The empirical outcomes for the observations (Fig. 6f) are similar to that shown in Fig. 4c, except for the differences in the smoothing. Empirically (black curves), early in winter there is little difference between SSWs and vortex returns. A zero crossing (and therefore vortex return and SSW) is a rare enough event in November and December that the likelihood of two occurring during the same period is very small. By late January in the observations, and early February in the GCM, the 20 days of separation requirement reduces the slope of the SSW probability curve. This reduces the  $\sim 1.2\% \text{ day}^{-1}$  maximum in vortex returns at the end of February (Fig. 6c) to  $\sim 0.8\%$  per day for SSWs (Fig. 6e).

In our statistical model, we calculate SSW probabilities by multiplying the probability of a vortex return on any given day by the chance that there were no SSWs in the 20 days prior. The red curves in Figs. 6e and 6f show the outcomes. The curves for the observations and the model are similar and both maximize at the beginning of March. This reinforces our earlier finding that the observed January maximum in terms of absolute SSW counts (Fig. 4a) must be interpreted with care owing to the small sample size. It suggests that, with more observations, the date of maximum SSWs would be shifted toward late winter as in the GCM. In terms of the annual SSW frequency, the statistical model produces  $0.59 \text{ SSW yr}^{-1}$  for the GCM and  $0.70 \text{ SSW yr}^{-1}$  for the NNR. This compares quite favorably against the empirically derived annual frequencies of 0.61 and 0.64 for GCM and NNR, respectively. For GCM and NNR, the maximum daily probability is  $\sim 1\%$  and occurs about 10 days after the empirical maximum, and the small underestimate in early winter is balanced by a similar small overestimate in late winter.

We also tested the sensitivity of empirical and statistical SSW frequencies to varying the number of days of separation (not shown). As expected, increasing the number of days of separation (NDSEP) makes the occurrence of multiple SSWs in the same winter less likely,

