Internal Variability of the Winter Stratosphere. Part I: Time-Independent Forcing

R. K. SCOTT

NorthWest Research Associates, Inc., Bellevue, Washington

L. M. POLVANI

Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York

ABSTRACT

This paper examines the nature and robustness of internal stratospheric variability, namely the variability resulting from the internal dynamics of the stratosphere itself, as opposed to that forced by external sources such as the natural variability of the free troposphere. Internal stratospheric variability arises from the competing actions of radiative forcing, which under perpetual winter conditions strengthens the polar vortex, and planetary wave breaking, which weakens it. The results from a stratosphere-only model demonstrate that strong internal stratospheric variability, consisting of repeated sudden warming-type events, exists over a wide range of realistic radiative and wave forcing conditions, and is largely independent of other physical and numerical parameters. In particular, the coherent form of the variability persists as the number of degrees of freedom is increased, and is therefore not an artifact of severe model truncation. Various diagnostics, including three-dimensional representations of the potential vorticity, illustrate that the variability is determined by the vertical structure of the vortex and the extent to which upward wave propagation is favored or inhibited. In this paper, the variability arising from purely internal stratosphere dynamics is isolated by specifying thermal and wave forcings that are completely time independent. In a second paper, the authors investigate the relative importance of internal and external variability by considering time-dependent wave forcing as a simple representation of tropospheric variability.

1. Introduction

It is by now widely accepted that short-term variability of the extratropical stratosphere can arise from variability external to the stratosphere, in particular from variability of the tropospheric circulation, a principle source of planetary-scale wave forcing. The most dramatic example of such short-term stratospheric variability is the so-called sudden warming, during which the radiatively driven cyclonic circulation of the winter stratospheric polar vortex undergoes a rapid wave-induced deceleration, in extreme cases even reversing to easterly flow, and polar temperatures increase by several tens of degrees in the space of a few days. Numerous previous numerical and observational studies, dating back to the pioneering work of Matsuno (1971), have linked these events to anomalously large tropospheric wave amplitudes; see, for example, Andrews et al. (1987) and, more recently, the review by Haynes (2005).

Another way in which stratospheric variability can arise is through the action of dynamical mechanisms internal to the stratosphere itself. Studies dating back to Holton and Mass (1976) have shown how vacillating states can arise in zonally and meridionally truncated quasigeostrophic channel models as periodic solutions to the equations of motion. The main requirement of these models is the competing actions of thermal relaxation to a zonally symmetric basic state together with a planetary-scale wave forcing at the lower boundary. Later studies explored the parameter space defined by the radiative and wave forcing amplitudes and described the transition to vacillating states variously in terms of a Hopf bifurcation from a steady solution branch (e.g., Yoden, 1987), or a catastrophe-like transition from one stable solution branch to another (Chao 1985). Christiansen (2000a) demonstrated the existence of further bifurcations leading ultimately to chaotic solutions of the system.

Corresponding author address: R. K. Scott, NorthWest Research Associates, Inc., P.O. Box 3027, Bellevue, WA 98009-3027.
E-mail: scott@nwra.com

Extending this line of investigation, Scott and Haynes (2000) considered the additional effect of more complete meridional structure in a primitive equation model, yet still within the context of a zonally truncated system, and showed that internal variability persisted over a wide range of forcing amplitudes. Variability has also been found in models with isotropic horizontal resolution. Pawson et al. (1995) found variability in a controlled general circulation model (GCM) simulation, although the character of the variability was not coherent and had a frequency spectrum similar to red noise. The variability obtained by Christiansen (1999, 2000b) in a different GCM, on the other hand, showed more coherence. Even in the latter study however, the variability is not particularly regular in time, and there is a marked decrease in regularity as horizontal resolution is increased from T21 to T42. In a shallow water model of a forced polar vortex, Rong and Waugh (2004) also found a reduction in regularity as horizontal resolution was increased from T42 to T85. Given these ambiguities it remains inconclusive whether coherent variability is an inherent feature of the dynamics of the stratosphere.

Our intent, therefore, is to isolate the dynamical processes involved in internal stratospheric variability in as comprehensive a model as possible that nevertheless allows a simple interpretation of the model evolution. Since we are interested in internal stratospheric variability the natural choice is a GCM dynamical core, run in stratosphere-only mode and with a minimal number of extra physical processes included. Such a model has the additional advantage of permitting a thorough exploration of parameter space. The work presented here extends the preliminary study of Scott and Polvani (2004), which illustrated coherent variability in a GCM dynamical core with a constrained troposphere. We note that, in contrast to recent work by Taguchi et al. (2001), we are not concerned here with the stratosphere–troposphere coupled system, but are rather interested in carefully isolating the dynamical processes within the stratosphere itself. In a subsequent paper, we will address the issue of how these processes are modified under the influence of time-dependent tropospheric forcing.

A major concern and motivation for the present work lies in the fact that variability in dynamical systems often loses coherence as the number of degrees of freedom of the system increases. This was illustrated by Cehelsky and Tung (1987) for the case of multiple flow regimes in the troposphere. The possibility therefore exists that the variability obtained in the Holton–Mass-type models arises purely as an artifact of the severe numerical truncation. Our earlier work (Scott and Pol-

vani 2004) provides a first indication that this is not so in the case of the winter stratosphere, but that coherent (i.e., regular or quasi periodic) variability persists at higher (isotropic) resolution. The present paper extends that work and considers robustness and sensitivity of variability, as well as the dynamical mechanisms involved.

A further motivation for the present work arises from recent studies indicating the possible influence of the stratospheric circulation on both the short and long time scale evolution of the troposphere (e.g., Baldwin and Dunkerton 1999; Polvani and Kushner 2002). For example, if all stratospheric variability is generated by mechanisms whose origin lies entirely within the troposphere, then the stratosphere in itself can never be considered as having any influence on the tropospheric circulation. In this case, it can only act as a mediator of tropospheric variability, though perhaps having an integrating effect in time (e.g., Norton 2003). On the other hand, the existence of internal stratospheric variability raises the possibility of actual downward influence, whereby the stratosphere itself exerts a direct influence on the tropospheric circulation. Finally, an important issue in climate change is the extent to which internal stratospheric variability depends on long-term changes in the radiative forcing of the stratosphere.

This paper presents a detailed account of the persistence of coherent internal variability at a resolution that approaches that which is able to provide a reasonable representation of the dynamical processes involved, and which is comparable to the resolution of current generation comprehensive general circulation models. We present results over a wide range of physical and numerical parameters and find that coherent variability exists under general conditions. Further, we present a detailed description of the vortex structure and dynamical process involved in the variability. Diagnostics based on zonal means, wave fluxes, and three-dimensional potential vorticity fields provide a consistent picture of the variability arising from the competing actions of radiative cooling and wave forcing. In the companion paper we consider how the internal variability is modified when variability is introduced into the external forcing through a simple modulation of the lower boundary wave forcing.

The outline of the paper is as follows. In section 2, we describe our approach and numerical methods. In section 3, we discuss an example of the variability obtained with a particular set of forcing parameters and describe the dynamical processes and the time-dependent evolution of the vortex structure. In section 4, we examine in detail the dependence of the variability on the main parameters defining the thermal and wave forcing. In

section 5, we consider the robustness of the variability and its dependence on other physical and numerical parameters, including the radiative relaxation profile and forcing wavenumber, and the horizontal and vertical resolutions, and sponge layer. Section 6 concludes.

2. Method

In this work, our aim is not to reproduce exact features of the stratospheric circulation, and their dependency on details of thermal transfer, chemical process, gravity wave drag, and so on. Rather, we wish to isolate a fundamental dynamical property, namely the ability of the stratosphere to generate variability in itself, to examine the dependence of this property on the main dynamical forcings, and to demonstrate that this property is largely insensitive to other details, both physical and numerical. To achieve such a goal within a numerical modeling study, we need a model configuration that allows us complete control over the relevant dynamical properties.

We use a semispectral, pressure coordinate model to solve the dry primitive equations in a spherical domain. The model is the Built on Beowolf (BOB) dynamical core developed jointly at National Center for Atmospheric Research (NCAR) and Columbia University; numerical details, implementation, and availability are documented fully in Rivier et al. (2002) and Scott et al. (2004). Here it is used as a stratosphere-only model, with no dynamically active troposphere. There are two sources of radiative/dynamical forcing: a relaxation to a prescribed radiative temperature equilibrium, representing radiative cooling over the winter pole; and a wave forcing at the lower boundary representing wave generation in the troposphere and at the ground. Details of each follow.

The radiative forcing is a standard Newtonian cooling to a radiative equilibrium temperature distribution, T_e, that represents a perpetual January state with colder temperatures over the winter pole. The distribution T_e is obtained from a superposition of an isothermal atmosphere at a temperature $T_0 = 240$ K with a pressure dependent cold anomaly over winter pole T_{PV}:

$$T_e = [1 - w(\phi)]T_0 + w(\phi)T_{PV}(p), \quad (1)$$

where the dependence on latitude ϕ is given by

$$w(\phi) = \{1 - \tanh[(\phi - \phi_0)/\delta\phi]\}/2 \quad (2)$$

and where the winter polar vortex equilibrium temperature is defined as

$$T_{PV}(p) = T_0(p/p_b)^{-R\gamma/g}. \quad (3)$$

In the above, $\phi_0 = 50°$, $\delta\phi = 10°$, and $p_b = 200$ hPa. The parameter γ defines the strength of the associated polar vortex and takes the values 0.5, 1, and 1.5 in the integrations presented below. The time scale, τ at which temperatures are relaxed to T_e is uniform throughout the stratosphere at $\tau = 10$ days. (Sensitivity of the results to τ is discussed in section 5.)

We discuss the wave forcing in the context of the lower boundary condition, which has the form of a specified, time-independent geopotential, $\Phi = \Phi_s$, of the type traditionally used in mechanistic modeling of the stratosphere. First, the zonal mean component, $\overline{\Phi}_s$ fixes the zonal mean velocity, $\overline{u}$, at the lower boundary (independently of T_e, which only determines $\partial\overline{u}/\partial z$) and, in the calculations presented below, has the form

$$\overline{\Phi}_s(\phi) = \begin{cases} U_0 a\Omega \dfrac{3\sqrt{3}}{4} \sin^4\phi & \phi \geq 0 \\ 0 & \phi < 0 \end{cases}, \quad (4)$$

where $a = 6371$ km is the planetary radius, and $\Omega = 2\pi$ day^{-1} is the planetary rotation rate. Through thermal wind balance, this form for $\overline{\Phi}_s$ gives a broad midlatitude jet with peak winds of approximately U_0 (neglecting the curvature term in the balance equation). We choose $U_0 = 30$ m s^{-1} as representative of the winds in the lower winter stratosphere, but neither this value nor the form of Φ_s significantly affect the results; they are included here simply for reproducibility.

Stationary planetary waves are forced at the lower boundary by specifying the zonally varying component of the geopotential:

$$\Phi'_s = gh_0 G(\phi) \cos(m\theta) \quad (5)$$

with

$$G(\phi) = \begin{cases} \sin^2\left(\pi\dfrac{\phi - \phi_1}{90° - \phi_1}\right) & \phi > \phi_1 \\ 0 & \text{otherwise} \end{cases}, \quad (6)$$

where g is the gravitational acceleration, θ is longitude, m is the zonal wavenumber of the forcing, and $\phi_1 = 30°$. The function G attains its maximum value of one at $\phi = 60°$. In the calculations presented below, we consider separately forcing wavenumbers $m = 1$ and $m = 2$ and forcing amplitudes h_0 between 0 and 1000 m, varied as an external parameter. Note that the wave forcing amplitude and phase in each calculation is completely time-independent, with the sole exception of an initial 10 day $\sin^2$ ramp on Φ_s to initialize the integration smoothly.

