Factors Controlling Nonlinearity in Mechanically Forced Stationary Waves over Orography

TODD D. RINGLER AND KERRY H. COOK

Atmospheric Science Program, Cornell University, Ithaca, New York

(Manuscript received 10 September 1996, in final form 21 April 1997)

ABSTRACT

The forcing of stationary waves by the earth's large-scale orography is studied using a nonlinear stationary wave model based on the quasigeostrophic equations. The manner in which wind speed, meridional temperature gradient, Ekman pumping parameter, linear damping, orographic shape, and meridional wind structure affect the validity of the linearized equations is examined and the nonlinear response is investigated.

A critical mountain height that separates the linear from the nonlinear regime is defined based on the linear quasigeostrophic potential temperature equation applied at the surface. The largest critical heights (those responses in which nonlinearity is least important) are obtained when the surface damping is weak or nonexistent. Also, relative maximums in mountain critical heights are obtained when the ratio of surface wind to surface wind shear does not vary in the meridional direction. These critical height results are validated using the fully nonlinear stationary wave model.

The nonlinearly balanced response to imposed orography is diagnosed at the surface and aloft. The nonlinear effects of eddy wind/orography interaction and nonlinear advection are found to be important only in the vicinity of the orography. The structure of the nonlinear response at the surface is found to be robust and is characterized (in the Northern Hemisphere) by a high and low situated to the northwest and southeast, respectively, of the mountain center. This orientation of the surface response leads to a stationary wave train that propagates preferentially toward the equator.

The system is sensitive enough to both the surface wind and meridional temperature gradient that the observed seasonal variations in the zonal mean circulation will significantly alter the character of the response. As the meridional temperature gradient decreases, the relative importance of nonlinearity increases while the amplitude of the response at the upper levels decreases. Therefore, this model indicates that summertime mechanically forced stationary waves should be weaker, but more nonlinear, than their wintertime counterparts.

1. Introduction

One of the many problems in climate dynamics is the task of reproducing and explaining the observed stationary wave pattern. Persistent planetary-scale anomalies are readily identified whenever observed fields are averaged over a few weeks or more. While the amplitude of stationary wave patterns is substantially less than that of the transient baroclinic eddies in middle latitudes, the long timescale of the stationary waves makes them of primary importance in controlling regional climate characteristics.

Any zonal asymmetry in the boundary conditions, such as orographic height, has the potential to produce a stationary response. Any mechanism that creates zonal asymmetry in the diabatic heating rate can also generate stationary waves. Furthermore, these planetary-scale anomalies can themselves produce zonal asymmetry in the transient eddies, which can feed back onto the planetary-scale flow.

Previous analytical and modeling studies of stationary waves demonstrate that large-scale topography is an important factor in determining stationary wave structure in the Northern Hemisphere midlatitudes. These studies have shown that ideas from β-plane linear perturbation theory, most notably the concept of the generation of Rossby waves by the obstruction of westerly flow impinging on mountains, are relevant for understanding observed stationary wave patterns. As discussed in the following section, a number of numerical models have been developed to take advantage of the linearity and/or stationarity of the problem, and these models provide a sound first-order understanding.

Several authors (see below) have suggested that in order to extend our basic understanding to more geophysically realizable systems, it is crucial to understand the limits of linear theory and to consider nonlinear processes. Orographic height, surface wind speeds, meridional temperature gradients, and attendant wind shears are among the factors that determine the degree of linearity of the atmosphere's response to topography.

Corresponding author address: Todd D. Ringler, Dept. of Atmospheric Sciences, Colorado State University, Fort Collins, CO 80523. E-mail: todd@atmos.colostate.edu

While the influence of some of these factors has been explored in limited ways, a thorough survey of the factors that lead to nonlinearity and a complete physical description of the nonlinear response have yet to be given. The purpose of this paper is to address these issues.