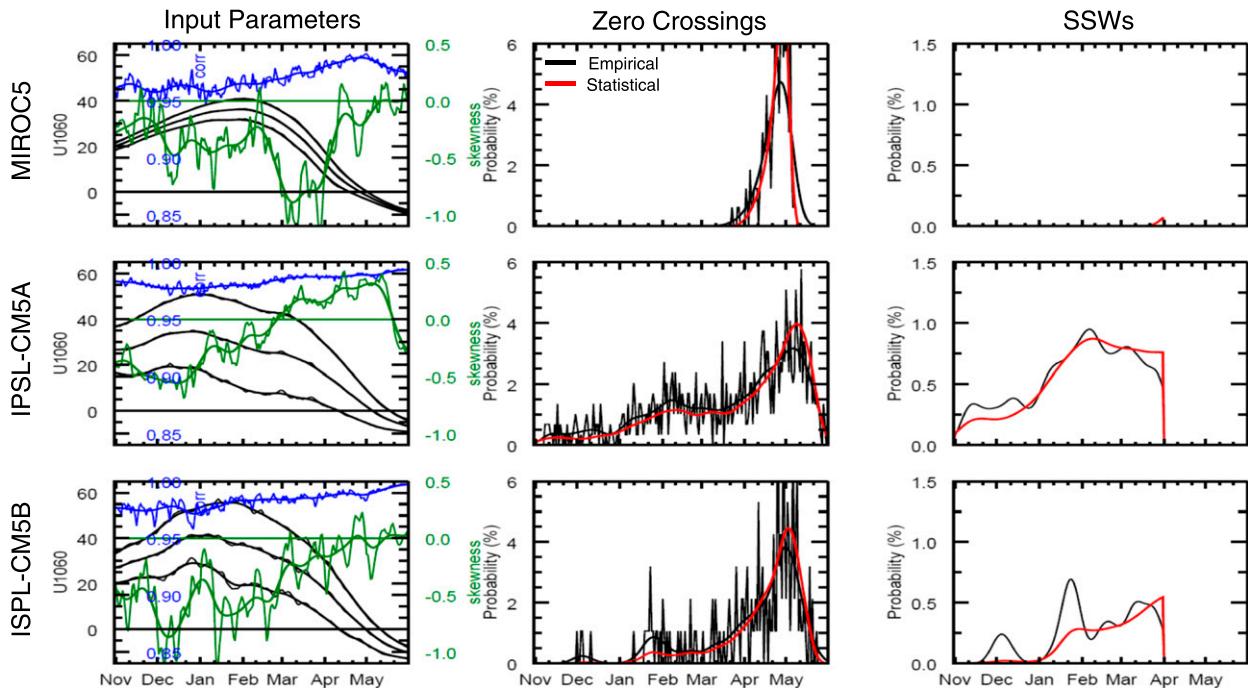


FIG. 7. Additional model results. The models are MIROC5 (historical run 2; 1850–2012), IPSL-CM5A (medium resolution; RCP4.5; 2006–2300), and IPSL-CM5B (low resolution; RCP8.5; 2006–2100). (left) Input parameters show the mean of U1060 plus or minus one standard deviation (black), the lag-1 correlation (blue), and the skewness (green).

reducing the number of late-winter SSWs (in particular March SSWs), reducing the number of overall SSWs, and increasing the relative frequency of midwinter SSWs. For example, increasing NDSEP from 20 to 40 days reduces the daily SSW frequency at the beginning of March from 0.93% to 0.85%.

#### g. Final warmings

Final warmings (Figs. 6g and 6h) are the last zero crossing of U1060 in winter. The area under the empirical (black) and statistical (red) final warming curves are by definition unity, as there is exactly one final zero crossing in every winter. In GCM and NNR, empirical final warmings appear at the beginning of March and maximize in early to mid-April, when mean U1060 crosses zero. Also, final warmings decrease sharply in May, when the probability of westerlies in U1060 becomes near zero.

In March, the near-correct estimate of zero crossings (Fig. 6a) and the overestimate of vortex returns (Fig. 6c) and SSWs (Fig. 6e) implies that the statistical model overestimates the chance of positive U1060 after a zero crossing. Thus, we expect our statistical model to produce final warmings later in the year than the empirical data. With the GCM input parameters, our statistical model produces a curve similar to the empirical data, but as expected, shifted late by approximately three days.

This shift is maintained through May because final warmings that do not occur earlier in the year must occur later.

#### h. Application to other models

To test whether the success of our statistical model still holds when applied to other datasets, we next use input parameters from select CMIP5 models. The particular models we choose (MIROC5, IPSL-CM5A, and IPSL-CM5B) are extreme outliers in terms of their SSW seasonality and annual frequency (Fig. 7) and therefore provide a good test case for our statistical model. The question is, is the unusual SSW distribution in these models reflected in changes to the input parameters for our statistical model, and can the statistical model therefore replicate the empirical results?

The SSW distribution of MIROC5 (Watanabe et al. 2010), a model that we choose because of its lack of empirical SSWs, is well reproduced by our statistical model. It produces no SSWs throughout the vast majority of winter, with a very small chance in late March. Also of note, zero crossings are narrowly clustered around mid-April, which is again well represented by our statistical model. From a brief analysis of input parameters it is clear that the mean and correlations of MIROC5 are similar to that of NNR and CM2.1 but that the standard deviation is much too small. Thus, in this

