Observational Evidence of a Stratospheric Influence on the Troposphere by Planetary Wave Reflection

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

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

(Manuscript received 15 January 2003, in final form 17 April 2003)

ABSTRACT
Recent studies have pointed out the impact of the stratosphere on the troposphere by dynamic coupling. In the present paper, observational evidence for an effect of downward planetary wave reflection in the stratosphere on Northern Hemisphere tropospheric waves is given by combining statistical and dynamical diagnostics. A time-lagged singular value decomposition analysis is applied to daily tropospheric and stratospheric height fields recomposed for a single zonal wavenumber. A wave geometry diagnostic for wave propagation characteristics that separates the index of refraction into vertical and meridional components is used to diagnose the occurrence of reflecting surfaces. For zonal wavenumber 1, this study suggests that there is one characteristic configuration of the stratospheric jet that reflects waves back into the troposphere—when the polar night jet peaks in the high-latitude midstratosphere. This configuration is related to the formation of a reflecting surface for vertical propagation at around 5 hPa as a result of the vertical curvature of the zonal-mean wind and a clear meridional waveguide in the lower to middle stratosphere that channels the reflected wave activity to the high-latitude troposphere.

1. Introduction
In the last decade there has been a growth in observational and model studies that suggest that variations in the stratospheric mean state caused by natural variability and external forcing might have a significant effect on the tropospheric climate through the dynamic link between the two atmospheric layers (e.g., Kodera 1993; Graf et al. 1994, 1995; Hartley et al. 1998; Shindell et al. 1999a,b; Hartmann et al. 2000; Robock 2000).

The fundamental mechanism of the dynamic troposphere–stratosphere coupling is the upward propagation of planetary waves that are generated in the troposphere by orography and heat sources. These waves then change the stratospheric mean flow when they grow enough to break and be absorbed. Changes in the tropospheric circulation can therefore have a substantial effect on the circulation of the stratosphere. Since wave propagation is on the whole upward from the tropospheric source to the stratospheric sink, an effect of the stratospheric jet on the troposphere is not as straightforward. The stratospheric basic state, however, has a direct effect on the propagation characteristics of the waves (e.g., Charney and Drazin 1961; Matsuno 1970). As a result, zonal-mean flow anomalies in the stratosphere will modify the waves and correspondingly their interaction with the mean flow.

Model studies indicate that the mean flow perturbations that result from the interaction with planetary waves tend to propagate downward and poleward, on a relatively slow timescale of a few weeks (e.g., Hines 1974b; Holton and Mass 1976; Kodera et al. 1996). This has been proposed to be the mechanism behind the time-lagged troposphere–stratosphere correlations in the features of the Northern Hemisphere annular mode (NAM) seen in observations (Baldwin and Dunkerton, 1999, 2001). Model studies by Plumb and Semeniuk (2003), however, point out that a downward migration of NAM-like anomalies does not necessarily imply, in and of itself, a controlling influence of the stratosphere on the troposphere. Rather, a downward migration of the wave–mean flow interaction region can yield a similar signal, with the stratosphere passively responding to the troposphere.

A mechanism for a downward dynamic influence that is based on the fact that localized potential vorticity (PV) anomalies induce geopotential height perturbations non-locally, was suggested by Hartley et al. (1998), Ambaum...
and Hoskins (2002), and Black (2002). These studies have linked changes in the strength of the lower-stratospheric polar vortex to NAM/NAO-like anomalies at the surface [North Atlantic Oscillation (NAO)]. Baum and Hoskins (2002) explain the dynamic link as a geostrophic and hydrostatic adjustment of the atmospheric column due to the associated zonal-mean PV anomalies in the lower stratosphere.

A different direct mechanism by which the stratosphere can affect the troposphere is by reflecting waves that propagate upward from the troposphere. Theoretically, reflection will directly affect the planetary-scale flow in the troposphere (by changing the longitudinal orientation of the planetary wave pattern, and by changing the wave amplitude), and this anomaly may further induce zonal-mean and eddy anomalies. Past studies that have suggested downward reflection include Hines (1974a), Sato (1974), Geller and Alpert (1980), and Schmitz and Grieger (1980), but none of these studies (some of which were based on idealized models) have proven that it has a significant effect on the real troposphere.

In general, it is difficult to find observational evidence for a significant effect of downward reflection on the tropospheric flow. Clearly, one reason is that dynamic processes in the troposphere are the main source for wave disturbances in this atmospheric layer. Also, the dynamic upward influence from the troposphere to the stratosphere strongly dominates. Nevertheless, Perlwitz and Graf (2001, hereafter PG) showed a maximum relationship between stratosphere and troposphere when the stratosphere leads by about 6 days when they isolated zonal wavenumber (ZWN) 1 from the total geopotential height fields. They found such a relationship for winter seasons when the polar vortex in the lower stratosphere is anomalously strong, but not for winter seasons with a weak polar vortex. Motivated by the idea that reflecting surfaces form when the zonal wind exceeds a certain critical value (Charney and Drazin 1961), PG concluded that downward reflection is the source of the maximum correlation when the stratosphere leads. Their analysis, however, did not examine the vertical wave propagation characteristics of the stratospheric basic state; hence there was no mechanistic explanation of the dependence of the results on the strength of the vortex. In an unrelated study, Harnik and Lindzen (2001, hereafter HL) developed and used a diagnostic for studying the geometry of reflecting surfaces during Southern Hemisphere winter. This diagnostic essentially separates the more commonly used index of refraction (e.g., Matsuno 1970) into vertical and meridional components.