The following section reviews much of the previous modeling work on orographically induced stationary waves, with an emphasis on results that point to the importance of nonlinearity. Section 3 contains a description of the model developed for this study and the place of this model in the hierarchy of models that have been applied to this problem. In section 4 a critical height is defined and applied to characterize the influence of various parameters on the degree of linearity of the solution. In section 5 the spatial structure of a prototypical nonlinear response is compared to the linear response. The effects of surface wind and meridional temperature gradient on the amplitude of the nonlinear response are discussed in section 6. In section 7 we draw some conclusions based on our findings.

2. Background

Steady-state models have been developed to exploit the fact that the time tendency of the planetary-scale flow is small compared to the magnitude of the forcing by neglecting time rates of change in the governing equations. Most of these models are based on a set of linear dynamical equations. Discrepancies between linear theory and observations indicate that some of the essential dynamics is missing from the linear framework.

Two types of approximation are made to study orographically induced stationary waves in the linear framework. One is to neglect the nonlinear terms, such as the advection of perturbation quantifies by the eddy wind, in the atmospheric domain. The other approximation comes about in setting the lower boundary condition to generate mechanical orographic forcing, which describes the forcing of vertical velocity by the obstruction of the low-level flow. In the traditional linear approximation, this vertical velocity is given by $w_{\text{lin}} = [u]\,\partial h/\partial x$, where $[u]$ is the zonal mean flow evaluated along a constant pressure surface and h is orographic height. The true mechanical forcing, which is evaluated along the lower boundary, is $w_{\text{true}} = \mathbf{v}\cdot\nabla h$ where $\mathbf{v}$ is the total horizontal flow field defined as $\mathbf{v} = [\mathbf{v}] + \mathbf{v}^*$ and the asterisk indicates the deviation from the zonal mean. Saltzman and Irsch (1972) showed that the portion of the vertical velocity that the linear theory neglects, $w_i = \mathbf{v}^*\cdot\nabla h$, is as large as the portion that it retains. This portion of the forced vertical velocity, w_i, is not a true nonlinear term since it is not the multiplication of two unknown terms. On the other hand, w_i is often referred to as the nonlinear mechanical forcing since it is the portion of the forced vertical velocity that is not accounted for by traditional linear theory.

In the past decade several attempts have been made to include some of the essential dynamics missing from stationary linear models without reverting to the full primitive equations. Chen and Trenberth (1988) developed a stationary wave model that retains the full mechanical forcing, $w_{\text{lin}} + w_i$, evaluated on a constant pressure surface while still using linear equations to govern the flow in the interior. They find that the interaction of the eddy meridional flow with topography is important for improving the simulation of observed stationary waves. Valdes and Hoskins (1991) constructed a nonlinear stationary wave model based on the primitive equations and found a substantially smaller response to mechanical forcing than that expected from linear theory. Cook and Held (1992) used an idealized general circulation model (GCM) and a linear stationary wave model to infer the importance of nonlinearity at different orographic heights. Since Cook and Held used a GCM that included a hydrologic cycle and radiation, the GCM captured not only the direct changes in the mechanical forcing due to orography, but also the indirect effects due to changes in the thermal forcing.

A number of authors (Trenberth and Chen 1988; Valdes and Hoskins 1991) discuss the concept of a critical height to separate the linear regime from the nonlinear regime. One potential pitfall of this construct is the suggestion that orographic height is the dominant parameter controlling linearity. This is certainly not the case since the ambient atmospheric conditions play a central role in deciding the degree of linearity. For example, Cook and Held (1988) showed that the response to a 4-km high Laurentide ice sheet over North America is essentially linear in the presence of the large ice-age meridional temperature gradient, but not when the present-day meridional temperature gradient is imposed.

Another possible misconception of critical height is that it marks the height at which the response evolves rapidly from a linear regime to a nonlinear regime. While this is possible, it is not the case in general. This work and that of others suggests that the transition is rather smooth, with a sizable ''fuzzy'' area in between the regimes.

Despite the danger of these misconceptions, the use of a critical height is useful for organizing discussion. The set of equations used in this study, the quasigeostrophic equations, makes the definition of a critical height particularly meaningful.