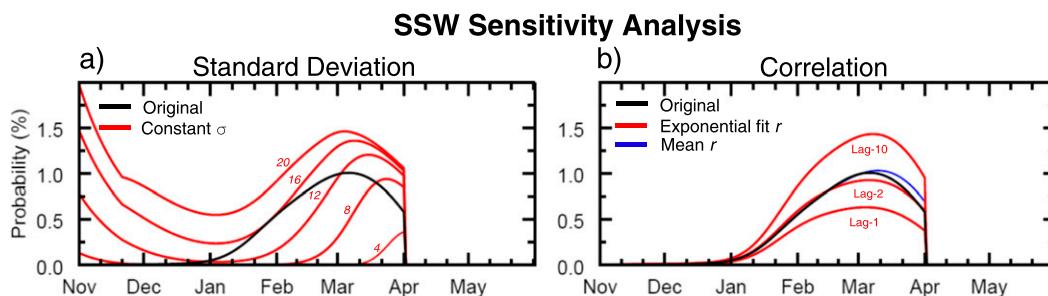


FIG. 8. Sensitivity analysis for CM2.1. Shown are the original results from our statistical model (black), and (left) the result when standard deviation is set to a constant (red; in  $\text{m s}^{-1}$ ) throughout the year. (right) Results from using constant lagged correlations when all lagged correlations are set to their mean using all days of the year (blue) and correlation is approximated by a least squares fit considering only days 0–1 (lag 1), days 0–2 (lag 2), and days 0–10 (lag 10) (red). Lag 1 overestimates and lag 10 underestimates the actual autocorrelations.

case the lack of variability in polar vortex strength directly relates to the lack of SSWs.

The opposite is true in IPSL-CM5A (Dufresne et al. 2013), which we choose because of its overall large number of SSWs compared to most other CMIP5 models. Our statistical model again represents the empirical number and seasonal distribution of SSWs and zero crossings well.

Finally, we examine IPSL-CM5B (Dufresne et al. 2013), which has a rare empirical January SSW maximum, as opposed to the usual February maximum in most other models. Again, the zero crossing calculations fit well, but the statistical model underestimates the actual number of SSWs in January. During that month, the large number of SSWs may be explained by the large U1060 standard deviation, as the mean U1060 of this model is higher than in any other dataset. Sampling uncertainty may also play a role, as this dataset contains only 95 years. Overall, the analysis of the three additional models suggests that the evolution of the U1060 standard deviation and thus of the strength of the stratospheric wave driving is also a strong predictor for the occurrence of SSWs.

#### i. Sensitivity analyses

The skewness in U1060 in early winter (Fig. 5f) indicates a limitation of our assumption of an unskewed normal distribution. This assumption results in an underestimate of events with  $U1060 < 0$  compared to the negatively skewed empirical distribution. Although the numerical algorithm we use to calculate integrals of multivariate normal distributions (Genz et al. 2004) does not incorporate skewness, we attempt to account for this underestimate by simply shifting the mean of our daily unskewed normal distributions toward smaller U1060. By doing so, we accurately represent the percent of values of U1060 below zero for any given day in our empirical dataset. We then use the same correlation and

standard deviation as before for our input parameters to refine our calculation of SSW probability. The new calculations (not shown) yield results that align very well with the distribution of SSWs in the GCM until late February, but the overestimate during March (Fig. 6e) still remains. This suggests that the neglect of skewness is an important cause for the issues in early winter. For the persistent overestimate in March, however, other reasons must be sought. We suspect that different autocorrelations between the winter and summer regimes of the polar vortex and that the input parameters being calculated from a combination of both are responsible.

In an attempt to better understand the interplay between polar vortex strength and stratospheric wave driving and its role in setting the timing of SSWs, we next ask how sensitive the statistical model is to variations in the prescribed seasonal changes in mean, standard deviation, and autocorrelation of U1060. The temporal evolution of U1060 is important because zero crossings and thus SSWs are more likely when the mean increases over time than when it decreases. Along with a higher mean (Fig. 5a) and lower standard deviation (Fig. 5b), this explains why November and December have far fewer zero crossings and SSWs than February and March.

To investigate the influence of standard deviation and hence wave driving alone, we rerun our statistical model using the original GCM-derived input parameters for mean and correlation but use constant standard deviations ranging from 4 to  $20 \text{ m s}^{-1}$  (Fig. 8a). As expected, the probability of SSWs decreases with decreasing standard deviation. For example, a constant standard deviation of  $4 \text{ m s}^{-1}$  results in SSWs that exclusively occur in March, a situation similar to that of the MIROC5 model (Fig. 7). An increasing standard deviation not only increases the SSW probability during all days, but it also shifts the late winter maximum to earlier dates. Of note is that the minimum probability of SSWs is always at the January maximum of mean U1060.

