Seasonal Variability of the Polar Stratospheric Vortex in an Idealized AGCM with Varying Tropospheric Wave Forcing

ADITI SHESHADRI AND R. ALAN PLUMB
Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts

EDWIN P. GERBER
Courant Institute of Mathematical Sciences, New York University, New York, New York

(Manuscript received 9 July 2014, in final form 30 January 2015)

ABSTRACT

The seasonal variability of the polar stratospheric vortex is studied in a simplified AGCM driven by specified equilibrium temperature distributions. Seasonal variations in equilibrium temperature are imposed in the stratosphere only, enabling the study of stratosphere–troposphere coupling on seasonal time scales, without the complication of an internal tropospheric seasonal cycle. The model is forced with different shapes and amplitudes of simple bottom topography, resulting in a range of stratospheric climates. The effect of these different kinds of topography on the seasonal variability of the strength of the polar vortex, the average timing and variability in timing of the final breakup of the vortex (final warming events), the conditions of occurrence and frequency of midwinter warming events, and the impact of the stratospheric seasonal cycle on the troposphere are explored. The inclusion of wavenumber-1 and wavenumber-2 topographies results in very different stratospheric seasonal variability. Hemispheric differences in stratospheric seasonal variability are recovered in the model with appropriate choices of wave-2 topography. In the model experiment with a realistic Northern Hemisphere–like frequency of midwinter warming events, the distribution of the intervals between these events suggests that the model has no year-to-year memory. When forced with wave-1 topography, the gross features of seasonal variability are similar to those forced with wave-2 topography, but the dependence on forcing magnitude is weaker. Further, the frequency of major warming events has a nonmonotonic dependence on forcing magnitude and never reaches the frequency observed in the Northern Hemisphere.

1. Introduction

One of the most obvious characteristics of the winter stratosphere, aside from the dominance of planetary-scale waves, is its high degree of variability, on both intraseasonal and interannual time scales. The most dramatic manifestations of this variability are the major stratospheric sudden warmings (SSWs) that occur on average about 0.6 times per winter in the Northern Hemisphere (e.g., Charlton and Polvani 2007) but only once in the observational era in the Southern Hemisphere, though similar events that fail to meet the “major” criterion are common in both hemispheres. Major SSW events can be classified into “displacement” and “splitting” events, in approximately equal numbers (Charlton and Polvani 2007). Similar final warmings (FWs), which occur at the end of winter at the final transition into the summertime circulation, occur in both hemispheres. These events show many similarities to SSWs, especially in being characterized by planetary wave amplification; their timing is variable around or after the spring equinox (being delayed by around 2 months with respect to equinox in the Southern Hemisphere), with the early events being the most active (Black and McDaniel 2007a, b; Hu et al. 2014b).

While some part of this wintertime variability is undoubtedly a reflection of variability in tropospheric wave forcing, it is clear from modeling studies that such variability can arise spontaneously as a consequence of the dynamical interaction between waves and the zonal flow in the deep stratosphere–troposphere system. Even the one-dimensional Holton–Mass model (Holton and Mass 1976) exhibits such behavior; depending on the
amplitude of steady wave forcing at the bottom boundary, the system can exhibit steady solutions or, at larger forcing amplitude, quasi-periodic or chaotic behavior (e.g., Yoden 1990). More realistic global models have revealed similar behavior (Christiansen 2000; Taguchi et al. 2001; Gray et al. 2003; Scott and Polvani 2004, 2006) and have demonstrated that a realistic degree of variability can be generated even when the tropospheric wave forcing is fixed in time.

Some full GCMs exhibit reasonable stratospheric variability, but shortcomings still exist. Full GCMs tend to underestimate the wave forcing in the Northern Hemisphere but slightly overestimate the wave forcing in the Southern Hemisphere (Eyring et al. 2010, chapter 4), while many such models underestimate the frequency of major warming events (Charlton et al. 2007). Experiments with idealized GCMs permit an investigation of parameter dependence that has proved useful to better understand stratospheric dynamics in models as well as the real atmosphere.

In mechanistic studies with global dynamical models run in “perpetual solstice” mode with simple planetary-scale surface topography and with the stratospheric state determined by Newtonian relaxation to a specified equilibrium temperature distribution, Taguchi et al. (2001) and Gerber and Polvani (2009) documented the dependence of their model’s climatology on the amplitude of the specified topography. Taguchi et al. (2001) found SSW-like variability appearing at modest topographic amplitude, becoming more frequent and more intense, with a weaker vortex, with greater amplitudes. Gerber and Polvani (2009), who also explored the role of different stratospheric vortex strengths, found similar behavior with forcing at zonal wavenumber 2 and were able to identify cases with a realistic (for the Northern Hemisphere) frequency of major SSW events. Unlike Taguchi et al., however, they were unable to find a realistic regime with wave-1 forcing; rather, at forcings below a certain value, variability was found to be too weak, with no SSWs, while at larger forcings, the vortex was completely destroyed and the model climatology was highly unrealistic.

Greater realism and, especially, the key aspects of the observed differences between variability in the Northern and Southern Hemispheres, have been achieved by imposing a seasonal cycle in the imposed equilibrium temperature distribution. Using a one-dimensional Holton–Mass-like model, Plumb (1989) found a transition from a cycle like that of the southern stratosphere, with greatest wave amplitudes in early and, especially, late winter with weak topographic amplitudes, to one like that of the Northern Hemisphere, with peak amplitudes throughout winter. Taguchi and Yoden (2002), using wavenumber-1 forcing, found similar behavior in a global model, highlighting the shift of maximum wave amplitudes and of zonal wind variability from late winter to midwinter as topographic amplitude was increased.