In the present paper we combine the statistical approach of PG and the dynamic diagnostics of HL using Northern Hemisphere data in order to determine the basic-state features under which reflection occurs and whether reflection occurs often enough to explain the statistical relationship between tropospheric and stratospheric waves. We concentrate in this study on ZWN 1, which is the dominant wavenumber in the stratosphere. The outline of the paper is as follows. In section 2, we describe the data as well as the statistical and wave geometry diagnostics that we use. In section 3, we examine the vertical wave propagation during the high-winter season (January–March) and show that the wave geometry can indeed explain the statistical signal in the troposphere as resulting from downward reflection of wave energy. In section 4, we investigate under what conditions reflecting surfaces form by comparing fall and high winter as well as by studying interannual variability. In section 5, we discuss the relation of our results to previous findings by PG and conclusions are presented in section 6.

2. Data and analysis approach

a. Description of datasets

This study is based on two datasets: the 4 times daily 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 NASA Goddard Space Flight Center (GSFC) Atmospheric Chemistry and Dynamics Branch.

NCEP–NCAR reanalysis daily mean geopotential height fields at 500, 50, 30, and 10 hPa for the period 1979–2002 are used to study the statistical relationship between tropospheric and stratospheric wave-1 fluctuations. The horizontal resolution is 2.5° × 2.5°. Since the NCEP–NCAR reanalysis does not extend above 10 hPa, we use the stratospheric analysis daily zonal-mean zonal wind and temperature fields, available at 18 levels (1000–0.4 hPa), to calculate the basic-state wave geometry. This dataset consists of rawinsonde and satellite data in the troposphere and only satellite retrievals in the stratosphere (above 70 hPa in the Northern Hemisphere). The horizontal resolution is 2° × 5° longitude–latitude. 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 at http://hyperion.gsfc.nasa.gov/Data_services/met/about_nmc_data.html.

b. Time-lagged SVD analysis

We want to isolate the space–time structures of the related tropospheric and stratospheric wave-1 fluctuations in Northern Hemisphere geopotential height fields, which reveal the process of vertical propagation. Since upward and downward propagation of wave energy is related to westward and eastward phase tilts with height of the wave fields, respectively, the related patterns representing upward propagation and downward reflection will differ from each other (see, e.g., the observed ex-
ample undergoing reflection in Fig. 7 of HL). Consequently, a statistical analysis based on patterns that are isolated from daily tropospheric and stratospheric height fields with no time lag, or from monthly/seasonal averages are not appropriate.

We therefore use a time-lagged SVD analysis—a singular value decomposition (SVD) analysis of the time series of two fields, determined for each time lag \( \tau \) (in days) separately [see Czaja and Frankignoul (2002) for a detailed description]. With this approach a stratospheric height field \( Z_{\text{STRAT}} \) at time \( t \) and a tropospheric field \( Z_{\text{TROP}} \) at time \( t + \tau \) are expanded into \( K \) orthogonal signals, as follows, plus noise:

\[
Z_{\text{STRAT}}(x, t) = \sum_{k=1}^{K} u_k(x) a_k(t) \quad (1)
\]

\[
Z_{\text{TROP}}(x, t + \tau) = \sum_{k=1}^{K} v_k(x) b_k(t + \tau). \quad (2)
\]

The covariance between \( a(t) \) and \( b(t + \tau) \) is the 4th singular value \( s_k \) of the covariance matrix between \( Z_{\text{STRAT}} \) and \( Z_{\text{TROP}} \), and the coupled modes are ordered with decreasing covariance for increasing \( k \) (Bretherton et al. 1992). The total squared covariance (SC) between the two fields can be determined by

\[
\text{SC} = \sum_{k=1}^{K} s_k^2.
\]

To study the direct impact of wave dynamics, both for upward and downward propagation, we use temporally unfiltered daily data but apply a strong spatial filtering in the zonal direction by isolating a specific zonal wavenumber. Note that the zonal-mean height field, which is dominant in the NAM-like spatial patterns (e.g., Thompson and Wallace 1998, 2000), is excluded.

Using the ZWN-1 fields, we carry out a series of 31 SVD analyses with varying time lags \( \tau \) (−15 days, \ldots, 0, \ldots, 15 days) between the time series of the tropospheric and the stratospheric fields. Since we keep the time series of the stratospheric field fixed, a positive (negative) time lag indicates that the time series of the stratospheric (tropospheric) field is leading. To detect at which time lag the maximal relationship between the two wave fields exists, the correlation coefficients of expansion coefficients of the first mode \( a_1 \) and \( b_1 \) for each of the 31 SVD analyses are calculated and combined in a graph \([r_{\text{SVD}}(\text{lag})]\). The same is done for higher modes. The significance of the correlation coefficient between the expansion coefficients is tested by determining their confidence interval taking into account the autocorrelation of the temporal expansion coefficients (Lau and Chan 1983).

We isolate the leading coupled modes on the basis of 90-day seasonal periods from 1979/80 to 2001/02. We study both high winter [January–February–March (JFM)] and fall [September–October–November (SON)]. The sample used to calculate \( r_{\text{SVD}}(\text{lag}) \) consists

\[
\begin{align*}
\text{of 2070 (}=90 \times 23) \text{ realizations. In addition, the SVD analysis for the whole period is used as the basis for analysis of subsamples of the data (composites). We select months (30-day periods) or seasons (90-day periods) of the temporal expansion coefficients for the relevant statistical calculations. These subsamples are selected on the basis of an index that varies monthly or interannually, respectively. We call the correlation coefficients determined this way \( r_{\text{SVD}} \) (i.e., \( r_{\text{SVD}} \) for composites).
\end{align*}
\]

We remove the mean seasonal cycle and apply a square root of cosine latitude weighting prior to the SVD analysis. In order to consider intraannual variability, we also remove the seasonal mean for each year individually. In this way, we exclude the influence of a trend on the correlation coefficients.

c. Wave geometry diagnostics

To determine whether a given basic state will reflect waves, we use the wavenumber diagnostic developed in HL, which we briefly describe here for stationary waves of ZWN 1.

