The Seasonal Cycle of Planetary Waves in the Winter Stratosphere

R. K. SCOTT AND P. H. HAYNES

Centre for Atmospheric Science, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, United Kingdom*

(Manuscript received 6 October 2000, in final form 11 July 2001)

ABSTRACT

A simple mechanistic model, with time-independent planetary wave forcing at the lower boundary, is used to investigate the observed differences between the Northern Hemisphere (NH) and Southern Hemisphere (SH) stratospheric evolution. Considered here are differences in both the zonal mean and the zonal wavenumber 1 structure of the winter vortex throughout its seasonal evolution. In particular, for small and intermediate forcing values the model robustly produces the midwinter minimum in geopotential wavenumber 1 amplitudes observed in the SH lower and middle stratosphere. This feature is considered from the different viewpoints of transitions between stable equilibria, wave transmission properties of the mean flow, and resonant type interactions with the lower boundary wave forcing, to assess the extent to which these viewpoints, previously studied in the context of models representing variation with height only, are relevant to a more complicated model representing variation in both height and latitude. Results show that resonance with the lower boundary forcing in early and late winter, leading to early and late winter wave maxima and hence to an associated midwinter minimum, is the dominant mechanism in the model considered. For large forcing amplitudes more typical of the NH, two possible winter and late winter evolutions are found to exist for the same forcing amplitudes. Which of these is selected depends on the details of the forcing in early winter. Note that there is no constant value of wave forcing amplitude that produces both a SH-like, undisturbed winter evolution as well a strong final warming, suggesting that time-varying lower boundary wave forcing is necessary to simulate these aspects of the SH evolution.

1. Introduction

This paper considers the seasonal evolution of the winter stratospheric circulation with particular emphasis on the differences between the Northern Hemisphere (NH) and Southern Hemisphere (SH) evolution. The major difference is that the NH polar vortex is weaker and more disturbed compared to that of the SH. In other words, the zonal mean zonal velocities are generally weaker, and the NH geopotential wave amplitudes are generally stronger than those in the SH. The NH polar vortex is also susceptible to more frequent and stronger sudden warming events, with the attendent reduction or even reversal of zonal mean zonal velocities and temperature increases. Such interhemispheric differences can be associated with a more disturbed NH troposphere resulting from stronger NH topography and land–sea surface temperature contrasts. As observations of the SH stratosphere have improved, a clear pattern of winter evolution has emerged, with twin maxima in the geopotential wave amplitude in early and late winter, and an intervening midwinter minimum, first noted by Labitzke (1980) and Hirota et al. (1983), and investigated further by Randel (1988) using an 8-yr (1979–86) climatology of daily geopotential data. In contrast, no such pattern has been observed in the NH. Shiotani et al. (1993) have also noted interesting interannual variability in the winter evolution in the SH.

The different strengths of the NH and SH polar vortices and their evolution throughout winter has well-known and important consequences for stratospheric chemistry and midlatitude ozone concentrations (see, e.g., McIntyre 1990; WMO 1999, Chapter 7, and references therein). Interest in the dynamics of the stratospheric circulation has been reinvigorated by the occurence of several anomalous cold NH winters in the 1990s and the difficulties of determining whether such anomalous winters are part of a systematic trend, perhaps triggered by increasing atmospheric CO_2, or simply a manifestation of interannual variability (e.g., Waugh et al. 1999).

In this paper two features of the seasonal evolution of the stratospheric circulation will be discussed. First, the twin maxima in geopotential wave amplitude ob-

* The Centre for Atmospheric Science is a joint initiative of the Department of Applied Mathematics and Theoretical Physics and the Department of Chemistry.

Corresponding author address: Richard Scott, Department of Applied Physics and Applied Mathematics, 200 Seeley W. Mudd Building, Columbia University, 500 W. 120th St., New York, NY 10027. E-mail: rks14@damtp.cam.ac.uk

served in early and late winter in the SH will be considered from different dynamical points of view (sections 2 and 4). We will use an idealized model to look at both wave transmission properties of the zonal mean flow as well as resonance properties of the mean flow with the lower boundary wave forcing. Second, different types of vortex evolution, namely SH-like undisturbed–quasilinear evolution and NH-like disturbed–nonlinear evolution, will be considered as possible stable evolution states arising from similar wave forcing amplitudes (sections 2 and 5). This latter discussion does not imply that it might be possible to obtain a SH-like vortex evolution with NH type forcing amplitudes, or vice versa. Rather, it suggests that, once established, a strong vortex may be less disturbed by subsequent strong wave forcing (which, in the SH, could result from anomalously strong baroclinic disturbances in the troposphere). Such a claim has natural theoretical justification from elementary considerations such as the Charney–Drazin criterion for wave propagation (Andrews et al. 1987) and from the refractive index properties of the zonal mean zonal velocities.

The model used in the main part of the study is a primitive equation model on a spherical domain with full latitude and height structure and a single wave-mean structure in the zonal direction. This simple zonal structure is capable of representing many important features of the large-scale circulation, as discussed in Haynes and McIntyre (1987) and references therein. This simple model allows for computationally inexpensive integrations of the annual cycle and therefore a thorough exploration of various parameter regimes. The wave forcing is represented by specification of the geopotential wave amplitude at the lower boundary and the seasonal cycle is represented by a time-varying radiative basic state, to which the zonal mean temperature field is relaxed using Newtonian cooling. In this configuration the model is a natural extension of the one-dimensional, height-only models used by Yoden (1990) and Plumb (1989) in similar studies, and is an important link between these and three-dimensional models.

The structure of the paper is as follows. In section 2 different mechanisms are outlined that have been suggested previously as possible explanations for the observed differences between the NH and SH evolution, in particular the SH midwinter wave minimum. These mechanisms are illustrated in the simplest possible wave-mean flow model, a channel model with no height or latitude structure and only one zonal wave. The idea here is not so much to compare the evolution of the atmosphere to this simple model but rather to make clear the concepts of resonance, quasi-static behavior, and multiple stable states that will be addressed later. In sections 3 and 4 the idealized stratospheric primitive equation model, which allows wave propagation in the height–latitude plane, is introduced. Particular attention is given to assessing the relevance of the concepts introduced in section 2, now combined with consideration of wave transmission properties, to the linear and quasi-linear regimes relevant to the SH. In section 5 the nonlinear response of the primitive equation model is examined when stronger wave forcing is used. Attention here is focused on the question of multiple stable states existing under the same wave forcing amplitude, and in particular how a given state can persist even when the wave forcing amplitude is changed. In section 6 we consider the robustness of the results presented in sections 4 and 5 to changes in the zonal resolution of the model by repeating certain integrations with higher zonal wavenumbers included. A summary and conclusions are given in section 7.

2. Main ideas and their illustration in a reduced model

We begin with a brief review of the signature of the SH midwinter wave minimum, and various suggested explanations. The minimum was first noted by Labitzke (1980) and Hirota et al. (1983) and investigated further by Randel (1988) using an 8-yr (1979–86) climatology of daily geopotential data. The data showed early and late winter maxima most prominently in the stationary and transient wavenumber one geopotential height amplitude in the middle and upper stratosphere, without any such twin maxima structure in the lower stratosphere or upper troposphere. Randel reports further that the wave maxima are in disagreement with lower stratospheric wave transmission properties as defined by the refractive index, which suggest strongest vertical wave propagation in August, during the observed wave minimum. The pattern can also be contrasted to the NH evolution, which has a single, stronger wave maximum occuring in January, shortly after midwinter.

Various numerical studies have been conducted of this midwinter minimum, which have led to conceptually different explanations of the underlying mechanisms. Plumb (1989) used a height-dependent beta-plane channel model containing only a single meridional mode, relaxing the zonal mean velocities towards an idealized, seasonally varying radiative basic state and forcing waves with zonal wavenumber 1 or 2 at different amplitudes. The essential characteristics of the midwinter minimum were successfully reproduced. In Plumb's (one-dimensional) model, the range of zonal velocities (U) allowing upward wave propagation is given by the Charney–Drazin (1961) criterion,

$$0 < U < U_c = \beta\left[k^2 + l^2 + \frac{f^2}{4N^2H^2}\right]^{-1}, \qquad (1)$$

where k and l are the zonal and meridional wavenumbers [Andrews et al. 1987, Eq. (4.5.15)]. For small forcing amplitudes the zonal mean zonal velocities in midwinter remained sufficiently strong to prevent upward wave propagation and the response remained in a linear regime close to the radiative basic state. Consequently, large wave amplitudes in middle and upper levels were

only obtained in early and late winter when the zonal mean zonal velocity permitted upward wave propagation by (1). For large forcing amplitudes, on the other hand, a nonlinear regime was obtained in which there was significant wave-mean flow deceleration, and large wave amplitudes were present throughout the entire winter. The two different regimes—linear for small forcing and nonlinear for large forcing—were likened to the observed SH and NH winter evolutions, respectively.

This essentially quasi-static explanation of the interhemispheric differences, based on wave transmission properties of the mean flow, was investigated further by Wirth (1991). By quasi-static we mean that the time lag between the seasonal cycle of the actual temperature and the seasonal cycle of the radiative equilibrium temperature is small. Note that the wave transmission properties, that is, those aspects of the mean flow that affect the propagation of waves, depend on both the strength and the latitudinal structure of the mean flow. The effect of the latter was absent in Plumb's (1989) one-dimensional model. Quantities such as the refractive index defined by Matsuno (1970) are a useful indication of the transmission properties but must be used with care since there is no scale separation between the waves and the mean flow. Wirth solved the linear steady-state quasi-geostrophic equations on the hemisphere, using a prescribed zonally symmetric basic-state and a stationary, wavenumber-1 lower boundary forcing based on climatological monthly mean winds and geopotential amplitudes. This procedure gave a maximum in stratospheric wave amplitudes in late winter but none in early winter. Further, since the refractive index in the lower stratosphere showed no seasonal variation compatible with the twin maxima in observed wave amplitude, it was suggested that the link between upper-level wave amplitudes and lower-level wave transmission properties of the mean flow may be weaker than previously thought.

An alternative approach was adopted by Wakata and Uryu (1987) and Yoden (1990). Using essentially the same model as Plumb (1989), they also observed near-radiative, linear behavior for weak forcing and disturbed, nonlinear behavior for strong forcing. With weak forcing, twin maxima in wave amplitudes were obtained, but were interpreted instead as transitions between separate stable solution branches of the steady system, in which the radiative basic-state wind strength was an external bifurcation parameter. With strong forcing, on the other hand, the bifurcation diagram, or configuration of the stable solution branches, was such that the response never reached the ''upper,'' near-radiative branch (Yoden 1990, Fig. 6).

The relevance of transitions between separate stable solution branches in a time-varying radiative basic state to the early and late winter SH dynamics can be seen most simply using a zero-dimensional representation of a wave-driven mean flow with relaxation to a basic state. The quasi-geostrophic shallow-water equations in a β channel, with only one meridional and one zonal Fourier mode, reduce under quasi-static assumptions to

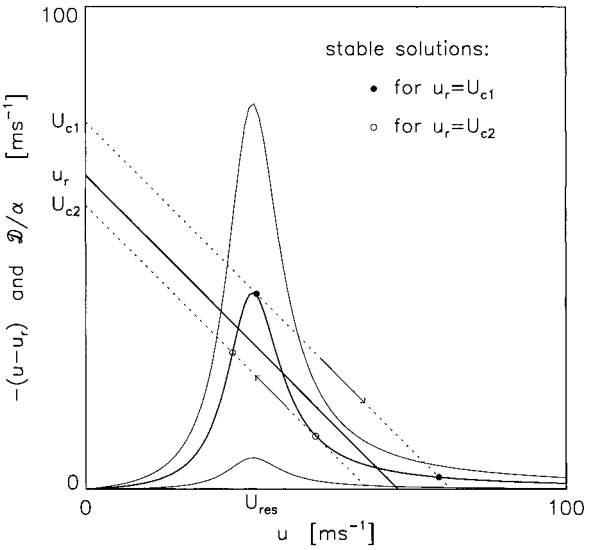

FIG. 1. The terms $\mathcal{D}/\alpha$ and $-(u - u_r)$ in (2). Straight lines are $-(u - u_r)$ for different values of u_r, labelled along the ordinate of (a), considered as an external parameter. Curved lines are $\mathcal{D}/\alpha$ for different values of forcing amplitudes h_0. The seasonal cycle is approximated by increasing u_r quasi-statically from a low value on the left of the u axis to a high value on the right and then decreasing it back again. As an example, stable solutions for $u_r = U_{c1}$ and $u_r = U_{c2}$ are shown by the solid and open circles, respectively. The curve for $\mathcal{D}/\alpha$ is based on Eq. (6.11) of Held (1983) with topography consisting of single Fourier component.

$$\frac{\partial \overline{u}}{\partial t} = -\alpha(\overline{u} - \overline{u}_r) - \mathcal{D}(\overline{u}; h_b), \qquad (2)$$

where $\overline{u}$ is the zonal mean zonal velocity, $\overline{u}_r$ is some radiative basic-state velocity, α is the relaxation rate, and $\mathcal{D}(\overline{u}; h_b)$ is the steady-state wave drag associated with a particular $\overline{u}$ and planetary wave boundary forcing h_b (Charney and DeVore 1979; Held 1983). The wave drag is assumed to have a single maximum, perhaps associated with resonance at $\overline{u} = U_{\text{res}}$, a large value of this maximum corresponding to NH conditions and a small value corresponding to SH conditions. The steady-state solutions are given by the intersections of the curves $\mathcal{D}(\overline{u}; h_b)/\alpha$ and $-(\overline{u} - \overline{u}_r)$ shown in Fig. 1 (after Held 1983). Thus, different values of $\overline{u}_r$, considered as an external parameter, lead to either one or three possible steady states, with bifurcation points at $\overline{u}_r = U_{c1}$ or $\overline{u}_r = U_{c2}$. When three steady states exist, the middle one is always unstable, since $\alpha + d\mathcal{D}/d\overline{u}$ is negative for that solution.