In fact, Scott and Haynes (2002), in similar calculations in which geopotential wave amplitude rather than topographic height was specified at the surface, found threshold behavior, with rather modest changes to the zonal winds below some critical forcing amplitude, changing suddenly to a state with large variability in early winter, followed by a complete collapse of the vortex and of its variability throughout late winter. This finding raises a question as to whether this transition corresponds to that found by Gerber and Polvani (2009) with wave-1 forcing: in perpetual-solstice mode, if such a vortex collapse occurred during the early part of the run, the collapse may be sustained indefinitely whereas, in the presence of an annual cycle, the vortex may be reinitialized every winter and thus produce a seasonal cycle like that found by Scott and Haynes.

The simulation of FWs has also been assessed in such models with seasonal forcing. Scott and Haynes (2002) did not find realistic final warmings in their strongly forced cases. However, Sun and Robinson (2009) and Sun et al. (2011) reported realistic FW events in a global model running from midwinter into summer. Like observed FWs, these events were associated with bursts of amplified planetary waves, which were demonstrated to play a major role in the events’ dynamics. Increasing tropospheric forcing in such models led to FWs becoming earlier on average but with greater variability in their timing.

Both SSWs and FWs have clear impacts on the tropospheric circulation. Fluctuations in vortex strength are followed by persistent perturbations to surface winds, manifesting themselves as “annular mode” anomalies (Thompson and Wallace 1998; Baldwin and Dunkerton 1999, 2001). Model studies indicate that synoptic-scale tropospheric eddies appear to play an important role in organizing and amplifying the stratospheric response into this form (e.g., Song and Robinson 2004; Kushner and Polvani 2004). Gerber and Polvani (2009) explicitly demonstrated a realistic tropospheric signal associated with SSWs in their global model. FWs are also observed to have a similar tropospheric impact, although its latitudinal structure differs somewhat from the annular-mode form (Black et al. 2006; Black and McDaniel 2007a,b; Hu et al. 2014b), and upper-tropospheric planetary-scale waves may play a role (Sun et al. 2011). Similar results have been found in simplified global models (Sun and Robinson 2009; Sun et al. 2011; Sheshadri et al. 2014).

In this paper, we revisit the seasonal behavior of the stratosphere in a global model, paying particular
attention to the dependence of seasonal and interannual variability on the amplitude and wavenumber of topographic forcing. The model is essentially the same as that used by Gerber and Polvani (2009), with modifications to the imposed equilibrium temperature distributions; the model setup is described in section 2. With wavenumber-2 topography (section 3) the model reproduces the essential characteristics of the observed stratosphere in the two hemispheres: with weak topographic forcing, maximum zonal wind variability occurs in late winter, with a late FW, while both of these characteristics shift toward earlier in the winter as the forcing is increased. Moreover, like Gerber and Polvani (2009), we find a range of topographic forcing amplitudes for which a realistic frequency of SSWs is generated. Unlike Scott and Haynes’s (2002) results with wave-1 forcing, however, we find that the evolution varies smoothly with the wave-forcing amplitude; there is no evidence of threshold behavior. Interestingly, consistent with the results of Taguchi and Yoden (2002), we do find evidence for saturation of wave amplitudes with stronger forcings: that is, stratospheric wave amplitudes do not increase once topographic forcing reaches a certain value.

As we shall describe in section 4, the story with wavenumber-1 forcing is somewhat different. The shift, with increased forcing, of the dynamical evolution toward early winter, though present, is weaker than in the wavenumber-2 case. Realistic FWs are produced, but we are unable to find any regime with a realistic frequency of major SSW events. While Gerber and Polvani (2009) could find no such events, some major SSWs do occur within a range of forcing amplitudes in these experiments, but never with the observed frequency.

Since the imposed equilibrium temperature in this model varies seasonally only in the stratosphere, any seasonal variations in the tropospheric circulation are obviously of stratospheric origin. As discussed in section 5, in some experiments we find a substantial tropospheric signal not only of SSWs and FWs, but also of the seasonal cycle itself. As found by Chan and Plumb (2009), we find this to be a consequence of the unrealistically long tropospheric annular-mode time scale in some experiments; in experiments with more reasonable time scales, this tropospheric signal disappears. Conclusions are presented in section 6.

2. Model setup

The model is dry and hydrostatic, solving the global primitive equations with T42 resolution. We use 40 hybrid levels, transitioning from a terrain-following $\sigma = p/p_s$ coordinate at the surface to pure pressure levels by 115 hPa. The hybrid levels are located at the same mean position as the levels in Polvani and Kushner (2002). We include bottom topography specified as in Gerber and Polvani (2009) in the Northern Hemisphere of the model in all experiments. The topography is specified by setting the surface geopotential height as follows:

$$\Phi_0(\lambda, \phi) = \begin{cases} g h_0 \sin^2 \left( \frac{\phi - \phi_0}{\phi_1 - \phi_0} \pi \right) \cos(m\lambda), & \phi_0 < \phi < \phi_1, \\ 0, & \text{otherwise} \end{cases}$$

where $\lambda$ and $\phi$ refer to longitude and latitude and $m$ and $h_0$ refer to the wavenumber and height of topography. The latitudes $\phi_0$ and $\phi_1$ are set to 25° and 65°N, so that the topography is centered at 45°N.