Having specified the main forcings and boundary conditions, all other numerical and physical modifications to the dynamical equations are kept to a mini-
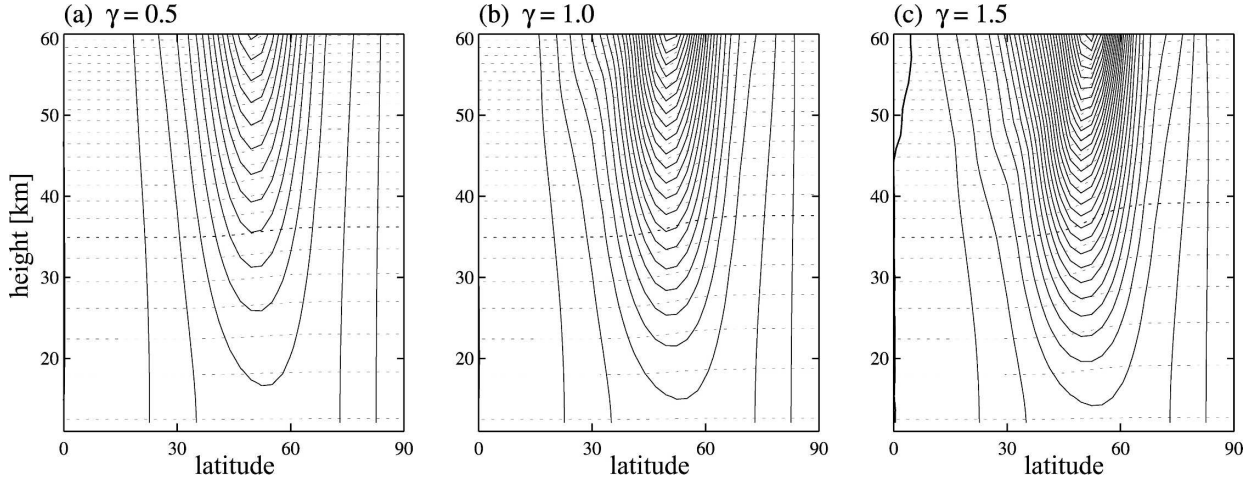
FIG. 1. Zonal mean potential temperature $\bar{\theta}$ (dotted) and zonal mean zonal velocity $\bar{u}$ (solid) at unforced equilibrium, i.e., the steady state obtained with wave forcing amplitude $h_0 = 0$, with radiative equilibrium profile given by (1)–(3) with (a) $\gamma = 0.5$, (b) $\gamma = 1.0$, and (c) $\gamma = 1.5$. Contour interval for $\bar{u}$ is 10 m s^{-1} and the zero contour is thick [just visible in the upper left corner of (c)]. Contour interval for $\bar{\theta}$ is 100 K and the 1000-K contour is bold.

mum. The model imposes a no-flux condition at the upper boundary $p = p_t$, which necessitates a sponge layer in the uppermost model levels to prevent spurious wave reflection. The sponge is restricted to above $p_{sp} = 0.5$ hPa, above the domain of interest, and damps the waves only at a rate $k_{sp} \max\{0, [(p_{sp} - p)/p_{sp}]^2\}$, with $k_{sp} = 2(\text{day})^{-1}$. This is an important departure. Traditionally, sponge layer–type damping has also been imposed on the zonal mean velocity as a form of gravity wave drag parameterization used to close the strong westerly jet in the mesosphere. However, recent studies (Shepherd et al. 1996; Shepherd and Shaw 2004) have demonstrated that such a zonal mean drag term does not respect the angular momentum constraint of the total atmosphere, and as such may lead to unwanted feedback and downward influence effects. To avoid these unwanted side effects, and since the mesospheric winds are outside our immediate domain of interest, we omit the zonal mean drag altogether. As we have verified, however, doing so has almost no effect on the results; similar variability was obtained both with and without a drag on the zonal mean velocity (see section 5b).

To avoid the build up of enstrophy at small scales, a horizontal diffusion is included using a scale-selective ∇^8 hyperdiffusion (uniform in pressure). A diffusion time scale of $\tau_h = 0.25$ days at the smallest scales is used for most of the calculations presented below. The sensitivity to τ_h as well as to the parameters τ, k_{sp}, are further discussed in section 5.

The horizontal resolution used for the results presented in the main part of the paper is T42, equivalent to a grid spacing of approximately 2.8°. There are $L + 1$

half-levels in the vertical between the lower boundary at pressure $p = p_b \equiv 200$ hPa and the upper boundary at $p = p_t \equiv 0.02$ hPa. The half-levels are located at pressures $p_{l+1/2} = p_t(p_b/p_t)^{l/L}$, $l = 0, \ldots, L$, with $L = 40$. The L full-levels are located at the geometric mean of consecutive half-levels. The vertical domain thus extends from a notional tropopause at 12 km to approximately 75 km. The T42L40 resolution used allows us to perform a comprehensive exploration of the two-dimensional parameter space described below, but is high enough that we are confident that the results obtained are not artifacts of excessive truncation. As confirmation, we separately doubled the horizontal and vertical resolution for a single case, results for which are discussed in section 5b.

The model is initialized in a state of rest and run in the above perpetual January configuration for an integration period of 1000 days. Figure 1 shows the steady equilibrium states obtained for values of $\gamma = 0.5$, 1.0, 1.5 in the absence of wave forcing (i.e., $h_0 = 0$), and shows the direct relation of γ to the strength of the polar vortex obtained. Since the only force acting on $\bar{u}$ is the very weak (at these large scales) hyperdiffusion, the distributions shown are very close to those in exact thermal equilibrium with T_e.

Finally, we wish to emphasize again that all forcing in the model, both at the lower boundary and in the interior of the stratosphere, is time-independent. In addition, since the model lower boundary is at 200 hPa, the troposphere and its accompanying synoptic variability are, by construction, excluded from our calculations. Finally, since the zonal mean profiles shown in Fig. 1 are stable to shear instabilities, this potential source of
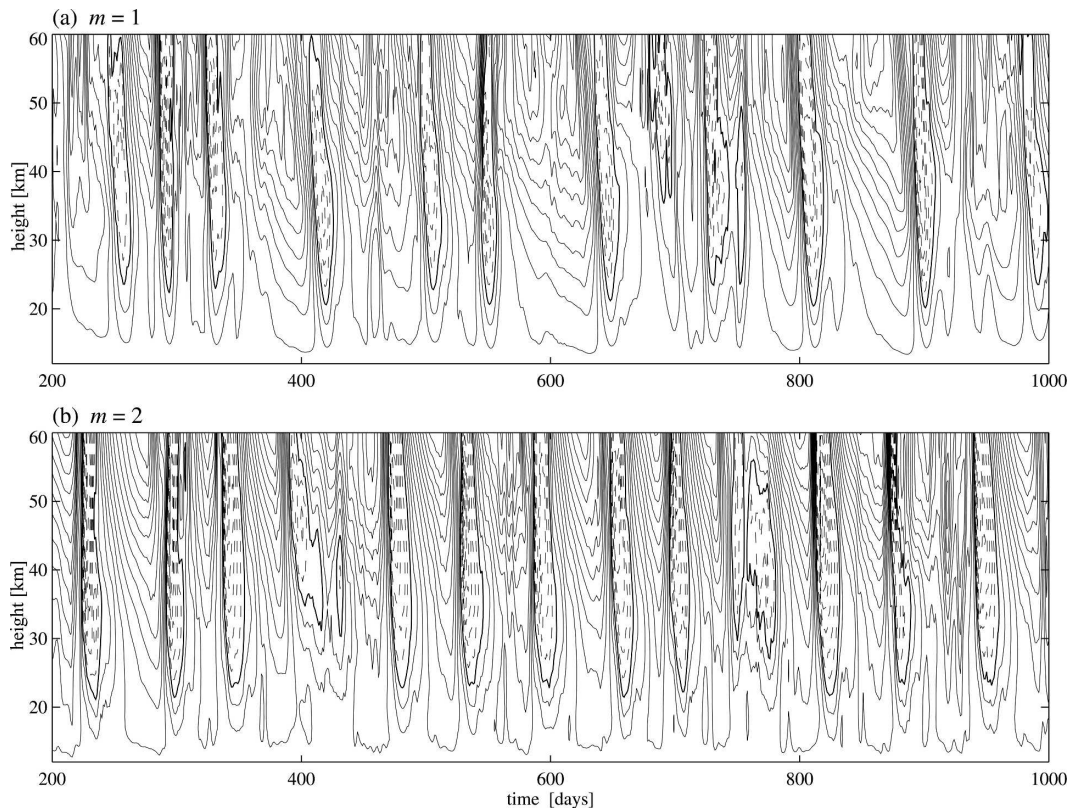
Fɪɢ. 2. Zonal mean zonal velocity $\bar{u}$ at $\phi = 60°$ lat (approximate latitude of the polar jet) for wave forcing amplitude $h_0 = 600$ m and radiative equilibrium $\gamma = 1.0$: (a) $m = 1$, (b) $m = 2$. Contour interval is 10 m s^{-1} and negative contours are dashed.

variability does not exist in our model. Hence any time-dependence in the model must result from the internal dynamics of the stratosphere itself.

3. General characteristics

We begin by presenting and discussing in some detail an illustrative example, chosen to be representative of the stratospheric variability obtained with the model configuration outlined above. Figure 2 shows the zonal mean zonal velocity at 60°, $\bar{u}(z, t)$ as a function of height and time for wave forcing at wavenumber $m = 1$ (Fig. 2a) and $m = 2$ (Fig. 2b). For this particular case the wave forcing amplitude is $h_0 = 600$ m and the strength of radiative forcing is $\gamma = 1$. The character of variability obtained is typical of that found at other parameter values, comprising regular, but not perfectly periodic, vacillation cycles, in which the polar vortex strengthens systematically before reversing throughout the domain in a major sudden warming type event (here defined loosely as the reversal of $\bar{u}$ throughout a large part of the domain).

Insofar as the variability consists of a series of sudden

warming events followed by a more gradual recovery of the vortex, the character of the variability in the two cases $m = 1$ and $m = 2$ is similar. In this zonal mean view, the main difference between the two cases is a lower frequency of the sudden warmings for $m = 1$, approximately one every 100 days compared with one every 60 days for $m = 2$. Further, for $m = 1$, the warmings are in general (though not always) less sudden than for $m = 2$, although in both cases we note that there is substantial variation in the strength and nature of sudden warmings; that is, the vacillations are not perfectly regular. We focus for now on the case $m = 2$ and note some fundamental properties of the variability; we return to details of the difference in variability between $m = 1$ and $m = 2$ in section 5 below.

First, an important observation needs to be made concerning the vertical flux of wave activity into the interior of the model domain: while the lower boundary condition wave forcing is constant in time, the flux of wave activity there is strongly time-dependent. This is illustrated in Fig. 3 for the simulation shown in Fig. 2b. Note the rapid increase in wave flux through the lower boundary that coincides with the sudden deceleration
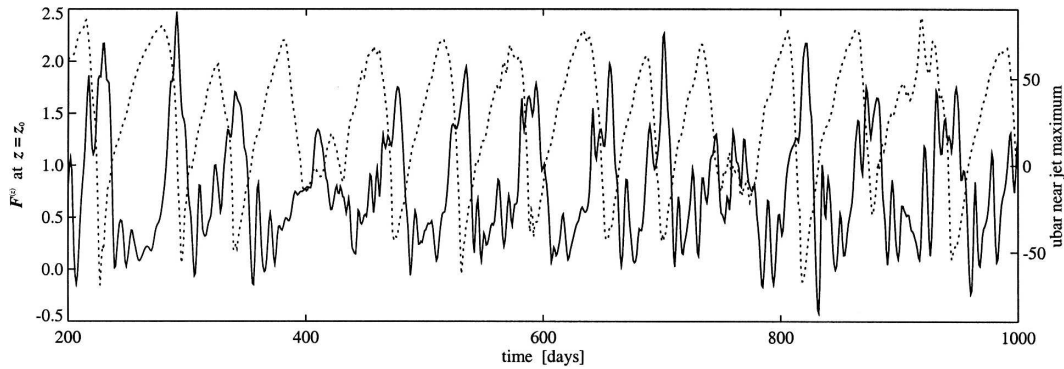
FIG. 3. Vertical EP flux at the lower boundary $F^{(z)}|_B$ averaged over latitude (solid) and $\bar{u}$ at $\phi = 60°$ lat and z = 41 km (dashed) for the case $h_0 = 600$ m, $\gamma = 1.0$.

of the upper-level vortex. This effect was noted by Dunkerton et al. (1981), who discussed both the increase in upward wave flux at the lower boundary and the associated increase in westward phase tilt of the waves. We note further that the effect is not necessarily related to the presence of an artificial lower boundary and that similar variability in the wave flux was obtained by the authors in a model with a carefully controlled troposphere (Scott and Polvani 2004). The implication is that a sudden burst of tropospheric wave forcing is not a prerequisite for the sudden increases in wave flux into the stratosphere often accompanying major sudden warmings, but that in fact the upward wave flux into the stratosphere can be controlled by the stratosphere itself. An analytic expression for the vertical wave flux at the lower boundary was given in Esler and Scott [2005, see discussion following Eq. (27) therein] for the case of an axisymmetric vortex subject to a pulse of wave forcing.