If $\overline{u}_r$ is varied quasi-statically then rapid transitions between the stable steady states can occur. This was investigated by Clark (1992) using the above zero-dimensional model with gradually decreasing $\overline{u}_r$ in the context of the SH final warming. Consider first the case $\overline{u}_r \gg U_{\text{res}}$ for the middle curve in Fig. 1, taken to be representative of midwinter conditions in the SH. As

$\overline{u}_r$ decreases the $\overline{u}$, remains on the rightmost stable solution branch close to $\overline{u}_r$. When $\overline{u}_r$ reaches U_{c2}, as defined in Fig. 1, this solution branch disappears and a rapid transition occurs to a solution branch on which the velocity is far from the radiative basic state. (Solutions that are close to radiative equilibrium lie close to the horizontal axis, those that are further from radiative equilibrium lie further away.) This rapid transition of u is associated with a rapid increase in wave amplitude as the system passes through resonance at $\overline{u} = U_{\text{res}}$ and the phase speed of the free mode of the system approaches the phase speed of the forcing.

A similar picture that can be interpreted as the SH early winter evolution is obtained by gradually increasing $\overline{u}_r$ from a value $\overline{u}_r \ll U_{\text{res}}$. Again, $\overline{u}$ remains on the leftmost solution branch until $\overline{u}_r$ reaches U_{c1}, as defined in Fig. 1, at which point this solution branch disappears and a rapid transition occurs. This time, however, the response approaches resonant conditions before the original stable solution is lost and the wave amplitude, and hence $\mathcal{D}$, increases, preventing $\overline{u}$ from following $\overline{u}_r$. The transition between states, therefore, is now from one of large wave amplitude and a velocity that is far from the radiative basic state, to one of small wave amplitude and a velocity that is close to the basic state.

When the above two cases are combined a crude representation of the seasonal cycle is obtained. The transitions between stable solution branches will then occur for intermediate values of forcing amplitude that give a wave drag, $\mathcal{D}$, whose peak lies below the line corresponding to the maximum value of $\overline{u}_r$; that is, rapid transitions will occur if there is a $\overline{u}_r$ with $u_r > U_{c1}$. For sufficiently large forcing, more representative of the NH, U_{c1} will be greater than $\overline{u}_r$ at all times during the cycle. Instead of transitions between stable solution branches, the response will then remain on the left slope of the peak of $\mathcal{D}/\alpha$, with large wave amplitude and velocity far from the basic state for large, midwinter values of $\overline{u}_r$. Thus, in some simple sense, the model captures the difference between the intermediately forced SH and strongly forced NH.

For radiative relaxation rates, (the motivation behind the velocity relaxation at rate α in the above model), typical of the stratosphere, the quasi-static assumptions of the above description may be less relevant. This was pointed out by Yoden (1990) for the height-dependent channel model, which contains conceptually equivalent steady-state behavior to the above. Despite the transient nature of the evolution, Yoden's study nonetheless exhibited traces of transitions between solution branches when an intermediate forcing was used, in contrast to behavior that essentially remained on one solution branch when stronger forcing was used. Further, Yoden's equivalent of the leftmost solution branch in Fig. 1 underwent a Hopf bifurcation in the cases of strong forcing that led to vacillating behaviour. In this respect the evolution obtained might be considered to be even closer to that observed in the NH winter.

In the following we investigate to what extent the above concepts of wave transmission, resonance, and multiple stable states are manifested in a more complicated model containing both height and latitude structure. In particular, section 4 considers the quasi-linear aspects of wave transmission and resonance and their relevance to the SH midwinter wave minimum, and section 5 considers the nonlinear aspects of multiple stable evolution states and their relevance to the different NH- and SH-like disturbed and quiescent vortex evolutions.

3. Model description

The model used in this study is a mechanistic primitive equation model developed by Saravanan (1992) and used previously by Scott and Haynes (1998), with pressure as the vertical coordinate and a spectral representation in the horizontal, using a spherical domain. The time stepping uses the semi-implicit scheme, treating terms associated with gravity waves implicitly, and uses a Robert time filter to damp the computational mode associated with the leap-frog scheme (Haltiner and Williams 1980). There are 31 pressure levels, equally spaced in log pressure $z = -H \ln(p/p_s)$, where $p_s =$ 100 000 Pa, spanning $z_B = 11$ km (20 660 Pa) to $z_T =$ 82.6 km (0.75 Pa) at intervals $\Delta z = 2.4$ km, and 31 latitudinal Legendre modes. The reference temperature profile is that of an isothermal atmosphere with $T_{\text{ref}} =$ 240 K, corresponding to a scale height H of 7 km. For most of the simulations carried out in this work only the zonal mean and wavenumber one components are retained in the zonal direction. The extent to which such severe truncation can represent the large-scale features of the nonlinear evolution is discussed in, for example, Haynes and McIntyre (1987) and references therein. In section 6 some simulations are presented that retain additional wave-two and wave-three components.

The lower boundary wave forcing is through a perturbation in the geopotential height field near the lowest model level. This is of the form

$$\Phi' = h_0 E(t) G(\phi), \tag{3}$$

where $E(t)$ is grown smoothly from $E(t = 0) = 0$ to $E(t > t_s) = 1$, with $t_s = 10$ days. The latitudinal structure is given by $G(\phi) = 4\hat{\mu}^2(1 - \hat{\mu}^2)$, where $\hat{\mu}(\mu_0 < \mu < 1) = (\mu - \mu_0)/(1 - \mu_0)$, $\hat{\mu}(0 < \mu < \mu_0) = 0$, and where $\mu = \sin\phi$. Choosing $\mu_0 = \sin(\pi/6)$ gives a maximum $G(\phi) = 1$ located at $\phi = 60°$ with $G(\phi < 30°) = 0$. This is close to the latitudinal location of the maximum in 300 mb stationary wavenumber one geopotential height amplitude illustrated in Randel (1988, his Fig. 7). Finally, h_0 is the constant forcing amplitude, which is varied as an external parameter between integrations.

The seasonal cycle is included by radiative relaxation to a seasonally varying, zonally symmetric potential temperature field $\theta_R(\phi, z, t)$, which is a sinusoidal superposition of summer and winter potential temperature

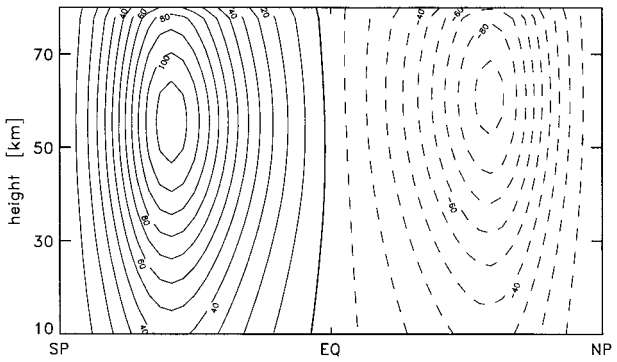

FIG. 2. The radiative basic state U_R given by (4) for the main set of experiments (sections 4 and 5) as a function of height and latitude. The contour interval is 10 m s^{-1} and dashed contours are negative.

fields $\theta_{RS}(\phi, z)$ and $\theta_{RW}(\phi, z)$, respectively, given by $\theta_R = r\theta_{RS} + (1 - r)\theta_{RW}$ with $r(t) = [1 + \cos 2\pi t/(365 \text{ days})]/2$. The fields θ_{RS} and θ_{RW} are calculated by first specifying velocity profiles U_{RS} and U_{RW}, and then calculating the geopotential and potential temperature fields that are in balance with these velocity profiles. The velocity fields are given by $U_{RS} = U_R(-\phi \geq 0)$ and $U_{RW} = U_R(\phi \geq 0)$, with

$$U_R(\phi, z) = \cos\phi[u_0 \tanh(b_0(\phi - \phi_0)) + J_1 + J_2], \quad (4)$$

for constants u_0, b_0, ϕ_0, and where the $J_i(\phi, z)$, are defined by

$$J_i = u_i \operatorname{sech}[b_i(\phi - \phi_i)] \operatorname{sech}(a_i(z - z_i)), \quad (5)$$

where $i = 1, 2$, for constants u_i, b_i, ϕ_i, a_i, z_i. The values of the constants for the velocity profile used in the simulations presented in sections 4 and 5 are:

$$\begin{aligned} &u_0 = 30 \text{ m s}^{-1}, && b_0 = 0.01, && \phi_0 = 20°, \\ &u_1 = 200 \text{ m s}^{-1}, && b_1 = 0.06, && \phi_1 = 60°, \\ &u_2 = -150 \text{ m s}^{-1}, && b_2 = 0.06, && \phi_2 = -60°, \\ &a_1 = 0.04, && z_1 = 55 \text{ km}, \\ &a_2 = 0.04, && z_2 = 60 \text{ km}. \end{aligned} \quad (6)$$

The velocity profile U_R is shown in Fig. 2. All simulations discussed in the remainder of the paper are integrations for 1 yr, initialized with a radiative equilibrium state corresponding to midsummer (corresponding to $t = 0$ above). (We note that the model responses described in sections 4 and 5 were robustly reproduced for different choices of the radiative basic state, including changes in the strength of the winter westerlies, the position of the zero wind line, and the strength of the vertical shear in the lower stratosphere.) When balancing the model, the zonal mean geopotential at the lower boundary and the interior temperature field are determined by requiring that $\mathbf{T}^{\mathrm{T}} \cdot \boldsymbol{\beta} \equiv 0$, where $\mathbf{T}$ is a vector of temperatures at each pressure level, and $\boldsymbol{\beta}$ is a binomial vector of alternating sign (e.g., for five levels $\boldsymbol{\beta} = (1, -5, 10, -10, 5)$). This is one way of ensuring that a two grid length wave in the temperature field is minimized and is a simple modification (for pressure coordinates) of the approach set out in Hoskins and Simmons (1975).

The radiative relaxation rate is given as a function of height by $\alpha(z) = \{1.5 + \tanh[(z - 35)/7]\} \times 10^{-6} \text{ s}^{-1}$, as in Holton (1976). A Rayleigh friction sponge layer is applied above $z = 50$ km by relaxing to a zero velocity field at rate $\kappa(z) = \{1.02 - \exp[(50 - z)/40]\} \times 5 \times 10^{-6} \text{ s}^{-1}$. Finally there is horizontal scale selective damping using a ∇^8 hyperdiffusion with a damping rate of four per day at the smallest scale (highest total wavenumber).

4. Linear and quasi-linear responses

In this section we examine the model response to constant wave forcing of various amplitudes and a seasonally varying basic state. We look in detail at the wave transmission properties of the mean flow during the seasonal cycle and at resonance-like behavior. Consideration of how the addition of latitudinal structure alters the multiple equilibria and the transitions between these obtained by Yoden (1990) with a channel model, is delayed until section 5 below.

a. Dependence on forcing amplitude

Figures 3 and 4 show latitude–time cross sections of wavenumber one geopotential height amplitude at 33- and 38-km altitude respectively, for wave forcing amplitudes $h_0 = 10$–340 m. In each, the geopotential wave amplitude has been normalized by the forcing amplitude. There is a clear distinction between linear, quasi-linear, and nonlinear regimes. For forcing amplitudes $h_0 = 10$ m and $h_0 = 20$ m there is a similar structure between the responses, with wave amplitude maxima in early and late winter whose magnitudes scale linearly with the forcing amplitude. For 40 m $\leq h_0 \leq$ 200 m the responses again show a similar structure but the magnitudes of the early and late winter peaks deviate slightly from a strictly linear scaling. In contrast, for 260 m $\leq h_0 \leq$ 340 m, there is a single dominant maximum around midwinter, with a strongly transient, nonlinear response. The differences between the weakly forced, quasi-linear responses and the strongly forced, nonlinear responses crudely resemble the observed differences between the quiescent SH and the disturbed NH in the real stratosphere. We note in passing, however, that although the geopotential wave amplitude in the $h_0 = 260$ m case shows a strong time dependence, large fluctuations are nevertheless absent from the zonal mean zonal velocity. That is, the vortex itself is relatively undisturbed and more closely resembles the evolution of the SH vortex, albeit farther from the radiative basic state than for much smaller h_0 (see also Fig. 11