Linear damping of the horizontal winds is applied in the planetary boundary layer and in a sponge above 0.5 hPa, exactly as in Polvani and Kushner (2002).

Newtonian relaxation forces temperatures toward a zonally symmetric equilibrium temperature $T_{eq}$. In the stratosphere, the radiative relaxation time scale is 40 days. Within the troposphere, there is no imposed seasonal variation; $T_{eq}$ is specified as in Polvani and Kushner (2002), with the parameter $\varepsilon = 10$ K providing an asymmetry between the two hemispheres (our setup makes the Northern Hemisphere colder). Within the stratosphere, a seasonal cycle in $T_{eq}$ is prescribed following Kushner and Polvani (2006), as follows:

$$T_{eq}^{strat}(\phi, p, t) = [1 - W(\phi, t)]T_{US}(p) + W(\phi, t)T_{PV}(p),$$

where $\phi$ is latitude, $p$ is pressure, $T_{US}(p)$ is the temperature defined by the U. S. Standard Atmosphere, 1976 (COESA 1976), expressed as a function of pressure, and $T_{PV}(p)$ is the polar vortex $T_{eq}$ prescription of Polvani and Kushner (2002). The lapse rate of $T_{PV}(p)$ is $\gamma$, which can be varied to produce different strengths of the polar vortex. Here we use a fixed lapse rate of 4 K km$^{-1}$ in all experiments. The weighting function is

$$W(\phi, t) = \frac{1}{2} \left[ A_S(t) \{1 + \tanh[(\phi - \phi_{OS})/\delta\phi_S]\} ight. + \left. A_N(t) \{1 + \tanh[(\phi - \phi_{ON})/\delta\phi_N]\} \right],$$

where $A_S(t) = \max\{0, 0, \sin[2\pi(t - t_0)/\Delta T]\}$ and $A_N(t) = \max\{0, 0, \sin[2\pi t/\Delta T]\}$, with $t_0 = 180$ days, $\Delta T = 360$ days, $\phi_{OS} = -50^\circ$, $\phi_{ON} = 50^\circ$, $\delta\phi_S = -10^\circ$ and $\delta\phi_N = 10^\circ$. Thus, at a given polar latitude, $T_{eq}^{strat}$ varies between polar summer and polar winter over a 360-day year. In
the original Polvani and Kushner (2002) study, there is a smooth transition from tropospheric to stratospheric specifications of $T_{eq}$ across 100 hPa. Hung and Gerber (2013, meeting presentation) reported a number of problems that result from the use of the Polvani and Kushner (2002) equilibrium temperature profile. They concluded that the culprit was a bias in the lower-stratospheric equilibrium temperatures, which were too warm. In this study, we deviate slightly from the Polvani and Kushner (2002) equilibrium temperature profile, in that we cause $T_{eq}$ to transition from tropospheric to stratospheric specifications at 200 hPa, rather than at 100 hPa. Thus, the cold anomaly that is the representation of the stratospheric polar vortex in this model comes into effect lower down in the stratosphere and addresses the bias in lower-stratospheric equilibrium temperatures reported by Hung and Gerber (2013, meeting presentation). Figure 1 shows a snapshot of $T_{eq}$ at the Northern Hemisphere winter solstice and can be compared with Fig. 1 in Kushner and Polvani (2006).

This change leads to a marked improvement in the seasonal cycle of lower-stratospheric zonal winds. The improvement is evident in Fig. 2, which compares the annual cycle of zonal-mean zonal winds at 10 and 50 hPa generated from the model using the $T_{eq}$ specifications of Polvani and Kushner (2002) in the top panels with the new $T_{eq}$ specifications that we use in the middle panels. These panels are from a model configuration without topography, and, as will be described in section 3, this configuration results in a Southern Hemisphere–like stratospheric seasonal variability. The bottom panels show the annual cycle of zonal-mean zonal winds at 10 and 50 hPa for the Southern Hemisphere from ERA-Interim (Dee et al. 2011) for comparison. The new equilibrium temperature specifications lead to a stronger polar vortex at both 10 and 50 hPa. In addition, the persistence of westerly winds in the summer at 50 hPa is reduced, so that the new $T_{eq}$ specifications result in a sharper seasonal cycle, with a clear seasonal formation and breakup of the polar vortex. In sections 3 and 4, we use this improved model setup to explore the climatology and seasonal and interannual variability of the polar vortex forced by different amplitudes of wave-1 and wave-2 topography.

Table 1 summarizes results from the experiments analyzed here. We integrated every model experiment for 35 years and analyzed the last 30 years. Experiment 4 (with 4000-m wave-2 topography), which resulted in the most Northern Hemisphere–like stratospheric variability (i.e., a frequency of sudden warming events that matches the observations) was run for a further 20 years, giving us a larger dataset for the analysis of sudden warming events. Figure 3 compares the latitude–pressure structure of the wintertime (averaged from December to February) winds for the experiments without topography and wave-1 and wave-2 topography with $h_0 = 4000$ m. Increasing the tropospheric wave forcing from the inclusion of increasing heights of topography of either wavenumber reduces the strength of midwinter westerlies.