The quasigeostrophic (QG) equation of conservation of PV, in spherical coordinates, linearized about a zonal-mean basic state, yields the following wave equation (e.g., Matsumo 1970):

\[
\frac{a^2 f^2 z}{N^2} \frac{\partial^2 \Phi}{\partial z^2} + \frac{f}{\cos \theta} \frac{\partial}{\partial \theta} \left[ \cos \theta \frac{\partial}{\partial \phi} \left( \Phi \frac{\partial}{\partial f} \right) \right] + \left[ \frac{a^2 q_\theta}{U} - \frac{1}{\cos \theta} \frac{\partial}{\partial \phi} + a^2 f^2 F(N^2) \right] \Phi = \text{damping} \quad (3)
\]

with \( q_\theta \), the meridional gradient of zonal mean PV, given by

\[
q_\theta = \frac{\partial q}{\partial \phi} = \beta - \frac{1}{a_z^2} \frac{\partial}{\partial \phi} \left[ \frac{1}{\cos \theta} \frac{\partial}{\partial \phi} \left( \frac{\partial(U \cos \theta)}{\partial \phi} \right) \right] \quad (4)
\]

and \( \Phi \) is the wave geopotential streamfunction, which is related to the wave geopotential height field \( \phi \): \( \phi = \Phi e^{i\lambda N} / \sqrt{N} \). Here \( \phi, \lambda, \) and \( z \) are latitude, longitude, and log pressure height coordinates, respectively; \( U \) is the zonal-mean wind, \( N \) is the Brunt–Väisälä frequency (a function of zonal-mean temperature), \( \rho \) is density, \( a_z \) is the radius of the earth, \( f \) is the Coriolis parameter, and \( \beta \) its meridional gradient, and \( F(N^2) \) is a function of \( N^2 \) (see HL for exact form).

The first two terms on the left-hand side of Eq. (3) represent vertical and meridional propagation, respectively. The square root of the term in the squared brackets is generally referred to as the index of refraction \( (n_0) \) (Matsumo 1970), which determines the characteristics of the solution, with a sinusoidal form (wave-prop-
agation) in regions where \( n_{ref}^2 \) is positive and an exponential form (wave evanescence) in regions where \( n_{ref}^2 \) is negative. The basic state affects \( n_{ref} \) primarily through the ratio of meridional gradient of PV to the zonal-mean wind.

To determine whether waves can propagate in the vertical or meridional directions, it is necessary to separate \( n_{ref} \) into these two directions. This is trivial if the equations are separable in latitude and height, but in general this is not the case. Harnik and Lindzen developed a method to separate \( n_{ref} \) into meridional and vertical parts, referred to as the meridional (\( l \)) and vertical (\( m \)) wavenumbers. In analogy to \( n_{ref}^2 \), waves propagate in the meridional (vertical) direction in regions where \( l^2 > 0 \) \( (m^2 > 0) \), they are evanescent in the meridional (vertical) direction in regions where \( l^2 < 0 \) \( (m^2 < 0) \), and surfaces of \( l^2 = 0 \) \( (m^2 = 0) \) reflect the waves in the meridional (vertical) direction. They showed that \( l \) and \( m \) can be deduced from the steady-state wave geopotential height by dividing the first and second terms on the left-hand side of Eq. (3) by \(-\Phi \), and taking the real part.

In this study, we calculate \( l \) and \( m \) for a given basic state (from observations) using the spherical QG model of HL except that the lower boundary is at the surface rather than at the tropopause. We have a lid at 105 km, with a sponge layer above about 70 km and equatorward of about 20° latitude, to represent thermal and momentum damping by gravity waves in the upper stratosphere and absorption at the tropical critical surface or radiation to the other hemisphere. We force the model by specifying the amplitude of \( \Phi/f \) at the surface to be constant with latitude and solve for its structure as a function of latitude and height.

We use \( l \) and \( m \) to mainly diagnose the existence and location of reflecting surfaces (where \( l \) or \( m \) vanish). Even though \( l \) and \( m \) are nonlinear functions of the basic state (they are nonlinear functions of \( \Phi \), which is a linear function of the basic state), we find that it makes little difference for the results presented here whether we calculate \( l \) and \( m \) and then average them or calculate them from the averaged basic state. All averages (both in time and space) of the wavenumbers are done by squaring the wavenumber, then averaging, and taking a square root. This way, the averaged wavenumbers remain pure real (propagating) or pure imaginary (evanescent) quantities.

3. Vertical wave propagation during the high-winter season January to March

Perlwitz and Graf (2001) showed that the Northern Hemisphere 500- and 50-hPa geopotential height fields of ZWN 1 are closely related when the time series of the stratospheric field leads the time series of the tropospheric field by about 6 days. This close relationship was found in composites of winter seasons (December–February) characterized by an anomalously strong polar vortex in the lower stratosphere. A characteristic feature of the associated regression patterns for this mode shows an eastward phase shift of the wave pattern at 50 hPa relative to 500 hPa. This feature is consistent with a downward propagation of wave energy into the troposphere, suggestive of downward reflection. The main goal of this section is to find evidence for this mechanism both by extending the statistical analysis to 30 and 10 hPa and by calculating \( l \) and \( m \) of the basic state. We concentrate on the high-winter season JFM for which we find a close relationship when the stratosphere leads the troposphere, even when we consider all years.