Next, we consider in detail the dynamical evolution of the vortex during a complete sudden warming cycle. For definiteness we choose the time period from t = 660 up to the strong sudden warming occurring around $t = 700$ from the time series shown in Fig. 2b, for which $h_0 = 600$ and $\gamma = 1$. Although we focus on a single vacillation cycle, we emphasize that the dynamical mechanism we describe is quite general. As we show, the vacillation cycle can be considered as resulting from the competing effects of radiative cooling of the vortex and wave forcing, with the effectiveness of the latter being determined by the vertical structure of the vortex.

a. Zonal mean description

Figure 4 shows snapshots of $\bar{u}$, illustrating the meridional structure of the zonally averaged vortex. At

$t = 665$ the upper vortex has recovered substantially from the previous warming event but a region of very weak and easterly winds persists throughout the middle stratosphere. Thus, the upper vortex is disconnected from the westerly winds at low levels. By $t = 675$ the upper and lower jets have coalesced but remain weak in the middle stratosphere until around $t = 685$; until then the vortex evolution is dominated by radiative cooling, with upward wave activity suppressed by the weak winds in the middle stratosphere. At 690, however, we begin to see the effect of wave breaking on the vortex more clearly, as the zero wind line and the jet maximum begin to contract poleward. This process intensifies through $t = 695$ until at $t = 700$ the effects of wave forcing on the vortex are evident in the complete reversal of the zonal winds.

To support this picture, we show in Fig. 5 the Eliassen–Palm (EP) flux $\mathbf{F}(\phi, z)$ and divergence $\nabla \cdot \mathbf{F}(\phi, z)$ over the same period. Between $t = 665$ and $t = 675$ waves are completely shut off from the upper vortex, as predicted from the distribution of $\bar{u}$, supporting the assertion that the evolution of the upper vortex at this time is dominated by radiative cooling. Wave activity entering the domain at this time is instead refracted equatorward, and is absorbed right around the region of zero winds near 30° and 15–20 km, this region subsequently extending upward as the lowermost vortex recovers and waves begin to penetrate higher into the domain (Fig. 5b). Throughout this period, note that the total amount of wave activity entering the domain remains comparatively low.

Up until $t = 685$, therefore, the effect of wave forcing on the vortex is relatively weak. At this point, however, two related dynamical effects come into operation. First, the waves begin to penetrate deeper within the stratosphere; and second, more wave activity begins to enter the domain. These two effects are most likely
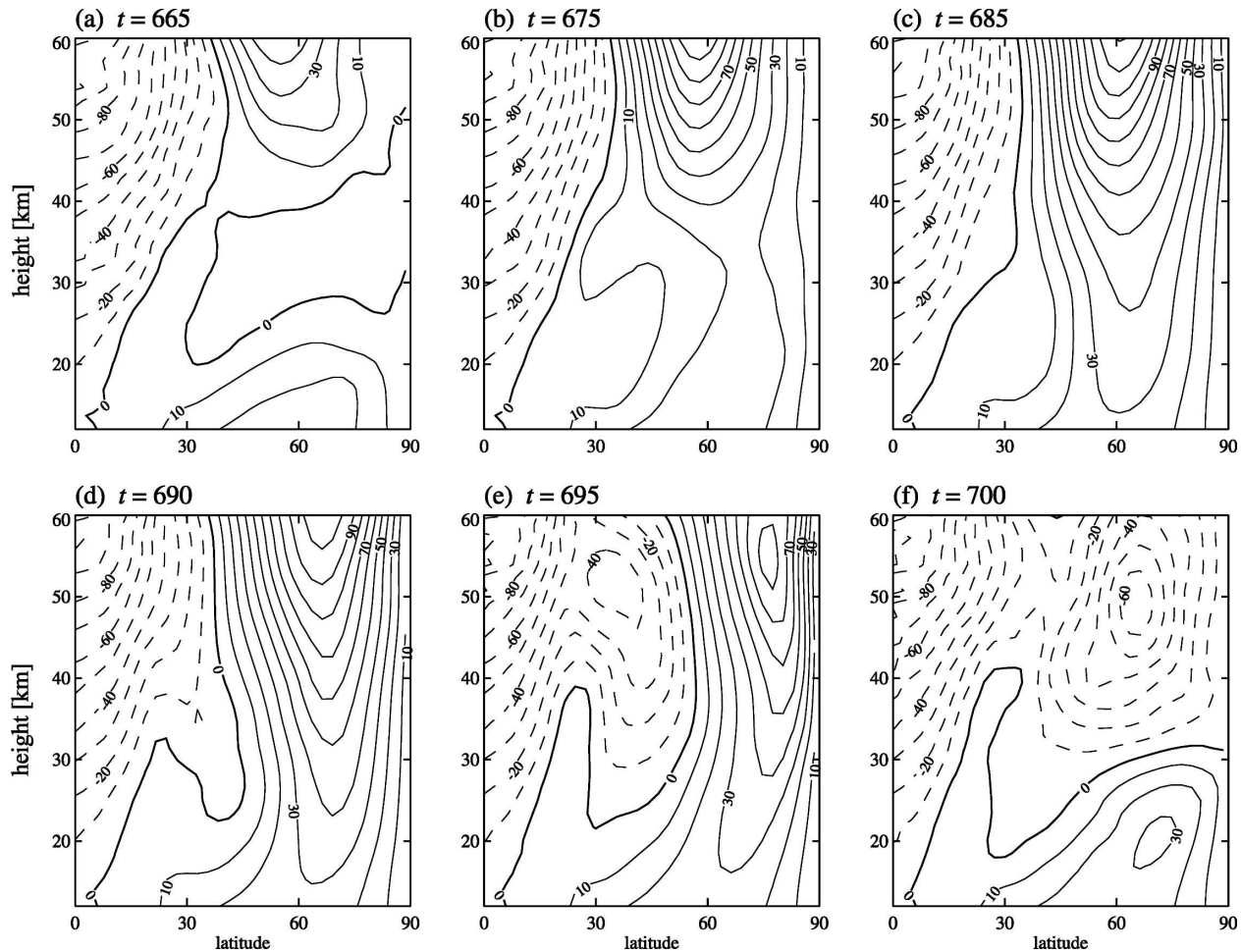
FIG. 4. Zonal mean zonal velocity $\bar{u}$ for wave forcing amplitude $h_0 = 600$ m and $\gamma = 1.0$ at times (a)–(f) $t = 665, 675, 685, 690, 695, 700$ days. Contouring is as in Fig. 1.

both manifestations of the change in $\bar{u}$ in the lower-middle stratosphere at this time, and possibly linked to the tendency for negative vertical shear to cause downward reflection of upward propagating waves (e.g., Harnik and Lindzen 2001). When the vortex becomes reestablished at all levels at $t = 685$, wave activity entering the domain increases suddenly and penetrates further into the domain. By $t = 695$ there is significant wave activity reaching the upper stratosphere and significant wave flux convergence there, leading to the strong deceleration of the zonal mean flow between $t = 695$ and $t = 700$. Note that the quantity shown here is the true EP flux[1] and its divergence and not these quantities rescaled by the inverse of density, as is often

shown. The plotted divergence thus represents the total force on the flow, not the force per unit mass, and the small values in the upper stratosphere have a large effect on the zonal velocity.

b. Potential vorticity description

To illustrate the three-dimensional evolution of the vortex, we now consider the Ertel potential vorticity (PV), which is materially conserved by the flow and which is a natural framework in which to consider the morphology of Rossby wave breaking. Following, Polvani and Saravanan (2000), we use the scaled PV introduced by Lait (1994), and nondimensionalize according to Eq. (4) in Polvani and Saravanan. With this nondimesionalization, the scaled PV is an order one quantity at all levels in the polar stratosphere.

Figure 6 shows the 0.8-level surface of the scaled PV at times from $t = 665$ to $t = 700$, corresponding to the
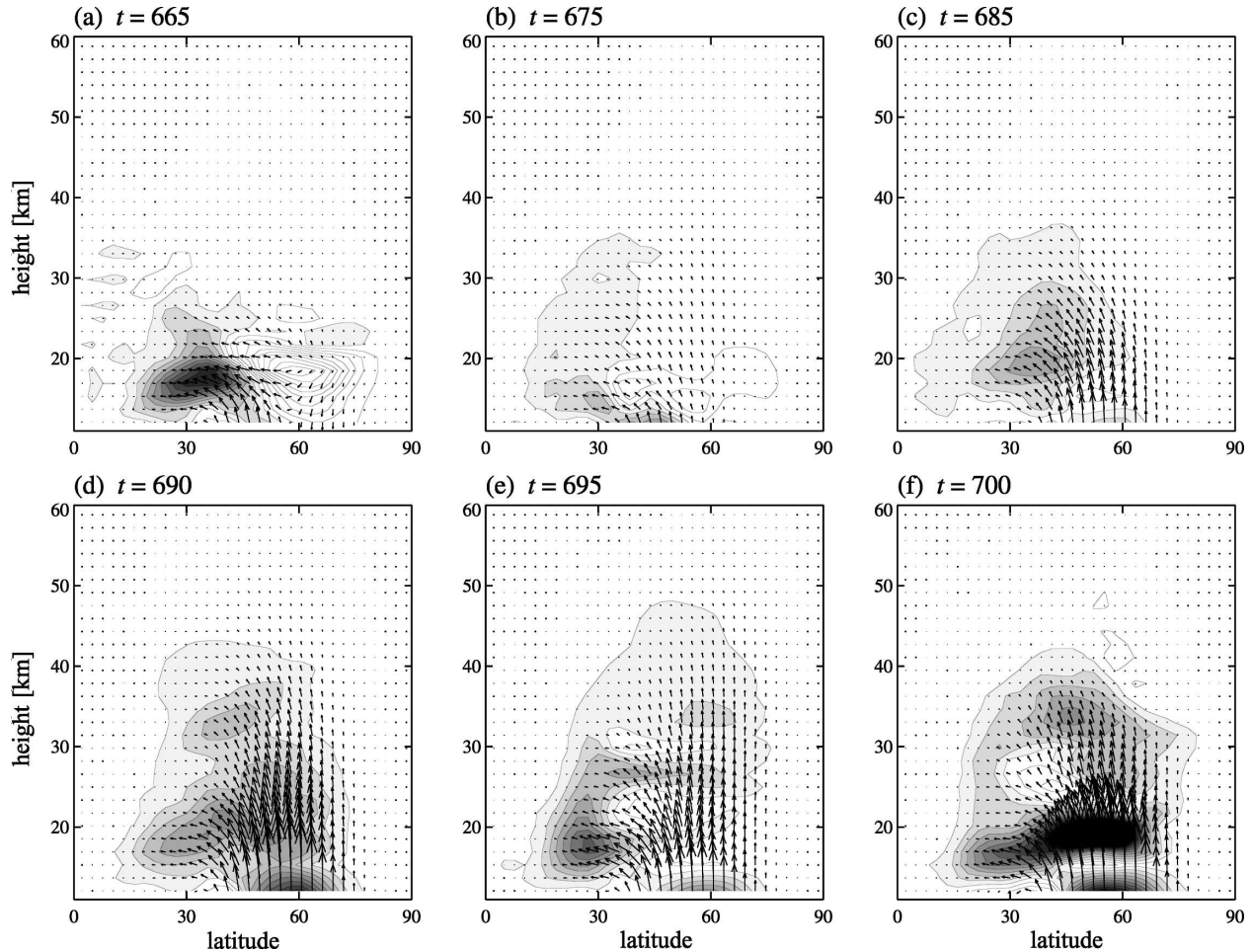
---

[1] The EP flux is related to the flux of wave activity under suitable Wentzel–Kramers–Brillouin (WKB) conditions, but is defined independently of any such approximation.

