Wave Reflection and Focusing prior to the Major Stratospheric Warming of September 2002*

NILI HARNIK
Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York

RICHARD K. SCOTT
Department of Applied Mathematics, Columbia University, New York, New York

JUDITH PERLWITZ
NASA Goddard Institute for Space Studies, and Center for Climate Systems Research, Columbia University, New York, New York

(Manuscript received 29 May 2003, in final form 11 May 2004)

ABSTRACT

Observations of the Southern Hemispheric winter conditions indicate that the major warming of September 2002 resulted from a combination of stationary wave-1 and traveling wave-2 forcing events and suggest that wave and mean-flow anomalies present earlier that winter may have also played a role. Quantities such as the location of the zero wind line, the strength and wave geometry of the vortex, and the horizontal and vertical wave fluxes all differed significantly from climatological values throughout much of the 2002 winter. An analysis of the anomalous features suggests the hypothesis that the persistence of a traveling wave 2 may have increased the likelihood of the combination with stationary wave 1, leading to the observed unprecedented increase in upward Eliassen–Palm flux preceding the warming.

The anomalous conditions of the 2002 winter began as early as mid-May of that year and consisted of a large burst of wave flux into the stratosphere and a strong deceleration of the vortex during its early stage of development. The low-latitude easterly anomaly that resulted from this (unprecedented) event appears to have enhanced the poleward focusing of wave activity in the mid- and upper stratosphere during the rest of the winter. The altered wave geometry of the 2002 vortex allowed internal reflection of traveling wave 2, which helps to explain its unusual persistence during the rest of the winter.

1. Introduction

The September 2002 stratospheric sudden warming was unprecedented in being the first major warming ever to be observed in the Southern Hemisphere (SH). The warming was marked by an upward burst of wave activity that was larger not only than any observed previously in the SH but also than any observed previously in the Northern Hemisphere (NH). During late September 2002, the observed Southern Hemisphere poleward heat flux at 150 hPa averaged over 45°–75° latitude, representative of the upward flux of wave activity through that surface, exceeded 80 K m s^{-1}. In contrast, in all years prior to the warming, the observed poleward heat flux had never exceeded 50 K m s^{-1} in the SH and 60 K m s^{-1} in the Northern Hemisphere.[1] This is remarkable, given the interhemisphere differences in planetary wave forcing, with larger topography and land–ocean contrasts generating larger amplitude planetary waves, and generally larger bursts of wave activity, in the NH than in the SH.

In this paper, we show that the large flux of wave activity in September 2002 in fact arose from a combination of two anomalously large amplitude waves: a traveling wave 2 and a stationary wave 1. A close inspection of the available observational data indicates that such a combination of large upward wave-1 and -2 fluxes has not occurred at any other time in recent years. It is natural to ask, therefore, what caused these two waves to occur simultaneously in 2002.

* Lamont-Doherty Earth Observatory Contribution Number 6619.

Corresponding author address: Dr. Nili Harnik, Dept. of Geophysics and Planetary Sciences, Tel Aviv University, Tel Aviv 69978, Israel.
E-mail: harnik@post.tau.ac.il

[1] These values are derived from data presented in the NASA GSFC Atmospheric Chemistry and Dynamics branch Web site, based on observations from 1979 to the present.

The observations indicate that the SH polar vortex was anomalous in other respects throughout the winter of 2002. For example, as early as May 2002, mid–upper stratospheric subtropical winds were anomalously easterly compared to the climatology. Later, in midwinter, the polar vortex was shifted farther poleward than usual, corresponding to an anomalously small vortex area. In addition to these mean-flow anomalies, observed waves also exhibited unusual characteristics. Heat fluxes were anomalously large in the midstratosphere during most of the winter, indicating greater wave activity. More specifically, an eastward-propagating zonal wavenumber 2 and anomalous wavenumber-2 heat fluxes were observed throughout much of the winter.

The above observations suggest a plausible scenario of events. The preconditioning of the vortex, namely, its poleward shift and reduction in area, almost certainly made it more susceptible to breakup in late September, as a simple consequence of the reduction in angular momentum. Moreover, as we show below, the preconditioning was also intimately connected with the persistence of large wave-2 amplitudes throughout much of the winter. The observed persistence of large wave-2 amplitudes partially answers why waves 1 and 2 occurred simultaneously; a strong pulse of wave 1 at almost any time during late winter would likely have led to such a combination.

2. Data and diagnostics

This study is based mainly on two datasets: the daily mean reanalysis of the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR; Kalnay et al. 1996) and the stratospheric analysis product compiled and distributed by the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC) Atmospheric Chemistry and Dynamics Branch. Both are used from 1979 to 2002. Most of the results presented in this paper are based on the reanalysis, which has a horizontal resolution of 2.5° × 2.5°. We have repeated much of the analysis using NASA GSFC satellite-based wind analyses, which show essentially the same results. We have also made use of European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) to calculate the meridional flux of Ertel potential vorticity (PV) as a way of estimating wave absorption.

To determine whether a given basic state will reflect waves, we use the *wave geometry* diagnostic developed in Harnik and Lindzen (2001). This diagnostic essentially separates the more commonly used index of refraction (e.g., Matsuno 1970) into vertical and meridional components (referred to as vertical and meridional *wavenumbers*). In analogy to the index of refraction, a real (imaginary) vertical wavenumber indicates vertical wave propagation (evanescence).

Harnik and Lindzen (2001) showed that the separation into vertical and meridional wavenumbers can be diagnosed from the steady-state wave solution of a quasigeostrophic model run with the observed basic state [see Eqs. (8) and (9) of Harnik and Lindzen 2001]. The basic state is specified using the NASA GSFC stratospheric analysis daily zonal-mean zonal wind and temperature fields. The data consist of rawinsonde and satellite data in the troposphere and only satellite retrievals in the stratosphere (above 100 hPa in the Southern Hemisphere). The horizontal resolution is 2° latitude × 5°longitude, and the data are available on 18 levels between 1000 and 0.4 hPa. Winds are calculated at GSFC from geopotential height using a balanced wind approximation (Randel 1987). The data start on 26 November 1978 and continue through the present. For more details, see the NASA GSFC Web site (http://hyperion.gsfc.nasa.gov/Data_services/met/about_nmc_data.html). We use this dataset for the wave geometry calculation since the NCEP–NCAR reanalysis does not extend higher than 10 hPa. See Harnik and Lindzen (2001) for details of this calculation.

3. The major warming event

We begin by reviewing the main features of the major warming itself, with emphasis on the vertical wave fluxes into and within the stratosphere immediately prior to the warming. A more detailed account of the nonlinear dynamical evolution of the polar vortex during the warming can be found in Charlton et al. (2005). Because in the SH zonal wavenumber 1 is typically quasi-stationary, whereas zonal wavenumber 2 typically propagates eastward (Leovy and Webster, 1976), we expect the sources and propagation characteristics of these waves in the stratosphere to be different. We therefore calculate the heat fluxes for zonal wavenumbers 1 and 2 separately.