Using the time-lagged SVD analysis, we investigate the leading coupled modes between wave-1 height fluctuations at a fixed tropospheric pressure level (500 hPa) and the three stratospheric pressure levels (50, 30, and 10 hPa), individually. We expect the analysis to show the following features if, indeed, downward reflection of waves is depicted: 1) The local maxima in \( r_{SVD}^{l=0} \) at positive time lags \( [r_{SVD}^{max}=0.5] \) occur at larger lags for higher stratospheric levels, consistent with longer propagation times. 2) The associated regression patterns for the coupled modes at \( r_{SVD}^{max}=0.5 \) show a continuous eastward phase shift with increasing height. We expect similar features for the local maxima in \( r_{SVD} \) at a negative time lag (but with a westward, rather than eastward phase tilt with height), consistent with upward wave propagation. Note that unlike an analysis of Eliassen–Palm (EP) flux anomalies, this approach will differentiate between a reduction in the amount of upward propagating waves, and downward reflection. A reduction in the amount of upward propagating wave activity in the total field will affect the value of SC at negative time lags (will most likely reduce it), and will not result in a peak in total SC at positive time lags.

A series of 31 time-lagged SVD analyses is conducted as follows: We keep the 90-day time period for the stratospheric wave-1 field fixed from 1 January to 31 March while the time series of the 500-hPa wave-1 field (Z-ZWN1500) is shifted by \(-15\) to 15 days. That means the tropospheric time series start on 17 December and 15 January for the first and last SVD analysis, respectively. In the following, we describe the results for Z-ZWN110 in detail, and point out differences in the results for Z-ZWN1500 and Z-ZWN110 heights when relevant.

First we look at the squared covariance between the Z-ZWN1500 and Z-ZWN110, as a function of time lag (Fig. 1). The total SC is maximized when Z-ZWN1500 leads Z-ZWN110 by 5–6 days. Another local maximum, which is about 60% weaker, is found when Z-ZWN1500 leads by about 6 days. It is also clear that the total SC of Z-ZWN1500 and Z-ZWN110 can be explained by two coupled modes in which the first coupled mode clearly...
dominates. This very small number of modes describing the vertical coupling between wave-1 fluctuations in the troposphere and stratosphere is not surprising, taking into account that a wave of a single zonal wavenumber is described at each latitude by two parameters (e.g., the amplitude and the phase). Note that both for Z-ZWN1500 and Z-ZWN110, the spatial patterns of the first two modes are similar but with a phase shift of 90° in the zonal direction.

In order to illustrate the strength of the relationship between the coupled modes at various time lags, the correlation coefficients \[ r_{SVD}(\text{lag}) \] between the temporal expansion coefficients of the leading mode for each of the 31 SVD analyses are determined. Figure 2 displays the results for the three stratospheric levels. The \[ r_{SVD} \] for Z-ZWN110 shows a maximum of about 0.55 at a time lag of -5 to -6 days (i.e., the troposphere leads). A second local maximum of about 0.36 is found when the stratosphere leads by 6-7 days. The two maxima are highly significant since the 99% confidence level is 0.12.

A similar double-peaked structure of \[ r_{SVD}(\text{lag}) \] is also found for the other stratospheric levels (Fig. 2, the dotted line for 50 hPa and solid line for 30 hPa). Consistent with a propagation of the wave signal, the time lag at which these peaks occur increases with the vertical distance between the tropospheric and stratospheric fields.

Table 1 summarizes the phase shift with altitude of wave 1 at 65°N of the associated regression patterns. The phase shift is determined for the coupled modes corresponding to \[ r_{SVD}\text{lag}<0 \] and \[ r_{SVD}\text{lag}>0 \] of Fig. 2. Figures 3a and 3b show the associated regression patterns of the leading coupled modes determined for the time lag of -6 and 6 days, respectively. As expected, the associated regression patterns for these two lags exhibit a different structure. For the leading mode at a lag of -6 days, the ridge of wave 1 is shifted westward at the 10-hPa level relative to the 500-hPa level. At 65°N the phase shift amounts to -132.5°. By comparing the phase shift between the three analysis series at 65°N (Table 1), a continuous westward phase tilt of the waves with increasing altitude is clearly found.3

The striking feature in the regression patterns for the +6 day lag is an eastward phase shift with increasing height (Fig. 3 and Table 1). The phase shift of the wave-1 ridge increases from 62.5°E at 500 hPa to 107.5°E at 10 hPa, consistent with a downward reflected wave. In addition, the 500-hPa wave-1 pattern associated with the wave-1 fluctuation at 10 hPa six days earlier also shows a large eastward phase tilt with increasing latitude (Fig. 3b, left panel). This meridional structure and its relation to the basic-state structure in the troposphere needs to be studied in further detail.

So far, we have found that the relationship between

3 We note that the phase shift between the various levels is the same for a range of time lags around the peaks. This suggests that the implied vertical phase shift is a robust qualitative representation of the wave vertical phase structure (and not an artifact of longitudinal wave propagation that may occur on time scales shorter than the time lag of the correlation peaks).
The phase difference \( d \lambda \) (in °E) at 65°N between the associated ZWN1 patterns at 500 hPa and various stratospheric levels (50, 30, and 10 hPa). Negative and positive values indicate westward and eastward phase shifts with height, respectively. The phase differences were determined from the associated patterns of the leading coupled mode at the time lags of \( r_{\text{max} \text{lag}=0} \) and \( r_{\text{max} \text{lag}=0} \) as indicated in Fig. 2.

<table>
<thead>
<tr>
<th>SVD analysis</th>
<th>( r_{\text{max} \text{lag}=0} )</th>
<th>( r_{\text{max} \text{lag}=0} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 hPa; 10 hPa</td>
<td>Lag (days) (-6)</td>
<td>Lag (days) (6)</td>
</tr>
<tr>
<td></td>
<td>(-132.5)</td>
<td>(107.5)</td>
</tr>
<tr>
<td>500 hPa; 30 hPa</td>
<td>Lag (days) (-5)</td>
<td>Lag (days) (5)</td>
</tr>
<tr>
<td></td>
<td>(-105.0)</td>
<td>(85.0)</td>
</tr>
<tr>
<td>500 hPa; 50 hPa</td>
<td>Lag (days) (-4)</td>
<td>Lag (days) (4)</td>
</tr>
<tr>
<td></td>
<td>(-80.0)</td>
<td>(62.5)</td>
</tr>
</tbody>
</table>

The time series of tropospheric and stratospheric wave-1 fluctuations, as given by a time-lagged SVD analysis, is maximized when either of the time series leads by several days. The peak, when the troposphere leads, is clearly due to upward propagation of planetary waves, which has been extensively studied (e.g., Randel 1988). Next, we use the wave geometry diagnostic of section 2c to examine the propagation characteristics of the basic state. Existence of a reflective basic-state configuration will indicate that downward reflection is a possible cause for the maximal relationship between the waves in the stratosphere and those in the troposphere a few days later.