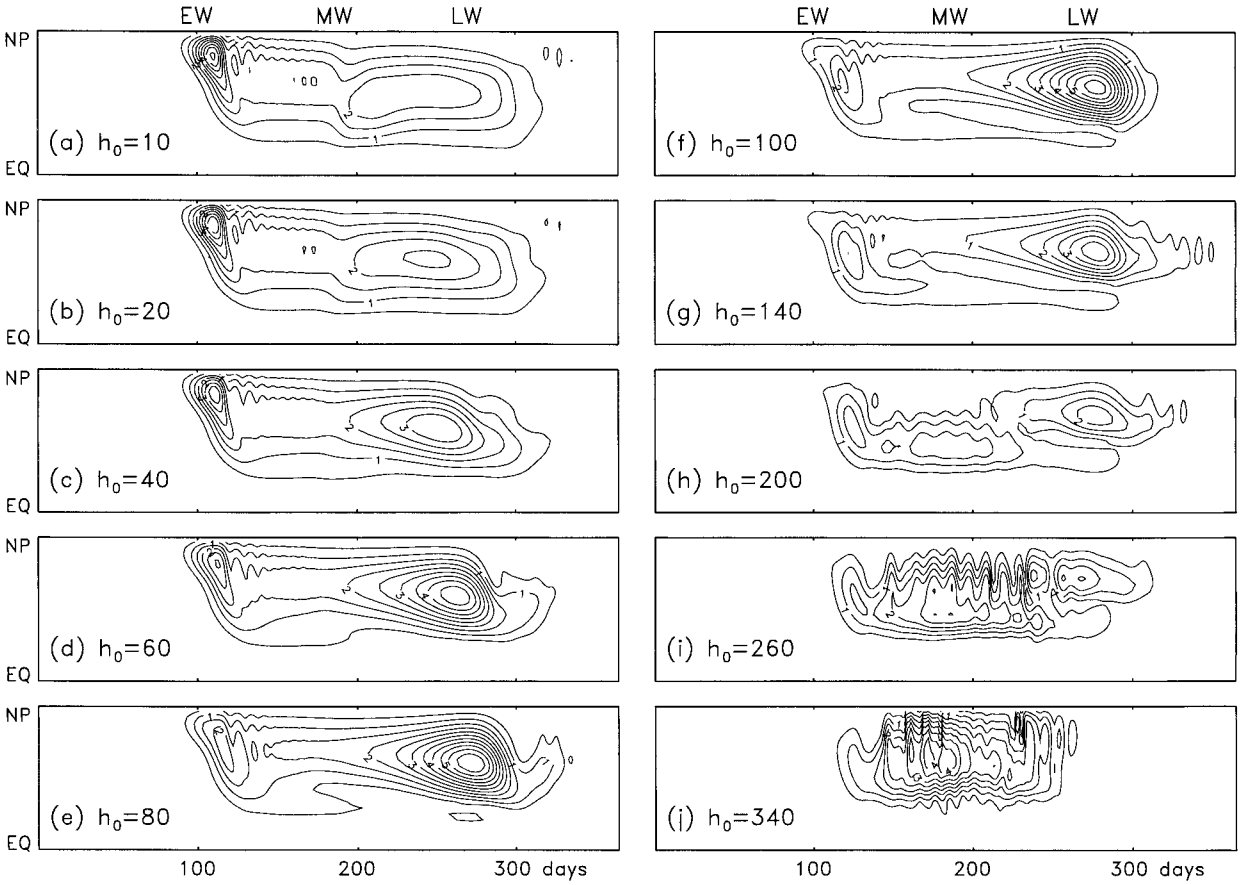

FIG. 3. Geopotential wave-one amplitude at the 33-km height level normalized by the wave-forcing amplitude h_0 as a function of latitude and time: (a) $h_0 = 10$ m, (b) $h_0 = 20$ m, (c) $h_0 = 40$ m, (d) $h_0 = 60$ m, (e) $h_0 = 80$ m, (f) $h_0 = 100$ m, (g) $h_0 = 140$ m, (h) $h_0 = 200$ m, (i) $h_0 = 260$ m, and (j) $h_0 = 340$ m. The contour interval is 0.5. Day zero corresponds to midsummer; EW, MW, LW, correspond to early, mid-, and late winter.

below). Only for $h_0 = 340$ m (and above) does the zonal mean zonal velocity more closely resemble the strongly transient evolution of the NH vortex. (The time-averaged general structure resembles that of Fig. 12 below.)

b. Resonance

For both of the linear forcing amplitudes $h_0 = 10$, 20 m, and even more clearly for the quasi-linear forcing amplitudes 40 m $\leq h_0 \leq$ 140 m, the double maxima structure, and attendant midwinter minimum, is clearly reproduced. Since the evolution in all cases is very similar to the linear evolution, consideration of the zero-dimensional model of section 2 suggests that the concept of the maxima arising from rapid transitions between stable steady states is inapplicable. In terms of the zero-dimensional model description of section 2, the wave drag in all the cases described above corresponds to the lowermost curve for $\mathcal{D}$ in Fig. 1 (i.e., weak wave forcing), and is such that only a single steady state exists at any time. Nevertheless, the curve has a maximum for a particular value of u associated with a resonance of the system, indicating a corresponding maximum in wave amplitudes. The same picture can be seen in the two-dimensional model response using a suitable analogue of the quantity $\mathcal{D}$, namely the meridionally averaged vertical component of EP flux through the lower boundary $\langle F^{(z)}|_B \rangle$ given by

$$\langle F^{(z)}|_B \rangle = \frac{1}{2\pi a^2} \int_0^{\pi/2} F^{(z)} \cos\phi \, d\phi|_{z=z_B}$$

$$= -\frac{1}{2\pi a^2} \int_V \rho_0 \mathcal{F} a \cos\phi \, dV = -\frac{1}{2\pi a^2} \mathcal{T}, \quad (7)$$

where $\mathcal{F} = (1/\rho_0 a \cos\phi)\nabla \cdot \mathbf{F}$ is the force per unit mass and $\mathcal{T}$ the total torque produced by the waves on the zonal mean flow. In the second line of (7), V denotes the domain of the model (i.e., the stratosphere) and the divergence theorem has been used. It can be seen from the second line of (7), that $\langle F^{(z)}|_B \rangle$ is analogous to the $\mathcal{D}$ of the zero dimensional situation.

Since the wave forcing in the model is held constant, any changes in $\langle F^{(z)}|_B \rangle$ must be attributed to the mean flow structure changing in such a way that more or less wave activity is able to enter the domain. As for the

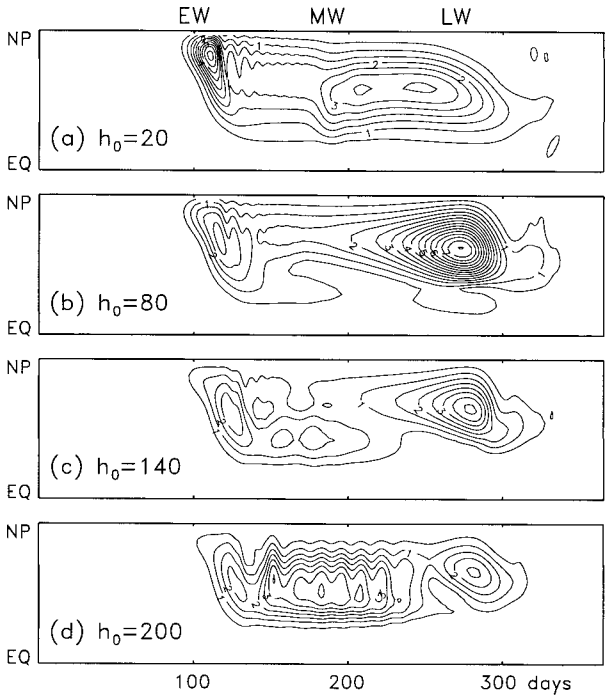

FIG. 4. Same as Fig. 3, but for the 38-km height level: (a) $h_0 = 20$ m, (b) $h_0 = 80$ m, (c) $h_0 = 140$ m, and (d) $h_0 = 200$ m.

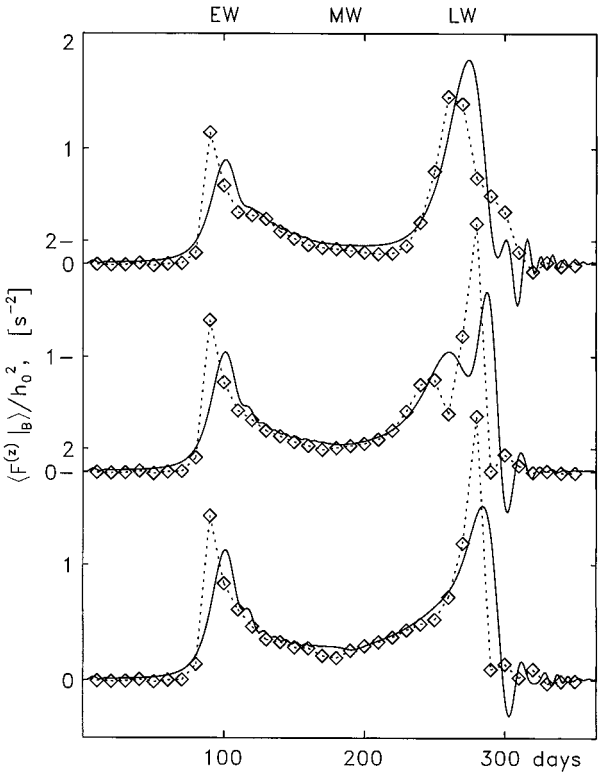

FIG. 5. Meridionally averaged vertical component of EP flux through the lower boundary $\langle F^{(z)}|_B\rangle$ normalized by the square of the wave-forcing amplitude h_0 as a function of time for $h_0 = 20$ m (lower), $h_0 = 60$ m (middle), $h_0 = 100$ m (upper). Solid lines denote the time-dependent, seasonally evolving cases; diamonds joined by dotted lines denote the values of the corresponding linear, steady-state simulations for that day and wave forcing amplitude. Day zero corresponds to midsummer.

zero dimensional system, one way in which this can happen is through a resonance, in which the phase speed of a free mode in the system approaches that of the wave forcing, which in this case is zero. This has been discussed in McIntyre (1982, section 5) and was considered in more detail by Clark (1992), for example. In the present model the nature of the lower boundary condition precludes deep free modes with a Lamb-mode-like structure in the vertical (Yoden 1987). Any free-mode phenomenon must therefore involve internal reflection, for example, between a reflecting critical layer (Dunkerton et al. 1981) and the lower boundary. Our understanding of the details of such free modes is still only partial. Nonetheless, if there is a resonance, the linear undamped response in the resonant configuration will consist of linear growth with time of both the wave amplitude and of the quantity $\langle F^{(z)}|_B\rangle$.[1] In the case with damping on the waves, the linear wave response will reach a steady state such that the resonant forcing is balanced by the wave damping. The closeness of the system to resonance and the extent to which the wave forcing projects onto the free mode in the system will both determine the final steady-state linear wave response.

Figure 5 shows the meridionally averaged vertical component of EP flux through the lower boundary as given by (7), normalized by h_0^2, for forcing values $h_0 =$ 20, 60, 100 m. These results suggest that for linear or quasilinear forcing values the system passes through a resonance in early and in late winter. Note, however that the concept of resonance is associated with a basic state that is constant in time, and so the extent to which it is applicable to a seasonally evolving basic state should be tested. To do so, we fix the zonal-mean state at its value on a certain day within the seasonal cycle and integrate the resulting forced–damped model until the wave response reaches a steady state. By fixing the zonal-mean state we are essentially linearizing the model about the day in question. The results are shown in Fig. 5, dotted curve with symbols, as a function of the day within the seasonal cycle from which the basic state is taken. The agreement between the steady-state results and those from the nonlinear seasonally evolving model, in particular the agreement of the early and late winter maxima, suggest that the quasi-static idea of resonance introduced above is, to a reasonable degree, applicable to the time-evolving system under consideration. The most significant difference between the time-evolving and the steady-state responses lies in the timing of the

[1] Note that $\langle F^{(z)}|_B\rangle$ is linear in wave amplitude a and scales as $h_0 a$, even though $F^{(z)}$ is quadratic in a in the interior of the domain.

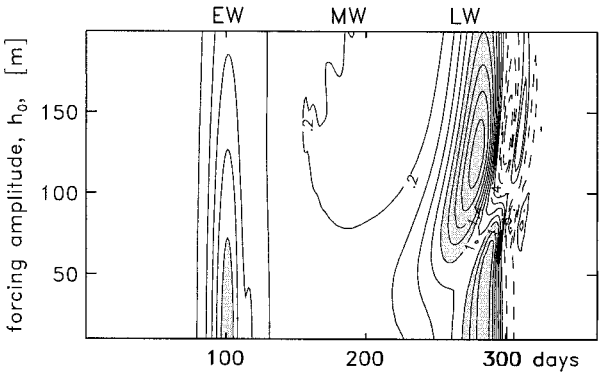

FIG. 6. Meridionally averaged vertical component of EP flux through the lower boundary $\langle F^{(z)}|_B\rangle$ normalized by the square of the wave-forcing amplitude, as a function of time and wave-forcing amplitude. The contour interval is 0.2 and values greater than 1.0 are shaded. Day zero corresponds to midsummer.

maxima, which should be interpreted simply as a lag of the evolving wave response to the wave forcing.

Before considering the relationship between $\langle F^{(z)}|_B\rangle$ and the geopotential wave amplitudes shown in Fig. 3, we note briefly the change in the structure of the time variation of $\langle F^{(z)}|_B\rangle$ as h_0 is increased. Of particular interest is the change in the structure of the late winter peak. From Fig. 6, it is seen that as the forcing is increased the late winter structure changes from a single peak on day 280 for $h_0 = 10$ m, to a double peak on days 260 and 290 for $h_0 = 60$ m, to a single peak again on day 270 for $h_0 = 100$ m (see also Fig. 5). It appears that as h_0 is increased, the resulting stronger mean flow changes during the winter produce a basic state that has an additional, slightly earlier resonance (day 260 for $h_0 = 60$ m). Upon further increase in h_0, this earlier resonance peak strengthens while the original weakens. The actual changes in the mean flow on day 260 (not shown) resulting from increasing h_0 from 10 m to 80 m are relatively small, confined to latitudes equatorward of 30°, with a maximum change of about -15 m s^{-1} at around 15° latitude, and a shift of the zero wind line of about 15° poleward. For all of the above forcing amplitudes, the general structure of the zonal mean zonal velocity broadly resembles that of Fig. 12a below.