FIG. 5. EP flux $\mathbf{F}$ (arrows) and EP flux divergence $\nabla \cdot \mathbf{F}$ (contours and shading) for wave forcing amplitude $h_0 = 600$ m and $\gamma = 1.0$ at times (a)–(f) $t = 665, 675, 685, 690, 695, 700$ days. Negative values of $\nabla \cdot \mathbf{F}$ (regions of EP flux convergence) are shaded. Note, actual EP flux and divergence values are plotted, not values weighted by the inverse of the density.

times just discussed. Although not always located in the region of steepest horizontal PV gradients, this surface is nevertheless broadly representative of the vortex edge throughout the domain. The time frames show the gradual growth and sudden destruction of the vortex. At $t = 665$ the lower and middle vortex are still strongly deformed as a result of the previous sudden warming, to the extent that it is barely present at these levels. At upper levels, on the other hand, the vortex has already begun to recover under radiative forcing. By $t = 675$ the lower vortex has partially recovered but is still weak. The claim made above that the upward wave propagation is inhibited at this time is thus supported by the three-dimensional PV distribution.

By $t = 685$ the lower vortex has grown further and the vortex now appears more coherent vertically. It is at this point that wave activity begins to propagate further into the domain. The result of the increased upward

wave propagation is visible by $t = 690$, when a large tongue of polar air is peeled off the main vortex, in a deep wave-breaking event. At this time it is also possible to estimate the vertical wavelength: the phase tilt is approximately $\pi/2$ over the entire domain spanning $z = 12$ km to $z = 60$ km, suggesting a wavelength 4 or 5 times the depth of the stratosphere. Given the lack of scale separation, we are thus reluctant to describe the wave dynamics here in terms of wave reflection or critical levels.

By $t = 695$ the wave breaking has intensified, consistent with the strong upward propagation at this time, and the whole middle and upper vortex becomes highly distorted, with polar air extending far into low latitudes. Finally, by $t = 700$ most of the middle and upper vortex has been completely destroyed and only the lowermost vortex retains any coherence. A few days later (not shown), even the lowermost vortex is largely destroyed
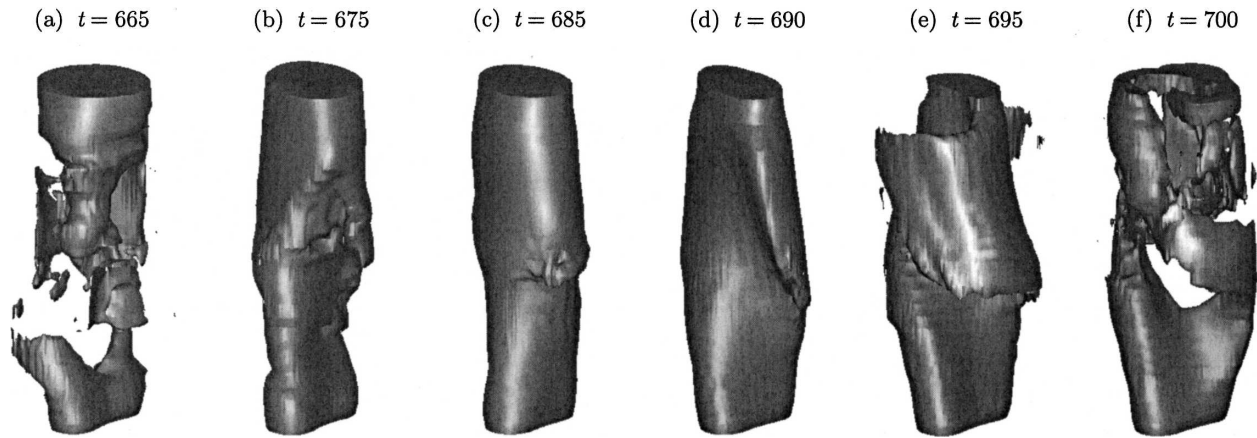
FIG. 6. The 0.8 level of scaled potential vorticity for the case $h_0 = 600$ m and $\gamma = 1.0$ at times (a)–(f) $t = 665, 675, 685, 690, 695,$ 700 days.

and resembles closely the state of the vortex at $t = 665$ shown in Fig. 6a.

Figure 7 shows in more detail the PV evolution on a single isentropic surface, $\theta = 1020$ K, around $z = 35$ km.

At $t = 665$, the PV over most of the high latitudes is homogenized and the strong PV gradients at 30° should be interpreted as the subtropical edge of the surf zone and a consequence of the PV mixing resulting from the
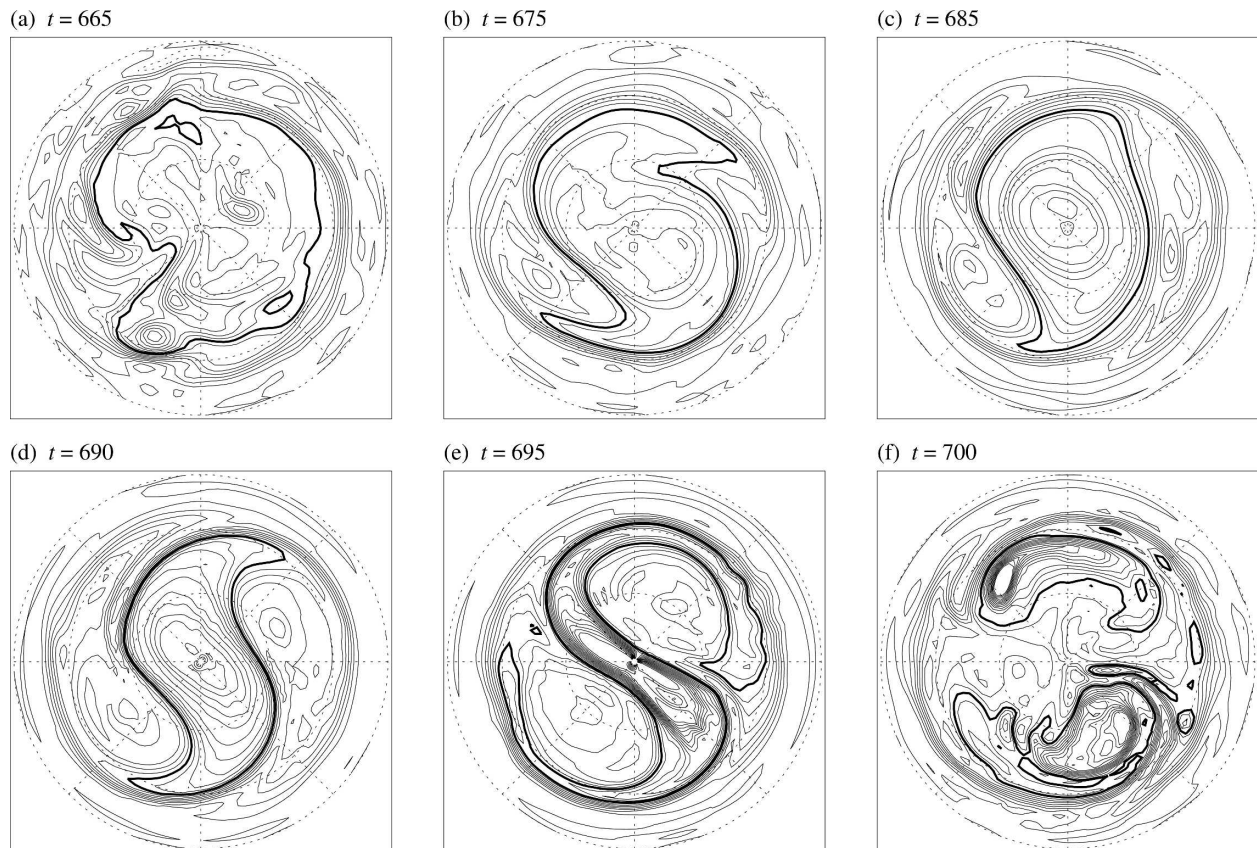

FIG. 7. Potential vorticity on $\theta = 1020$ K isentropic surface (approximately 35 km altitude) and at times (a)–(f) $t = 665, 675, 685, 690, 695, 700$ days. The heavy contour corresponds to the 0.8-PV level shaded in Fig. 6 and the contour interval is 0.1. An equal area, Lambert projection is used with lat lines at 0°, 30°, and 60° shown dotted.

preceding sudden warming; the actual polar vortex at this time is not well-developed. Subsequently, the vortex grows steadily at this level, forced radiatively, and remains predominantly zonally symmetric with only small wave disturbances. Note that although there are substantial contour displacements already at $t = 675$, these occur in regions of relatively weak horizontal PV gradients, and so are not indicative of large wave activity. Indeed, large displacements of contours in weak PV gradients are inevitable, and have been discussed in various setting (e.g., Legras and Dritschel 1993; Legras et al. 2001; Rong and Waugh 2004). Such erosion of weak PV gradients is also a crucial process in the steepening of the vortex edge, which can be observed around $t = 685$. In contrast, we know that strong contour displacements coinciding with the steep gradients of the vortex edge, such as the distinctive wave-2 shape developing at $t = 690$, must be associated with significantly increased wave activity. The disturbance develops further over the next five days, until it completely distorts the vortex core, extruding polar air to lower latitudes. The vortex eventually splits, and forms two corotating satellites of high PV in a homogenized surf zone extending across the pole.

4. Dependence on forcing parameters

As already mentioned, the dynamical behavior described above is quite generally observed in the simulations carried out in this study. Rather than further document the details of the evolution of different sudden warmings, we proceed instead to examine how the general characteristics of the variability depend on the basic forcing properties of the system. We consider two model parameters as being fundamental to the behavior of the system, namely the radiative cooling, γ, and the wave forcing amplitude, h_0. Since these represent the main forcings of the system, we treat them separately and in more detail than the other physical parameters discussed in section 5 below. This approach is consistent with previous studies considering internal variability in highly truncated models.

In agreement with the results obtained in previous studies with highly truncated models, we find that for a given value of γ the system exhibits a whole range of behavior as the forcing amplitude h_0 is varied, from steady states at small h_0 to strongly time-dependent states at large h_0, such as those illustrated in Fig. 2. In fact, strong variability is found over a large range of wave forcing amplitudes and is generally characterized by recurring major warmings at intervals of approximately 40–100 days. To quantify this variability, we introduce some simple diagnostics based on the zonal

mean zonal wind at $\phi = 60°$ and $z = 40$ km, $u_{jet}(t)$, roughly coincident with the polar jet maximum in the upper stratosphere.

A simple, yet informative measure of the variability can be obtained by calculating the frequency spectrum of the time series of $u_{jet}(t)$. Figure 8 shows the power in each Fourier mode, $E(\omega) = |\hat{u}(\omega)|^2$ normalized by γ^2, for various values of γ and h_0 covering the parameter space of interest. The normalization accounts for the fact that strength of the zonal mean flow increases approximately linearly with increasing γ. We do not show $E(\omega)$ for $h_0 \leq 200$ since all simulations quickly approach a steady state at these low wave forcing amplitudes. Note also that the case discussed in section 3 above ($\gamma = 1$, $h_0 = 600$) appears in the central row. Peaks in $E(\omega)$ represent variability dominated by a particular frequency. The height of the peak indicates the strength of the variability, in the sense of the magnitude of the changes to u_{jet}, for example, from strong westerly to strong easterly in the case described above. The width of the peak indicates the coherence or regularity of the variability: a very narrow peak corresponds to almost periodic variability at the frequency of the peak; a broader peak corresponds to a greater distribution of dominant frequencies and less coherent variability.