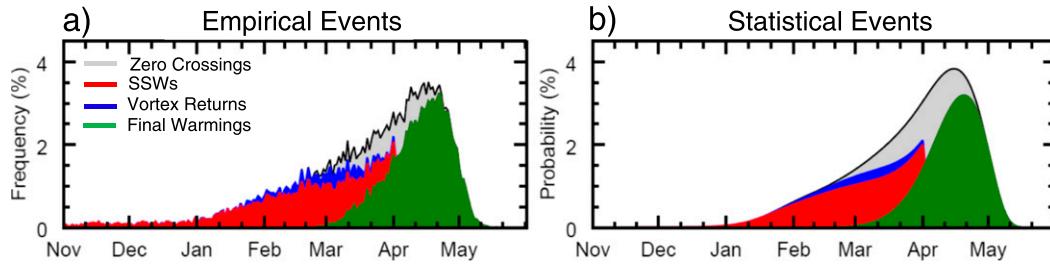


FIG. 9. CM2.1 event statistics.

We next explore the impact of variations in the autocorrelation of U1060. To this end, we modify our original correlation matrix  $r(t, \tau)$  to become the same for all dates  $t$  and therefore simply  $r(\tau)$ . Using the annual mean correlation at each lag (Fig. 8b, blue curve) has very little impact on the outcome. We then make an exponential least squares fit of  $r(\tau)$  to the previously determined mean lagged correlations. Because we suspect that the outcomes from our statistical model are dominated by the large correlations at short lags, we investigate an increasing number of lags (lag 1, lag 2, lag 10) in fitting the annual mean autocorrelation function, with each fit starting at lag 0. The overall shapes of all three correlation-modified model outcomes (Fig. 8b, red curves) are similar and maximize SSWs on the same day as our original model. The lag-10 fit underestimates the actual autocorrelation at short lags, leading our statistical model to overestimate the number of zero crossings and the overall probability of SSWs. Lag 1 leads to the opposite. Using lag 2 most closely resembles our original statistical model, confirming our original assumption that our model is most sensitive to autocorrelation at short lags. Thus, autocorrelation only impacts the magnitude of SSW probability and the annual frequency of SSWs, but it does not impact the seasonal distribution.

#### 4. Conclusions

The seasonal distribution of SSWs is investigated, motivated by the fact that observations have an earlier date (January) in the maximum number of SSWs than most climate models (February–March). We argue that SSWs in the observations are undersampled and that the observed seasonal distribution of SSWs is statistically nonrobust. We resample a long climate model simulation many times to mimic the observations and find that the modeled and observed SSW distributions are statistically indistinguishable. There is also evidence for an end-of-February SSW maximum in the observations, just like in the model, when the number of days over which averages are taken are less than a month. We thus conclude that the best theoretical estimate for the time

of the maximum number of SSWs is at the end of February. We expect that in the future, when the observed SSW sample becomes sufficiently large, SSWs will occur more often in late winter than in midwinter, causing a shift in the seasonal SSW distribution in observations. This also suggests that the observed seasonal cycle of SSWs cannot be used to constrain models and their stratospheric circulation performance.

This work also tries to shed light on the factors that control the seasonality of SSWs. We hypothesize that between January and March the seasonality is largely determined by the interplay between the decreasing strength of the polar vortex, favoring SSWs in late winter or early spring, and the decreasing strength of the eddy driving of the vortex, favoring more midwinter SSWs. To investigate this idea, we construct a simple model that is based on input from the climatological statistics of the stratospheric polar vortex winds and the assumption that the winds follow a multivariate normal distribution. The climatological input parameters are the 1) strength, 2) variability, and 3) autocorrelation of the zonal mean zonal wind at 10 hPa and 60°N (U1060). The statistical model applies the well-known “WMO” criterion to these inputs to calculate the daily varying probability of SSWs and SSW-related events (zero crossings, vortex returns, and final warmings). Some information about actual SSWs is indirectly contained in our statistical model, as we derive the input climatologies from all years, including years with SSWs.

Overall, the statistical model closely reproduces the actual data of our long climate model simulation (Fig. 9). Most importantly, just like the actual climate model data, the statistical model points to a late-winter SSW maximum. Driving the same statistical model with observations also leads to a late-winter maximum, confirming that the observed midwinter SSW maximum is mainly due to the small number of observed SSWs and the resulting sampling uncertainty. The success of the statistical model also verifies our hypothesis that the seasonality of SSWs can be largely explained from the seasonal variations of the polar vortex strength and stratospheric wave driving.