Figure 4 shows the JFM vertical and meridional wavenumbers, calculated from the monthly mean basic state for each of the JFM months during this time period.

**Fig. 3.** Heterogeneous regression patterns (gpm) of the leading coupled mode of the Z-ZWN1_{500} and Z-ZWN1_{10} when (a) the Z-ZWN1_{500} leads the Z-ZWN1_{10} by 6 days and when (b) the Z-ZWN1_{10} leads the Z-ZWN1_{500} by 6 days. These maps are determined by regressing the time series of the Z-ZWN1_{500} (Z-ZWN1_{10}) height fields onto the temporal expansion coefficients of the leading mode of Z-ZWN1_{500} (Z-ZWN1_{10}).
The shading indicates regions of wave evanescence. Looking at the meridional wavenumber below 10 hPa, we see two waveguides, a high-latitude waveguide centered around 60°N, which is very well defined between 100 and 20 hPa, and a subtropical–midlatitude one centered around 30°N, which is clearest below 40 hPa (Fig. 4b). The high-latitude waveguide is clearly the one relevant for upward propagation of waves to the stratosphere, consistent with the observation that planetary waves peak at high latitudes in the stratosphere (e.g., Geller et al. 1983). Looking at the vertical wavenumber (Fig. 4a) in the region of this high-latitude meridional waveguide, we see a hint of an evanescent region (poleward of 60°N) above 10 hPa (dashed lines denote small values). Since the wavenumbers are an average over all high-winter months, this hint of evanescence suggests there may be a clear reflecting surface during some but not all of the months, as a result of the large interannual variability in the stratosphere during Northern Hemisphere winter.

4. Variability in the basic state and its effect on wave statistics and wave geometry

The results of the previous section suggest that during JFM the observed statistical signal of downward influence may possibly be explained by a wave geometry configuration of a reflecting surface above 10 hPa, with a well-defined high-latitude waveguide in the lower stratosphere along which the waves propagate. In this section, we determine more explicitly what configuration of the stratospheric basic state gives reflection.

In their study of wave reflection in the Southern Hemisphere, HL found that reflecting surfaces formed towards late winter as part of the climatological seasonal cycle and occasionally in mid winter as a result of strong wave deceleration. According to Eq. (3), the basic state affects $n_{ref}$ primarily through the ratio of meridional gradient of PV to the zonal-mean wind. In the cases they analyzed, HL found that $l$ assumed a structure that was broadly similar to the meridional structure of $q_y$, with meridional reflecting surfaces occurring where $q_y$ was small or negative as a result of large meridional curvature [second term on the right-hand side of Eq. (4)]. Vertical reflecting surfaces for wave 1 formed when the vertical curvature [third term on the right-hand side of Eq. (4)] caused $q_t$ to become small or negative (see Fig. 8 of HL), which in practice occurred when winds decreased with height in the upper stratosphere.

Assuming these relations hold for the Northern Hemisphere as well, we look for periods during which the jet at high latitudes peaks in the middle or upper stratosphere. Unlike the Southern Hemisphere, the climatological jet does not shift downward in the Northern Hemisphere, but during winter (starting around December), when the planetary waves are largest, it does exhibit a large intraseasonal and interannual variability.

Starting with the seasonal variability, Fig. 5 compares...
the winter (JFM) and fall (SON) zonal-mean wind averaged over 1979/80–2001/02. The main difference in the stratosphere is in the strength of the winds at high latitudes (Fig. 5c). In the lower stratosphere, the JFM jet is stronger by about 7 m s\(^{-1}\) at 30 hPa, while in the upper stratosphere (above 5 hPa) the jet is stronger in fall.

Note that the maximum jet strength does not change much between the seasons—in both the jet reaches about 30–35 m s\(^{-1}\) at 1 hPa, and it seems to peak above the domain of observations. The relevant difference for wave reflection therefore seems to be the vertical curvature at high latitudes in the upper stratosphere rather than the strength of the zonal wind.

To examine whether these differences in the basic state are important for the wave dynamics, we repeat the time-lagged SVD analysis for the SON data (Fig. 6). We keep the time series of the stratospheric field fixed, from 1 September to 29 November for each year (1979–2001). As we find during JFM (Fig. 2), all three stratospheric levels have a maximum in the correlation \( r_{\text{SVD}} \) when the tropospheric field leads the stratosphere by several days. The most striking difference is the lack of a second local maximum when the stratosphere leads, which suggests there is no significant downward reflection of wave 1 during fall. A more subtle difference is the shorter time lag at which the maximum correlation occurs—it occurs 2 days earlier in SON compared to JFM, suggesting the planetary waves propagate upward faster during fall. We discuss this difference later.