A comparison of Fig. 3 and Fig. 5 for $h_0 \leq 60$ m indicates that the timing of the early winter peak in $\langle F^{(z)}|_B\rangle$ closely corresponds to the early winter peak in geopotential amplitude: the peaks occur at around 105 days and 110 days, respectively. This point in the seasonal cycle is only a few days after the change from summer easterlies to winter westerlies and the low wave forcing amplitudes are insufficient to alter significantly the mean flow before the occurrence of the resonance. As a result, the structure of the peaks in both $\langle F^{(z)}|_B\rangle$ and in the geopotential are similar over this range of h_0.

In late winter the relation between $\langle F^{(z)}|_B\rangle$ and the geopotential amplitude is more complicated. For linear forcing, $h_0 \leq 20$ m, the late winter peak in $\langle F^{(z)}|_B\rangle$ occurs around day 285, *after* the maximum in geopotential amplitude, which occurs around day 250. The lack of a corresponding peak in the geopotential amplitude after day 285 can be explained by considering the evolution of the mean flow, recalling that a given amount of wave activity will be associated with larger geopotential wave amplitudes on a zonal-mean flow that is stronger (Simmons 1974). Thus, although the early and late winter peaks in $\langle F^{(z)}|_B\rangle$ both occur when the mean flow is near equinox conditions, by the time the waves have reached the middle stratosphere, the early winter flow has increased in strength, resulting in a stronger geopotential wave amplitude response, whereas the late winter flow has reduced in strength, resulting in a weaker response. For quasi-linear forcing 40 m $\leq h_0 \leq$ 200 m the stronger late winter peak in geopotential amplitude is associated with the appearance of the second late winter peak in $\langle F^{(z)}|_B\rangle$ on day 260 for $h_0 = 60$ m, as described above and illustrated in Fig. 6. Since this peak occurs slightly earlier in the winter evolution when the mean flow is still relatively strong there is a stronger response in the geopotential wave amplitude.

c. Wave transmission properties

The relation between the peaks in wave activity entering the domain and the geopotential wave amplitude can be further clarified by considering the wave transmission properties of the mean flow. At the same time we consider whether Plumb's (1989) channel model behavior, namely, the direct influence of the seasonal evolution of wave transmission properties on the midwinter minimum in geopotential wave amplitude, can also be seen in the present, latitudinally dependent model. In Figs. 7–9 below we present Eliassen–Palm (EP) flux vectors and divergence on selected days, normalized by the value of wave activity entering the domain $\langle F^{(z)}|_B\rangle$ on that day. The fluxes normalized in this way give a clearer indication of the relative wave transmission properties of the flow on different days than the unnormalized fluxes.

The linear forcing case of $h_0 = 20$ m, shown in Fig. 7, indicates markedly different transmission properties in early winter to those in mid- and late winter (where we identify early, mid-, and late winter with the maxima and minima in $\langle F^{(z)}|_B\rangle$), as shown in Fig. 6). In particular, the confinement to high latitudes of the flux vectors explains the latitudinal structure of the geopotential wave amplitude response to the early winter peak in $\langle F^{(z)}|_B\rangle$ seen in Figs. 3 and 5. Between midwinter and late winter, there is very little difference in the wave transmission properties in midlatitudes, except for a weakening of flux vectors on day 285. This weakening explains further the lack of a late winter maximum in the geopotential wave amplitude response to the late winter peak in $\langle F^{(z)}|_B\rangle$. Similarly the constancy of the flux vectors between days 180 and 260 is consistent

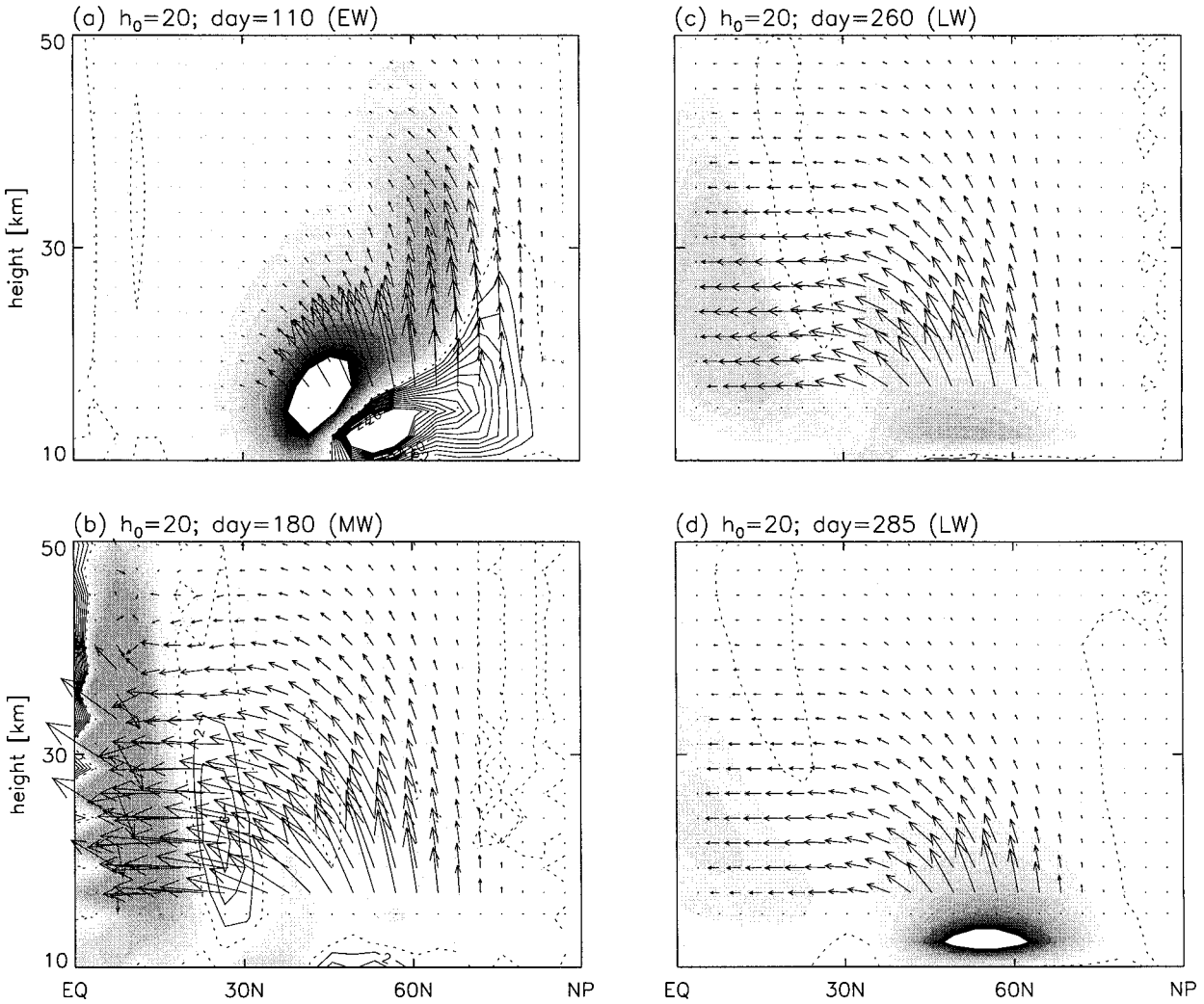

FIG. 7. Height–latitude cross-sections of EP flux and divergence normalized by the meridionally averaged vertical component of EP flux through the lower boundary for $h_0 = 20$ m: (a) day 110, (b) day 180, (c) day 260, and (d) day 285. The graphical conventions are as in Dunkerton et al. (1981); contour and shading interval is 2, in units of $2\pi a^3 \rho_s \times 10^{-12}$ m^{-1}; negative values (convergence) are shaded and zero contour is dotted.

with only small changes in wave amplitude over the same time, the gradually increasing $\langle F^{(z)}|_B \rangle$ and the gradually decreasing zonal-mean basic state.

For the weak quasilinear forcing case of $h_0 = 80$ m, shown in Fig. 8, only small changes in the above transmission properties are found. Relative to the previous case using $h_0 = 20$ m, wave transmission here is slightly weaker in early winter (not shown), slightly weaker in midwinter (Fig. 8a), and slightly stronger in late winter (Fig. 8b). These changes explain both the slight weakening of the early winter peak in geopotential wave amplitude and the formation of the strong late winter peak seen in Fig. 3. That the late winter geopotential wave amplitude response should be so strongly altered by a modest change in wave transmission properties of the mean flow can be explained by the earlier occurrence of the late winter peak in $\langle F^{(z)}|_B \rangle$, and the stronger zonal-mean basic state at the time of the peak. Although the midlatitude wave transmission properties are less favourable to upward propagation in midwinter than in early and late winter, it is clear from the above discussion that, unlike in Plumb's (1989) model, they are insufficient on their own to explain the geopotential amplitudes seen in Fig. 3 and Fig. 4.

For the strong quasilinear forcing case of $h_0 = 200$ m, shown in Fig. 9, qualitatively different transmission properties are found. We emphasize that this case is quasilinear in the sense that the winter vortex is still relatively undisturbed and slowly evolving in time, resembling more the SH vortex than the NH vortex. The most obvious change resulting from the increased forcing is the enhanced vertical wave transmission in midlatitudes, with significant wave activity reaching the midlatitude middle and upper stratosphere in midwinter

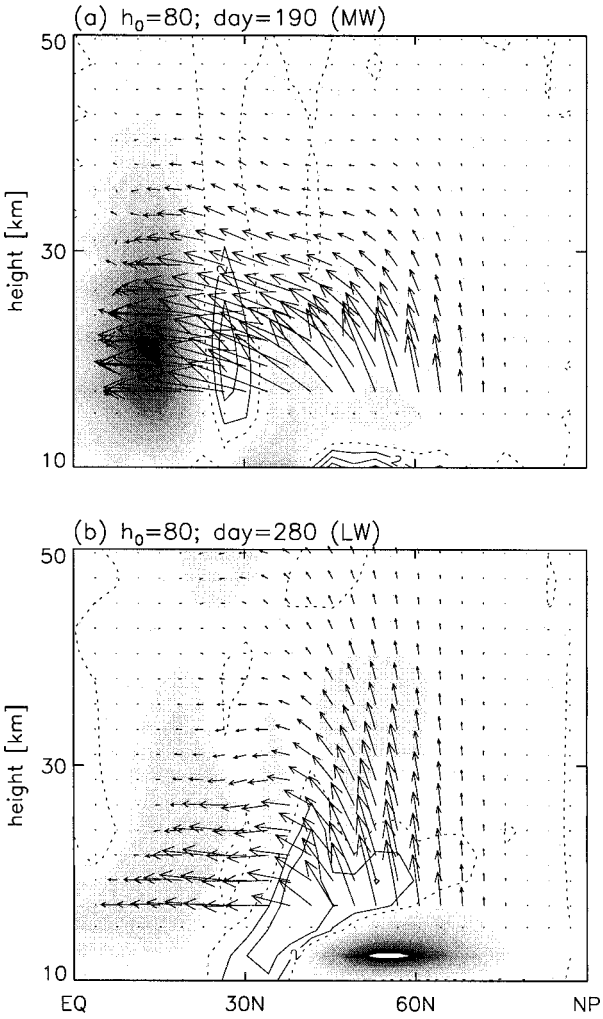

FIG. 8. Same as Fig. 7, but for h_0 = 80 m: (a) day 190 and (b) day 280.

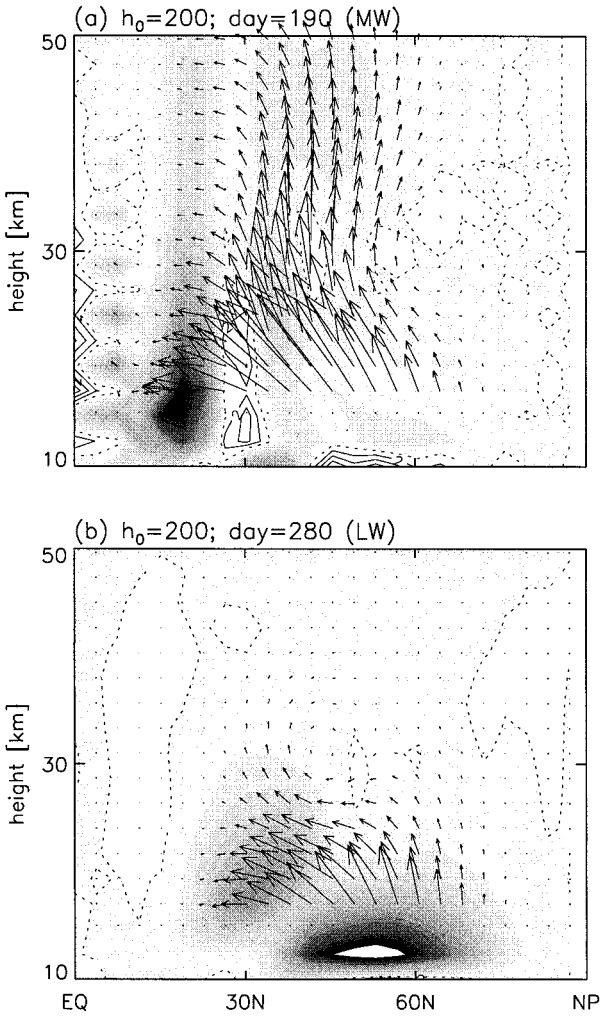

FIG. 9. Same as Fig. 7, but for h_0 = 200 m: (a) day 190 and (b) day 280.