Our analysis of the statistical model and its sensitivity to variations in input parameters leads to the following conclusions:

- 1) The seasonally varying strength of the polar vortex is most important for the seasonality of SSWs. It determines how much energy is required to produce a wind zero crossing. A weak vortex is more likely to lead to an SSW than a strong vortex; hence SSWs favor late winter. However, toward early spring, when the vortex becomes very weak, it is increasingly difficult to fulfill the 10-day rule for a vortex return, which then again decreases the probability for SSWs. The temporal evolution of the vortex strength is also important: it is positive during November and December, decreasing the probability of a zero crossing. Our results agree well with Jucker et al. (2014), who show that the overall frequency of SSWs in a model is highly anticorrelated with the climatological strength of the polar vortex.
- 2) The seasonally varying effects of breaking waves, indirectly included in our statistical model by the U1060 standard deviation, represents the available energy to perturb the vortex. The U1060 standard deviation is a good proxy for the net convergence of eddy activity, which includes effects from seasonal variations in the wave driving at the lower stratosphere and from the refraction of the waves by the background flow. Empirically, eddy activity maximizes during January (Fig. 3a), a time during which the vortex is still too strong to produce many SSWs. Heading toward spring, the eddy activity decreases, but at a rate slower than the decreasing strength of the polar vortex.
- 3) The late-winter conditions are optimal for SSWs. Then, there is still ample supply of wave energy, but the energy needed for a vortex breakdown is relatively small due to the rapidly decreasing strength of the winds.
- 4) From the fact that SSWs reach their maximum frequency at the end of February, which is long after the January maximum in planetary wave activity, it appears that wave driving is less important for SSWs than the vortex strength. This provides some evidence for the ideas of Jucker (2016) that stratospheric wave driving is not a major limiting factor for SSWs and that SSWs are not necessarily related to bursts of upward wave fluxes from the troposphere.
- 5) The autocorrelation of the stratospheric winds can be interpreted as the “stiffness” of the polar vortex and thus also determines the amount of energy needed for a vortex breakdown. The seasonal variations of autocorrelation are generally small and so is their influence on SSWs. Autocorrelation has little impact

on the shape of the SSW distribution and the date of the SSW maximum; it only affects the overall annual frequency of SSWs.

Based on these findings, it now becomes clearer why early November experiences so few SSWs, despite a vortex that is similar in strength to that at the end of February, the time of maximum SSWs (Fig. 5a). One plausible explanation is differences in wave energy available for a vortex breakdown. From Fig. 3, the lower-stratospheric wave energy is 20% larger at the end of February ( $-4.5 \times 10^5 \text{ kg m s}^{-4}$ ) than at the beginning of November ( $-3.7 \times 10^5 \text{ kg m s}^{-4}$ ). The remaining differences must be related to seasonal changes in the refractive properties of the background flow, leading to 40% more eddy convergence in February ( $3.5 \text{ kg m s}^{-4}$ ) than in November ( $2.5 \text{ kg m s}^{-4}$ ). This and nonlinearities in the dynamics then lead to the almost 300% differences in the standard deviation of U1060 (Fig. 3) between the end of February ( $\sigma \sim 14 \text{ m s}^{-1}$ ) and the beginning of November ( $\sigma \sim 5 \text{ m s}^{-1}$ ). The gradual strengthening of the polar vortex during November and December is another contributing factor, as it further reduces the likelihood for zero crossings. Variations in autocorrelations play only a minor role.