Figure 1 shows the observed 100-hPa heat flux, calculated for zonal waves 1 and 2 (black solid and dashed lines, respectively) along with their sum (shading) and the total heat flux (for all wavenumbers; thick gray line), for the period 1 May–10 October 2002. A number of points are worth noting. First, wave episodes, characterized by bursts of upward wave flux, are typically of either one zonal wavenumber or the other during most of the winter. In late September of 2002, however, both wavenumbers 1 and 2 were large and coincident, resulting in an anomalously large upward wave flux at that time. Furthermore, we also see that the sum of waves 1 and 2 accounts for almost all of the total heat flux during most wave events, including the initial stages of the large burst of poleward heat flux in September.

Figure 2 shows scatterplots of the wave-2 versus wave-1 heat fluxes at 100 and 30 hPa, using daily data from all of the years 1979–2002, with 2002 in black. The September 2002 large event is clearly seen at 100 hPa as

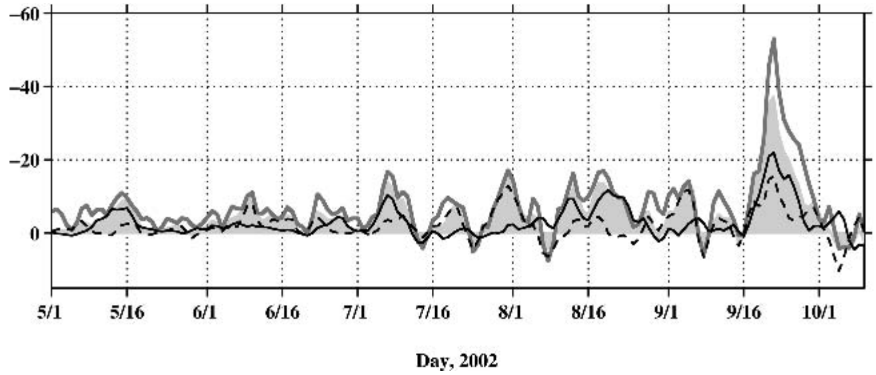

FIG. 1. The 1 May–10 Oct 2002 daily time series of 100-hPa heat fluxes (averaged over 30°–90°S), for wave 1 (solid black), wave 2 (dashed), total (thick gray line), and the sum of waves 1 and 2 (shaded). Units are K m s^{-1}. Note that the vertical axis is flipped, with negative values (upward wave flux) up.

the diagonal extension of the black dots to values that are in the top percentiles of each individual wavenumber. Looking at each wavenumber individually, the September 2002 wave event appears large but not unprecedented (in fact, wave 2 is not the largest observed, and wave 1 barely exceeds previous values). However, the *simultaneous* occurrence of top-percentile values of heat fluxes has never before occurred for both wavenumbers, and the diagonal extension seen in 2002 is clearly anomalous. Note that because of their different propagation characteristics, the two waves do not show the same coincident growth at 30 hPa (or 10 hPa; not shown) as they do at 100 hPa.

Figure 1 raises the obvious question of why such strong pulses of both wavenumbers occurred simultaneously in September 2002. One possibility, of course, is that it was simply a chance occurrence, initiated by some freak combination of tropospheric wave sources, and implying that the major warming was just a random event. In the next section, however, we present observational evidence that indicates that the chances of strong pulses of both wavenumbers occurring simultaneously may have been increased by anomalous conditions of the vortex throughout the winter and, moreover, that those conditions may have been determined by events even earlier.

4. Anomalous conditions prior to the warming

In this section, we describe how the general conditions throughout the winter of 2002 were anomalously

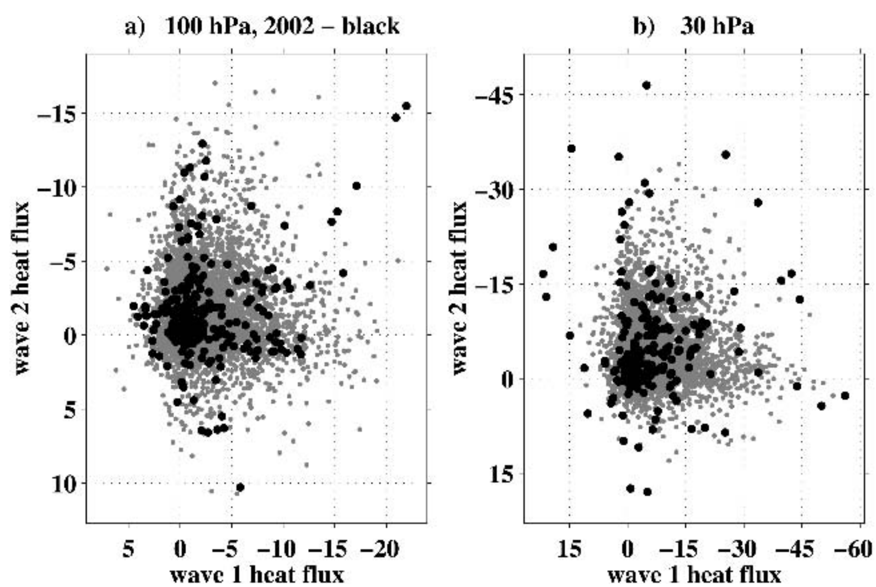

FIG. 2. Scatterplots of wave 2 vs wave 1 daily heat fluxes (averaged over 30°–90°S), at (a) 100 and (b) 30 hPa, for 1 May–10 Oct. The 1979–2001 values are in gray, and 2002 values are in black. Units are K m s^{-1}. Stronger upward propagation (more negative values) are upward and to the right.