3. Stratospheric seasonal variability in the presence of wave-2 topography

Stratospheric seasonal variability is illustrated here with histograms of the 10-hPa monthly mean zonal-mean zonal wind at 60° [cf. similar figures of Naujokat (1981) and Taguchi and Yoden (2002) showing frequency distributions of polar temperature]. The seasonal variability of the observed Southern and Northern Hemisphere stratospheres (1979–2008 from ERA-Interim) is shown in Fig. 4 for reference. The summertime winds are similar in both hemispheres; winds are weakly easterly and show little variability. Significant differences between the hemispheres, however, appear in the winter winds. The midwinter westerlies in the Antarctic polar vortex are much stronger than those in the Arctic, and most of the variability in the Antarctic winds occurs at the end of the winter and into spring, in contrast to the midwinter variability evident in the Arctic.

Zonal wind histograms for experiment 1 (without topography) and experiments 2–4 (with increasing amplitudes $h_0$ of wave-2 topography) are shown in Fig. 5. The strength of the midwinter westerlies in the model experiments reduces with increased $h_0$. The experiment
without topography most resembles the observed southern stratosphere, with most of the variability occurring after the spring equinox and little variability through the winter. Although the maximum winds do not reduce significantly between the experiments with $h_0 = 0$ and $h_0 = 2000$ m, the latter exhibits somewhat more wintertime variability. The experiments with 3000- and 4000-m topography begin to resemble the real Northern Hemisphere, with weaker westerlies and significant variability in the strength of westerlies around midwinter. These experiments also show reduced mean westerlies and much less variability in late winter and
early spring, similar to the results of Taguchi and Yoden (2002) with wave-1 forcing. However, unlike the threshold behavior reported by Scott and Haynes (2002), also with wave-1 forcing, this shift toward early winter occurs rather smoothly as $h_0$ is increased.

The climatological evolution of the geopotential height amplitude of wave 2 in the middle stratosphere for these experiments is shown in Fig. 6. Even without topographic forcing, the amplitudes reach 500 m, forced presumably by the nonlinear interaction of baroclinic,

<table>
<thead>
<tr>
<th>Expt</th>
<th>Topography</th>
<th>Peak 50-hPa wind at 60°N (m s$^{-1}$)</th>
<th>Mean SFW timing (day of year)</th>
<th>Standard deviation of SFW timing (days)</th>
<th>No. of SSWs (yr$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>—</td>
<td>28</td>
<td>160</td>
<td>19.4</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2000-m wave 2</td>
<td>32</td>
<td>141</td>
<td>21.3</td>
<td>0.17</td>
</tr>
<tr>
<td>3</td>
<td>3000-m wave 2</td>
<td>19</td>
<td>98</td>
<td>20.5</td>
<td>0.2</td>
</tr>
<tr>
<td>4</td>
<td>4000-m wave 2</td>
<td>13</td>
<td>88</td>
<td>31</td>
<td>0.62*</td>
</tr>
<tr>
<td>5</td>
<td>2000-m wave 1</td>
<td>30</td>
<td>155</td>
<td>21.2</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>3000-m wave 1</td>
<td>28</td>
<td>149</td>
<td>17.5</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>4000-m wave 1</td>
<td>23</td>
<td>114</td>
<td>10.6</td>
<td>0.033</td>
</tr>
<tr>
<td>8</td>
<td>5000-m wave 1</td>
<td>13</td>
<td>97</td>
<td>15</td>
<td>0.1</td>
</tr>
<tr>
<td>NH</td>
<td>—</td>
<td>24</td>
<td>103 (13 Apr)</td>
<td>15.1</td>
<td>0.61**</td>
</tr>
<tr>
<td>SH</td>
<td>—</td>
<td>55</td>
<td>335 (1 Dec) or 155 days after 1 Jul</td>
<td>12.9</td>
<td>0.033</td>
</tr>
</tbody>
</table>

* This value is from a 50-yr integration, 20 years longer than the other experiments.
** This value is from a larger set of years (1958–2013).

Fig. 3. Latitude–pressure structure of December–February winds (m s$^{-1}$) for (top left) experiment 1 (no topography), (top right) experiment 4 (4000-m wave-2 topography), and (bottom) experiment 6 (4000-m wave-1 topography). The white patch at the bottom indicates the extent of topography. The contour interval is 5 m s$^{-1}$.
synoptic-scale eddies in the troposphere (Scinocca and Haynes 1998). However, note that the broad peak in wave amplitudes, occurring between the winter solstice and spring equinox, does not coincide with the peak in zonal wind variability for this case of $h_0 = 0$ m (cf. Fig. 5). With topographic forcing, the amplitudes are greater; while wave growth occurs at about the same time (i.e., about 1 month before the winter solstice), the timing of peak amplitudes and of their subsequent collapse drifts systematically earlier as $h_0$ is increased, now mirroring the similar shift in mean zonal wind behavior seen in Fig. 5. One remarkable feature of Fig. 6, which is also evident in the results of Taguchi and Yoden (2002), is the apparent saturation of wave amplitudes, which do not show any marked increase as $h_0$ is increased from 2000 to 4000 m. At small $h_0$, there is no evidence of the early winter resonance described by Scott and Haynes (2002), nor of the springtime peak seen in observations and found in models both by Scott and Haynes and by Taguchi and Yoden.