SSWs are rare and extreme circulation events, originating from nonlinear wave–mean flow interaction. Based on these complicated dynamics, it is surprising how well our statistical model, with its simple assumption of unskewed normality, reproduces the features seen in the actual data. It indicates that there is nothing too special about SSWs and that they simply form the tail of a rather continuous distribution of varying polar vortex strengths. This is similar to the ideas of Coughlin and Gray (2009), who find a certain degree of SSW continuity as long as the polar vortex is within its warm state. However, the same authors also hypothesize that the polar stratosphere can be described by two well-separated states, contradicting to some extent our findings. Our idea of continuity is also somewhat at odds with the widely used definition for SSWs (CP07), which is based on the zero crossing of the zonal mean zonal wind. From a dynamical perspective, the zero-wind line is highly significant as it forms the critical level for upward-propagating planetary waves (Charney and Drazin 1961) and presumably marks the level for further wave breaking below. However, our results indicate that, from a statistical perspective, the zero-wind threshold is arbitrary and that several different thresholds could be used to define extreme, SSW-like events.

*Acknowledgments.* We thank the two anonymous reviewers and Martin Jucker for their constructive

comments. Part of this work was funded by the National Science Foundation under Grant 1446292. We are grateful for support by the Center for High Performance Computing at the University of Utah.

## APPENDIX

### Vortex Return Calculations

Our vortex return calculations are based on a series of nested integrals. The equation integrated is the multivariate normal distribution [(1)]; however, the dimensionality of the problem changes depending on the start date of positive U1060. To represent days of positive (negative) U1060, the lower limit of integration in (1) is chosen to be  $-\mu_k/\sigma_k(-\infty)$ , where  $k$  is the day of interest, while the upper limit is  $\infty(-\mu_k/\sigma_k)$ . A vortex return on day  $n$  requires 10 consecutive days of positive U1060 on any day after a zero crossing. We begin by calculating  $P(\text{VR}^*)$ , the chance of having 10 consecutive days of positive U1060 on any day after the initial zero crossing (day  $n$ ), up to 22 May. We use 22 May as end date as 10 days of positive U1060 starting on that date would end on 31 May.

The calculation of  $P(\text{VR}^*)$  changes slightly depending on how many days after day  $n$  the period of 10 consecutive days of positive U1060 begins. To represent the zero-crossing returning on day  $n + 1$ , we use a 12-dimensional integral, requiring that U1060 is positive

on day  $n - 1$ , negative on day  $n$ , and positive on days  $n + 1$  through  $n + 10$ . For the 10 consecutive days of positive U1060 from day  $n + 2$  to  $n + 11$ , we must also consider that day  $n + 1$  has a negative U1060. Therefore, we use a 13-dimensional integral with a positive U1060 on day  $n - 1$ . We use a similar technique for the 10 consecutive days of positive U1060 starting at later days ( $n + 3 \dots$ ), while maintaining the day before the 10 positive values as negative and ignoring the days in between.

Beginning with day  $n + 12$ , we must consider  $P(\text{VR}^\#)$ , the probability that there have already been prior 10 consecutive days of positive U1060. We begin with the same 13-dimensional integral as before to determine the probability for this to happen between days  $n + 2$  and  $n + 11$ . We then subtract a 22-dimensional integral, which calculates the probability that day  $n - 1$  had a positive U1060, day  $n$  a negative U1060, day  $n + 1$  through  $n + 10$  a negative U1060, the day before these 10 consecutive days had a negative U1060, and the 10 days of interest again showed a positive value. Starting on day  $n + 13$  (the 23 dimensional integrals are unnecessary on day  $n + 12$ ) through 22 May, we also subtract as many 23 dimensional integrals as necessary, implying that the first 10 consecutive days of positive U1060 happened after day  $n + 1$  but before the day of interest. Using the notation described in the methodology section, the chance that there are multiple 10-day periods with positive U1060 after a day of negative U1060 before day  $n + 31$  can be written as

$$P(\text{VR}^\#)_{n+31} = [P(\text{ZC})_n \cup U_{n+1}^+ \cup U_{n+2}^+ \dots \cup U_{n+10}^+ \cup U_{n+30}^- \cup U_{n+31}^+ \dots \cup U_{n+40}^+] \\ + [P(\text{ZC})_n \cup U_{n+1}^- \cup U_{n+2}^+ \dots \cup U_{n+11}^+ \cup U_{n+30}^- \cup U_{n+31}^+ \dots \cup U_{n+40}^+] \dots \\ + [P(\text{ZC})_n \cup U_{n+19}^- \cup U_{n+20}^+ \dots \cup U_{n+29}^+ \cup U_{n+30}^- \cup U_{n+31}^+ \dots \cup U_{n+40}^+].$$