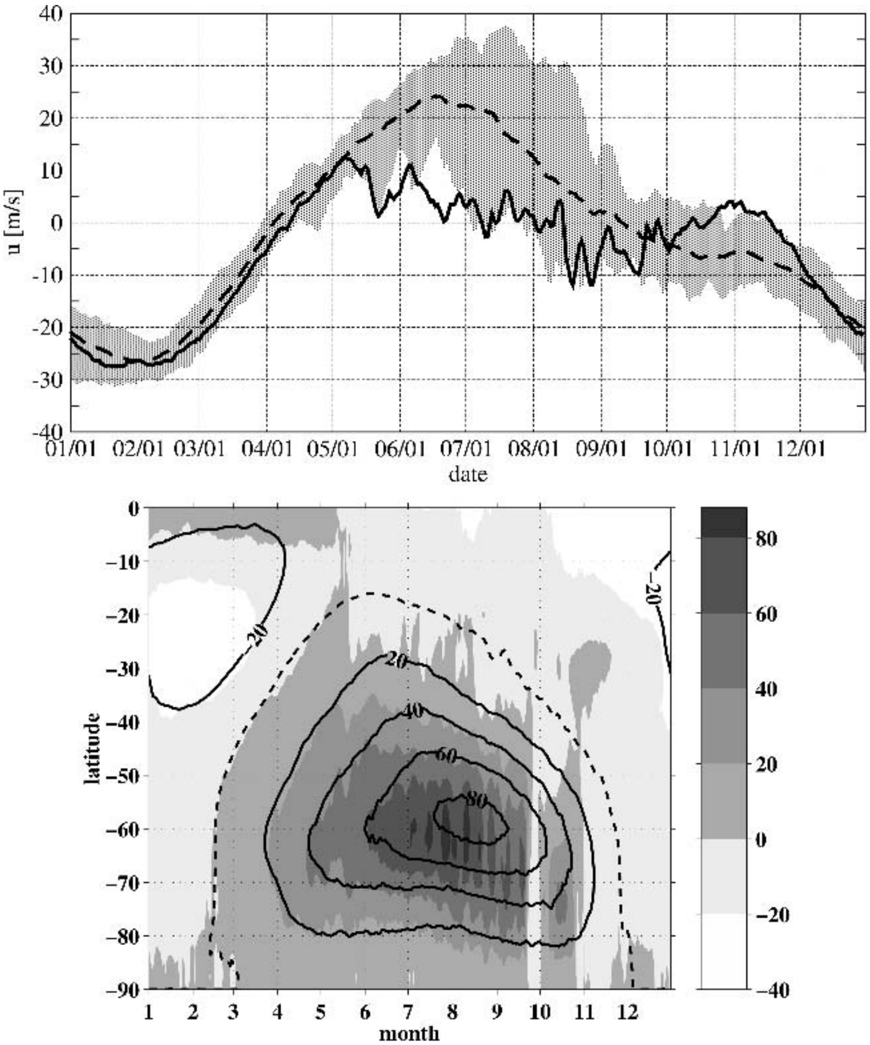

FIG. 3. (top) Time evolution of zonal-mean wind at 10 hPa, 30°S for 2002 (solid) and the 1979–2001 mean (dashed). The shading represents the range of observed values for 1979–2001. (bottom) Latitude–time plot (using daily data) of zonal-mean wind at 10 hPa (m s^{-1}). The 1979–2001 mean in contours (zero line dashed), where the 2002 mean is shaded, with negative values bright and positive values gray.

different from the climatological conditions. We consider the anomalous characteristics of both the wave propagation and the basic-state mean flow on which they propagate.

a. General conditions during the winter

Striking zonal-mean wind anomalies in the winter of 2002 were observed as early as May. Figure 3 shows the zonal-mean wind at 10 hPa at 30°S (referred to below as U_{us}),[2] along with the mean and the range of maximum and minimum wind values observed for 1979–2001. Up until around mid-May, the zonal-mean wind in the subtropics is seen to closely follow the climatology. At that time, however, the zonal-mean flow decelerated rapidly to values lower than those previously observed at this time of year, with anomalously low values of around 0.5 m s^{-1} persisting for much of the winter, decreasing to around -10 m s^{-1} in mid-August. Compared to other years, U_{us} was unprecedentedly low in early winter (mid-May–June) and below average throughout most of the rest of the winter. While the anomalies are even more striking closer to the equator (see Gray et al. 2005 for further discussion of the equatorial wind anomalies), we chose 30°S for this presentation based on the time lag correlations shown later (Fig. 7). Figure 3 (bot-

[2] An average over 30–5 hPa, using the NASA GSFC satellite-based wind analyses yields similar results, suggesting that the anomalies span the upper stratosphere.

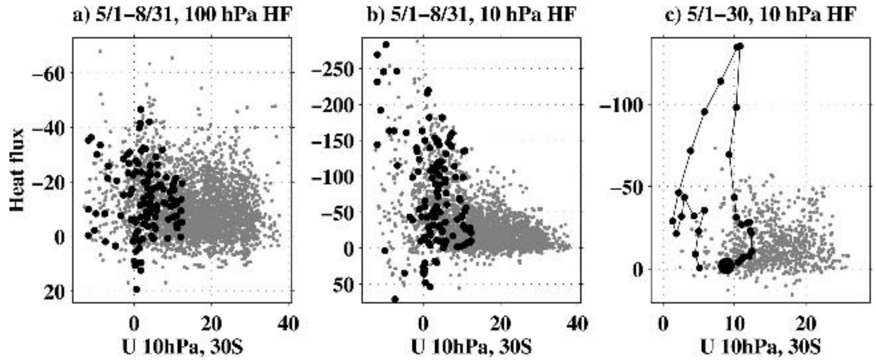

FIG. 4. The 1 May–31 Aug daily values of the (a) 100 and (b) 10 hPa heat flux averaged between 30° and 90° vs the zonal-mean wind at 10 hPa, 30°S. Gray dots are for 1979–2001, black dots are for 2002. (c) The same as in (b), only for 1–30 May. Units for heat fluxes are K m s^{-1}, and for winds are m s^{-1}. Note that the heat flux axis is upside down, so that stronger upward wave propagation (more negative heat flux) is upward.

tom) also shows that the mid-May deceleration is not confined to the subtropics but extends well into the midlattitudes.

Previous studies (Holton and Tan 1980; see also the review by Baldwin et al. 2001, and references therein) have suggested that subtropical winds associated with a particular phase of the quasi-biennial oscillation (QBO) affect the propagation of midlatitude planetary waves, with greater poleward focusing of wave activity when the QBO is in an easterly phase. We now show that the anomalous low-latitude, midstratospheric easterlies that developed in May 2002 were also associated with a poleward focusing of wave activity that persisted throughout the winter. Before discussing the causality of this association in section 4b, we first examine several other diagnostics directly related to the poleward focusing of waves.

Although scatterplots of U_{us} versus the momentum flux, a measure of latitudinal wave propagation, indicate that strong poleward wave fluxes are weakly associated with small U_{us}, momentum fluxes suffer from being a rather noisy diagnostic. A clearer picture of the wave–mean-flow relation is obtained from scatterplots of U_{us} versus the heat flux, a measure of vertical wave propagation. These are shown in Fig. 4 using the heat flux averaged over 30°–90°S at 100 (Fig. 4a) and 10 hPa (Fig. 4b) for the period 1 May–31 August of 1979–2002. The year 2002 is plotted in black (note that the major warming is not included). There is a clear tendency for stronger 10-hPa heat flux when U_{us} is small. In particular, 2002 shows anomalously large heat flux and low U_{us} values. In contrast, the heat flux at 100 hPa does not exhibit such a clear relation to the middle stratospheric subtropical winds, and the 2002 heat flux at 100 hPa is not as anomalously large as at 10 hPa.