Figure 7 shows the vertical and meridional wavenumbers calculated by averaging over the wavenumbers of the monthly mean basic states for each of the SON months during this time period. We see that, consistent with the statistical analysis, waves can propagate upward throughout the stratosphere. There is an evanescent region near the tropopause, at all latitudes, which we expect is narrow enough for part of the wave energy to tunnel through (part will reflect back). This might be a contributing factor to observed wave amplitudes being smaller during fall compared to winter (see, e.g., the climatology by Randel 1992). Such a region of evanescence for vertical propagation should decrease the correlation between the stratospheric and tropospheric wave-1 fluctuations for both propagation directions. In addition, the meridional wavenumber does not show the clear high-latitude waveguide in the lower/mid stratosphere (Fig. 7b), that is found during JFM between 100 and 20 hPa (Fig. 4b). Instead, the high-latitude and subtropical waveguides merge to one wide waveguide above 100 hPa. Note that, when waves reflect downward, a well-defined high-latitude meridional wave-
guide helps to produce a strong statistical signal in the troposphere, since without it the downward reflected waves can disperse in the meridional direction and get absorbed in the subtropical zero wind line. 

Figures 4 and 7 suggest that the vertical wavenumber averaged over the high latitudes (58°–74°N) and the meridional wavenumber averaged over the lower stratosphere (13.5–24 km) provide useful measures for the reflective configuration of the stratospheric basic state for ZWN 1. We use these measures to illustrate interannual variations during both seasons. We calculate the wavenumbers from the time-mean basic state for each of the 3-month periods. Figure 8 shows the time–height plot of the high-latitude average of the vertical wavenumber, as well as the time–latitude plot of the lower-stratosphere meridional wavenumber for SON (Figs. 8a,b) and JFM (Figs. 8c,d). We see that during SON there are no reflecting surfaces for vertical propagation, while in JFM reflection forms during about half of the years. In addition, during SON of most years there is one wide meridional waveguide extending from the subtropics to high latitudes. In JFM, on the other hand, except for a few years, there are two waveguides, a narrow subtropical one and a wider high-latitude one.

The large degree of interannual variability during high winter motivates us to define an index in order to isolate years with reflecting surfaces. Since the present results, as well as those of HL, suggest that reflection forms when the stratospheric jet has a peak in the stratosphere, we define our index to be the average vertical wind shear in the high-latitude upper stratosphere. We use $U(2–10) = \langle U \rangle (2 \text{ hPa}) - \langle U \rangle (10 \text{ hPa})$, where $\langle U \rangle$ is the zonal-mean wind, averaged over 58°–74°N and over time. We expect reflection to form when the index is negative. Figure 9 shows the interannual variations of the index for JFM and SON. We see that during SON the index is always positive (long-term mean amounts to about 10 m s$^{-1}$). During JFM, on the other hand, the index exhibits a much larger interannual variability and oscillates between positive and negative values with the mean value being near zero. We note that during most of the 1990s, the index was negative. During these years, we also see a formation of reflecting surfaces in the upper stratosphere (Fig. 8c), consistent with our results so far.

Due to the high variability in the Northern Hemisphere stratospheric winter circulation, the vertical wind

\[ \frac{f^2}{H \alpha_z} \frac{\partial U}{\partial z} \overline{f^2} \frac{\partial^2 U}{\partial z^2} \overline{\alpha_z}, \]

where $H$ is the density scale height. A simple scale analysis suggests that the vertical curvature at the jet maximum will contribute positively to the PV gradient, while the negative shear above the jet peak and the curvature above that (near the implied jet minimum, which is often above the domain of observations) will both contribute negatively. The negative shear above the jet peak, therefore, appears to be the main cause for the reflecting surface, which is located in a region where the curvature is smallest. Note also that the actual value of $U$ contributes positively to $\alpha_z$, so a small or negative $q_v$ is necessary for small or negative $\alpha_z$.
Fig. 8. (a) The vertical wavenumbers, calculated from the SON mean basic states for the years 1979–2001, averaged over 58°–74°N. (b) The meridional wavenumbers, calculated from the SON mean basic states for the years 1979–2001, averaged over 13.5–24 km. (c),(d) Same as (a),(b), respectively, but for JFM 1980–2002. Vertical wavenumber contours (units of $10^{-5}$ m$^{-1}$) are shown at 0.01 (thick line); 2, 4 (dashed); and 6, 8, 10, 15, 20 (solid). Meridional wavenumber contour interval is 1 rad$^{-1}$, and the 0.01 line is thick.

shear in the high-latitude upper stratosphere varies considerably also within individual winter seasons. This is illustrated in Fig. 10, which shows $U(2–10)$ based on monthly means, for the winter months (JFM) from 1980 to 2002. To isolate more clearly the configuration for the formation of reflecting surfaces in the upper stratosphere for this season, we study composites for all JFM months with an index higher (lower) than 0.5 $\pm 0.5$ standard deviations ($\sigma$). From Fig. 10 it can be seen that 22 (21) high-winter months are included in the composite for the positive (negative) index $U(2–10)$.

First, we study the composites of the basic-state zonal-mean wind (Fig. 11). We note that basic states for the positive index (Fig. 11b and dashed line in Fig. 11c) exhibit a vertical structure similar to the basic state of fall (Fig. 5b and dashed line in Fig. 5c). The negative
index composite, on the other hand, shows a maximum around 30 hPa (Fig. 5a and solid line in Fig. 5c).

Next, we examine the wave propagation characteristics illustrated by the average vertical and meridional wavenumbers (Fig. 12). We see a clear difference between the positive and negative index states. During negative index months a reflecting surface forms below 5 hPa poleward of 50°N, along with a very clear meridional waveguide below 10 hPa between 50° and 80°N. During the positive index months, on the other hand, waves can propagate all the way up through the stratosphere.

Finally, we examine the relationship to the troposphere by calculating the correlations using only the months with a positive or negative index. That means that for each time lag, we use the daily temporal expansion coefficients determined on the basis of the SVD analyses series between Z-ZWN1500 and Z-ZWN1 10 for the whole winter season (JFM) and select data for months for which $U(2\text{--}10)$ is $0.5 \sigma$ ($-0.5 \sigma$). Then, we determine the correlation coefficients $r_{\text{SVD--C}}$ between the selected time series for these composites (Fig. 13).