This temporal shift of wave behavior is also evident in the magnitude of wave activity propagation into the stratosphere. Figure 7 shows the mean seasonal cycle of eddy heat flux $\overline{vT'}$ at 50 hPa, a measure of the vertical component of Eliassen–Palm flux in the lower stratosphere, for the same four wavenumber-2 experiments. Unlike what is seen in geopotential height amplitudes in Fig. 6, here we do see a clear late winter/spring peak in the heat fluxes for $h_0 = 0$ m; the peak shifts systematically toward midwinter with increasing $h_0$. It is also evident that the latitude of the maximum heat fluxes migrates with the seasonal cycle; we will revisit this in the section on final warming events.

a. **SSW events**

All of our experiments exhibit some level of wintertime variability. For our purposes, we adopt the definition of Charlton and Polvani (2007), and define “major” SSWs as those involving a reversal of the zonal-mean westerlies at 10 hPa and 60°. Events that occur within 20 days of the final warming (to be defined below) are excluded.

In observations, the Arctic vortex exhibits major SSWs about every other year on average [in the combined ERA-40 (Uppala et al. 2005) and ERA-Interim record from 1958 to 2013, there are 34 major SSWs, giving a mean frequency of 0.61 events per winter],
FIG. 5. Histograms of 10-hPa zonal-mean zonal wind at 60°N for experiments (top left) 1, (top right) 2, (bottom left) 3, and (bottom right) 4. The ordinate shows the number of occurrences of a specific value of zonal-mean zonal wind.
Gerber and Polvani (2009) found that the most realistic frequency of such events (major SSWs every 200–300 days in their perpetual-January integrations) occurred when their model was forced with 3000-m wave-2 topography. In our model, as seen in Table 1, the inclusion of wave-2 topography of modest amplitude leads to infrequent major SSWs (experiments 2 and 3). The frequency of such events increases monotonically with $h_0$ (in the range explored here). We find that major SSWs occur with a realistic Northern Hemisphere–like frequency when $h_0 = 4000$ m (experiment 4, for which the average occurrence is 0.62 per winter).

We now examine both “weak” and “strong” vortex events from experiment 4. We identify weak and strong vortex events based on $1\sigma$ and $2\sigma$ annular-mode anomalies (anomalies of the principal component time series corresponding to the first EOF of geopotential height at 10 hPa that exceed one or two standard deviations). For each year, we identify the largest annular-mode anomalies; that is, we count at most one weak vortex event and one strong vortex event every year. Among this set of “largest” weak and strong vortex events, we group them into $1\sigma$ and $2\sigma$ events. For each year, we identify the largest annular-mode anomalies; that is, we count at most one weak vortex event and one strong vortex event every year. Among this set of “largest” weak and strong vortex events, we group them into $1\sigma$ and $2\sigma$ events. In the 50 model years, there were 50 weak vortex events that met the $1\sigma$ criterion, among which 30 events also met the $2\sigma$ criterion. There were fewer strong vortex events: 41 events met the $1\sigma$ criterion, among which only 6 events also met the $2\sigma$ criterion. The top row of Fig. 8 shows composites of geopotential height anomalies normalized by their standard deviations for a 60-day period centered on the 20 weak vortex events that met the $1\sigma$ criterion, among which 30 events also met the $2\sigma$ criterion. There were fewer strong vortex events: 41 events met the $1\sigma$ criterion, among which only 6 events also met the $2\sigma$ criterion. The top row of Fig. 8 shows composites of geopotential height anomalies normalized by their standard deviations for a 60-day period centered on the 20 weak vortex events that met the $1\sigma$ criterion, among which 30 events also met the $2\sigma$ criterion. The top row of Fig. 8 shows composites of geopotential height anomalies normalized by their standard deviations for a 60-day period centered on the 20 weak vortex events that met the $1\sigma$ criterion, among which 30 events also met the $2\sigma$ criterion.
anomalies at 95 hPa normalized by their standard deviations and composited on all 50 weak vortex events. There is a burst of poleward heat flux anomalies just prior to the weak vortex events, followed by equatorward heat flux anomalies just after. The wavenumber-2 component accounts for most of these heat flux anomalies.

Figure 9 shows geopotential height anomalies normalized by their standard deviations for a 60-day period centered on the 41 strong vortex events, with their associated heat flux anomalies at 95 hPa, also normalized by their standard deviations. Composites of 2σ strong vortex events are not shown, because of their small numbers. Just as in the case of weak vortex events, we see anomalously equatorward/poleward heat flux anomalies just before and after strong vortex events. Our analysis of weak and strong vortex events can be compared with the studies of Baldwin and Dunkerton (2001) and Gerber and Polvani (2009). However, strong vortex cases were not analyzed in the perpetual-winter configuration of the model by Gerber and Polvani (2009) because events were not well captured by any particular threshold; the vortex tended to slowly build up and decay, without a marked, eventlike structure.