Given the large variability in the forcing of waves from the troposphere, a more robust relation may be found by looking at time averages. Figure 5 shows the values of U_{us} (dashed) and the heat flux at 10 hPa (solid) averaged over the months May–August for each of the years in the period 1979–2002. We normalize the 10-hPa heat flux by the heat flux at 100 hPa and the density ratio, so as to isolate the increase in upward wave propagation for a given wave flux entering the stratosphere. Only the months May–August are used to avoid the large contribution to the 2002 mean values from the major warming itself. The observations clearly show anomalously large heat fluxes and anomalously low winds in 2002. Furthermore, over the whole time series, the two quantities are strongly anticorrelated (−0.83). Table 1 shows the correlations of the May–August mean heat flux at 100, 30, and 10 hPa, with the mean U_{us}. We see very large correlations[3] (all except wave 2 at 100 hPa are highly significant) that indicate more upward wave activity fluxes during years when the subtropical winds are anomalously easterly.

Note that while the correlation with the 100-hPa heat fluxes is high, the correlation is still higher at 30 and 10 hPa. In addition, the 2002 100-hPa heat fluxes were not as anomalous as the 10-hPa heat fluxes, normalized or otherwise (Figs. 4 and 5): at 100 hPa, the 2002 fluxes were the third largest after 1996 and 1992, whereas at 10 hPa, they were significantly larger than all other years. This suggests a change in the net vertical Eliassen–Palm (EP) flux divergence into the 100–10-hPa layer. A wave activity budget of the atmospheric box spanning 30°–90°S and 100–10 hPa would require the momentum flux out of this box (at 30°S) and/or the net wave absorption to change. Figure 6 shows a latitude–height plot of the meridional potential vorticity flux (indicative of the EP flux divergence) averaged over May–August for 2002 and the climatology. We see much larger fluxes during 2002. An average over the 100–10-hPa level suggests

[3] Note that a negative heat flux indicates upward wave propagation; hence a positive correlation indicates stronger upward propagation with weaker winds.

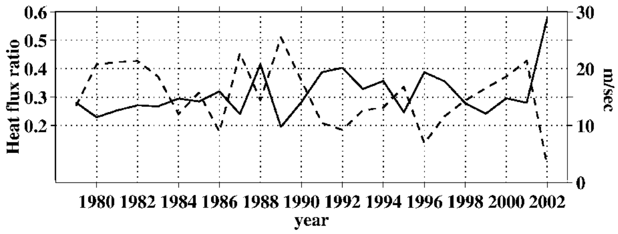

FIG. 5. The yearly time series of the 1 May–31 Aug mean zonal-mean wind at 10 hPa, 30°S (m s^{-1}, dashed) and the normalized 10-hPa heat fluxes (averaged between 30° and 90°). The heat flux is normalized by the value at 100 hPa, times the density ratio ($10\langle v'T'\rangle_{10}/100\langle v'T'\rangle_{100}$, solid).

that the increase is as large as 100%. Note that we find an increase both in the wave absorption and in the vertical flux divergence; hence a decrease in momentum flux is needed to close the wave activity budget.

b. *Causality of the mean-flow and wave anomalies*

It is very difficult to clearly establish the causality of wave and mean-flow anomalies. This is because the waves affect the mean flow and the mean flow affects the waves. One way to disentangle the causality, at least to some extent, is to calculate time lag correlations. Figure 7 shows the time lag correlations of the zonal-mean wind at 10 hPa, 0°–90°S with the heat flux values averaged between 30° and 90°S, at 10 (Fig. 7a) and 100 hPa (Fig. 7b). Daily data for 1 May –31 August of 1979–2002 are used. As expected, we see large correlations (correlations larger than 0.14 are significant at least at the 99% level if the autocorrelation of the time series is considered) for positive time lags (i.e., when the heat flux is leading). This is consistent with waves decelerating the mean flow, so that stronger upward propagation (more negative heat flux) leads to weaker winds. An upward-equatorward wave propagation is consistent with the maximum correlations being more equatorward for the 10-hPa heat fluxes, the time lag being larger for 100-hPa heat fluxes, and the maximum correlations shifting equatorward for larger time lags. We also see, however, a significant correlation between the subtropical winds and the 10-hPa heat fluxes when the wind is leading the waves. This suggests that not only do the waves have a strong effect on the mean flow (which one would expect), but also that the mean flow has a significant effect on the waves, namely, that weaker low-latitude winds *cause* greater upward wave fluxes. The absence of a correlation with the 100-hPa heat fluxes at negative lags supports this argument, because one would expect that the wind would affect wave fluxes at the same level but not below. This effect of the mean flow on the waves is consistent with the Holton–Tan mechanism and suggests a mechanism of why the conditions in 2002 remained so anomalous throughout the winter.

TABLE 1. The correlations between the zonal-mean wind at 10 hPa at 30°S and the total, wave-1, and wave-2 heat fluxes at 100, 30, and 10 hPa (averaged over 30°–90°S). Parentheses denote values that are not significant (using the 95% confidence level). The 99% confidence level is 0.52.

	100 hPa	30 hPa	10 hPa
Total	0.80	0.86	0.87
Wave 1	0.65	0.71	0.75
Wave 2	(0.3)	0.57	0.67

The scatterplots shown in Fig. 4, as well as Fig. 5, are consistent with the subtropical easterly anomalies acting to deflect wave activity from the equator and focus it into the upper-stratosphere high latitudes instead. That is, for a given amount of wave flux entering the stratosphere from below (represented, e.g., by the 100-hPa heat flux), a larger amount of wave activity reaches the mid- to upper stratosphere (represented by the increased 10 hPa heat flux) when the subtropical winds are anomalously easterly. Of course, the scatterplots could also be explained by the waves propagating upward and decelerating the mean flow at each level as they reach it, but only provided the effects of the deceleration were felt quickly, within a day. For example, since waves at 10 hPa have a quick effect on the 10-hPa winds, the deceleration there is rapid and large heat fluxes are therefore always associated with weak winds. Waves at 100 hPa, however, take a few days to reach 10 hPa, which means that their effect on the 10-hPa winds