3. Model description and numerical methods

a. Motivation

The model used in this study is a stationary wave model, so it solves for the time-mean response forced by specified boundary conditions. Since several different types of stationary wave models already exist, this model is compared with previously developed models to define the context of this study.

The most sophisticated stationary wave model is based on the nonlinear primitive equations in a spherical geometry (Ting and Held 1990; Valdes and Hoskins 1991). While these models are exceptionally powerful for diagnosing and understanding the GCM results, problems in obtaining convergent, and therefore accurate, solutions have limited their utility. At the other end of the spectrum are models such as the one developed by Ashe (1979). This was a two-level hemispheric model with coarse horizontal resolution based on the linear balance equations of Lorenz (1960). Since the scale of the nonlinear response can be different than the scale of the forcing, this model may have been of insufficient horizontal and vertical resolution to confidently determine the extent to which nonlinearity may be important.

The model used here is of intermediate complexity in the hierarchy of stationary wave models. It is based on the quasigeostrophic equations on the midlatitude β plane, so it is fully nonlinear but has a restricted domain. The quasigeostrophic equations are more tractable than the primitive equations numerically, and we find that problems with obtaining convergent solutions are minimal. At the same time, the quasigeostrophic system retains much of the nonlinearity that is interesting in studying stationary waves, namely, nonlinear advection, eddy wind–orography interaction, and orography–diabatic heating interaction. In addition, the concise dynamical framework of conservation of potential vorticity allows fairly straightforward explanations of the often complicated nonlinear responses. Large portions of the parameter space can be explored without the prohibitive computational expense required by a primitive equation model.

b. Governing equations

1) NONLINEAR QUASIGEOSTROPHIC EQUATIONS

The model is governed by two equations, one applied at the surface and another to describe the interior. These equations are derived from the standard quasigeostrophic system.

Following Held (1983), the quasigeostrophic thermodynamic equation has the standard form of

$$\frac{\partial}{\partial t}\left(\frac{\partial \psi}{\partial z}\right) + J\left(\psi, \frac{\partial \psi}{\partial z}\right) = \frac{-N^2}{f_o} W + \frac{g}{f\Theta} Q. \qquad (1)$$

The sole independent variable of the system is the streamfunction ψ. The vertical coordinate is $z = H_s \ln(p_s/p)$, where $H_s = R\Theta_s/g$ is the scale height at the mean surface pressure, p_s; $\Theta(z)$ is the basic-state potential temperature profile, g is the acceleration of gravity, N^2 is the Brunt–Väisälä frequency, and Q is the diabatic heating rate in kelvins per second. The lower boundary condition on the vertical velocity, $W = dz/dt$, is

$$W(z = 0) = -\frac{f_o}{g}\frac{\partial \psi}{\partial t} + J(\psi, h) + \alpha \nabla^2 \psi, \qquad (2)$$

where h is the height of the surface orography. Here, $J(\psi, h) = \mathbf{v} \cdot \boldsymbol{\nabla} h$ is the full kinematic boundary condition and $\alpha \nabla^2 \psi$ represents the effects of Ekman pumping. The first term on the rhs of (2) goes to zero in the steady-state response.

The quasigeostrophic potential vorticity equation, which is derived from the quasigeostrophic vorticity equation, governs the interior:

$$\frac{\partial}{\partial t}\nabla^2 \psi + J(\psi, \nabla^2 \psi + \beta y) = \frac{f_o}{\rho}\frac{\partial (\rho W)}{\partial z}, \qquad (3)$$

where $\rho(z) = \exp(-z/H_s)$. Eliminating W between in Eqs. (1) and (3) gives

$$J(\psi, q) = \frac{g f_o}{\rho \Theta}\frac{\partial}{\partial z}\left(\frac{\rho Q}{N^2}\right) + D, \qquad (4)$$

where

$$q = f_o + \beta y + \nabla^2 \psi + \frac{f_o^2}{\rho}\frac{\partial}{\partial z}\left(\frac{\rho}{N^2}\frac{\partial \psi}{\partial z}\right) \qquad (5)$$

and D represents dissipation.