We then find the overall probability of a vortex return  $P(\text{VR})_n$  by removing the sum of these probabilities from  $P(\text{VR}^*)$ , following

$$P(\text{VR})_n = P(\text{VR}^*)_n - P(\text{VR}^\#)_{n+12} - P(\text{VR}^\#)_{n+13} \dots \\ - P(\text{VR}^\#)_{22\text{May}}.$$

## REFERENCES

- Albers, J., and T. Birner, 2014: Vortex preconditioning due to planetary and gravity waves prior to sudden stratospheric warmings. *J. Atmos. Sci.*, **71**, 4028–4054, doi:10.1175/JAS-D-14-0026.1.
- Azzalini, A., 1985: A class of distributions which includes the normal ones. *Scand. J. Stat.*, **12**, 171–178, <http://www.jstor.org/stable/4615982>.
- , and A. Dalla Valle, 1996: The multivariate skew-normal distribution. *Biometrika*, **83**, 715–726, doi:10.1093/biomet/83.4.715.
- Black, R., B. McDaniel, and W. Robinson, 2006: Stratosphere–troposphere coupling during spring onset. *J. Climate*, **19**, 4891–4901, doi:10.1175/JCLI3907.1.
- Butchart, N., and Coauthors, 2011: Multimodel climate and the variability of the stratosphere. *J. Geophys. Res.*, **116**, D05102, doi:10.1029/2010JD014995.
- Butler, A., D. Seidel, S. Hardiman, N. Butchart, T. Birner, and A. Match, 2015: Defining sudden stratospheric warmings. *Bull. Amer. Meteor. Soc.*, **96**, 1913–1928, doi:10.1175/BAMS-D-13-00173.1.
- Charlton, A., and L. Polvani, 2007: A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks. *J. Climate*, **20**, 449–469, doi:10.1175/JCLI3996.1.
- , and Coauthors, 2007: A new look at stratospheric sudden warmings. Part II: Evaluation of numerical model simulations. *J. Climate*, **20**, 470–488, doi:10.1175/JCLI3994.1.
- Charlton-Perez, A., L. Polvani, J. Austin, and F. Li, 2008: The frequency and dynamics of stratospheric sudden warmings in the 21st century. *J. Geophys. Res.*, **113**, D16116, doi:10.1029/2007JD009571.