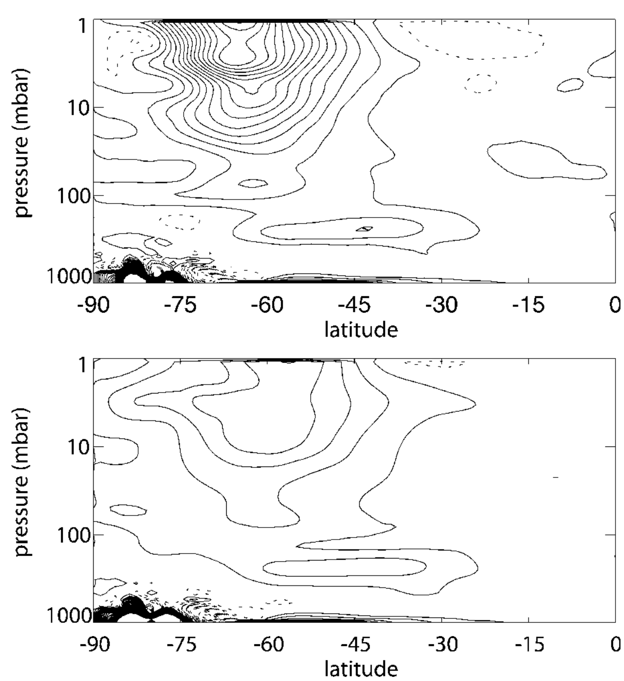

FIG. 6. Latitude–height plots of the meridional flux of Ertel PV, averaged over May–Aug for (a) 2002 and (b) all other years during 1979–2001. Contour interval is the same in the two plots, and negative values are dashed. The calculation was done using ECMWF data.

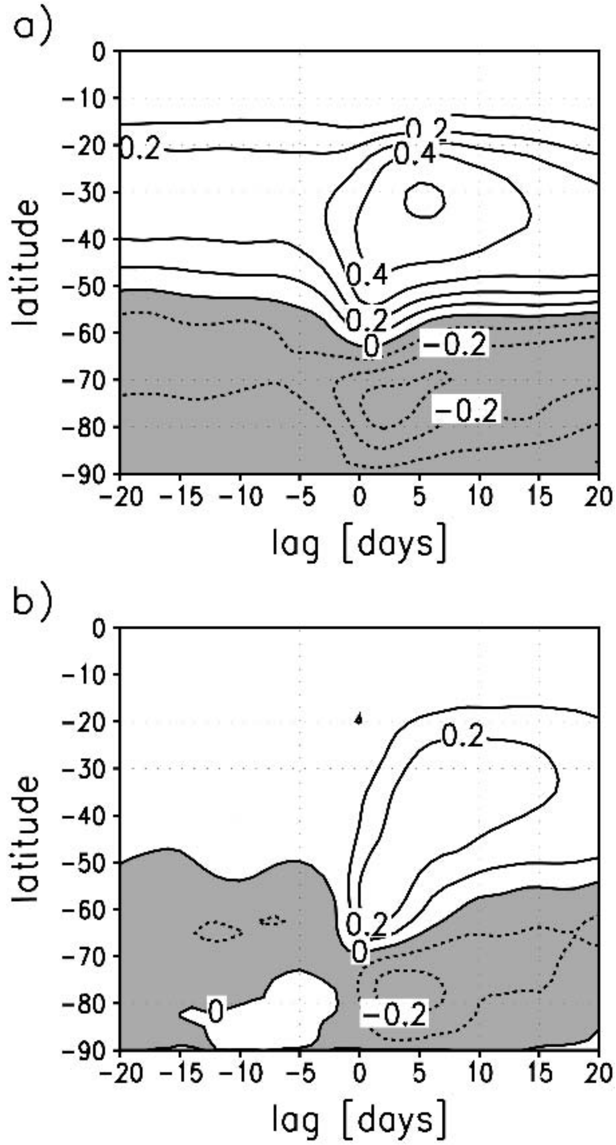

FIG. 7. Time lag correlations of the zonal-mean wind at 10 hPa, 0°–90°S with the heat flux values averaged between 30°–90°S, at (a) 10 and (b) 100 hPa. Daily data for 1 May–31 Aug 1979–2002 are used. Positive lags indicate that the heat flux is leading. The contour interval is 0.1, and negative values are shaded. Correlations larger than 0.14 are significant at the 99% level, if the autocorrelation of the time series is considered.

is not immediate and large heat fluxes are not always associated with weak winds. We suggest, therefore, that both effects, the wave focusing by the subtropical winds and the deceleration of the winds by the waves, contribute to the shape of the scatterplots.

c. Persistence of wave anomalies throughout the winter

The anomalous propagation of waves into the mid- to upper stratosphere will of course affect the polar jet structure. Indeed, the latitude–time plot of 10-hPa zonal-mean wind (Fig. 3, bottom) suggests that while initially 2002 follows the climatological values quite closely, the abrupt wind anomalies that develop in mid-May mostly in the subtropics are followed by a poleward shift of the vortex during June, with episodic weakening preceding the breakup associated with the major warming (see also Fig. 2 of Newman and Nash 2005). In this section, we discuss the effects of changes in the jet structure on the propagation of planetary-scale waves during 2002.

Looking at the heat fluxes, one of the striking features we find in 2002 is the relatively strong downward reflection of planetary waves. This is shown clearly in Fig. 4, as the occurrence in 2002 of large equatorward (i.e., positive) heat fluxes, indicating downward propagation of wave activity. We note that for all years, an equatorward heat flux is seen only for small or negative U_{us}, both at 100 and 10 hPa. Further examination of the heat fluxes shows that 2002 had relatively many days with anomalous positive and negative values, in particular wave 2 (not shown).

Given the suggestion in the data that there was anomalous wave reflection in 2002, we examine the wave propagation characteristics using the wave geometry diagnostic of Harnik and Lindzen (2001; see section 2 of the current paper). It is important to note that the wave propagation characteristics are strongly dependent on the zonal wavenumber and wave period (e.g., Charney and Drazin 1961). For example, reflection of wave 1 in both hemispheres occurs when the upper-stratospheric shear becomes negative either as a result of wave deceleration or as part of the climatological seasonal cycle in Southern Hemispheric late winter (Harnik and Lindzen 2001; Perlwitz and Harnik 2003). On the other hand, we find that zonal wave 2 generally has a reflecting surface somewhere in the stratosphere, even when the upper-stratospheric shear is positive, consistent with the rule that vertical propagation is reduced for waves of smaller horizontal scale (Charney and Drazin 1961). We therefore calculate the propagation characteristics of waves 1 and 2 separately.