Let us start by concentrating on the middle row, which corresponds to $\gamma = 1$. We see, first of all, that the case discussed above does not represent the most coherent variability obtained in this framework. At $\gamma = 1$, the variability obtained with $h_0 = 800$ is both more regular and more vigorous, as the time–height plot of $\bar{u}$ verifies; see Fig. 9b. Note also the increase in frequency with increasing h_0, particularly obvious between $h_0 = 600$ and $h_0 = 800$, when there is a single peak (cf. also Fig. 2b and Fig. 9b). At $h_0 = 400$ (still $\gamma = 1$) while the dominant frequency is clearly lower, there are now two strong peaks, indicating intermittency in the flow. In this case, approximately every second sudden warming is reduced in intensity: the spectral peak at $T = 80$ days corresponds to the type of variability discussed above, consisting of recurring sudden warmings and vortex regrowth, while the larger peak at $T \sim 160$ days arises from the lower frequency modulation. At $h_0 = 300$ (not shown) there is a spectral peak at $T = 100$ days but also significant power at low frequencies, which corresponds to intermittency as illustrated in the time–height plot shown in Fig. 9a. Here a quiescent regime between $t = 400$ and $t = 650$ (and, to a lesser extent, between $t = 150$ and $t = 350$) interrupts the otherwise regular sequence of sudden warmings. Despite these variations in dynamical behavior, however, we emphasize that coherent internal variability is present over a very wide range of forcing amplitudes, from $h_0 = 300 - h_0 = 800$.
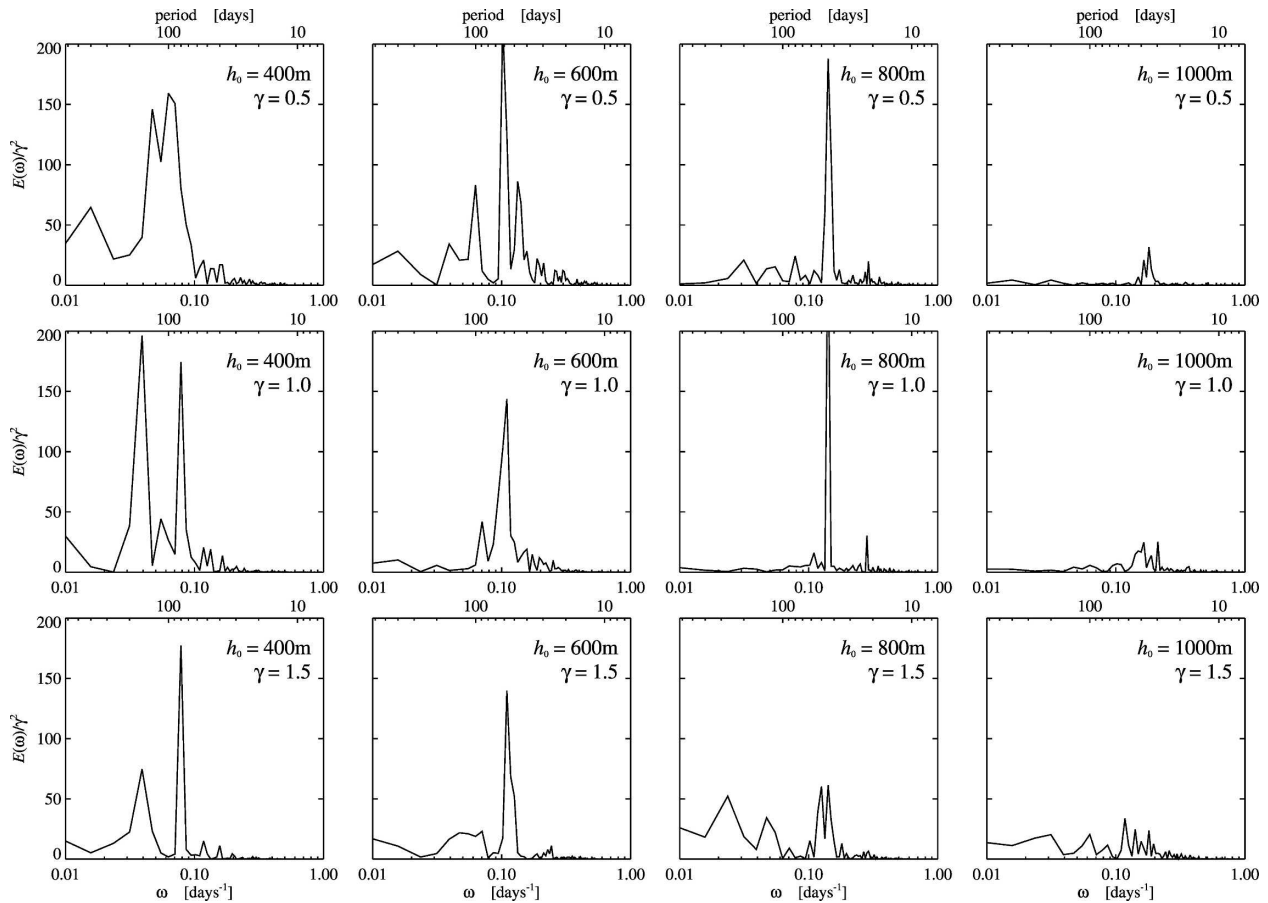
FIG. 8. Frequency power spectra of $\bar{u}$ at latitude $\phi = 60°$ and altitude $z = 41$ km, normalized by γ^2, for the cases (left to right) $h_0 = 400$–1000 m and (top to bottom) $\gamma = 0.5$, 1.0, and 1.5.

At still larger forcing amplitudes, for $h_0 = 1000$, the coherent variability is strongly attenuated. Inspection of time–height plots of $\bar{u}$ (not shown) indicate that the reason for this is that the vortex remains persistently weak in the middle stratosphere, and does not recover sufficiently to allow the burst of upward wave propagation and wave-induced deceleration of the upper vortex observed at lower forcing amplitudes. A manifestation of this can be seen in the time averaged $\bar{u}$ in the meridional plane, shown in Fig. 10c, which reveals predominantly easterly winds in the middle stratosphere. A systematic reduction in the time averaged winds with increasing h_0 can be seen from the other panels (Figs. 10a,b), although part of the reason the winds are so much stronger for $h_0 = 300$ is simply the prolonged period of vortex stability seen in Fig. 9a.

A similar pattern is found for other values of γ, shown in the top and bottom rows of Fig. 8. In particular, note that the dominant period of the response appears independent of γ, decreasing steadily with h_0 in each case from around $T = 100$ days below $h_0 = 400$ to around $T = 45$ days for $h_0 = 800$. All γ cases also exhibit intermittence at $h_0 = 400$ (and below), with a broad or double peak structure in $E(\omega)$. Note that the apparent loss of coherence for $\gamma = 1.5$ and $h_0 = 800$ is again due to intermittence, evidence for which can be seen in the broader and double peak structure, as well as the power at very low frequencies. In fact, plots of $u(z, t)$ for $\gamma = 1.5$ and $h_0 = 1000$ are similar to Fig. 9a, but squashed in time corresponding to the high frequency dynamics.

Finally, we summarize all simulations considered here by considering the following basic quantities, again based on the time series $\bar{u}_{jet}(t)$ at $\phi = 60°$ and $z = 40$ km: the time mean, the standard deviation, the average extrema or envelope,[2] and the dominant frequency. We present these in Fig. 11 as a function of wave forcing amplitude, (with simulations at $h_0 = 0$, 100, 200, . . . , 1000) for each of $\gamma = 0.5$, 1.0, and 1.5.

―――――――――

[2] An approximate envelope to $\bar{u}_{jet}$ is defined by local maxima and minima; the quantity plotted in Fig. 11 is the time average of this envelope.

(a) $h_0 = 300$
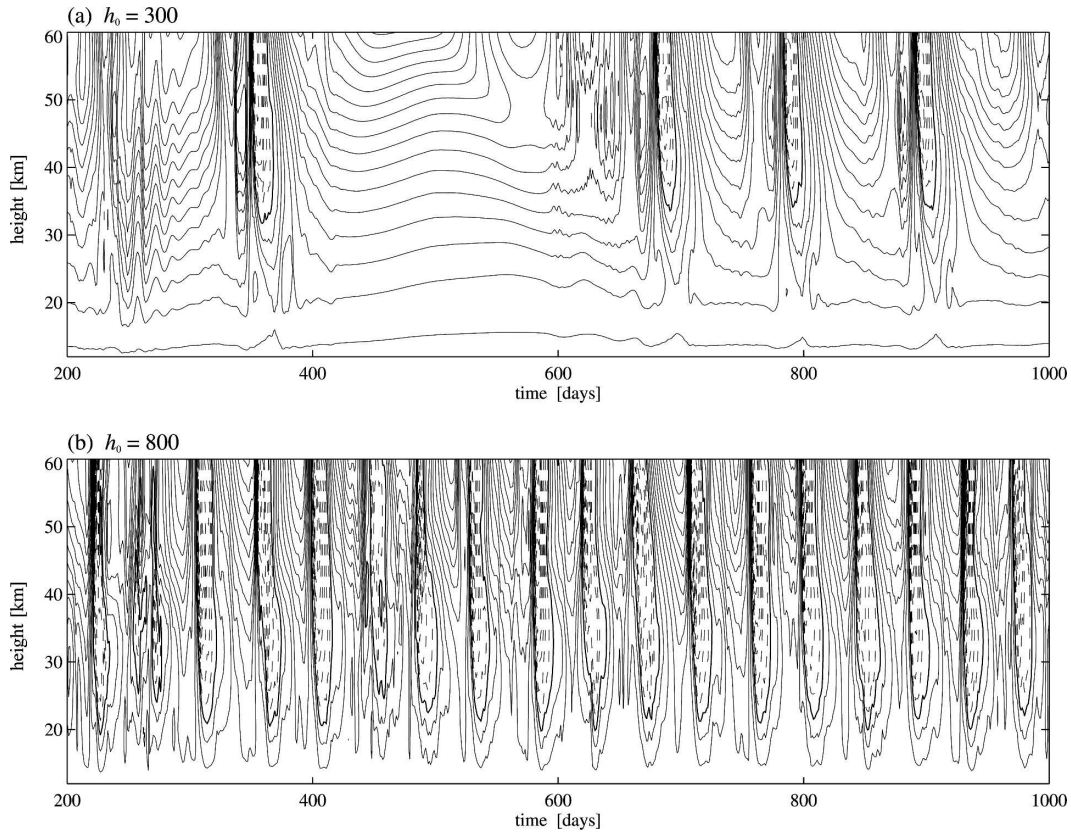

(b) $h_0 = 800$

FIG. 9. Same as in Fig. 2b but for wave forcing amplitude (a) $h_0 = 300$ m and (b) $h_0 = 800$ m.