(Fig. 9a). Another clear change is the weaker vertical wave transmission in late winter. The geopotential amplitudes in Figs. 3 and 4 reflect, these changes with stronger wave amplitudes in midwinter and weaker wave amplitudes in late winter: although the early and late winter maxima are still present there is now clear evidence of a midwinter maximum, which eventually dominates at higher forcing values. For h_0 = 200 m, the midwinter maximum is of similar magnitude to the early and late winter maxima, despite significantly more favourable midwinter wave transmission properties, as seen in Fig. 9a. We emphasize again that the transmission properties are insufficient on their own to explain the behavior. It is necessary to take into account the amount of wave activity entering the domain $\langle F^{(z)}|_B\rangle$, as shown in Figs. 5 and 6. Recall also that the EP fluxes and divergences in Figs. 7–9 have been normalized by $\langle F^{(z)}|_B\rangle$ evaluated on the corresponding day.

While considering EP fluxes and divergences, it is interesting to note the structure of the wave evolution immediately following the early winter resonant peak. Using the linear forcing case of h_0 = 20 m as representative,[2] we have plotted in Fig. 10 the EP flux and divergence on days 115–130 with the same scaling as the day-110 fields of Fig. 7a. The strong dipolar structure in the EP flux divergence on day 110 shows a steady southward translation over the following 20 days. On day 120 a second similar but weaker structure emerges at 60°N and follows the first south. The dipolar structure in the divergence, and the orientation of the flux field, resembles a slowly southward moving Rossby-wave packet, although the group velocity of the packet is somewhat slower than that normally associated with freely evolving wave packets on a steady basic state (Karoly and Hoskins 1982, Figure 6b).

[2] In early winter, the evolution is fairly insensitive to the wave forcing: the feature of interest here was obtained almost exactly for forcing values up to h_0 = 100 m and approximately for forcing values values up to h_0 = 160 m.

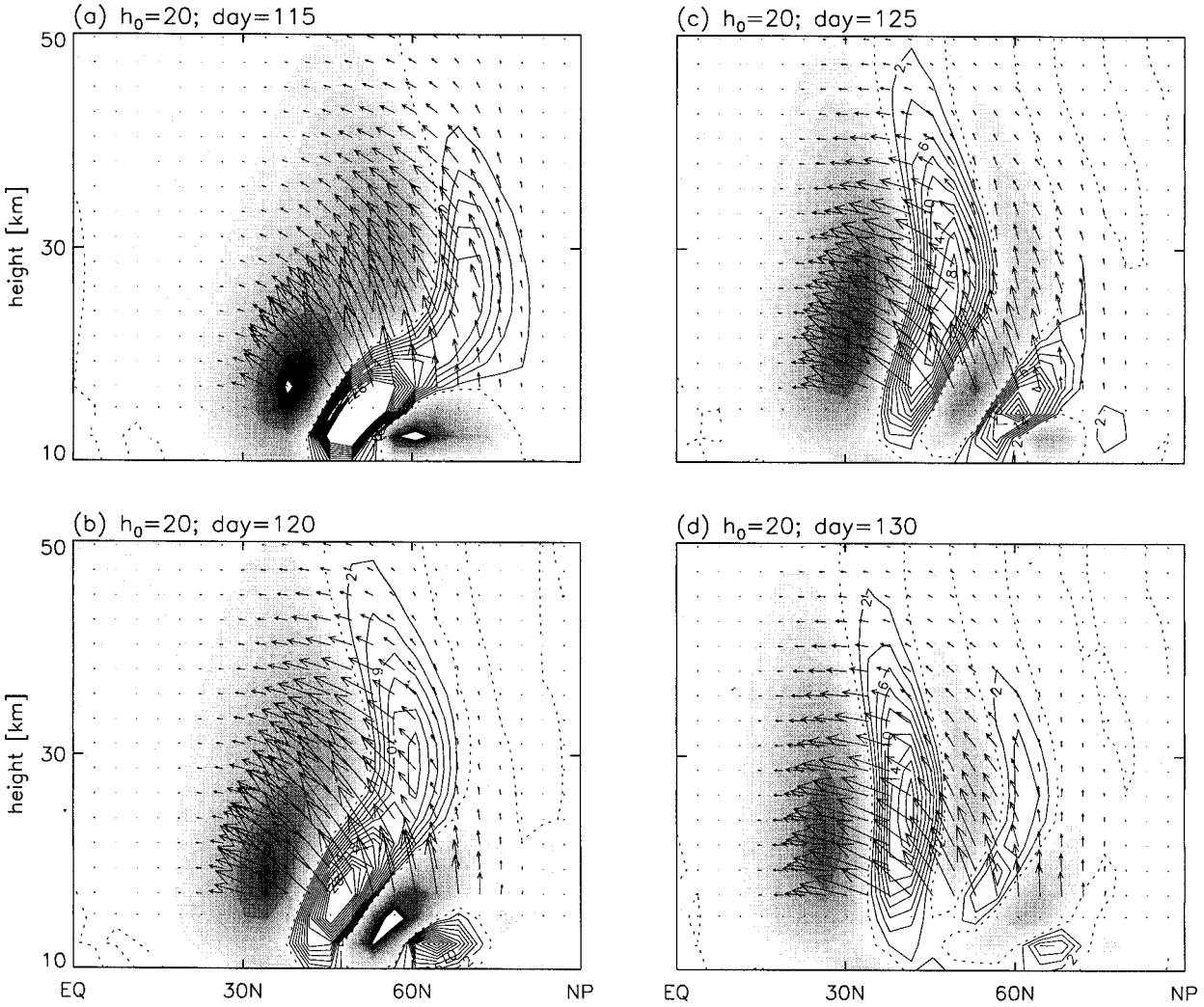

FIG. 10. Same as Fig. 7, but for $h_0 = 20$ m: (a) day 115, (b) day 120, (c) day 125, and (b) day 130; and normalized by the value of $\langle F^{(z)}|_B \rangle$ on day 110, as for Fig. 7a.

We conclude this section by putting the above results into the context of the real stratosphere. From the above discussion it appears, first, that our simple model can reproduce the observed midwinter minimum in geopotential wave amplitudes that is observed in the SH winter stratosphere, and second, that the dominant mechanism involved is a resonance with the lower boundary, rather than, as suggested by Plumb (1989), changing wave transmission properties of the zonal mean flow. However, we should keep in mind that both the position of our lower boundary and also the condition applied there are artificial. Thus our results fall short of explaining the actual observations of the real stratosphere. What we have achieved, rather, is to illustrate a possible mechanism that *may* be present, to a greater or lesser degree, in the real stratosphere. It remains an open question whether a real stratosphere would exhibit resonant behavior with wave forcing generated from a real troposphere. It has even been suggested (McIntyre 1982) that in late winter the SH stratosphere and troposphere together may become resonant with the ground, involving the strong wave-one topography of the Antarctic subcontinent. Such a suggestion could be tested most naturally with a sigma coordinate model, the lower boundary then being a material surface.

5. Nonlinear responses

In this section we continue to explore the model response beyond the quasilinear regime as the forcing amplitude is increased beyond $h_0 = 200$ m. We present the general dependence of the response on the forcing amplitude followed by a more detailed investigation of multiple flow responses observable with a single forcing amplitude.

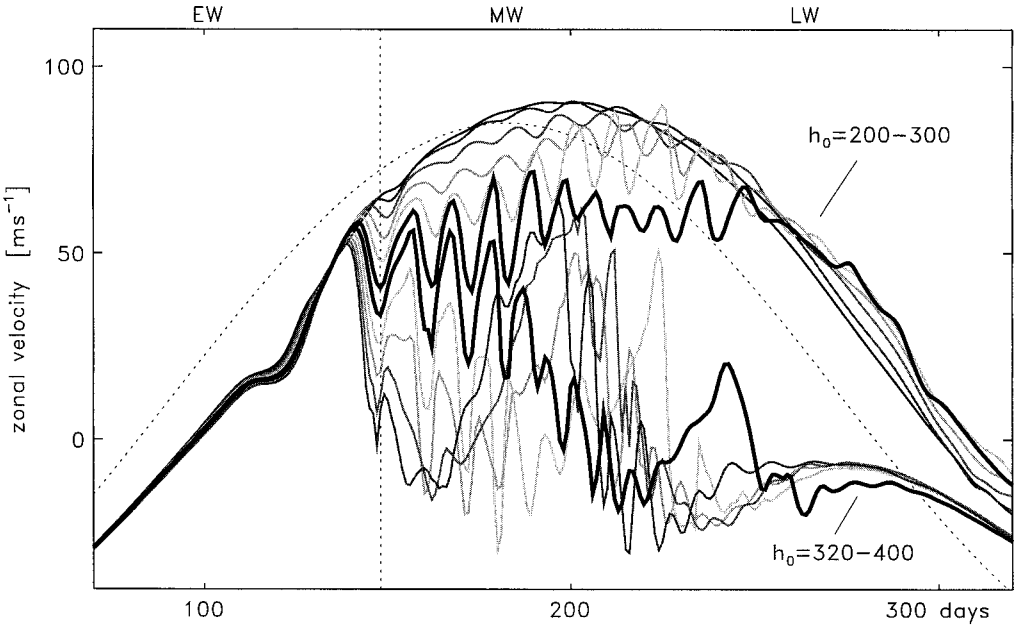

FIG. 11. Zonal mean zonal velocity at 60°N and the 38-km height level as a function of time for forcing amplitudes, reading down the dotted line from top to bottom, h_0 = 200, 220, 240, 260, 280, 300 (thick), 320 (thick), 340, 360, 380, and 400 m. The dotted curve denotes the radiative equlibrium velocity at the same location.

a. Dependence on forcing amplitude

Figure 11 shows the zonal mean velocity at 60°N and at a height of 38 km for forcing values h_0 = 200–400 m at 20-m increments. The upper dark-to-light grey scale corresponds to h_0 varying from 200 to 300 m and the lower light-to-dark scale corresponds to h_0 varying from 320–400 m. A clear transition in the response can be seen when the forcing is increased from h_0 = 300 m to h_0 = 320 m, the transition being most pronounced in late winter. For a forcing amplitude of h_0 = 300 m and below, the winter evolution is relatively undisturbed, with strong westerlies persisting into late winter and spring, and with a smooth change to summer easterlies following the sinusoidal shape of the radiative seasonal cycle (dotted). For forcing amplitude of h_0 = 320 m and above, on the other hand, the winter evolution is strongly disturbed, with large departures from the radiative basic-state and strong transient behavior. Further, the winter westerlies are wiped out long before the seasonal springtime transition to easterlies.

Before going on to a more thorough analysis of the two types of evolution mentioned above, it is worth recalling another feature of nonlinear models of the type introduced in section 2: namely the possibility of multiple stable equilibria. Here, by multiple stable equilibria we mean multiple solutions, which may be steady or unsteady, that are obtainable under identical steady forcing conditions. In the following we use the term stable equilibrium to refer to a solution obtained with steady (wave and radiative) forcing, and the term stable regime to refer to a stable equilibrium type solution obtained with a seasonally evolving radiative basic state (i.e., unsteady radiative forcing). The relevance of transitions between multiple stable regimes to the seasonal cycle of planetary waves in the stratosphere when nonconstant radiative forcing is included was investigated by Yoden (1990) using a simple, height-dependent channel model. There, the transitions were induced by slowly varying a bifurcation parameter that represented the seasonal cycle. In the context of our latitude- and height-dependent model, Yoden's bifurcation parameter, a single number representing a constant vertical wind shear, becomes a latitude- and height-dependent field. It thus becomes difficult in our case to consider rigorously any sort of bifurcation diagram.

One alternative is to look for traces of rapid transitions in the zonal mean velocities at a particular latitude–height location, such as those depicted in Fig. 11. In doing so it is necessary to restrict attention to those periods of the evolution when the changes in the mean flow are otherwise slow and result directly from the changes in the seasonally evolving radiative basic state, at least on one side of the transition. Such considerations applied to Fig. 11 then lead to two possible candidates: the first is the sudden deceleration followed by rapid acceleration at around days 110–120, indicated by the kink in each graph at this time; the second is the rapid deceleration around days 140–150.

First, consider the kink around day 110. It is associated with the early winter resonant peak described in section 4. The peak in the wave activity entering the domain contributes to the retardation of the mean flow from the seasonally evolving radiative basic state, and the following reduction in wave activity allows the re-

covery of the mean flow towards the basic state. This picture is the same as that described in section 2 in the context of the zero dimensional model. Note that the changes in the mean flow during this period are small, so if the kink is related to a transition between two flow regimes, these must be relatively similar. Another point to note is that the kink is observed over a wide range of forcing amplitudes. Under perpetual conditions corresponding to those at the time of the kink, and large forcing, we would not expect such regimes to persist, so they could not be termed stable in the normal sense. However, within the scope of a seasonally evolving basic state this is not essential. On the other hand, Scott and Haynes (2000) observed two nearby and steady stable equilibria in a similar model under perpetual conditions for small forcing values. These were distinquished by their different wave transmission properties and latitudinal potential vorticity (PV) gradients. It is therefore possible that in the present study there is a transition in the structure of the zonal mean velocities, beyond that induced by the seasonal cycle.

Returning to Fig. 11, consider the second candidate mentioned above, the rapid deceleration of the zonal mean zonal velocity around day 140. Again we argue that the notion of transitions between multiple regimes is inappropriate. First, the model of the seasonal cycle discussed in section 2 gives transitions between multiple regimes in early winter that are accelerations, following the upper arrow in Fig. 1 from the stable solution on the left to the stable solution on the right (solid circles). These transitions are opposite in sense to the deceleration seen around day 140 in Fig. 11. Second, there is a steady change in the nature of the deceleration with increasing forcing amplitude, both in the timing of the event and in the magnitude, with an earlier and stronger deceleration for larger forcing amplitude. We conclude therefore that the deceleration is simply a result of the wave forcing on the zonal-mean velocity.