The wave geometry calculation requires us to specify a wave frequency. While wave 1 is stationary (zero frequency), wave 2 in the Southern Hemisphere typically propagates eastward (first noted by Harwood 1975; Hartmann 1976; Leovy and Webster 1976). To determine a reasonable value for the wave period, we look at a longitude–time plot of the wave-2 component of the geopotential height field at 10 hPa for 1 June–16 October (Fig. 8). We see a very clear eastward propagation, which lasts for about 2.5 months (July–mid-September), with an average period on the order of 12 days. A similar plot at 30 hPa looks very much like that at 10 hPa, with the same phase propagation speed. This sort of eastward propagation is characteristic of zonal wave 2 in the Southern Hemispheric stratosphere and has been noted in various past studies (e.g., Manney et

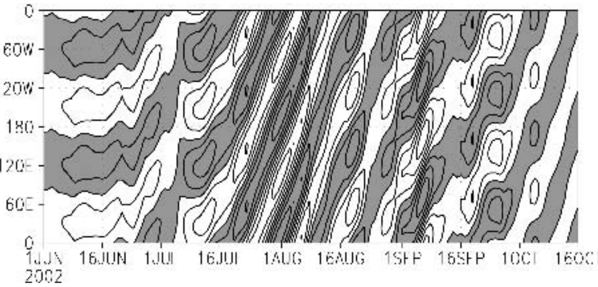

FIG. 8. Longitude–time plot of the zonal wave-2 component of the geopotential height field at 10 hPa, 60°S for 1 Jun–16 Oct 2002. Negative values are shaded, and the contour interval is 300 m.

al. 1991, and references therein). The long persistence of the eastward-propagating wave 2 during 2002 of almost 3 months is unusual but not unheard of (in 1983, a similar persistence of about 2.5 months was observed; Shiotani et al. 1990).

Figure 9 shows a 5-day running mean (daily values show a similar but slightly noisier picture) of the vertical wavenumber of Harnik and Lindzen (2001) for a stationary wave 1 and a traveling wave 2 with a period of 12 days for 3 June–28 September.[4] We see a reflecting surface exists throughout the winter for wave 2, and for wave 1 it forms intermittently, with a gradual downward progression toward late winter.

While it is clear that a reflective basic-state configuration formed for both wavenumbers during 2002, the question remains as to why there were more instances of strong downward propagation during 2002 compared to the climatology (as Fig. 4 seems to suggest). As mentioned above, episodic reflection of wave 1 occurs when upper-stratospheric shear becomes negative due to wave-induced deceleration (Harnik and Lindzen 2001; Perlwitz and Harnik 2003). The enhanced wave propagation and wave drag described above acting, as it did in 2002, on a more poleward-shifted vortex is very

[4] For this calculation, we used the observed zonal-mean wind and temperature fields for 1 June–30 September, using the NASA GSFC analysis (see section 2).

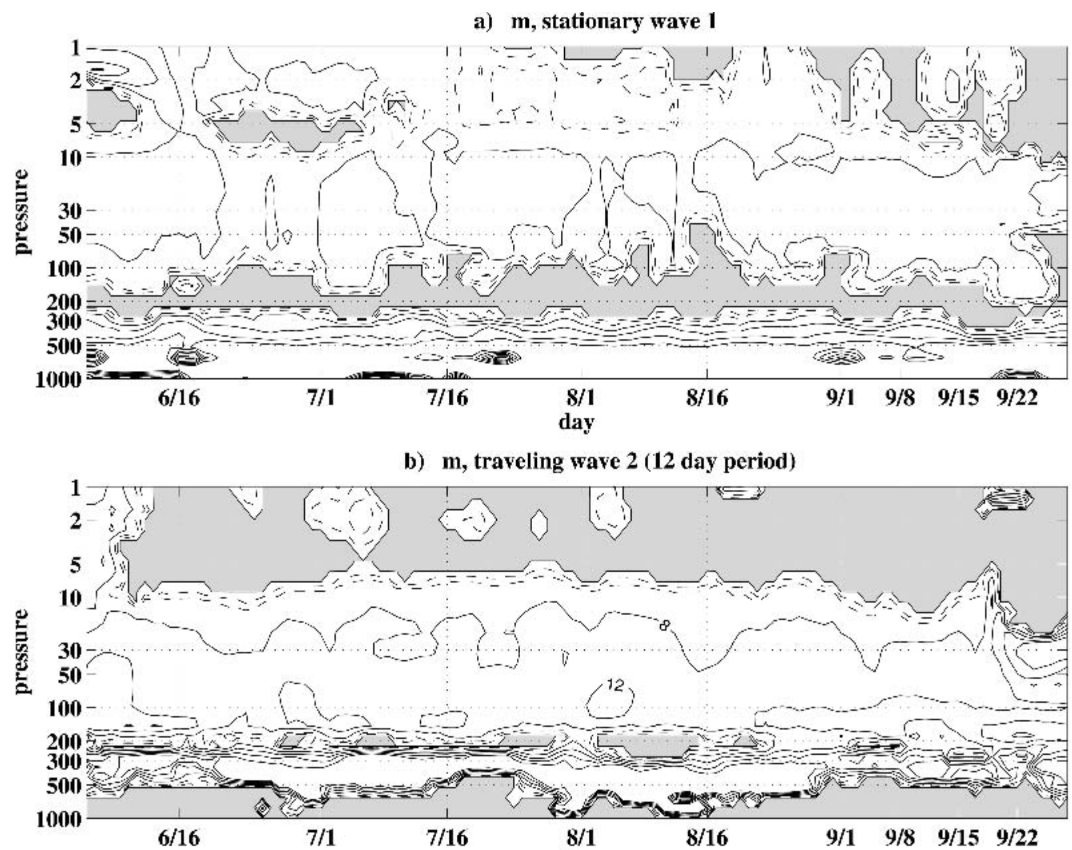

FIG. 9. The 5-day running mean time–height plots of the vertical wavenumber of Harnik and Lindzen (2001), averaged between 58° and 74°S, for (a) stationary zonal wavenumber 1 and (b) wave 2 with a 12-day period. Wavenumbers are calculated from the 1 Jun–30 Sep 2002 daily zonal-mean wind and temperature fields. Contours (with units of 10^{-5} m^{-1}) are at 2, 4 (dashed), and 8–20 in jumps of 4 (solid). Imaginary values (evanescence regions) are shaded.

likely the cause of such enhanced upper-stratospheric deceleration.

For wave 2, which is generally reflected downward, the issue is what controls the height at which the reflecting surface will form. Because both the jet and the zero wind line were shifted anomalously poleward in 2002, the PV distribution had a narrower latitudinal length scale, causing wave modes to have a higher meridional wavenumber. Furthermore, we expect higher meridional wavenumbers to have a lower vertical wavenumber. Because there is always a region where the vertical wavenumber squared goes through zero for wave 2, a reduction in the vertical wavenumber will tend to shift the reflecting surface (i.e., the level above which the vertical wavenumber is imaginary) downward.