The second governing equation of the model is the lower boundary condition of the quasigeostrophic system. Substituting (2) into (1) gives

$$J(\psi, s) = \frac{g}{f_o \Theta_s} Q - \frac{\alpha N^2}{f}\nabla^2 \psi + D, \qquad (6)$$

where

$$s = \frac{\partial \psi}{\partial z} + \frac{N^2}{f} h. \qquad (7)$$

Both q and s are only functions of externally supplied parameters and the sole independent variable, ψ. Orography, h, enters the system only at the surface and has been included in the Jacobian on the lhs of (6). An equation similar to (6) is applied at the top boundary, with $z = Z_t$ and without orography or Ekman pumping. Results reported here use the model with $Q = 0$ in (4) and (6). The effects of diabatic heating are considered in a companion paper.

The dissipation, D, includes linear damping and hyperdiffusion of the form

$$D = -\gamma a^* - \mu \nabla^4 a^*, \qquad (8)$$

where

$$a^* = \begin{cases} q - [q], & 0 < z < Z_t \qquad (9a) \\ \dfrac{\partial \psi}{\partial z} - \left[\dfrac{\partial \psi}{\partial z}\right], & z = 0. \qquad (9b) \end{cases}$$

Therefore, the dissipation acts to restore the system toward the zonal mean condition.

Equations (4) and (6) show that, in the absence of heating and dissipation, q and s are conserved quantities in the interior and at the lower boundary, respectively. In addition to conserving potential vorticity in the in-

terior, the quasigeostrophic system conserves a quantity proportional to the parcel's entropy at the surface. Multiplying (7) by $f_o H_s/R$ gives

$$S = \Theta_e + \frac{\partial \Theta}{\partial z} h, \tag{10}$$

where Θ_e is the parcel's potential temperature minus the basic-state potential temperature, $\Theta(z = 0)$. Thus, Θ_e contains information about the meridional potential temperature gradient. Just as q is conserved in the model interior, S is conserved at the surface.

We refer to S as ''surface entropy'' and use this quantity to understand the nonlinear response to imposed orography. This quantity was derived in a slightly different form by Tung (1983). Physically, we can think of the surface entropy as being composed of two parts. The first is the potential temperature of a parcel located along the lower boundary, $z = 0$. The second part is the potential temperature anomaly created by a parcel being elevated on the orography to a height, h. The conservation of S following the geostrophic flow necessitates a strong interplay between the mechanical forcing and the surface potential temperature response.

2) Linear equations and residuals

In this study, both linear and nonlinear solutions are generated from the governing equations (4) and (6). To solve for the linear response to orography, (4) and (6) are linearized about the zonal mean flow to give

$$[u]\frac{\partial q^*}{\partial x} + v^*\frac{\partial [q]}{\partial y} = \frac{g f_o}{\rho \Theta}\frac{\partial}{\partial z}\left(\frac{\rho Q^*}{N^2}\right) + D, \qquad 0 < z < Z_t \tag{11}$$

$$[u]\frac{\partial s^*}{\partial x} + v^*\frac{\partial [s]}{\partial y} = \frac{g}{f_o \Theta} Q^* - \frac{\alpha N^2}{f_o}\nabla^2 \psi^* + D, \qquad z = 0. \tag{12}$$

A linear operator is formed and the linear solution, ψ^*, readily obtained. Once this linear solution has been obtained, an estimate of the error involved in neglecting eddy–eddy interactions can be obtained post facto by evaluating

$$E_i = J(\psi^*, q^*) \qquad 0 < z < Z_t \tag{13a}$$

$$E_s = J(\psi^*, s^*) = J\left(\psi^*, \frac{\partial \psi^*}{\partial z} + \frac{N^2}{f_o} h\right), \qquad z = 0, \tag{13b}$$

where $E_{i(s)}$ is defined as the linear error in the interior (surface). When only mechanical forcing is present, the imposed forcing at the surface is given as

$$L_s(x, y) = [u]\frac{N^2}{f}\frac{\partial h}{\partial x}. \tag{13c}$$

The residual (13b) can be compared to the imposed forcing (13c) to assess the validity of the linear approximation.