Finally, we note that Yoden (1990) also found evidence of rapid transitions between stable regimes in late winter, a deceleration that qualitatively resembled the observed SH stratospheric final warming (e.g., Mechoso et al. 1988). The simple model of section 2 again illustrates the main features, with the lower arrow in Fig. 1 indicating a transition from the stable solution on the right to the one on the left (open circles). However, on the basis of Fig. 11, there is no signature of a rapid late winter transition in our latitude–height model in the current parameter regime, except possibly for the case $h_0 = 300$ m. Instead, there is a smooth late winter transition to summer easterlies ($h_0 \leq 300$ m), or else the winter vortex is destroyed too early ($h_0 \geq 320$ m). One possible explanation for the absence in the model of a realistic final warming is that the applied lower boundary wave forcing is constant throughout the winter. Randel (1988) shows a marked increase in stationary wave-one at 300 mb and 60°S between August and October, which could contribute to enhanced upward wave propagation at this time.

Although there is no evidence in our model of transitions between stable regimes during the course of the seasonal cycle, Fig. 11 nevertheless suggests that multiple stable regimes may exist. In particular, we consider the two responses to the forcing amplitudes $h_0 = 300$ m and $h_0 = 320$ m as representative of the two stable regimes, which we refer to in the following as E1 and E2. It is interesting to note that Shiotani et al. (1993) also found evidence of two distinct regimes in the observed SH winter evolution in a 10-year dataset of winds and geopotential wave amplitudes. These were characterized by low-latitude and high-latitude jet years according to the location of the maximum of the zonal mean zonal velocity. Although, the differences in zonal structures of the low-latitude and high-latitude jet years and the differences in E1 and E2 are quite different (compare Shiotani et al. 1993, their Fig. 4, e.g., with Figs. 12 and 13 below), in both cases we note that it appears to be the early winter wave evolution that controls the selection of a particular regime. In the following we investigate the range of forcing values over which the regimes E1 and E2 can be observed, that is, are stable over the timescale of the seasonal cycle.

The different structures of the zonal mean velocities of the two regimes are shown in Fig. 12. These have been averaged over days 180–250 to facilitate comparison, since each regime is rapidly varying throughout winter. The structures of the zonal mean velocities are relatively insensitive to the time interval used in the averaging. We see that the regime obtained with forcing amplitude 300 m (i.e., E1) has a well-formed polar vortex with a strong maximum in zonal wind around 70°N in the upper stratosphere (Fig. 12a). It is instructive to consider also the zonal mean PV, scaled by $\theta^{9/2}$ to remove the exponential increase with height (see Lait 1994, for details), where θ is the potential temperature. Figure 13a shows that E1 consists of strong meridional gradients of scaled PV north of 60°N, a broad surf zone of weaker gradients in middle latitudes, and a narrow region of strong gradients in the subtropics. Both the zonal-mean velocity and the zonal-mean scaled PV of E1 should be contrasted to the corresponding fields for the regime obtained with forcing amplitude 320 m, E2, shown in Figs. 12b and 13b. Although the subtropical region of strong PV gradients is still present, there are no strong PV gradients at high latitudes and the ''midlatitude surf zone'' now extends all the way to the pole. Similarly, the zonal-mean velocities are substantially weaker in middle and high latitudes and there is no clear vortex structure. An additional diagnostic, the volume of the vortex, was used recently by Polvani and Saravanan (2000) to quantify the three-dimensional effect of wave breaking on an idealized polar vortex. The same diagnostic applied here again illustrates the difference between the two states: the contour of scaled PV of value 40 marks the vortex edge near 60° throughout the

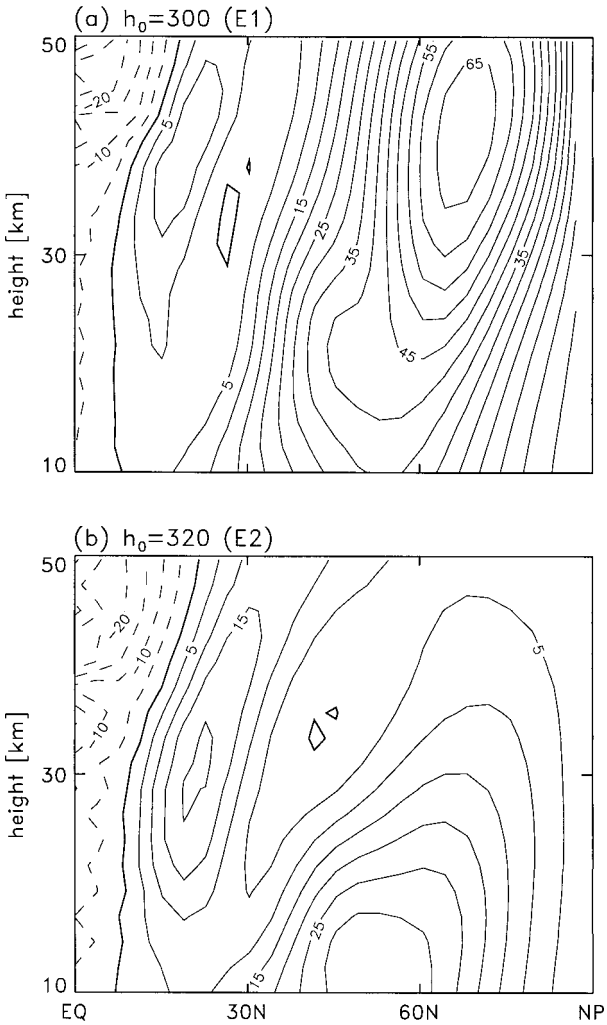

FIG. 12. Zonal mean zonal velocity averaged over days 180–250 as a function of latitude and height: (a) $h_0 = 300$ m (state E1), and (b) $h_0 = 320$ m (state E2). The contour interval is 5 m s^{-1} and dashed contours are negative.

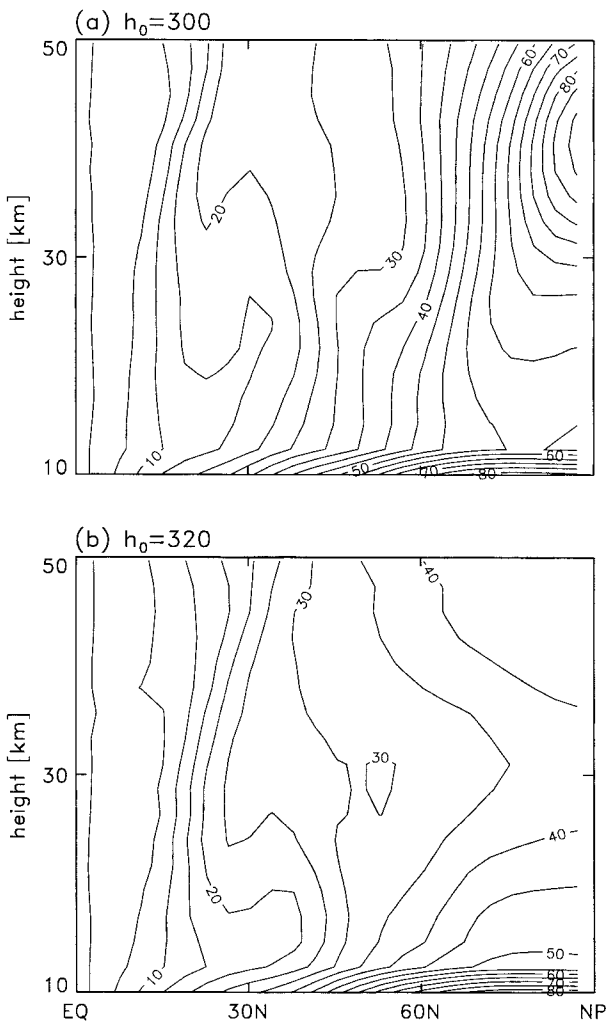

FIG. 13. Zonal-mean scaled potential vorticity [Ertel PV divided by $(\theta/\theta_0)^{9/2}$ with $\theta_0 = 420$ K] averaged over days 180–250 as a function of latitude and height: (a) $h_0 = 300$ m (state E1), and (b) $h_0 = 320$ m (state E2). The contour interval is 5, in units of $10^{-6} \times$ K kg^{-1} m^2 s^{-1}.

whole of the stratosphere in E1, but does not even span the vertical in E2.

b. Multiple flow regimes

The above demonstrates that for two slightly different forcing amplitudes, two very different responses are obtained. This sensitivity is interesting but it falls short of verifying the existence of multiple stable regimes, that is, multiple responses that can be obtained for a single forcing amplitude. To address this we now present some additional experiments, in which the early winter forcing amplitude is either larger or smaller than the final forcing amplitude used throughout the remainder of the seasonal cycle. In doing so it is our intention to start the winter evolution in state E1, then to increase the forcing to an amplitude that is associated (by Fig. 11) with state E2, and determine whether E1 persists or E2 develops. The same will also be done in reverse, starting with E2 and looking for the development of E1 after reducing the forcing amplitude at a certain time.

We first consider simulations in which the state E1 is initially established by integrating the model with no wave forcing, following which we increase the forcing to a large final value h_{final} after a certain time t_{delay}. The time dependence of the forcing amplitude is illustrated by the solid line in Fig. 14. We want to determine whether there is some value of t_{delay} for which the state E1 persists, even when h_{final} is large. By large we mean of sufficient amplitude to result, in the absence of any time delay, in the state E2, that is, amplitudes of 320 m and higher.

The results of several simulations, with various values of t_{delay} and h_{final} are shown in Fig. 15, which displays the zonal mean zonal velocity at 60°N and 38 km as in Fig. 11. For the present, ignore the dotted lines. The

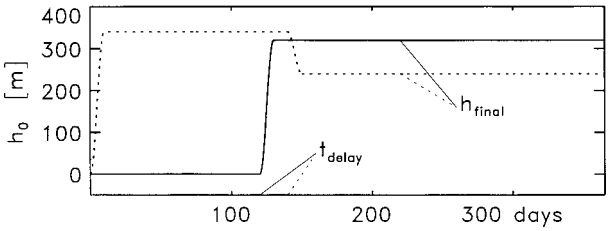

FIG. 14. Examples of the time dependence of the forcing amplitude used in simulations in which the forcing is delayed by t_{delay} days (solid) or is reduced to its final amplitude h_{final} after stronger early winter forcing (dotted).

thick solid lines in each panel represent values of t_{delay} of 120, 140, and 160 days. Simulations using $t_{delay} = 0$ (shown previously in Fig. 11) have been replotted here using a thin solid line. All the simulations with $h_{final} =$ 300 m (Fig. 15a) remain in the state E1, which should be expected since a forcing amplitude of $h_0 = 300$ m, applied with no delay, also results in E1 as seen in Fig. 11. However, from Fig. 15b we see that all simulations using $t_{delay} > 0$ with $h_{final} = 320$ m also remain in the state E1. This should be contrasted with the E2 response obtained with $t_{delay} = 0$, shown by the thin line in Fig. 15b. We thus find two different mid-to-late winter responses, E1 and E2, that can be obtained for the same value of mid-to-late winter forcing amplitude, the only difference between the simulations being in the early winter forcing amplitude. Similarly, in Fig. 15c with $h_{final} = 340$ m we see that those simulations with $t_{delay} = 140$ and 160, also remain in the state E1. In this case we note that a delay of 120 days is insufficient to retain E1, and that the stronger final forcing amplitude is capable of pulling the model response into E2.

The vertical structure of the zonal mean zonal velocity at 60°N for states E1 and E2 is shown in Figs. 16a,b for the undelayed forcing cases with $h_0 = 300$ m and $h_0 = 320$ m (corresponding to the heavy lines of Fig. 11), and in Fig. 16c, for the delayed forcing case with $t_{delay} = 140$ days and $h_{final} = 340$ m. It is particularly interesting that even after strong sudden warming type events around days 235 and 270, the evolution returns to the relatively undisturbed state of E1, rather than experiencing further vortex reduction and entering the disturbed state of E2. This can be explained by consideration of polar stereographic plots of PV (not shown), which indicate that the vortex shrinks and is pushed strongly off the pole but otherwise remains intact.