Given that the upward wave flux decreases with height due to equatorward refraction, a lower reflecting surface will reflect more wave activity downward. In addition, a calculation of the EP fluxes shows reduced equatorward propagation during 2002 compared to the climatology (both for waves 1 and 2), consistent with the findings of sections 4a and 4b. This is illustrated in Fig. 10, which shows the July–August mean jet, along with the wave-2 EP fluxes, and the region of vertical evanescence for an eastward-propagating wave 2 (shaded; to illustrate the vertically reflecting surfaces) for the 1979–2001 climatology (Fig. 10a) and 2002 (Fig. 10b). The dashed line represents the critical surface, that is, the location at which a traveling wave 2 with a period of 12 days propagates at the same speed as the zonal-mean wind.[5] While in the climatology waves refract equatorward, in 2002 they propagate more vertically and are reflected downward by the more prominent reflecting surface. The reduced equatorward propagation and increased vertical propagation are clear from the EP fluxes. Note that the EP fluxes represent a superposition of upward and downward, poleward and equatorward propagation. While there was more downward propagation in 2002, the upward propagation was also much stronger, with the result that there is a small net increase in vertical EP flux.

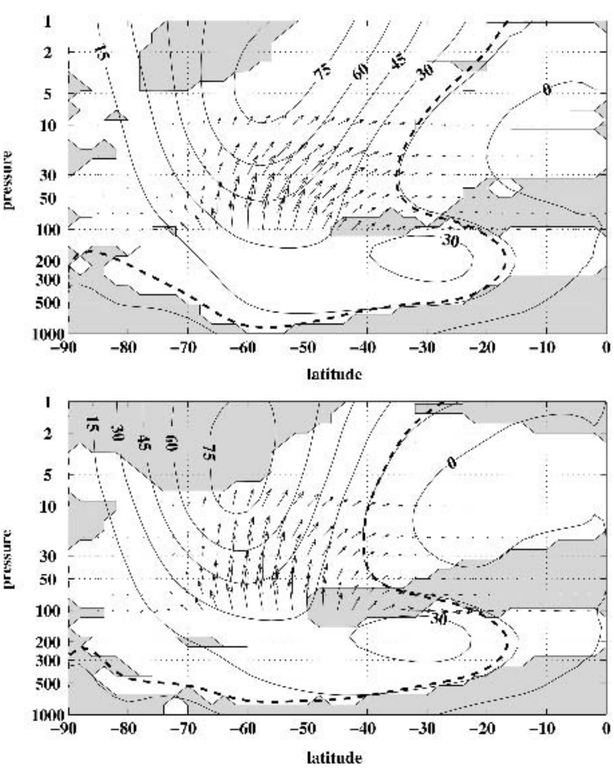

FIG. 10. The Jul–Aug zonal-mean wind (m s^{-1}; solid contours), wave-2 EP fluxes (arrows), vertical evanescence regions for wave 2 with a 12-day period (shading), and the corresponding critical surface (thick dashed contour), for (a) 1979–2001 and (b) 2002.

As described above, the more poleward location of the critical surface throughout the middle stratosphere in 2002 results in a larger meridional wavenumber and, consequently, a smaller vertical wavenumber and a vertically reflecting surface in the upper stratosphere that is both more pronounced and extends farther downward. Thus, whereas a zonal wave 2 propagating on the climatological basic state would be refracted equatorward, the same wave propagating on the 2002 basic state would propagate more vertically, experience the vertically reflecting surface, and be reflected back downward. The reduced equatorward propagation and confinement of wave 2 to the midlatitudes helps explain why large wave-2 amplitudes persisted throughout most of the 2002 winter.

d. The onset of anomalous conditions

An obvious question these observations raise is what caused the wind anomaly in early winter to begin with. Figure 4c shows the scatterplot of 10-hPa heat flux and U_{us} for the month of May. Note that in 2002 there was an unprecedentedly large wave event reaching 10 hPa, with heat fluxes almost double those observed in any previous May. To illustrate the time evolution, the daily values of May 2002 are connected in order of occurrence, with 1 May plotted using a large dot. On 1 May, U_{us} was around 10 m s^{-1}, and the increase in heat fluxes preceded a deceleration of the winds to the anomalously low values (lowest in the record for May) observed later on. As can be seen in Fig. 3 (bottom), the winds at higher latitudes also experienced a rapid deceleration at this time. Such rapid deceleration is reminiscent of the "early winter warmings" that have been observed previously in the SH (e.g, Farrara et al. 1992; also Juckes and O'Neill 1988 for NH events), but never

[5] Although observed wave-2 periods show some variation between years, for simplicity we have also used a period of 12 days to calculate the vertical wavenumber and the critical surface of the climatology for comparison with 2002.

before had the wave fluxes or the deceleration induced by the warming attained the extreme values observed in May 2002.

One possible reason for the subsequent persistence of the easterly anomalies in the subtropical mid- to upper stratosphere might be due to the long memory of momentum anomalies in the subtropics (Scott and Haynes 1998). Another possible reason is the continued wave forcing by the unusually strong wave activity that appears to be intimately connected with the weaker state of the polar vortex throughout the winter. In this scenario, the strong early winter warming in May may have kicked the winter stratosphere into a weak vortex and large wave amplitude regime from which it was unable to recover. What caused the unprecedented large-wave event and the subsequent wind deceleration in May 2002 remains to be understood.

5. Summary and conclusions

We have presented a description of the 2002 SH winter and how this differed significantly from other winters in terms of the zonal-mean flow and wave-1 and -2 disturbances. Our goal has been to highlight some of the features that may have contributed to the remarkable and unprecedented major sudden warming event that took place on 23 September of that year. Because our analysis is restricted to looking at anomalies and correlations in the data, we cannot present direct proof of the causality of events. Rather, we have combined evidence of clear anomalies in the data with simple dynamical arguments, leading to a consistent and plausible explanation of the chain of events.

From our analysis, it appears that the main event that laid the foundation for all of the anomalies that followed, including the major warming itself, occurred as early as 15 May, with a large burst of upward wave activity into the stratosphere. The primary effect of this wave event was the deceleration of the low-latitude winds in the mid- to upper stratosphere. The easterly anomaly that formed as a result was significantly stronger than has been observed in any other year at that point in the seasonal cycle, and the zero wind line was correspondingly shifted significantly farther poleward than normal. Because of the long memory of low-latitude momentum anomalies, the anomaly persisted into midwinter.

The poleward position of the zero wind line in turn may have resulted in enhanced poleward focusing of planetary wave activity (indicated by anomalously poleward EP fluxes), analogously to that described by Holton and Tan (1980) in relation to the easterly phase QBO. This is supported in the data by the time lag correlations of the zonal-mean wind and heat fluxes, which show significant correlations when the subtropical winds lead the heat fluxes. Consistently, anomalously strong wave fluxes in the middle stratosphere were found. The enhanced poleward wave focusing is likely to have contributed to the further erosion of the main vortex, and to a further persistence of subtropical easterly anomalies throughout most of the winter, resulting in a weaker and more poleward vortex by late winter prior to the major warming. This is consistent with a previous modeling study by Scott and Haynes (2002), which found that details of the mid- to late winter vortex evolution were crucially dependent on the strength of the early winter wave-forcing amplitude. Forcing in early winter is thus able to set vortex conditions before the vortex has had a chance to develop to its full strength under radiative effects. Observational evidence for similar regime selection in the SH, again according to the early winter evolution, was presented by Shiotani et al. (1993).