c. Solving the fully nonlinear equation set

In constrast to the linearly balanced solutions, which are obtained by expanding (4) and (6) about the zonal mean flow field, to obtain approximate nonlinear balanced solutions, (4) and (6) must be expanded about a fully three-dimensional flow. The technique used to obtain nearly balanced nonlinear solutions has two steps. The first is an iteration procedure in which, for a given forcing, the error is reduced to an acceptable level. This process consists of initially finding the response in the zonal mean basic state. Then this response is added to the zonal mean basic state and the system is expanded about this (new) three-dimensional flow. The iteration cycle of finding the response, updating the previous solution with the response, and finding the response to the new state, continues until the system converges to yield a nonlinearly balanced response for a given mountain height. The second step involves the ''continuation'' over orographic height. Once a nonlinearly balanced solution is obtained for a given mountain height, the orographic height is increased and the iteration cycle is performed again using the balanced solution obtained at the previous orographic height as the first guess for the new solution.

In the iteration procedure, a solution of the form

$$\psi^{i+1}(x, y, z) = \psi^i(x, y, z) + \delta(x, y, z) \tag{14}$$

is assumed so that the solution at iteration $i + 1$ is the sum of a known solution at iteration i and a correction term, δ. During the iteration procedure, the solution is relaxed with a relaxation coefficient of $\tau = 0.20$. This iteration technique was successfully used by Haupt et al. (1993) to solve for steady-state modon solutions in a 2D barotropic flow; we have extended the method to 3D flow. Equation (14) is substituted into (4) and (6) and those equations are expanded in δ. As in Haupt et al. (1993), neglecting all terms $O(\delta^2)$ leads to a system of equations that can be written in the form

$$H\delta = F. \tag{15}$$

The actual forms of H and F are written out in appendix A. Note that F is known since it contains only terms based on the previous iteration and is the error in the system. Therefore, if $F \to 0$ the solution is exactly nonlinearly balanced.

Using a 2D Fourier series in the horizontal and finite differences in the vertical, H and F are discretized according to

$$\psi^i = \sum_{k=1}^{K} \sum_{l=-L}^{L} \tilde{\psi}^i(k, l, z_d) e^{i(kx+ly)} \tag{16a}$$

and

$$\delta = \sum_{k=1}^{K} \sum_{l=-L}^{L} \tilde{\delta}(k, l, z_d) e^{i(kx+ly)}, \tag{16b}$$

where k and l refer to the zonal and meridional wavenumbers, respectively, and subscript d denotes a given vertical level. The $k = 0$ wavenumber is truncated, so the zonal mean flow is not altered in the course of the iteration. Since the solution field is real, only half of the wavenumber space needs to be retained in (16), therefore the $-k$ wavenumbers are also truncated. (The $-k$ wavenumbers can be recovered since they are the complex conjugate of the $+k$ wavenumbers.)

In all the results presented, either 16 or 24 waves (denoted as C16 and C24, respectively) are retained in both k and l ($K = L = 16$ or 24). A circular truncation is applied to the expansion; for example, when $K = L = 24$, all wavenumbers that satisfy $k^2 + l^2 > 24^2$ are truncated. Since we are also interested in the far-field stationary wave response, the resolved domain is 20 000 km by 20 000 km. This results in a collocation grid spacing of 416 km by 416 km at C24 resolution.

In the vertical there are 30 levels between 0 and 4 scale heights. Two-thirds of the levels are below 1.5 scale heights to better capture the vertical structure of the stationary wave. Two-sided finite differences are used in the interior and one-sided finite differences are used at the boundaries.