First consider the dependence of the diagnostics on h_0. There is a certain similarity across all values of γ. All γ cases have steady states for $h_0 = 200$ and there is a sudden transition to vacillating states at $h_0 = 300$. Further, all γ show a systematic decrease in time mean $\overline{u}$ with increasing h_0, although the standard deviation and envelope remain roughly constant around mean. This means that the intensity of variability is roughly independent of h_0, while the average vortex strength decreases with h_0. Finally, all γ cases show a systematic increase in dominant frequency with increasing h_0 (the period decreases from about 100 days at $h_0 = 300$ to
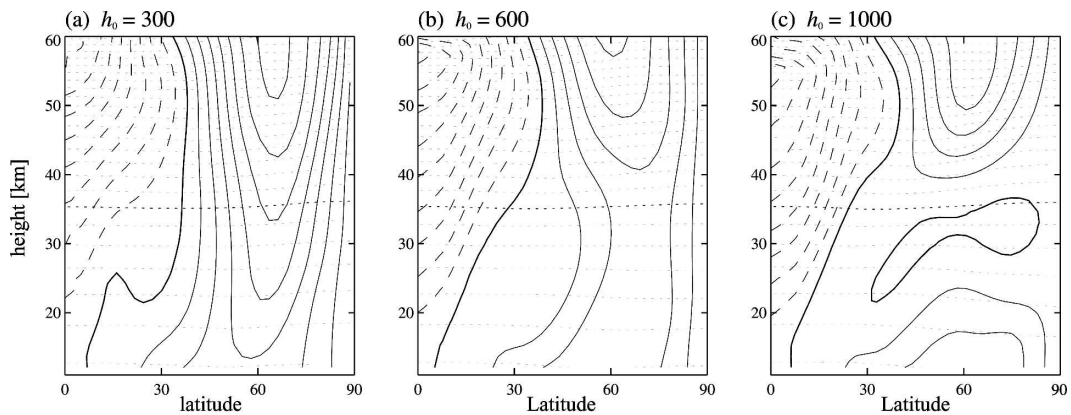
(a) $h_0 = 300$ (b) $h_0 = 600$ (c) $h_0 = 1000$

FIG. 10. Zonal mean potential temperature $\overline{\theta}$ and zonal mean zonal velocity $\overline{u}$ averaged from $t = 400$ to $t = 1000$ days for wave forcing amplitude (a) $h_0 = 300$ m, (b) $h_0 = 600$ m, and (c) $h_0 = 1000$ m and $\gamma = 1.0$. Contouring is as in Fig. 1.
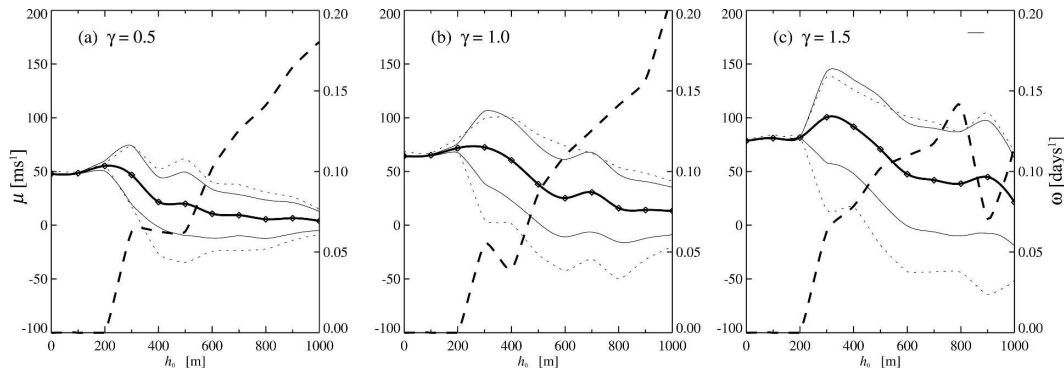
FIG. 11. Measures of jet speed (defined as $\bar{u}$ at $\phi = 60°$ and $z = 41$ km) mean and variability as a function of forcing amplitude $h_0 = 0 \ldots 1000$ m: mean (heavy solid; each diamond indicates a simulation); standard deviation (light solid, drawn relative to mean); envelope (light dotted), defined as the average maximum and minimum jet speed; median frequency (heavy dashed), defined as the frequency for which $E(\omega)$ (as shown in Fig. 8) is maximum. (a) $\gamma = 0.5$, (b) $\gamma = 1.0$, (c) $\gamma = 1.5$.

about 40 days at $h_0 = 1000$); that the maximum frequency is not exactly monotonic appears to be a consequence of the fact that it is not always well defined (e.g., at $\gamma = 1.5$, $h_0 \geq 900$, or $\gamma = 1$, $h_0 = 400$).

Concerning the dependence on γ, we note the following. First, there is a general increase in time mean $\bar{u}$ with increasing γ, a simple consequence of the colder equilibrium temperature profile. Second, there is an increase in the envelope or extreme departures from the mean with increasing γ, particularly the large negative values seen for $\gamma = 1.5$ at higher forcing amplitudes. There is also an increase in the standard deviation. Both these quantities indicate an increase in vigor of variability as the radiative cooling becomes stronger. Finally, we note that the weaker standard deviation for $\gamma = 0.5$ is likely related to the fact that the variability is less regular in these cases: this is not readily apparent from Fig. 8 but visible in height–time plots of $\bar{u}$ (not shown). In these cases, the vortex is generally weaker, and there appears less scope for sudden changes in the wave transmission properties of the flow of the form described above.

5. Sensitivity and robustness

Since this paper is concerned with the robustness of internal variability, we must now investigate the effect of the other main physical and numerical parameters on the variability. After the parameters γ and h_0, the radiative relaxation rate, τ, and the forcing wavenumber, m, are arguably the next most influential physical parameters, and we examine the dependence on these next. Again we are interested in whether the variability persists over a broad range of these forcing parameter values, and, if it does, which parameters determine

qualities of the variability such as the peak frequency and magnitude of sudden warmings. Similarly we wish to establish that the results presented here are not sensitive to the numerical procedure chosen to carry out the computation of the physical model, including the particular resolution used and any unphysical parameterizations such as the sponge layer; this is considered in section 5b.

a. Physical parameters

1) RADIATIVE RELAXATION RATE

We first examine the sensitivity to the radiative relaxation time scale, τ. To do so, we consider three variations of the foregoing simulations, identical in all respects apart from the value of τ, which is (i) halved to $\tau = 5$ days, (ii) doubled to $\tau = 20$ days, and (iii) set to a height-dependent value $\tau = \tau_H(z)$ representative of the height dependence of radiative time scales in the stratosphere (Holton 1976), where $\tau_H(z)$ is given by

$$\tau_H(z) \equiv \{1.5 + \tanh[(z - 35)/7]\}^{-1} \times 10^6 \text{ s} \quad (7)$$

with z in kilometers; τ_H varies from around 20 days below 35 km to around 4 days above 35 km.

The results for the series of simulations with $\gamma = 1$, $h_0 = 0, 100, 200, \ldots, 1000$ are summarized in Fig. 12. We see, first of all, that in terms of time mean $\bar{u}$ and standard deviation the general variability obtained in all cases is generally similar to that obtained above (cf. Fig. 11b). In particular, strong variability is once again found over broad range of h_0.

Comparing more closely with the $\tau = 10$ case discussed above, we find that the dominant frequency of the variability is higher for $\tau = 5$ and lower for $\tau = 20$;
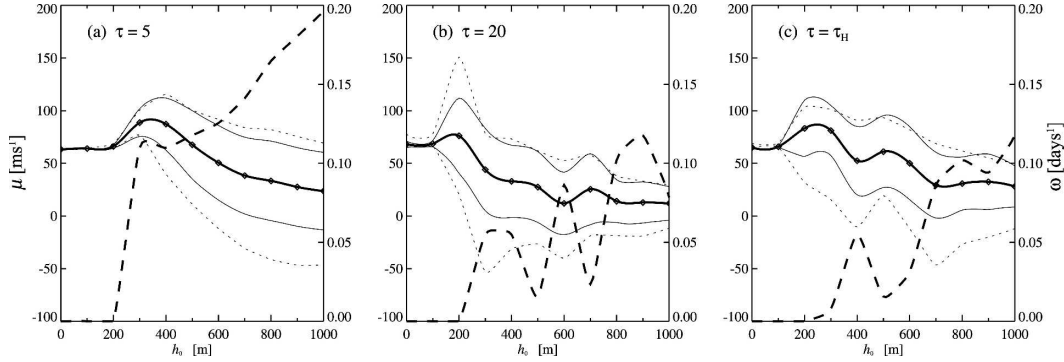
FIG. 12. Same as in Fig. 11b but with different values of thermal relaxation τ.

that is, the frequency of the variability decreases with increasing radiative time scale. In addition, the dominant frequency for the case $\tau = \tau_H$ is closer to that for the case $\tau = 20$ than for $\tau = 5$. Since τ_H is approximately 20 days in the lower stratosphere and 4 days in the upper stratosphere, this suggests that it is the lower stratospheric radiative time scale that is important in determining the frequency of the variability. Conversely, the time mean $\bar{u}$ for $\tau = \tau_H$ is closer to that for $\tau = 5$; hence we may conclude, perhaps unsurprisingly, that it is the upper stratospheric radiative time scale that is important in determining the upper-stratospheric jet strength.

Finally, we find intermittency in certain forcing ranges, which for $\tau = 20$ and $\tau = \tau_H$ is reflected in the nonmonotonic dependence of the dominant frequency on the wave forcing amplitude. In particular, the dominant frequency as determined by the spectral analysis is not always the frequency of a sudden warming–vortex growth cycle, but sometimes represents lower frequency modulations. Having said this, however, we emphasize that even in these cases coherent variability is still evident. The main point of this exercise is to demonstrate that the variability is robust and persists across a wide range of radiative time scales.

Since the radiative profile $\tau_H(z)$ is closer to the actual vertical profile of radiative time scales in the stratosphere, we show as a final example, the variability of the zonal mean winds for this τ, with $\gamma = 1$ and the two cases $h_0 = 300, 800$; see Fig. 13. Note the remarkable regularity of the variability. Comparing this with Fig. 9 we see that the variability is in fact more regular with $\tau = \tau_H$ than with the constant $\tau = 10$. Also, we see directly that the frequency of sudden warmings is lower in the case $\tau = \tau_H$, especially for $h_0 = 800$. We may even speculate that it is a consequence of the vertical variation of τ_H that leads to the stronger regularity of the variability: on the one hand, the longer relaxation time scale in the lower stratosphere allows a lower frequency

variability; on the other hand the shorter radiative time scale in the upper stratosphere means that the vortex more nearly approaches its equilibrium state, presumably resulting in less variation between warming cycles. In any case, the approximation of constant τ used in section 4 does not appear to be a significant shortcoming, and we can say with confidence that coherent internal variability is also a feature of more realistic relaxation rates.

2) FORCING WAVENUMBER

As was seen in Fig. 2, the character of the variability obtained is roughly similar between forcing wavenumber $m = 1$ and $m = 2$. In this section, we wish to establish whether the variability obtained with $m = 1$ is also robust over a large range of forcing parameters. To this end, we have repeated the main series of simulations presented in section 4, namely with $\gamma = 0.5, 1, 1.5$, $h_0 = 0, 100, \ldots, 1000$, with identical parameters as before but forcing instead with zonal wavenumber $m = 1$. To summarize the results we show in Fig. 14 the mean, standard deviation, envelope of extrema, and dominant frequency as a function of h_0 for each of the three values of γ.