While the focus of our work is the interplay between the stratospheric vortex and stratospheric wave fluxes, an enhancement of tropospheric wave forcing during the winter will also contribute to the erosion of the vortex. While the heat fluxes at 100 hPa were not as anomalous as they were at 10 hPa, they were nonetheless anomalously large (Newman and Nash 2005). Moreover, Newman and Nash show evidence of anomalous tropospheric wave amplitudes, with the most striking anomalies in the lower tropospheric midlatitudes and in the subtropical tropopause region.

An additional striking feature in the stratosphere was the coherent persistence of a traveling wave 2 throughout much of the winter, as indicated, for example, by phase diagrams. We hypothesize that this persistence was facilitated by the formation of a relatively well-defined (compared to the climatology) wave cavity, with a lower vertically reflecting critical surface, a more poleward zero wind line, and a narrower waveguide. This wave geometry was most likely a consequence of poleward shifting and weakening of the jet by the anomalous poleward wave focusing.

By September, two conditions prevailed that could have made the major warming possible. First, there were large traveling wave-2 amplitudes persisting on account of the wave geometry of the mean flow, and second, the vortex was weaker than normal. Whether by chance or by preconditioning of the vortex providing favorable conditions for enhanced upward propagation, the warming itself occurred when a large wave-1 disturbance combined with the existing traveling wave 2. It was the combination of these two wave events that gave rise to the dramatic increase in wave flux into the stratosphere; individually the wave fluxes were large but not unprecedented. In all of the years analyzed here, such a simultaneous combination of wave 1 and wave 2 has never been observed at any other time.

Acknowledgments. The daily mean NCEP–NCAR reanalysis were obtained from the National Oceanic and Atmospheric Administration Climate Diagnostics Center. We thank Dr. A. J. Miller of the Climate Prediction Center at NCEP and Dr. Paul A. Newman of the Atmospheric Chemistry and Dynamics Branch (Code 916) at NASA GSFC for producing the stratospheric analyses. We also thank the Goddard DAAC for distributing the data as part of NASA's Mission to Planet Earth program. The ECMWF ERA-40 data used in this study have been obtained from the ECMWF data server. NH is supported by National Oceanic and Atmospheric Administration Grant UCSIO-CU-02165401-SCF and by a Stork-Doherty Lectureship. Part of the work was done while NH was supported by a Lamont-Doherty Earth Observatory postdoctoral fellowship. RKS received support from the David and Lucile Packard Foundation. JP's research is supported by the NASA Climate and NASA Atmospheric Chemistry and Analysis programs. We thank Alan Plumb for a very useful conversation that encouraged us to study the 2002 warming in more detail. We thank Ted Shepherd and two anonymous reviewers for helpful comments and suggestions.

REFERENCES

Baldwin, M. P., and Coauthors, 2001: The quasi-biennial oscillation. *Rev. Geophys.,* **39,** 179–229.

Charlton, A., A. O'Neill, and W. L. P. Berrisford, 2005: The splitting of the stratospheric polar vortex in the Southern Hemisphere, September 2002: Dynamical evolution. *J. Atmos. Sci.,* **62,** 590–602.

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.

Farrara, J. D., M. Fisher, C. R. Mechoso, and A. O'Neill, 1992: Planetary-scale disturbances in the southern stratosphere during early winter. *J. Atmos. Sci.,* **49,** 1757–1775.

Gray, L., W. Norton, C. Pascoe, and A. Charlton, 2005: A possible influence of equatorial winds on the September 2002 Southern Hemisphere sudden warming event. *J. Atmos. Sci.,* **62,** 651–667.

Harnik, N., and R. S. Lindzen, 2001: The effect of reflecting surfaces on the vertical structure and variability of stratospheric planetary waves. *J. Atmos. Sci.,* **58,** 2872–2894.

Hartmann, D. L., 1976: The structure of the stratosphere in the Southern Hemisphere during late winter 1973 as observed by satellite. *J. Atmos. Sci.,* **33,** 1141–1154.

Harwood, R. S., 1975: The temperature structure of the Southern Hemisphere stratosphere August–October 1971. *Quart. J. Roy. Meteor. Soc.,* **101,** 75–91.

Holton, J. R., and H.-C. Tan, 1980: The influence of the equatorial quasi-biennial oscillation on the global circulation at 50 mb. *J. Atmos. Sci.,* **37,** 2200–2208.

Juckes, M. N., and A. O'Neill, 1988: Early winter in the northern stratosphere. *Quart. J. Roy. Meteor. Soc.,* **114,** 1111–1125.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Leovy, C. B., and P. J. Webster, 1976: Stratospheric long waves: Comparison of thermal structure in the Northern and Southern Hemispheres. *J. Atmos. Sci.,* **33,** 1624–1638.

Manney, G. L., J. D. Farrara, and C. R. Mechoso, 1991: The behavior of wave 2 in the Southern Hemisphere stratosphere during late winter and early spring. *J. Atmos. Sci.,* **48,** 976–998.

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

Newman, P. A., and E. R. Nash, 2005: The unusual Southern Hemisphere stratosphere winter of 2002. *J. Atmos. Sci.,* **62,** 614–628.

Perlwitz, J., and N. Harnik, 2003: Observational evidence of a stratospheric influence on the troposphere by planetary wave reflection. *J. Climate,* **16,** 3011–3026.

Randel, W. J., 1987: The evaluation of winds from geopotential height data in the stratosphere. *J. Atmos. Sci.,* **44,** 3097–3120.

Scott, R. K., and P. H. Haynes, 1998: Internal interannual variability of the extratropical stratospheric circulation: The low-latitude flywheel. *Quart. J. Roy. Meteor. Soc.,* **124,** 2149–2173.

——, and ——, 2002: The seasonal cycle of planetary waves in the winter stratosphere. *J. Atmos. Sci.,* **59,** 803–822.

Shiotani, M., K. Kuroi, and I. Hirota, 1990: Eastward traveling waves in the Southern Hemisphere stratosphere during the spring of 1983. *Quart. J. Roy. Meteor. Soc.,* **116,** 913–927.

——, N. Shimoda, and I. Hirota, 1993: Interannual variability of the stratospheric circulation in the Southern Hemisphere. *Quart. J. Roy. Meteor. Soc.,* **119,** 531–546.