Since the Fourier series forms an orthogonal basis, an equation for each $\tilde{\delta}(k_\alpha, l_\alpha, z_d)$ is obtained by taking the inner product of (15) with a chosen wavenumber pair, (k_α, l_α), written as

$$(e^{-i(k_\alpha + l_\alpha)}, H\delta) = (e^{-i(k_\alpha + l_\alpha)}, F). \tag{17}$$

This results in the matrix formation

$$\mathbf{H}\delta = \mathbf{F} \tag{18}$$

($\mathbf{H}$ and $\mathbf{F}$ refer to the discretized forms of H and F). Since the horizontal expansion is global, all horizontal wavenumbers on a given vertical level are coupled. In contrast, the finite differencing of the vertical discretization results in a local expansion where the quantities on any given z_d are coupled only to quantities on z_{d-1} and z_{d+1}. The local vertical coupling allows the matrix operator, $\mathbf{H}$, to be cast as a block tridiagonal and solved by block LU factorization (Golub and van Loan 1993).

After this iteration step has produced a convergent solution, the continuation procedure is used to obtain the nonlinear response to larger forcing. Beginning in a linear regime with very small orography, orographic heights are slowly increased. [A similar continuation process was used by Ting and Held (1990).] At each orographic height, the matrix system is iterated until a convergent solution is obtained. At each orographic height the iteration cycle is performed to obtain the convergent solution consistent with the new forcing.

In this matrix formulation of the fully nonlinear equations, information about both the orography and flow field is contained in the matrix operator, $\mathbf{H}$. Therefore, every time the orographic height is increased, $\mathbf{H}$ should be recomputed. However, we find that the accuracy of the solution is not compromised if $\mathbf{H}$ is recomputed less frequently, so we take 25-m height increments and recompute $\mathbf{H}$ at 500-m intervals. Nonlinear equations can produce multiple solutions and these can be path dependent. No evidence of multiple solutions within this model has been found.

d. Convergence criterion

A criterion for determining when the solution is sufficiently close to a balanced, nonlinear solution must be chosen. Since $\mathbf{F}$ [Eq. (18)] is the deviation from a balanced flow field for each (k, l, z_d), some norm of $\mathbf{F}$ would give a measure of the system error. However, there are two difficulties with using $\mathbf{F}$ to construct the convergence criterion. First, $\mathbf{F}$ has different units at the surface and the interior so that a suitable criterion for each would be required. Second, we have found, as has Haupt (1995, personal communication), the minimum system error that we have been able to obtain increases as the amplitude of the orographic forcing increases.

Instead, the convergence criterion is based on the root-mean-square (rms) of the correction, δ. At each level, the mass-weighted rms of δ is computed. The mass weighting is necessary since in vertically propagating waves $\rho_o|\psi^{*2}|$ remains constant, not $|\psi^{*2}|$. When the mass-weighted rms of δ drops below a specified value at all vertical levels, the solution is considered balanced. (The specified value is chosen so that the system error is at least three orders of magnitude smaller than the forcing.) The following example demonstrates that this is an acceptable convergence criterion.

Choosing a realistic wind profile and parameter settings (discussed below), we have obtained nonlinear-balanced responses to orographic heights ranging from 25 to 4000 m with the above convergence criterion. Plotted in Fig. 1 against orographic height are the rms of the nonlinear mechanical forcing $J(\psi, N^2/f_o h)$; the linear model error, E_s, defined in (13b); and the nonlinear model error, $\mathbf{F}$, defined in (18). The nonlinear model error is nearly four orders of magnitude less than the nonlinear mechanical forcing for all orographic heights. Thus, these solutions can be thought of as nonlinearly balanced, quasigeostrophic solutions.

The linear model error, E_s, is of the same order as the forcing for orographic heights greater than 0.5 km and even exceeds the forcing at higher orographic heights. Thus, the nonlinear model is more than three orders of magnitude more accurate than the linear model.