Major SSW events do not occur every year, either in the model or in the observed Northern Hemisphere stratosphere. The distribution of the occurrence of these events or, more generally, of winters that are highly disturbed, leads to the question: does the stratosphere retain some “memory” of the previous winter? The model has no external sources of memory (such as sea surface or land conditions), but the tropical zonal winds may retain memory from year to year (the low-latitude fly-wheel effect; Scott and Haynes 1998). In the 50 model years, there were 31 years in which the winter was disturbed enough to produce at least one major SSW. In the left panel of Fig. 10, we compare the distribution of the interval between years in which there were major SSWs (shown as blue bars) to the distribution that results from a time series of 1 million randomly generated 0s and 1s, with the probability of occurrence of a “1” fixed at 31/50 = 0.62 (shown as gray bars). With the random number generator, we explicitly specify that there is no memory from one “year” to the next. We sample the randomly generated 0s and 1s in batches of 50 years to generate 20000 batches of 50 years (i.e., we use a Monte Carlo method) and plot the 95th and 5th percentiles of the distribution of intervals between 1s as gray dots in the left panel of Fig. 10. For comparison, we follow a similar procedure for midwinter warming events identified from the ERA-40 [obtained from
Charlton and Polvani (2007)] from 1958 to 2002, combined with those we identify from ERA-Interim data from 2003 to 2013; these results are shown in the right panel of Fig. 10. The dearth of major SSWs in the 1990s (the interval of 8 years in the right panel of Fig. 10) is the only potentially significant departure from a random process with no interannual memory; an 8-yr gap occurs in only 2.7% of the random 56-yr batches. In our idealized model, however, the occurrence of major SSWs shows no sign of being significantly different from the null hypothesis of events occurring at random with a given probability. Thus, there is no evidence of winter-to-winter stratospheric memory in our model.

b. Final warming events

In our analysis of FW events in the model, we follow Black and McDaniel (2007a,b) in focusing attention on the 50-hPa level. Examining FW events in the Northern Hemisphere, Black and McDaniel (2007a) defined the timing of these events as the final time that the zonal-mean zonal wind at 50 hPa and 60°S reached the value of 10 m s⁻¹ (Black and McDaniel 2007b). In our case, each choice of the height and wavenumber of topography results in a different stratospheric climatology in the model. To apply a uniform definition of FW events across all model experiments, we define the timing of these events as the day that the zonal-mean zonal wind at 50 hPa and 60° falls below 25% of its annual maximum for the last time. For comparison, we apply the same definition to 30 years from ERA-Interim data (1979–2008) (shown for the Arctic and Antarctic vortices in the last two rows of Table 1).

As we have seen, experiment 1 (with no topography) results in a Southern Hemisphere–like stratospheric seasonal variability with a relatively smooth seasonal cycle of zonal winds and no major SSW events.1 FW events occur on average shortly before the summer solstice (i.e., more than one radiative relaxation time constant after equinox). As seen in Table 1, there is a monotonic shift toward earlier final warming events with increasing $h_0$, consistent with the results of Sun and Robinson (2009) and Sun et al. (2011). In experiment 4 (with $h_0 = 4000$ m, for which we see Northern Hemisphere–like stratospheric variability), the FW happens 72 days earlier (compared to experiment 1) on average (i.e., close to the vernal equinox), although the FW date becomes more variable from year to year. However, more modest values of $h_0$ do not cause an increase in the variability in timing of FW events.

The synoptic evolution of final warming events from the experiments without topography and with different amplitudes of wave-2 topography is illustrated in Fig. 11. Each row of Fig. 11 shows a typical example of a final warming event from experiments 1, 2, and 3 (experiment 4 is similar to experiment 3 in this respect and is therefore not shown). In the cases with no topography and 2000-m wave-2 topography, as the vortex goes through the final warming, it weakens, wanders off the pole, and disintegrates into multiple segments. With higher amplitudes of wave-2 topography, the final warming is always a split of the polar vortex, an example of which is shown in the third row of Fig. 11.

---

1 Kushner and Polvani (2005) reported the occurrence of one strong midwinter warming event in an 11 000-day perpetual-January integration in a model setup using the $T_{eq}$ specifications of Polvani and Kushner (2002), with a lapse rate of 2 K km⁻¹ and no topography. We did not see such an event in our 30-yr integration.
Figure 12 shows the 50-hPa zonal-mean zonal wind at 60°N for a 60-day period composited on FW events for experiments 1–4 (the experiments without topography and increasing heights of wave-2 topography). The change in winds as the vortex goes through the final warming is larger for higher-tropospheric wave forcing, with the exception of the 4000-m wave-2 case, in which the vortex is already very weak as it approaches the final warming. The transition in winds as the vortex goes through the final warming is largest in the case of the 3000-m wave-2 topography. Both the cases with 3000- and 4000-m wave-2 topography result in weak easterlies following the final warming.

Hu et al. (2014a), in their analysis of the timing of midwinter warming events and final warming events in the NCEP–NCAR reanalysis, suggested that late spring final warming events tend to be preceded by major SSWs, while early FW events do not. Similarly, in experiment 4, we find that final warming events occur on average 11 days later in years with major SSW events. However, we note that this difference is less than half the standard deviation in the timing of FWs in this experiment.

Figure 13 shows the seasonal cycle of \( vT' \) at 50 hPa composited on final warming events for all the wave-2 model experiments. The heat fluxes become strong around the time when the lower-stratospheric winds start to become variable and remain strong until the final warming, when they collapse. The burst of heat fluxes at the final warming is more intense with increasing tropospheric wave forcing but is less pronounced in experiment 4, for which SSW events occur about every other year and for which the average zonal winds are weaker. The latitude of the maximum heat fluxes migrates slightly poleward up to the final warming, after which the planetary-scale fluxes collapse and smaller-scale waves (which maximize at lower latitudes) then dominate, leading to a sharp equatorward shift of the locus of maximum heat fluxes.