Returning to the dotted lines in Fig. 15, which show results for $t_{delay} = 100$ days, we note that all three final forcing amplitudes $h_{final} = 300$, 320, 340 m produce what appears to be yet another evolution state. This is characterized by early winter zonal mean zonal velocities that are significantly weaker than those of both E1 and E2. It is unclear the extent to which this new state persists through the rest of the winter, or whether it approaches one of the states E1 (as in Fig. 15b) or E2 (as in Figs. 15a,c). Thus it is difficult to determine the

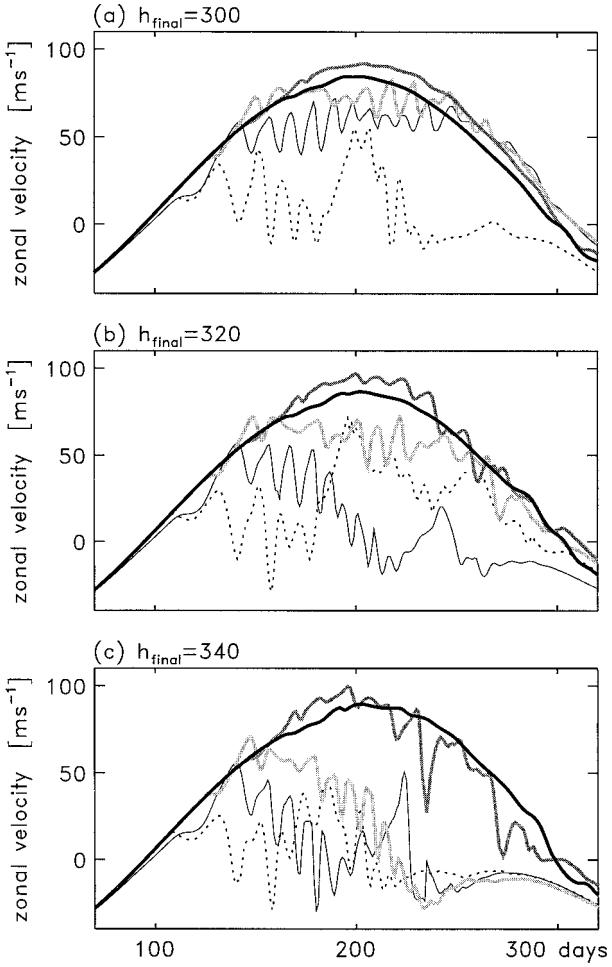

FIG. 15. Zonal mean zonal velocity at 60°N and the 38-km height level as a function of time: (a) $h_{final} = 300$, (b) $h_{final} = 320$, and (c) $h_{final} = 340$. Within each panel the forcing is delayed with delay $t_{delay} = 160$ days (thick black), $t_{delay} = 140$ days (thick gray), $t_{delay} = 120$ days (thick light gray), $t_{delay} = 100$ days (dotted). The corresponding evolution with constant forcing amplitude h_{final}, that is, $t_{delay} = 0$ days, taken from Fig. 11, is also included (thin solid).

extent to which this state is a stable evolution state on the timescale of the seasonal evolution. It seems likely, however, that the strong deceleration of the early winter vortex is linked to the transient growth of the lower boundary wave forcing occurring at the same time as the zonal-mean basic state passes through the resonance. This is supported by the sensitivity of the response to the delay of the forcing, with a disturbed early winter only resulting from t_{delay} in the range 95 to 105 days. A more detailed investigation is beyond the scope of the present paper but is planned as future work. Such an interaction of wave transience with a developing vortex structure could have implications for ''early winter warmings'' of the type reported by Juckes and O'Neill (1988) and Farrara et al. (1992).

Finally we consider the reverse of the above, that is, simulations in which the state E2 is established with a

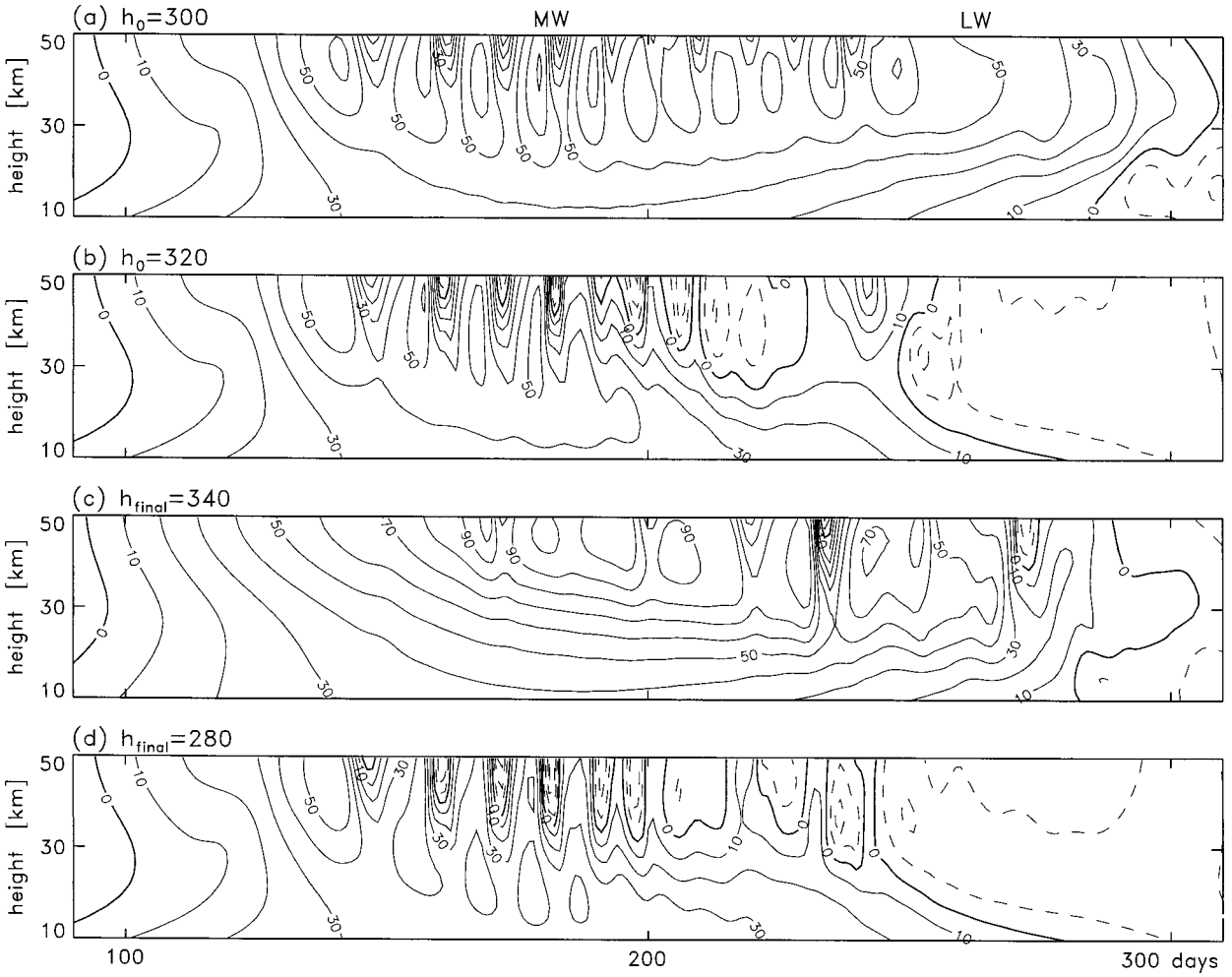

FIG. 16. Zonal mean zonal velocity at 60°N as a function of time and height: (a) constant forcing amplitude h_0 = 300 m (i.e., t_{delay} = 0), (b) constant forcing amplitude h_0 = 320 m (t_{delay} = 0), (c) final forcing amplitude h_{final} = 340 after t_{delay} = 140 days with zero initial forcing, and (d) final forcing amplitude h_{final} = 280 after t_{delay} = 150 days and initial forcing $h_{initial}$ = 340.

large forcing amplitude $h_{initial}$ = 340 m before decreasing the forcing to a smaller final value h_{final} after a certain time t_{delay}. The time dependence of the forcing amplitude is illustrated by the dotted line in Fig. 14. Now, we want to determine whether there is some time delay t_{delay} for which the state E2 persists even when the final forcing amplitude h_{final} is small. Again, by small we mean small enough that, in the absence of strong forcing in early winter, it would result in the state E1, namely, from Fig. 11, amplitudes of 300 m and lower.

The results of several simulations, with various values of t_{delay} and h_{final} are shown in Fig. 17, which again displays the zonal mean zonal velocity at 60°N and 38 km as in Fig. 11. The thick solid lines in each panel represent values of t_{delay} of 140, 150, and 160 days, and the thin solid line represents t_{delay} = 0 days. For each of the final forcing amplitudes h_{final} = 240, 260, and 280 m, using t_{delay} = 140 days always results in the evolution state E1 returning by midwinter. That is, the initial forcing amplitude $h_{initial}$ = 340 m is not present for enough of the winter evolution to ensure that the state E2 is established. On the other hand, a delay of t_{delay} = 160 days is long enough to establish the state E2 and this state then persists throughout the rest of the winter evolution, for each of the three final forcing amplitudes. For the intermediate value of t_{delay} = 150 days, the state E2 persists only for the final forcing amplitude of h_{final} = 280 m. The main point is again that it is possible for each of the states E1 and E2 to persist throughout the mid- and late winter evolution under identical forcing amplitudes, with the selection of the particular state determined by the forcing amplitudes in early winter only. For comparison with the representative states E1 and E2 obtained with the forcing conditions corresponding to Fig. 11 (h_0 = 300- or 320-m constant throughout the whole winter), the state E2 obtained with initial and final forcing amplitudes $h_{initial}$ = 340 m and h_{final} = 280 m, and delay t_{delay} = 150 days is shown in Fig. 16d.

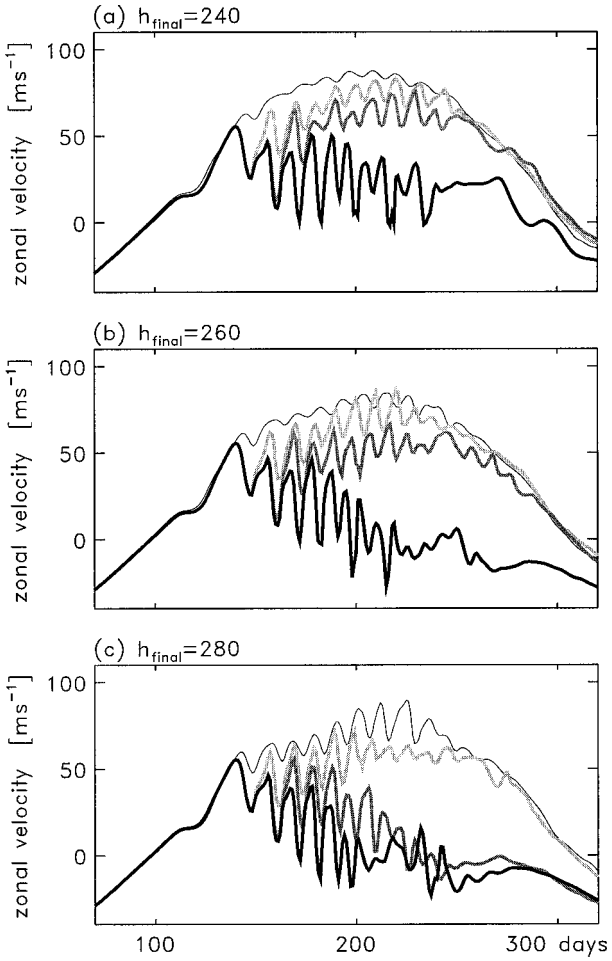

FIG. 17. Zonal mean zonal velocity at 60°N and the 38-km height level as a function of time: (a) $h_{final} = 240$, (b) $h_{final} = 260$, and (c) $h_{final} = 280$. Within each panel the forcing is initially $h_0 = 340$ m and is reduced to h_{final} after a time $t_{delay} = 160$ days (thick black), $t_{delay} = 150$ days (thick gray), $t_{delay} = 140$ days (thick light gray). The corresponding evolution with constant forcing amplitude h_{final}, taken from Fig. 11, is also included (thin solid).

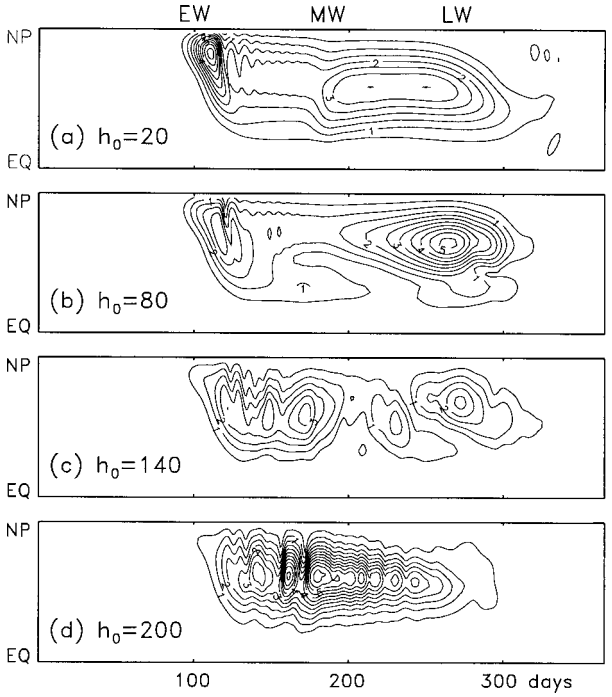

FIG. 18. Geopotential wavenumber one amplitude at the 38-km height level normalized by the wave-one forcing amplitude h_0 as a function of latitude and time (cf. Fig. 4) from a model that retains the first three zonal wavenumbers: (a) $h_0 = 20$ m, (b) $h_0 = 80$ m, (c) $h_0 = 140$ m, and (d) $h_0 = 200$ m. The contour interval is 0.5.

6. Sensitivity to the inclusion of more zonal wavenumbers

We finish with a brief examination of the robustness of the features described in sections 4 and 5 to the inclusion of higher zonal wavenumbers in the model equations. In doing so we retain the simple (and constant, i.e., $t_{delay} = 0$) wavenumber one form of the lower boundary forcing, but simply allow higher zonal wavenumbers to develop as a result of the nonlinearity of the equations. We consider only the effect of including wavenumbers up to three, since we expect this to give the most significant difference from the wavenumber one only integrations.

We expect that the inclusion of wavenumbers two and three will only make a small difference to the linear and quasi-linear regimes discussed in section 4, in which nonlinearity is weak. This is verified in Fig. 18, which shows the geopotential wave-one amplitude at 38 km, and which should be compared to Fig. 4. For $h_0 = 20$ m we see that the response is indeed very close to linear. For larger values of h_0 larger differences between the two evolutions are seen, but the qualitative picture is the same. These differences essentially consist of a smaller late winter maximum for quasi-linear forcing ($h_0 = 80$ m, $h_0 = 140$ m) a partial midwinter maximum already for $h_0 = 140$ m, and a larger midwinter maximum for $h_0 = 200$ m when the extra wavenumbers are included. Figure 19 shows the wavenumbers two and three geopotential wave amplitudes for the $h_0 = 80$ m and $h_0 = 140$ m cases. For $h_0 = 80$ m there is stronger nonlinearity in late winter than in early winter, whereas for $h_0 = 140$ m there is stronger nonlinearity throughout most of the winter.