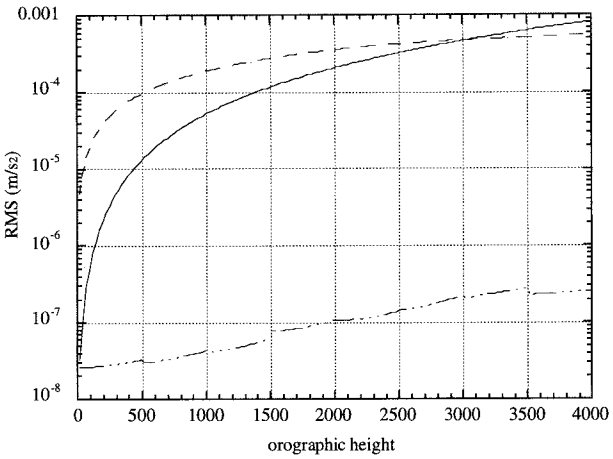

FIG. 1. A comparison of the magnitude of model error and the magnitude of mechanical forcing for a range of orographic heights. Dashed line represents magnitude of the total mechanical forcing as obtained from solving the nonlinear equations. The solid line represents magnitude of error that occurs if the linear approximation is made, and the dashed–dotted line gives error of the nonlinearly balanced flow.

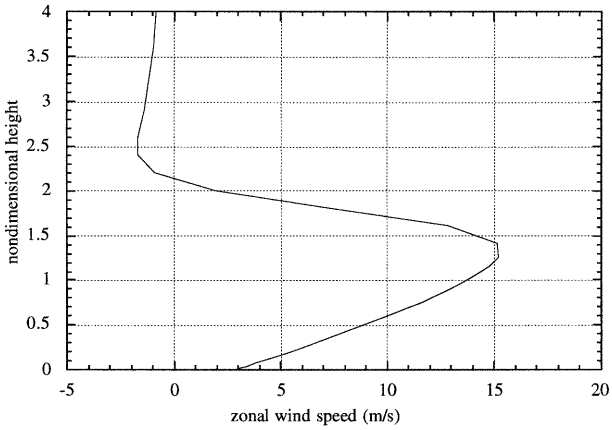

FIG. 2. A common vertical profile of the zonal mean wind used in this study.

e. Parameter settings and basic-state wind profiles

Since the system is forced, some dissipation is required to prevent resonance from corrupting the solution. Part of this dissipation takes the form of linear damping as given in (8). Below $z = 1.2H_s$, the linear damping coefficient, γ, is held uniform at 0.1 days^{-1}. Above $z = 1.2H_s$, γ is increased linearly to 1.0 day^{-1} at $z = 4.0H_s$ to damp vertically propagating waves, which could reflect off the top boundary. A sponge layer is located around the domain boundary with $\gamma = 1.0$ day^{-1} to prevent horizontally propagating wave energy from reentering the domain. Through specification of the domain size, the sponge layer is far enough from the forcing so that it does not effect the response. A scale-selective ∇^4 diffusion with a coefficient of $\mu = 10^{16}$ m^4 s^{-1} is also used. To account for the no-slip boundary condition at the lower surface, Ekman pumping with a value of $\alpha = 150$ m is included at the surface. The scale height, H_S, is set to 8500 m in all experiments.

The β plane is centered at 45° latitude. The meridional extent of the domain is 20 000 km, which is much larger than is valid under the β-plane approximation. Our analysis is, therefore, restricted to the near-field response. The far-field response is examined only to understand the qualitative nature of the stationary wave propagation.

All the orographic shapes specified have the form

$$h(x, y) = h_0 \exp\left(\frac{-x^2}{w_x^2} + \frac{-y^2}{w_y^2}\right) \tag{19}$$

with $w_x = w_y = 1500$ km unless otherwise stated.