4. Stratospheric seasonal cycle in the presence of wavenumber-1 topography

Figure 14 shows the seasonal cycle and variability of zonal winds at 60°N, 10 hPa, for a range of forcing amplitudes with wavenumber \( m = 1 \) topography. As for the \( m = 2 \) case, for stronger forcing the model exhibits weaker mean winds, increased variability, and a general drift of these features toward earlier in the winter, although not as markedly so as for the \( m = 2 \) case. In fact, the evolution through the winter of wave amplitudes, shown in Fig. 15, indicates a slight discrepancy in late winter between cases with forcing below or above \( h_0 = 3000 \) m, but otherwise the timing of the seasonal evolution of wave amplitudes is insensitive to forcing amplitude. Just as in the \( m = 2 \) case, geopotential height amplitudes at 10 hPa appear to saturate at values near 800 m with sufficiently large (3000 m) forcing amplitude. The one exception to this statement is for the case with \( h_0 = 5250 \) m, for which substantially larger amplitudes are reached in midwinter; with greater \( h_0 \), however, the
magnitudes revert to the saturated values. We currently have no explanation for this nonmonotonic behavior.

Another significant difference between these experiments and those with \( m = 2 \) forcing is that, for \( m = 1 \) forcing, we are unable to find a regime with a realistic (Northern Hemisphere) frequency of major SSWs. This mirrors the inability of Gerber and Polvani (2009) to find any such events in their perpetual-solstice experiments. We do find some events, however, for \( 4000 \leq h_0 \leq 5500 \) m, though with lower frequency—no more than 0.3 events per year—than observed in northern winter. Some of these are displacement and others splitting events, with more displacements than splits. Strangely, we find no major SSWs for even stronger forcing \( h_0 \geq 6000 \) m; in such cases, the stratospheric state at 10hPa exhibits a very high degree of interannual variability—in some winters the vortex never becomes strong—with violent midwinter disturbances that, however, fail to meet the major SSW criterion of a wind reversal at 60°N. This fact may just be an indication of the arbitrariness of this criterion (Coughlin and Gray 2009), rather than having any more fundamental implications.

![Examples of the evolution of 50-hPa geopotential height (m) leading up to and immediately after final warming events. The 50-hPa geopotential height is shown 5 days before, the day of, and 5 days after the final warming. These are examples of final warming events from experiments (top) 1 (no topography), (middle) 2 (2000-m wave-2 topography), and (bottom) 3 (3000-m wave-2 topography). Experiment 4 (4000-m wave-2 topography) is similar to experiment 3. The contour interval is 53 m.](https://example.com/figure11.png)
Reasonably realistic FW events are produced in these $m = 1$ experiments. As is evident from Table 1, they happen about a month earlier for $h_0 = 4000$ m than for $h_0 = 3000$ m, though the shift toward earlier FWs is less dramatic than for the $m = 2$ case, as is suggested by Figs. 14 and 15. Figure 16 illustrates the synoptic evolution of FW events for $h_0 = 3000$ m (other experiments are similar), compared with the case without topography.

The two are similar in showing no clear wavenumber preference after the event: in both cases, the vortex weakens, meanders off the pole, and disintegrates into multiple segments as it goes through the final warming.

5. Impact of the stratospheric seasonal cycle on the troposphere

Figure 17 shows the seasonal variation of 515-hPa winds for experiments 1 (no topography), 2 (2000-m wave-2 topography), 3 (3000-m wave-2 topography), 5 (2000-m wave-1 topography), and 6 (3000-m wave-1 topography) as climatological averages. Since (as noted in section 2) the seasonal cycle in imposed equilibrium temperature is confined to the stratosphere, any seasonal variations in the tropospheric circulation must be of stratospheric origin. At 515 hPa, the zonal flow is dominated by the model’s subtropical jet near 30° latitude. In the absence of topography, however, the jet exhibits strong seasonal variation, shifting poleward by almost 10° and weakening in spring, evidently revealing a strong seasonal coupling from the stratosphere. As Fig. 17 shows, this seasonal variation becomes substantially weaker with increasing surface topography.

It seems that this behavior arises because of the impact of topography on the time scale $\tau$ of the model’s annular mode, which is shown on each frame of the
FIG. 14. Histograms of 10-hPa zonal-mean zonal wind at 60°N for experiments (top left) 5, (top right) 6, (bottom left) 7, and (bottom right) 8. The ordinate shows the number of occurrences of a specific value of zonal-mean zonal wind.
figure for each experiment (we determined the annular mode as the first EOF of daily zonal-mean zonal wind at 850 hPa; \( \tau \) was then defined from the principal component autocorrelation function as the best least squares fit to an exponential decay.) Increasing the topography produces a marked reduction in \( \tau \), consistent with the arguments of Gerber and Vallis (2007). In turn, this leads us to expect a significant reduction in the sensitivity of the tropospheric jet to perturbation from the stratosphere, since Ring and Plumb (2008) and Gerber et al. (2008) found that the response of the tropospheric jet to external forcing increases with \( \tau \), as suggested by the fluctuation–dissipation relationship. Our interpretation of this behavior as a consequence of decreasing \( \tau \), rather than as a direct consequence of the presence of topography, is supported by experiments in which, following Chan and Plumb (2009), we fix topography but change \( \tau \) by altering the distribution of equilibrium temperature in the troposphere; results (not shown here) show the same reduction in tropospheric seasonality with reduced \( \tau \). Note from Fig. 17 that, even in the presence of large topography, \( \tau \) exceeds values appropriate to the observed atmosphere by a factor of 2 or 3, implying that even for these cases the potential for coupling with the stratospheric seasonal cycle is likely exaggerated compared with the real atmosphere.