For stronger forcing amplitudes we expect nonlinearity to be stronger and hence the possibility of significant deviations from the pattern of responses described in section 5. In fact we see from Fig. 20 that the same qualitative pattern is robustly reproduced, albeit with the transition between different regimes occuring at lower values of h_0 (compare Fig. 11). There is a clear distinction between the quasi-linear responses characterized by a relatively strong vortex throughout the winter ($h_0 \leq 220$ m) and the nonlinear responses characterized by a strongly disturbed vortex, most notable about a month after midwinter ($h_0 \leq 240$ m). In particular note

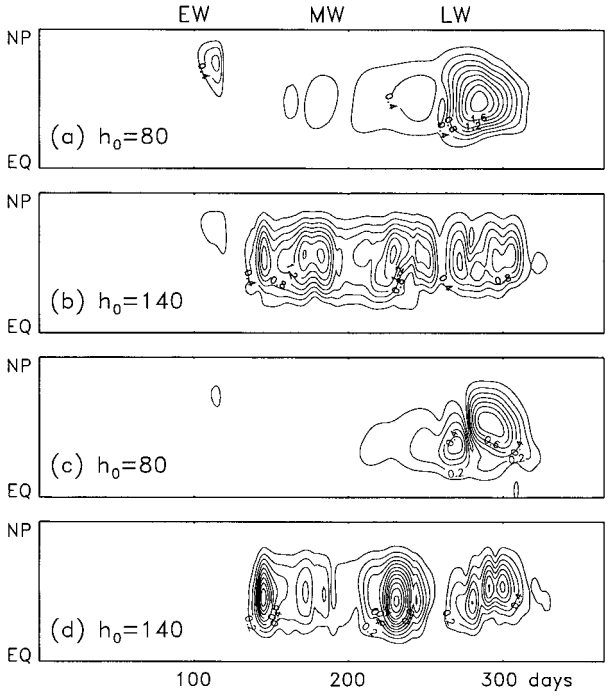

FIG. 19. Same as Fig. 18, but for the geopotential wavenumber two amplitude (a) h_0 = 80 m, (b) h_0 = 140 m, and geopotential wavenumber three amplitude (c) h_0 = 80 m, and (d) h_0 = 140 m. The contour interval is 0.2 in (a, b) and 0.1 in (c, d).

that, as for the wavenumber one only simulations reported earlier, no choice of forcing value produces a realistic final warming. For quasi-linear forcing there is a smooth transition in late spring from winter westerlies to summer easterlies, whereas for nonlinear forcing the vortex is destroyed too early (around a month after midwinter) and does not persist long enough for the notion of a final warming to be appropriate.

7. Conclusions

We have used a simple mechanistic model to examine features of the wintertime stratosphere evolution. Particular attention was given to how the inclusion of latitudinal structure in the model altered the behavior observed in previous studies using channel models with structure only in the vertical. The model exhibits behavior qualitatively similar to that of the real atmosphere, with Southern Hemisphere-like evolution for weaker forcing amplitudes and Northern Hemisphere-like evolution for stronger forcing amplitudes. The SH-like evolution showed both a stronger winter polar vortex than that of the NH evolution, as well as early and late winter maxima in geopotential wave amplitude throughout much of the stratosphere, similar to that documented by Randel (1988). The NH-type evolution showed strong transient events resembling stratospheric sudden warmings. Finally, two evolution states were identified that could both be obtained over the same range of forcing amplitudes, that is, the model exhibited two quasi-stable regimes. All the above features were

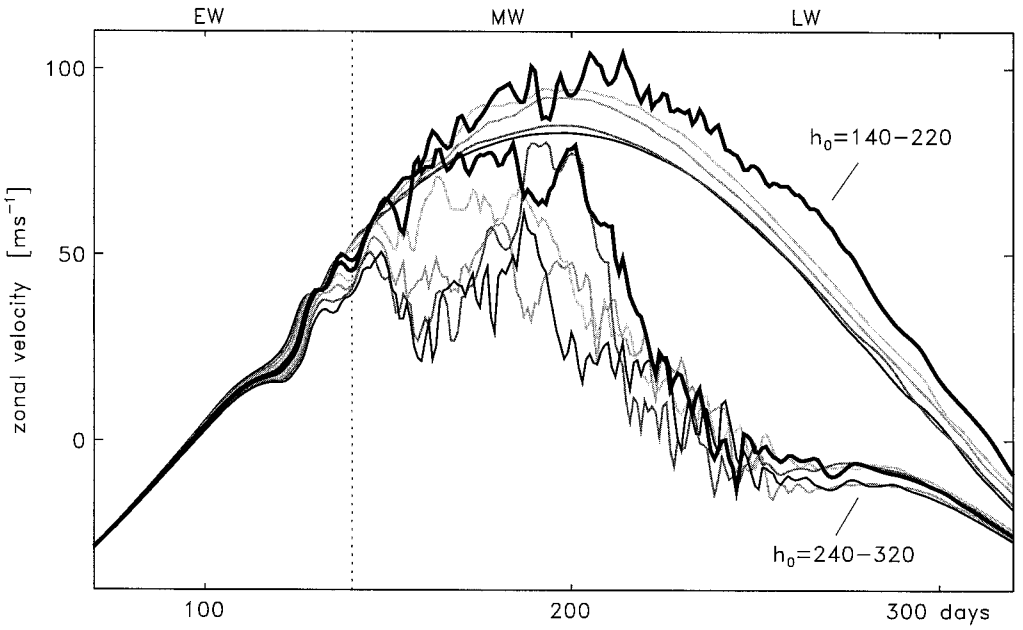

FIG. 20. Zonal mean zonal velocity at 60°N and the 38-km height level as a function of time from a model that retains the first three zonal wavenumbers for forcing amplitudes; reading down the dotted line from top to bottom, h_0 = 140, 160, 180, 200, 220 (thick), 240 (thick), 260, 280, 300, 320 m. (cf. Fig. 11).

found to be insensitive to inclusion of extra zonal wavenumbers in the model equations.

For small forcing amplitude a linear or quasi-linear model response was obtained. The stratospheric wave evolution in these cases resembled that of the SH, with early and late winter maxima in geopotential wave amplitude and an intervening midwinter minimum. This pattern was interpreted as resulting from a resonance of the system with the wave forcing in early and late winter, in analogy with that observed in a simple zero-dimensional wave–zonal-mean model. Forced–damped linear experiments confirmed this interpretation, and demonstrated that the implicit quasi-static assumptions were reasonable. Analysis of the wave transmission properties of the zonal-mean flow showed that conditions were more favorable to midlatitude upward wave propagation in early winter, but that there was little difference between those in midwinter and those late in winter. This is to be contrasted to the investigation of Plumb (1989) using a height-dependent channel model that found weaker midwinter wave amplitudes associated with the prevention of vertical wave propagation by strong midwinter zonal mean zonal velocities. Our results emphasize that when latitudinal structure is included, stronger midwinter zonal mean zonal velocities alone are insufficient to reduce upward wave propagation; full consideration must be taken of the height–latitude structure of the mean flow.

For larger forcing amplitude a nonlinear model response was obtained that more closely resembled the NH stratosphere wave evolution, with large geopotential wave amplitudes throughout winter. Two different evolution states, E1 and E2, were obtained for the same forcing values: in E1 the winter vortex remained relatively strong, in E2 it was significantly reduced by the action of the waves on the mean flow. Which winter evolution a particular model simulation selected was found to be dependent on the wave forcing amplitude in early winter, during the development stage of the polar vortex.

Despite the model exhibiting different evolution states there was no evidence of rapid transitions between the two, either in the sudden-warming setting of Chao (1985), or in the seasonal-evolution setting of Yoden (1990) described in section 2. In particular, the model failed to exhibit a realistic SH-like final warming in the late winters of the evolution state E1, even for wave-forcing amplitudes that were capable of producing the much more disturbed evolution state E2. One possible reason for this is the constant-amplitude wave forcing used in our study. In addition to the midwinter minimum in stratospheric wave amplitudes, Randel (1988) also reported an increase in the wave amplitudes in late winter in the lower stratosphere. Our results suggest that transient wave forcing might be needed to reproduce a realistic final warming in a SH-like evolution.

Acknowledgments. We thank the referees, particularly Charles McLandress, for valuable suggestions and detailed comments on the first version of the paper. Financial support for this work was provided by the U.K. Natural Environmental Research Council and the Met Office through a CASE studentship. The work of the Centre for Atmospheric Science is supported by the U.K. Natural Environment Research Council, in part through the U.K. Universities Global Atmospheric Modelling Programme, and by the Isaac Newton Trust. Richard K. Scott is currently supported by the European Commission under Contract EVK2-1999-00015.

REFERENCES

Andrews, D. G., J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics.* Academic Press, 489 pp.

Chao, W. C., 1985: Sudden stratospheric warmings as catastrophes. *J. Atmos. Sci.,* **42,** 1631–1646.

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.

——, and J. G. DeVore, 1979: Multiple flow equilibria in the atmosphere and blocking. *J. Atmos. Sci.,* **36,** 1205–1216.

Clark, P. D., 1992: Internal variability of the middle atmosphere. Ph.D. dissertation, University of Cambridge, 364 pp.

Dunkerton, T. J., C.-P. F. Hsu, and M. E. McIntyre, 1981: Some Eulerian and Lagrangian diagnostics for a model stratospheric warming. *J. Atmos. Sci.,* **38,** 819–843.

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.

Haltiner, G. J., and R. T. Williams, 1980: *Numerical Prediction and Dynamic Meteorology.* Wiley, 477 pp.

Haynes, P. H., and M. E. McIntyre, 1987: On the representation of Rossby wave critical layers and wave breaking in zonally truncated models. *J. Atmos. Sci.,* **44,** 2359–2382.

Held, I. M., 1983: Stationary and quasi-stationary eddies in the extratropical atmosphere: Theory. *Large-Scale Dynamical Processes in the Atmosphere,* R. P. Pearce and B. J. Hoskins, Eds., Academic, 127–168.

Hirota, I., T. Hirooka, and M. Shiotani, 1983: Upper stratospheric circulations in the two hemispheres observed by satellites. *Quart. J. Roy. Meteor. Soc.,* **109,** 443–454.

Holton, J. R., 1976: A semi-spectral numerical model for wave, mean-flow interactions in the stratosphere: Application to sudden stratospheric warmings. *J. Atmos. Sci.,* **33,** 1639–1649.

Hoskins, B. J., and A. J. Simmons, 1975: A multi-layer spectral model and the semi-implicit method. *Quart. J. Roy. Meteor. Soc.,* **101,** 637–655.

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

Karoly, D., and B. J. Hoskins, 1982: Three dimensional propagation of planetary waves. *J. Meteor. Soc. Japan,* **60,** 109–123.

Labitzke, K., 1980: Climatology of the stratosphere and mesosphere. *Philos. Trans. Roy. Soc. London,* **A296,** 7–18.

Lait, L. R., 1994: An alternative form for potential vorticity. *J. Atmos. Sci.,* **51,** 1754–1759.

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

McIntyre, M. E., 1982: How well do we understand the dynamics of stratospheric warmings? *J. Meteor. Soc. Japan,* **60,** 37–65.

——, 1990: Middle atmospheric dynamics and transport: Some current challenges to our understanding. *Dynamics, Transport and Photochemistry in the Middle Atmosphere of the Southern Hemisphere: San Francisco NATO Workshop,* A. O'Neill, Ed., Kluwer, 1–18.

Mechoso, C. R., A. O'Neill, V. D. Pope, and J. D. Farrara, 1988: A study of the stratospheric final warming of 1982 in the Southern Hemisphere. *Quart. J. Roy. Meteor. Soc.,* **114,** 1365–1384.

Plumb, R. A., 1989: On the seasonal cycle of stratospheric planetary waves. *Pure Appl. Geophys.,* **130,** 233–242.

Polvani, L. M., and R. Saravanan, 2000: The three-dimensional structure of breaking Rossby waves in the polar wintertime stratosphere. *J. Atmos. Sci.,* **57,** 3663–3685.

Randel, W. J., 1988: The seasonal evolution of planetary waves in the Southern Hemisphere stratosphere and troposphere. *Quart. J. Roy. Meteor. Soc.,* **114,** 1385–1409.

Saravanan, R., 1992: A mechanistic spectral primitive equation model using pressure coordinates. University of Cambridge Model Documentation, DAMTP.

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 ——, 2000: Internal vacillations in stratosphere only models. *J. Atmos. Sci.,* **57,** 3233–3250.

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

Simmons, A. J., 1974: Planetary-scale disturbances in the polar winter stratosphere. *Quart. J. Roy. Meteor. Soc.,* **100,** 76–108.

Wakata, Y., and M. Uryu, 1987: Stratospheric multiple equilibria and seasonal variations. *J. Meteor. Soc. Japan,* **65,** 27–42.

Waugh, D. W., W. J. Randel, S. Pawson, P. A. Newman, and E. R. Nash, 1999: Persistence of the lower stratospheric polar vortices. *J. Geophys. Res.,* **104,** 27 191–27 201.

Wirth, V., 1991: What causes the seasonal cycle of stationary waves in the southern stratosphere? *J. Atmos. Sci.,* **48,** 1194–1200.

WMO, 1999: Scientific assessment of ozone depletion, 1998. WMO Global Ozone Research and Monitoring Project Rep. 44.

Yoden, S., 1987: A new class of stratospheric vacillations in a highly truncated model due to wave interference. *J. Atmos. Sci.,* **44,** 3696–3709.

——, 1990: An illustrative model of seasonal and interannual variations of the stratospheric circulation. *J. Atmos. Sci.,* **47,** 1845–1853.