The main effect of separating the data into positive and negative index states is that the peak at stratospheric lead time disappears for the positive index case. In addition, this peak is more pronounced in the negative index months, compared to the analysis for all JFM (Fig. 2). These differences are consistent with the wave geometry.

Separating the data also affects the structure of $r_{\text{SVD--C}}$ for tropospheric lead times with the maximum correlation being smaller and at a shorter time lag for the nonreflective state (positive index case). Note that the maximum correlation for the nonreflective state is also much smaller than for SON (Fig. 6), which is also nonreflective. A possible reason might be the large abundance of major sudden warmings during the positive index years (see discussion in section 5), which will affect the wave structure in the stratosphere.

The difference in propagation time between the reflective and nonreflective states (which is also found when comparing JFM to SON) cannot be explained as the change in the vertical group velocity of the waves due to the changes in the basic state, since a rough estimate suggests vertical propagation should be faster for the reflective state. A possible explanation, which needs further testing, might be that downward reflection of waves causes the wave phase lines to be more vertical and results in an apparent increase in vertical propagation time (see Harnik 2002).

5. Discussion

In the previous sections we have combined statistical and dynamic diagnostics to study the formation of re-

It is easy to show (e.g., Karoly and Hoskins 1982) that the vertical group velocity of a Rossby wave, under Wentzel–Kramers–Brillouin (WKB) conditions is: $C_r \propto U^2 \nu q / N^2$. 


Fig. 9. The time series of seasonal mean $U(2\text{--}10)$ for (a) JFM and (b) SON. The shading indicates the region of the long-term average $\pm 0.25$ std devs.

Fig. 10. The time series of monthly mean $U(2\text{--}10)$ of JFM. The shading indicates the region of the long-term average $\pm 0.5$ std devs.
reflecting surfaces in the Northern Hemisphere stratosphere and their impact on tropospheric waves. Results from the two types of diagnostics consistently show evidence of reflection during high winter (JFM), but not during fall (SON). There is considerable interannual variability during JFM, so that reflection does not form during all winter seasons, but it does form frequently enough to have a significant statistical signal when using all high-winter seasons during 1979–2002. Both diagnostics, however, show much clearer evidence for reflection and its effect on the troposphere when we separate the data corresponding to months/seasons with reflective and nonreflective basic states. The basic-state configuration that shows reflection is characterized by a decrease with height of the zonal-mean wind in the upper stratosphere. Correspondingly, the mean wind shear between 2 and 10 hPa at high latitudes (averaged over 58°–74°N) is a good index for separating reflective from nonreflective basic states, with negative values indicating reflection and positive values indicating no reflection. A representation of the reflective state by the vertical wind shear is consistent with the findings of HL for the Southern Hemisphere.

These results seem to differ from previous findings of PG, who divided the 90-day winter seasons into strong and weak polar vortex conditions based on the 50-hPa index and found the statistical evidence for downward reflection only for the strong vortex case. Perlwitz and Graf (2001) motivated the separation into strong and weak vortex states by the idea that reflecting surfaces form when the zonal wind exceeds a certain critical value (Charney and Drazin 1961). We have found, on the other hand, that in the real atmosphere reflecting surfaces for wave 1 do not form when the winds exceed a certain critical value but, rather, when the meridional gradient of PV becomes weak or negative, as a result of the vertical curvature associated with negative shear in the upper stratosphere (see footnote 4).

To resolve this apparent inconsistency, we define another index, \( U(30) \), which represents the strength of the polar night jet in the lower stratosphere (the zonal mean wind at 30 hPa averaged over 58°–74°N) and compare the corresponding basic states and the statistical diagnostics for \( U(30) \) and \( U(2–10) \). Looking at the seasonal means (JFM), we find that the two indices agree very well with each other during a majority of the years, with their correlation coefficient being 0.74. Figure 14 shows the zonal-mean wind profiles averaged over high latitudes for the positive and negative composites (based on 0.25\( \sigma \)) for both indices, as well as the correlations \( r_{\text{SVD-c}} \) for the reflective configurations [negative \( U(2–10) \) and high \( U(30) \)]. We see that the two indices reflect similar variations in the basic state such that, when the JFM vortex has a negative shear in the upper stratosphere, it is also anomalously strong in the lower stratosphere and vice versa (Fig. 14a). Consistently, \( r_{\text{SVD-c}} \) for both indices captures the reflection very clearly (Fig. 14b). This suggests that PG succeeded in separating the basic states into reflective and nonreflective configurations simply because a strong lower-stratospheric polar night jet coincides with a negative vertical wind shear in the upper stratosphere, at least on seasonal timescales.

The fact that both indices pick out similar basic-state configurations for the seasonal means is consistent with the notion that reflection and absorption are more or less mutually exclusive (when waves are reflected downward, a large portion of their energy returns to the troposphere rather than being absorbed in the stratosphere), so the vortex tends to be stronger during seasons with a lot of reflection due to less wave deceleration. We note that for the seasonal-based index (Fig. 9a), a major sudden warming (as defined by Labitkze and Naujokat
2000) occurred during all JFM with a positive $U(2−10)$ and did not occur during any of the JFM with negative $U(2−10)$ except for the year 1986 in which a major warming occurred at the end of the season (in March). This is consistent with results of Giannitis (2001), who studied the mechanisms by which wave amplitudes are limited in the upper stratosphere using a nonlinear quasigeostrophic model. Giannitis found that wave-mean flow interactions limited wave amplitudes, either by forming a reflecting surface or by forming a critical surface that then absorbed the waves (a sudden warming), and the initial strength of the zonal-mean wind was the main factor determining which of these two processes occurred.