The similarities with the $m = 2$ series in all quantities are clear. As with the $m = 2$ series, as h_0 increases there is a change from a steady to time-dependent vortex around $h_0 = 300$, with a reduction in the zonal mean jet maximum and an increase in the variance and envelope. Throughout the range of forcing amplitudes between $h_0 = 500$ and $h_0 = 800$ the dominant frequency of the variability generally increases with h_0, but otherwise the behavior is roughly independent of h_0. As h_0 is increased still further, the variability as measured by the envelope of extrema and standard deviation begins to weaken as we approach a more continually disturbed vortex state. These features are all in common with the $m = 2$ case discussed above.
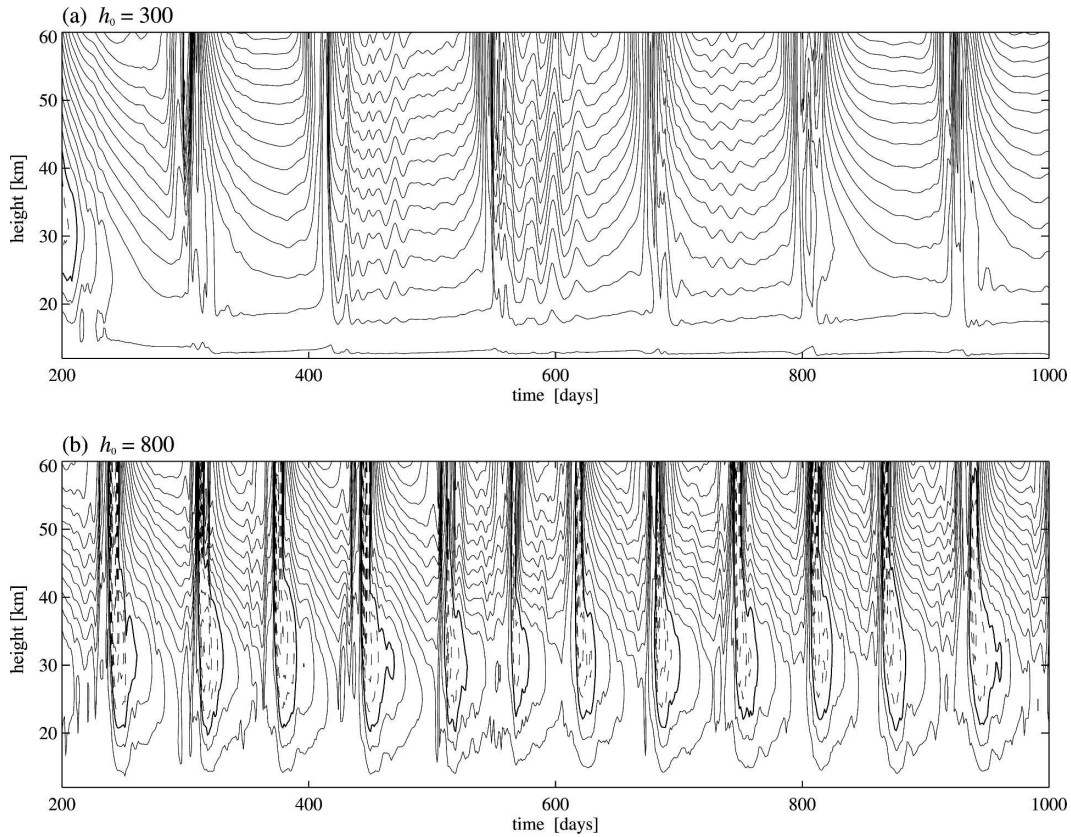
FIG. 13. Same as in Fig. 9 but with a radiative relaxation time scale τ_H given by (7): (a) $h_0 = 300$ m and (b) $h_0 = 800$ m.

b. Numerical parameters

1) SPONGE LAYER

Sponge layers in mechanistic models have traditionally served a dual purpose. On the one hand, they are designed to damp waves in the upper part of the domain and prevent wave reflection from the artificial upper boundary. In this capacity, they can be consid-

ered as a necessary numerical device in the absence of a proper radiation condition. Only the wave components of the flow are damped by the sponge. The sponge layer used in the present study falls into this category: it is a purely numerical device to prevent the reflection of waves from the artificial upper boundary. As a pure wave sponge, therefore, we considered first the sensitivity of the variability described above to the
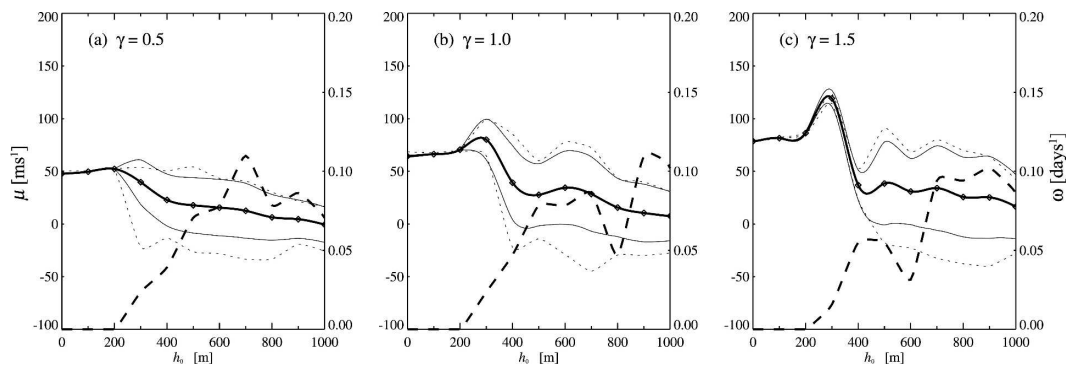

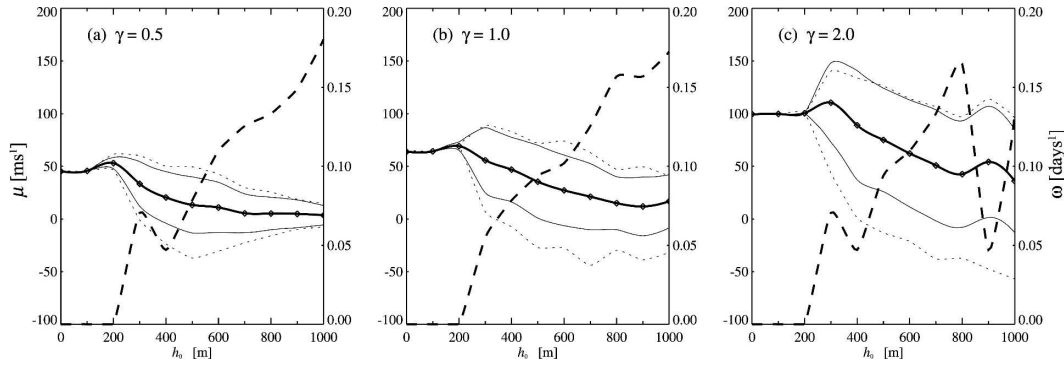FIG. 14. Same as in Fig. 11 but with wavenumber-1 forcing.

FIG. 15. Same as in Fig. 11 but with the sponge layer acting on both the waves and the mean flow.

strength of the sponge. In fact, the effect of doubling or halving the strength of the sponge has practically no effect on the variability obtained and even the sensitivity of details of individual cases to the strength of the sponge was small.

Sponge layers have also been used in another capacity, namely, when they are applied also as a drag on the zonal mean flow, they serve as a type of very basic gravity wave drag parameterization. In this capacity, they are used to control the strength of the zonal mean flow in the mesosphere and to close the polar vortex in the absence of adequate resolved gravity wave drag. Recently, however, the use of a drag on the zonal mean flow has been criticized as violating the angular momentum constraint of the atmosphere and leading to possible feedback effects at altitudes below that of the sponge itself (e.g., Shepherd and Shaw 2004).

Notwithstanding this criticism it is of interest to consider the effect of a zonal mean sponge on the variability discussed above, if only because such sponges have been used extensively in previous mechanistic modeling studies. We use the exact same form as specified in section 2, but impose it also on the zonal mean flow. We emphasize that including such a sponge represents a large perturbation to the model configuration: even apart from possible feedback effects, the equilibrium zonal mean flow profiles change dramatically as a result of the zonal mean drag and the polar vortex is strongly attenuated above a level of about 45 km. We therefore repeated the full series of simulations described in section 4 with this additional zonal mean sponge but with an otherwise identical model configuration. The results are summarized in Fig. 15. Given the significant differences in the structure of the zonal mean flow above 45 km, the similarity between Fig. 15 and Fig. 11 is perhaps surprising, these figures being based on the zonal velocity at 40 km. It suggests that the structure of the flow above the stratopause is not important to the main dynamics; in particular whether the vertical shear in the

upper stratosphere is positive or negative appears to have very little influence on the character of the internal stratospheric variability.

2) VERTICAL AND HORIZONTAL RESOLUTION

One of the aims of this paper is to establish whether coherent modes of internal stratospheric variability, of the type found in low-order models, persist as the number of degrees of freedom is increased. We have established this so far with the above simulations at T42 horizontal resolution with $L = 40$ vertical levels. It is natural to inquire, however, whether the coherent variability discussed above persists at still higher resolutions. Although computational resources prohibit a full investigation of parameter space at higher resolution, we have nonetheless repeated the case $\gamma = 1$, $h_0 = 600$ examined in section 4 at higher (and lower) vertical and horizontal resolution.

We first consider the effect of changing the vertical resolution. Figure 16 shows the zonal velocity as a function of height and time for the case $h_0 = 600$, $\gamma = 1$ shown in Fig. 2b, but with $L = 80$ vertical levels. Although the details of the individual sudden warming events are different, a consequence of the chaotic nature of the evolution, we see that there is very little difference in the character of the variability. Coherent variability certainly persists at the higher vertical resolution and, moreover, quantities such as the dominant frequency and strength of the sudden warmings are very similar to those of the $L = 40$ case. In fact, there is very little difference between the $L = 80$ case and the $L = 20$ case (not shown), and we may conclude that, as far as the vertical resolution is concerned, these results have converged numerically.

Given that the dynamics here is dominated by layerwise two-dimensional motion, and that typical vertical scales are large the above result may not be too surprising. In contrast, horizontal scales develop very steep gradients in potential vorticity (the controlling dynami-
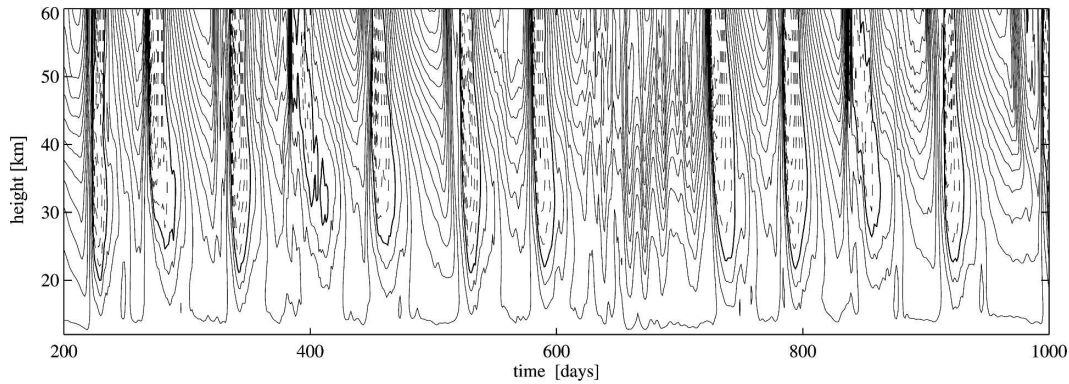
FIG. 16. Same as in Fig. 2b but with a vertical resolution of L80; horizontal resolution is T42.

cal quantity; see e.g., Fig. 7) and so we may expect a stronger dependence on the horizontal representation of the flow. For example, it was shown recently that the horizontal resolution of the vortex edge can have a strong influence on upward wave propagation (Scott et al. 2004). One way in which the effective horizontal resolution is modified is through the hyperdiffusion used to dissipate enstrophy at the smallest scales. However, it turns out that this is a relatively weak effect, and respectively doubling and halving the hyperdiffusion coefficient had almost no impact on the variability.

A more effective test is obtained by altering directly the horizontal resolution, while using the smallest hyperdiffusion allowable for numerical accuracy.[3] We therefore consider the effect of doubling and halving the horizontal resolution to T85 and T21, respectively, and using a hyperdiffusion that acts at the smallest scale on a time scale of $\tau_h = 1$ day in the case of T21, and $\tau_h = 1/24$ days $= 1$ h in the case of T85. These values were determined experimentally as sufficient for numerical accuracy and stability.

Figure 17 shows $\bar{u}(z, t)$, again for the case $\gamma = 1$, $h_0 = 600$, with identical parameters as in section 3 but with a horizontal resolution of T21 and T85. Comparing with the corresponding run at T42 (Fig. 2a) we see that the variability is more regular at T21 and less regular at T85. However, even at T85 the variability retains significant coherence, again being characterized by a sequence of sudden warmings separated by periods of gradual vortex recovery. Part of the loss of regularity is

due to intermittent behavior already discussed: for example, a relative weak vortex persists from $t = 350$ to $t = 500$ and is followed by a period of coherent variability characterized by repeated sudden warmings from $t = 500$ to $t = 700$. Further, the dominant frequency of the variability and the mean and variance of the jet are all comparable with those of the lower resolution cases. We conclude, therefore that, while we may not have reached true numerical convergence with respect to horizontal resolution, it nonetheless appears that the character of the variability does not change significantly with increasing horizontal resolution. Future work will consider in more detail the effect of horizontal resolution on details of wave breaking and the strength of the residual circulation.