**FIG. 15.** The climatological evolution of the wave-1 geopotential height amplitude (m) at 10 hPa, 60°N for experiment 1 (no topography), experiment 5 (2000-m wave-1 topography), experiment 6 (3000-m wave-1 topography), experiment 7 (4000-m wave-1 topography), 4500-m wave-1 topography, experiment 8 (5000-m wave-1 topography), 5250-m wave-1 topography, 5500-m wave-1 topography, and 6000-m wave-1 topography.

**FIG. 16.** Evolution of 50-hPa geopotential height (m) leading up to and immediately after final warming events. The 50-hPa geopotential height is shown 5 days before, the day of, and 5 days after the final warming. These are examples of final warming events from experiments (top) 1 (no topography) and (bottom) 5 (3000-m wave-1 topography). Experiment 6 (4000-m wave-1 topography) is similar to experiment 5 (not shown). The contour interval is 53 m.
Conclusions

One of the more surprising outcomes of these experiments is the different response of the model’s stratosphere to wave-1 and wave-2 forcings. With wave-2 forcing, the response evolves smoothly from Southern Hemisphere–like behavior, with strong midwinter westerlies and with eddy heat fluxes and zonal wind variability that peak in spring, to a Northern Hemisphere–like state in which zonal winds are much reduced on average, and both eddy heat fluxes and zonal wind variability maximize in midwinter, as the wave forcing (in the form of imposed surface topography) is increased. This behavior is similar to that found, using wave-1 forcing, by Taguchi and Yoden (2002). Given sufficiently strong forcing (topographic height of 4000 m) the model produces a realistic frequency of major warmings (as compared with the observed Northern Hemisphere). All of these major warming events take the form of vortex splits, which is not surprising, given the wave-2 forcing. We do not, however, find a climatological springtime amplification of the wave geopotential magnitude (as opposed to the heat flux) with weak forcing, in contrast both to Taguchi and Yoden’s results and to observed behavior in the Southern Hemisphere. One striking characteristic of our results is that the stratospheric wave amplitudes saturate, in the sense that increasing topographic height beyond a modest magnitude does not lead to increased geopotential wave amplitude in the middle stratosphere; there are suggestions of the same behavior in Taguchi and Yoden (2002, their Fig. 3). Unlike the wave-1

![Fig. 17. Seasonal cycle of 515-hPa winds from experiments (top left) 1 (no topography), (top right) 2 (2000-m wave-2 topography), (middle left) 3 (3000-m wave-2 topography), (middle right) 5 (2000-m wave-1 topography), and (bottom) 6 (3000-m wave-1 topography) shown as calendar averages. The tropospheric annular-mode time scale for each experiment is also indicated. Experiments with higher heights of topography of both wavenumbers are similar to experiment 3 (not shown).]
experiments of Scott and Haynes (2002), we find no suggestion of threshold behavior in the response of the system to different levels of topographic forcing nor of resonant behavior in early winter when the topographic forcing is weak.

The behavior of the model stratosphere’s response to wave-1 forcing is rather different. The zonal winds do weaken, and the pattern of their variability shifts somewhat earlier in the winter as topography is increased, but less markedly so than with wave-2 forcing. More strikingly, the dependence on topographic forcing is nonmonotonic. For peak topography $h_0$ up to about 5000 m, the winter vortex becomes increasingly disturbed, and some major warmings occur [in contrast to Gerber and Polvani (2009) who were unable to find any such events in an almost identical model run in perpetual-solstice mode] but never with the frequency observed in the Northern Hemisphere. However, their frequency then decreases with further increases in forcing, although the winter vortex remains highly disturbed. Our difficulty in reproducing the observed frequency with wave-1 forcing is curious, since about half of the observed midwinter warming events in the Northern Hemisphere are displacement events (Charlton and Polvani 2007). When major warmings do occur in the runs with wave-1 forcing, some are displacement events, while some are splits.

Full GCMs tend to underestimate the wave forcing in the Northern Hemisphere but slightly overestimate the wave forcing in the Southern Hemisphere (Eyring et al. 2010, chapter 4), while many such models underestimate the frequency of major warming events (Charlton et al. 2007). Reaching an adequate understanding of the dependence of the modeled stratosphere on tropospheric forcing should help in understanding the behavior of climate models; however, while progress is being made, more clearly remains to be done.

Acknowledgments. We thank Daniela Domeisen and Lantao Sun for modeling advice and Paul O’Gorman and Marianna Linz for useful discussions. We are grateful to two anonymous reviewers for their comments, which greatly improved the manuscript. This work was supported by the National Aeronautics and Space Administration through Grant NNX13AF80G. EPG’s contribution was supported by a grant from the National Science Foundation to New York University.

REFERENCES