While there is more physical basis for using the $U(2−10)$ index, one might be tempted to use the $U(30)$ index for a more comprehensive study of the effects of downward reflection and its variability because the obser-
Fig. 13. The r_{SVD,C} (lag) based on the subset of the temporal expansion coefficients determined from the time-lagged SVD analyses between Z-ZWN1500 and Z-ZWN1100 for JFM. The subset is chosen for negative or positive $U_{(2-10)}$ index months. See text for details. A positive time lag indicates that the stratospheric field is leading.

Fig. 14. (a) Composites of zonal-mean zonal wind averaged between 58° and 74°N (m s$^{-1}$) for the JFM seasonal mean index $U_{(2-10)}$ and $U_{(30)}$ (based on 0.25σ). (b) The $r_{SVD,C}$ (lag) based on the subset of the temporal expansion coefficients determined with time-lagged SVD analyses between Z-ZWN1500 and Z-ZWN1100 for JFM. The subsets are chosen for the JFM seasonal mean indices $U_{(2-10)} < -0.25$ and $U_{(2-10)} > 0.25$. A positive time lag indicates that the stratospheric field is leading.

This study has concentrated on ZWN 1, but we also applied the statistical and wave geometry diagnostics for wave 2, which is also observed in the stratosphere. While we find reflecting surfaces for ZWN 2 during fall and high winter, we do not find evidence for a second $r_{SVD}$ peak with the stratosphere leading the troposphere. A possible reason for the differences from wave 1 is that ZWN 2 takes only about 2 days to propagate into the midstratosphere (Randel 1988), making it harder to statistically separate the downward from the upward propagation. In addition, we find very pronounced regions of evanescence for upward propagation around the tropopause (which is even more pronounced than for wave 1 during SON, Fig. 7a). Partial reflections from this region might further interfere with the statistical signal.

6. Conclusions

This study gives clear evidence that reflecting surfaces develop in the stratosphere during some years in winter in the Northern Hemisphere and that there is a significant statistical signal of the stratospheric wave 1 related to the tropospheric wave 1 several days later. The combination of these results strongly supports the idea that downward reflected waves reach the troposphere and modify the structure of planetary waves there.

The analysis of the basic states that shows both the existence of reflecting surfaces and a strong downward statistical signal indicates that there is one basic configuration of the stratospheric jet that reflects waves back to the troposphere—when the polar night jet peaks in the high-latitude midstratosphere. The strong reflec-
tion signal in the troposphere during these times is a result of two coincident features—the formation of a reflecting surface for vertical propagation at around 5 hPa and a clear meridional waveguide in the middle and lower stratosphere that channels the reflected wave activity to the high-latitude troposphere (essentially preventing dispersion of the waves in the meridional direction before they reach the midtroposphere).

The vertical reflection surface forms as a result of the vertical wind curvature associated with this jet structure, while the meridional waveguide forms as a result of the increased meridional curvature. This reflective configuration is not found when looking at the climatological seasonal cycle in the Northern Hemisphere (the winds increase with height throughout the stratosphere during all winter months), but it is found during individual months in winter, as one manifestation of the interannual variability of the zonal-mean flow. We have further shown that a crude measure of the vertical wind shear in the upper stratosphere (e.g., the difference in zonal-mean wind between 2 and 10 hPa) is a good index for isolating seasons or individual months with a reflective configuration.

Given that a strengthening of the vortex in the lower stratosphere is often congruent with the jet peak forming at around 20 hPa (so that vertical wind shear is negative in the upper stratosphere), an index based on the strength of the vortex in the lower stratosphere also isolates differences in the vertical propagation characteristics of the basic states (as was found by PG), at least for seasonal means. This would seem to suggest that we have more downward reflection during high phases of the NAM in the lower stratosphere (as represented by the strength of the vortex there). This also implies a very basic positive feedback since an increase in the amount of downward reflection will weaken the deceleration exerted by the waves on the mean flow because less wave activity will end up being absorbed in the stratosphere. We note that downward reflection of waves back to the troposphere will result in a reduction of the upward EP fluxes in the upper troposphere/lower stratosphere (see Harnik 2002). A reduction of upward wave fluxes at the tropopause region during high phases of the NAM has indeed been observed (e.g., Shindell et al. 2001) and is attributed to an increased refraction of tropospheric waves toward the equator (rather than upward propagation to the stratosphere) as a result of the stronger jet (e.g., Limpasuvan and Hartmann 2000; Hu and Tung 2002; Lorenz and Hartmann 2003). Our results suggest that during reflective years, the effects of downward reflection on tropospheric and tropopause EP fluxes cannot be ignored. In addition, it has to be studied in more detail how stratosphere–troposphere coupling by wave reflection relates to the more commonly discussed zonal-mean downward coupling. In any event, to investigate the possible implications of changes in the reflective properties of the basic state for climate change impact in the troposphere, models have to be able to resolve full stratospheric dynamics. In particular, it is likely that models with a lid at 10 hPa will not be able to capture this mechanism.

Recent studies stress the implication of the stratosphere–troposphere coupling for the prediction of wintertime weather (Baldwin and Dunkerton 2001; Thompson et al. 2002), especially due to the observed evidence of a slow downward progression of NAM-like circulation anomalies. Downward reflection, which changes the planetary-scale flow in the troposphere on synoptic timescales, can plausibly affect synoptic-scale systems, and hence needs to be studied in more detail.

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. We thank David Rind, Richard Seager, James Hansen, Drew Shindell, and two reviewers for very useful comments on the manuscript. JP’s research is supported by the NASA Climate and NASA Atmospheric Chemistry and Analysis programs. NH is supported by National Oceanic and Atmospheric Administration Grant UCSIO-CU-02165401-SCF and by a Lamont-Doherty Earth Observatory postdoctoral fellowship.

REFERENCES
Graf, H.-F., J. Perlwitz, and I. Kirchner, 1994: Northern Hemisphere