6. Conclusions

The main purpose of this paper was to determine the extent to which coherent modes of internal stratosphere variability exist and can be found over a wide range of forcing parameters. We further wished to characterize the dominant dynamical processes responsible for the variability and determine which physical parameters influenced features of the variability such as its dominant frequency and intensity. In so doing, we have adopted the expedient of using a stratosphere-only model, with a simple lower boundary condition to represent wave forcing from the troposphere. This is of course an artificial construction that does not provide a complete representation of the troposphere–stratosphere coupled system, but which does have the advantage of allowing the dynamics of the stratosphere to be considered to some extent in isolation.

Coherent variability exists over the full range of forcing parameters that gives rise to a realistic polar vortex, including the strength of radiative cooling and amplitude of wave forcing. The ranges over which variability

---

[3] Note that it is important when choosing the hyperdiffusion coefficient that the PV remain well-resolved, in the sense that plots such as Fig. 7 do not exhibit noise on the scale of the smallest resolved wave; we stress in particular that it is not sufficient that integrated or averaged quantities such as pressure, temperature, or zonal mean fields be smooth. See Polvani et al. (2004) for further discussion.
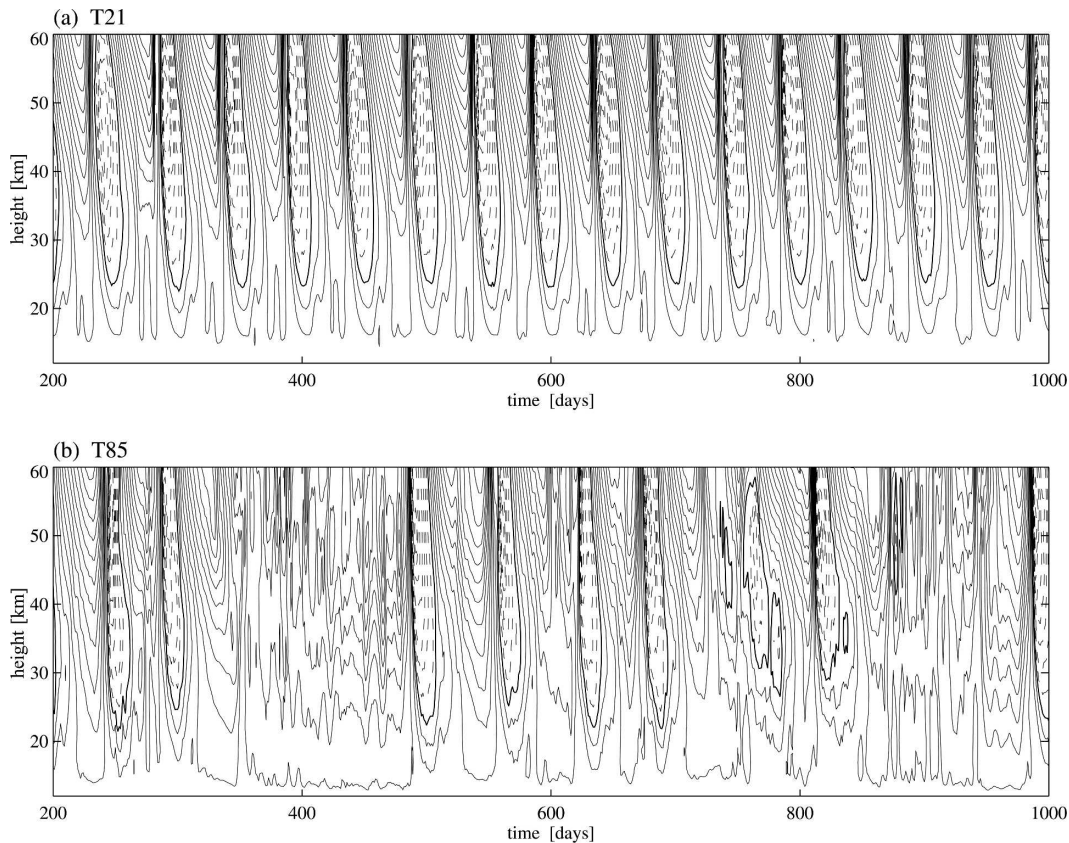
FIG. 17. Same as in Fig. 2b but with a horizontal resolution of (a) T21 and (b) T85; vertical resolution is L40.

was found more than span physically reasonable values. Further, the existence of coherent variability does not depend on the radiative time scale or wavenumber of the lower boundary forcing, nor on details of the numerical scheme used to solve the physical system. In particular, coherent variability appears to persist as the number of degrees of freedom increases, although current computational resources are insufficient for a complete analysis.

The variability arises because of the competing effects of radiative cooling, which acts to strengthen the vortex, and wave forcing, which acts to destroy it. It is thus characterized by a cycle of gradual intensification of the vortex followed by a rapid reduction in circulation resembling a stratospheric sudden warming. Vertical structure is important for this cycle: when the vortex is destroyed at low levels, further upward wave propagation is prevented and the upper vortex strengthens radiatively. When the lower vortex also eventually recovers upward wave propagation resumes. Importantly, when this happens, the total wave flux into the domain through the lower boundary increases dramatically, despite the fact that the lower boundary geopotential wave amplitude is fixed. Thus, the lower

stratosphere appears to act as a valve, preventing or allowing upward wave flux into the domain according to the state of the vortex, and, in our case, independently of the actual lower level geopotential wave amplitudes. Note that a similar dependence of the upward wave flux at the tropopause on the state of the stratosphere was observed in simulations which included a constrained (time-independent) troposphere (Scott and Polvani 2004). The variability in wave flux does not, therefore, appear to be simply an artifact of the choice of lower boundary condition.

Christiansen (1999) described similar vacillations in a GCM in terms of a descending critical level. Although the wave breaking in Fig. 7 begins at upper levels and progresses downward, it is not clear here to what extent the concept of a critical level applies, given the deep vertical structure of the wave. On the other hand, waves do appear to be prevented from propagating upward when the vortex is destroyed at low levels. Moreover, the dependence of the vacillation frequency on the radiative and wave forcing parameters that is obtained here agrees qualitatively with Christiansen's predictions based on scaling arguments, which gives a vacillation time scale proportional to the radiative time scale

and decreasing with increasing wave forcing amplitude. The dependence of time scale on forcing amplitude found here is also consistent with earlier studies using the quasigeostrophic channel model (Holton and Mass 1976; Yoden 1987) and zonally truncated primitive equation model (Scott and Haynes 2000).

The above results have implications for the potential role for stratospheric influence on the troposphere. If internal stratospheric variability did not exist, then the only way in which the stratosphere could be considered as having an influence on the troposphere would be in its capacity as a mediator of tropospheric variability. We have shown here that this is not the case. Further, the fact that the wave flux entering the model domain is controlled entirely by the state of the lower vortex supports the possibility, previously put forward by Hartmann et al. (2000), that the lower stratosphere may also control the amount of wave flux leaving the troposphere, and hence may influence other aspects of the tropospheric circulation. We will return to this question in Part II, where we examine the relative importance of internal variability and variability arising from time-dependence in the wave forcing.

REFERENCES

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

Baldwin, M. P., and T. J. Dunkerton, 1999: Downward propagation of the Arctic Oscillation from the stratosphere to the troposphere. *J. Geophys. Res.,* **104,** 30 937–30 946.

Cehelsky, P., and K. K. Tung, 1987: Theories of multiple equilibria and weather regimes—A critical reexamination. Part II: Baroclinic two-layer models. *J. Atmos. Sci.,* **44,** 3282–3303.

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

Christiansen, B., 1999: Stratospheric vacillations in a general circulation model. *J. Atmos. Sci.,* **56,** 1858–1872.

——, 2000a: Chaos, quasiperiodicity, and interannual variability: Studies of a stratospheric vacillation model. *J. Atmos. Sci.,* **57,** 3161–3173.

——, 2000b: A model study of the dynamical connection between the arctic oscillation and stratospheric vacillations. *J. Geophys. Res.,* **105,** 29 461–29 474.

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.

Esler, J. G., and R. K. Scott, 2005: On the excitation of transient Rossby waves on the polar stratospheric vortex and the barotropic sudden warming. *J. Atmos. Sci.,* **62,** 3661–3682.

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

Hartmann, D. L., J. M. Wallace, V. Limpasuvan, D. W. J. Thompson, and J. R. Holton, 2000: Can ozone depletion and global warming interact to produce rapid climate change? *Proc. Natl. Acad. Sci. USA,* **97,** 1412–1417.

Haynes, P. H., 2005: Stratospheric dynamics. *Annu. Rev. Fluid Mech.,* **37,** 263–293.

Holton, J. R., and C. Mass, 1976: Stratospheric vacillation cycles. *J. Atmos. Sci.,* **33,** 2218–2225.

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

Legras, B., and D. G. Dritschel, 1993: Vortex stripping and the generation of high vorticity gradients in two-dimensional flows. *Appl. Sci. Res.,* **51,** 445–455.

——, ——, and P. Caillol, 2001: The erosion of a distributed two-dimensional vortex in a background straining flow. *J. Fluid Mech.,* **441,** 369–398.

Matsuno, T., 1971: A dynamical model of the stratospheric sudden warming. *J. Atmos. Sci.,* **28,** 1479–1494.

Norton, W. J., 2003: Sensitivity of northern hemisphere surface climate to simulation of the stratospheric polar vortex. *Geophys. Res. Lett.,* **30,** 1627, doi:10.1029/2003GL016958.

Pawson, S., A. Meyer, and S. Leder, 1995: Internal variability in a perpetual January integration of a troposphere-stratosphere-mesosphere GCM. *Quart. J. Roy. Meteor. Soc.,* **121,** 369–397.

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.

——, and P. J. Kushner, 2002: Tropospheric response to stratospheric perturbations in a relatively simple general circulation model. *Geophys. Res. Lett.,* **29,** 1114, doi:10.1029/2001GL014284.

——, R. K. Scott, and S. J. Thomas, 2004: Numerically converged solutions of the global primitive equations for testing the dynamical core of atmospheric GCMS. *Mon. Wea. Rev.,* **132,** 2539–2552.

Rivier, L., R. Loft, and L. M. Polvani, 2002: An efficient spectral dynamical core for distributed memory computers. *Mon. Wea. Rev.,* **130,** 1384–1390.

Rong, P.-P., and D. W. Waugh, 2004: Vacillations in a shallow-water model of the stratosphere. *J. Atmos. Sci.,* **61,** 1174–1185.

Scott, R. K., and P. H. Haynes, 2000: Internal vacillations in stratosphere-only models. *J. Atmos. Sci.,* **57,** 2333–2350.

——, and L. M. Polvani, 2004: Stratospheric control of upward wave flux near the tropopause. *Geophys. Res. Lett.,* **31,** L02115, doi:10.1029/2003GL017965.

——, L. Rivier, R. Loft, and L. M. Polvani, 2004: BOB: Model implementation and users guide. NCAR Tech. Note NCAR/TN-456+IA, 30 pp.

Shepherd, T. G., and T. A. Shaw, 2004: The angular momentum constraint on climate sensitivity and downward influence in the middle atmosphere. *J. Atmos. Sci.,* **61,** 2899–2908.

——, K. Semeniuk, and J. N. Koshyk, 1996: Sponge layer feedbacks in middle-atmosphere models. *J. Geophys. Res.,* **101,** 23 447–23 464.

Taguchi, M., T. Yamaga, and S. Yoden, 2001: Internal variability of the troposphere–stratosphere coupled system simulated in a simple global circulation model. *J. Atmos. Sci.,* **58,** 3184–3203.

Yoden, S., 1987: Bifurcation properties of a stratospheric vacillation model. *J. Atmos. Sci.,* **44,** 1723–1733.