- , and Coauthors, 2013: On the lack of stratospheric dynamical variability in low-top version of the CMIP5 models. *J. Geophys. Res. Atmos.*, **118**, 2494–2505, doi:10.1002/jgrd.50125.
- Charney, J. G., and P. G. Drazin, 1961: Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *J. Geophys. Res.*, **66**, 83–109, doi:10.1029/JZ066i001p00083.
- Christiansen, B., 2000: A model study of the dynamical connection between the Arctic Oscillation and stratospheric vacillations. *J. Geophys. Res.*, **105**, 29 461–29 474, doi:10.1029/2000JD900542.
- Coughlin, K., and L. Gray, 2009: A continuum of sudden stratospheric warmings. *J. Atmos. Sci.*, **66**, 531–540, doi:10.1175/2008JAS2792.1.
- Delworth, T., and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *J. Climate*, **19**, 643–674, doi:10.1175/JCLI3629.1.
- Domeisen, D., A. H. Butler, K. Fröhlich, M. Bittner, W. A. Müller, and J. Baehr, 2015: Seasonal predictability over Europe arising from El Niño and stratospheric variability in the MPI-ESM seasonal prediction system. *J. Climate*, **28**, 256–271, doi:10.1175/JCLI-D-14-00207.1.
- Dufresne, J., and Coauthors, 2013: Climate change projections using the IPSL-CM5 Earth System Model: From CMIP3 to CMIP5. *Climate Dyn.*, **40**, 2123–2165, doi:10.1007/s00382-012-1636-1.
- Genz, A., F. Bretz, and Y. Hochberg, 2004: Approximation to multivariate t integrals with application to multiple comparison procedures. *Recent Developments in Multiple Comparison Procedures, Lecture Notes—Monograph Series*, No. 47, Institute of Mathematical Statistics, 24–32.
- Holton, J. R., and C. Mass, 1976: Stratospheric vacillation cycles. *J. Atmos. Sci.*, **33**, 2218–2225, doi:10.1175/1520-0469(1976)033<2218:SVC>2.0.CO;2.
- Jia, L., and Coauthors, 2017: Seasonal prediction skill of northern extratropical surface temperature driven by the stratosphere. *J. Climate*, **30**, 4463–4475, doi:10.1175/JCLI-D-16-0475.1.
- Jucker, M., 2016: Are sudden stratospheric warmings generic? Insights from an idealized GCM. *J. Atmos. Sci.*, **73**, 5061–5080, doi:10.1175/JAS-D-15-0353.1.
- , S. Fueglistaler, and G. K. Vallis, 2014: Stratospheric sudden warmings in an idealized GCM. *J. Geophys. Res. Atmos.*, **119**, 11 054–11 064, doi:10.1002/2014JD022170.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.
- Kidston, J., A. Scaife, S. Hardiman, D. Mitchell, N. Butchart, M. Baldwin, and L. 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., and L. M. Polvani, 2004: Stratosphere–troposphere coupling in a relatively simple AGCM: The role of eddies. *J. Climate*, **17**, 629–639, doi:10.1175/1520-0442(2004)017<0629:SCIARS>2.0.CO;2.
- Limpasuvan, V., D. L. Hartmann, D. W. J. Thompson, K. Jeev, and Y. L. Yung, 2005: Stratosphere–troposphere evolution during polar vortex intensification. *J. Geophys. Res.*, **110**, D24101, doi:10.1029/2005JD006302.
- Matsuno, T., 1971: A dynamical model of the stratospheric sudden warming. *J. Atmos. Sci.*, **28**, 1479–1494, doi:10.1175/1520-0469(1971)028<1479:ADMOTS>2.0.CO;2.
- Polvani, L., and D. Waugh, 2004: Upward wave activity flux as a precursor to extreme stratospheric events and subsequent anomalous surface weather regime. *J. Climate*, **17**, 3548–3554, doi:10.1175/1520-0442(2004)017<3548:UWAFAA>2.0.CO;2.
- Schoeberl, M. R., and D. F. Strobel, 1980: Numerical simulation of sudden stratospheric warmings. *J. Atmos. Sci.*, **37**, 214–236, doi:10.1175/1520-0469(1980)037<0214:NSOSSW>2.0.CO;2.
- Seviour, W. J. M., L. J. Gray, and D. M. Mitchell, 2016: Stratospheric polar vortex splits and displacements in the high-top CMIP5 climate models. *J. Geophys. Res. Atmos.*, **121**, 1400–1413, doi:10.1002/2015JD024178.
- Sigmond, M., J. Scinocca, V. Kharin, and T. Shepherd, 2013: Enhanced seasonal forecast skill following stratospheric sudden warmings. *Nat. Geosci.*, **6**, 98–102, doi:10.1038/ngeo1698.
- Staten, P., and T. Reichler, 2014: On the ratio between shifts in the eddy-driven jet and the Hadley cell edge. *Climate Dyn.*, **42**, 1229–1242, doi:10.1007/s00382-013-1905-7.
- Thompson, D., M. Baldwin, and J. Wallace, 2002: Stratospheric connection to Northern Hemisphere wintertime weather: Implications for prediction. *J. Climate*, **15**, 1421–1428, doi:10.1175/1520-0442(2002)015<1421:SCTNHW>2.0.CO;2.
- Watanabe, M., and Coauthors, 2010: Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. *J. Climate*, **23**, 6312–6335, doi:10.1175/2010JCLI3679.1.
- Wilks, D., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. Academic Press, 627 pp.
- Yoden, S., 1987: Bifurcation properties of a stratospheric vacillation model. *J. Atmos. Sci.*, **44**, 1723–1733, doi:10.1175/1520-0469(1987)044<1723:BPOASV>2.0.CO;2.