The specified basic-state wind profiles are consistent with those observed in the extratropics (Fig. 2). This profile is similar to that used by Held et al. (1985). The surface wind, surface wind shear, tropopause height, and tropopause strength are adjustable.

f. Limitations of quasigeostrophic framework

One often-stated quasigeostrophic requirement is that the ratio of the orographic height to the scale height be much less than one; that is, $h/H_s \ll 1$. This requirement arises from making one assumption and having to meet one quasigeostrophic constraint. The assumption is that the mechanical forcing is well described by the linear forcing, $[u]\ \partial h/\partial x$. The constraint that must be met is that the magnitude of the forced divergence be $O(\mathrm{Ro}U/L)$ where Ro the Rossby number, U is the characteristic velocity, and L the characteristic length scale. As the orographic height increases, the mechanical forcing saturates and is substantially less than the linear forcing estimate. Therefore, the assumption is too strong and the quasigeostrophic framework is valid for orographic heights larger than those for which linear theory is valid.

Another limitation of the quasigeostrophic framework concerns the implementation of the lower boundary condition on the $z = 0$ surface instead of the $z = h$ surface. When the zonal wind shear is large, the mechanical forcing is in error since quasigeostrophic theory uses $u(z = 0)$ instead of $u(z = h)$ as the forcing wind [see da Silva (1989) for more discussion on this topic]. Although this is a quantitative error, it does provide a conceptual simplification. When the surface is elevated, the magnitude of the surface wind is dependent on the local wind shear, therefore the two are not separable. This separation is preserved as a conceptual tool in the quasigeostrophic framework.

4. Mechanical orographic forcing—Limits of linearity

In this section the boundary between linear and nonlinear dynamics is examined from within the framework of linear theory. This provides a convenient way of

exploring the parameter dependence of the critical height, that is, the mountain height at which linearity is no longer valid. As shown below, the linear analysis gives results consistent with a direct comparison between linear and nonlinear solutions.

a. Review and definition of critical height

Several scaling arguments have been put forth to predict whether the response to a given orography will be linear or nonlinear. Trenberth and Chen (1988) scale the linear, quasigeostrophic, potential vorticity equation to determine the mountain height at which the perturbation streamfunction produced by air deflected around a mountain is equal to the perturbation induced by the air mass rising over the mountain. They argue that if the streamfunction field is perturbed less by the air going around the mountain than over it, the flow will be substantially deflected. In this case, the traditional linear boundary condition (13c) will no longer be valid. The critical height defined in this manner is only a function the earth's radius, mountain meridional scale, atmospheric scale height, and latitude. That the critical height is a not a function of the surface wind, meridional temperature gradient, or the rotation rate of the earth suggests that it does not capture all the relevant dynamics (see Lenters et al. 1995).

The scaling by Valdes and Hoskins (1991) is based on the perturbation potential temperature equation at the surface. They suggest that meridional deflection of the flow will dominate when

$$\left| L_y \frac{\partial[\Theta]}{\partial y} \right| < \left| h \frac{\partial[\Theta]}{\partial z} \right|, \tag{20}$$

where L_y is the mountain meridional scale. Thus, meridional deflection of the air mass will be important whenever the meridional slope of the orography is greater than the characteristic slope of the zonal mean isentropes. This scaling assumes that the meridional advection of the mean meridional temperature balances the orographic forcing (see Cook and Held 1992).

The scaling of Valdes and Hoskins can, alternatively, be obtained by taking the inviscid, adiabatic form of (12) forced solely by orography. By then assuming that the response decays vertically, the meridional advection of $[\Theta]$ balances the orographic forcing, and the basic-state winds are meridionally uniform, (20) is obtained by hypothesizing that linearity fails when $v^* \approx [u]$ at the surface. (This criterion is generous toward linear theory since it would probably fail well before eddy velocities approach $[u]$.)

As discussed in section 2, here we use the full linear solution to a given mountain height in a given atmospheric environment to define a critical height. One can assess the validity of linear theory by computing the linear forcing error (13b) post facto and comparing this error to the specified forcing (13c). The forcing error is an estimate of the forcing that would arise due to the interaction of the linearly perturbed flow with itself and with the orography. This is only an error estimate because the perturbed flow in the nonlinear case would not be the same as in the linear case.

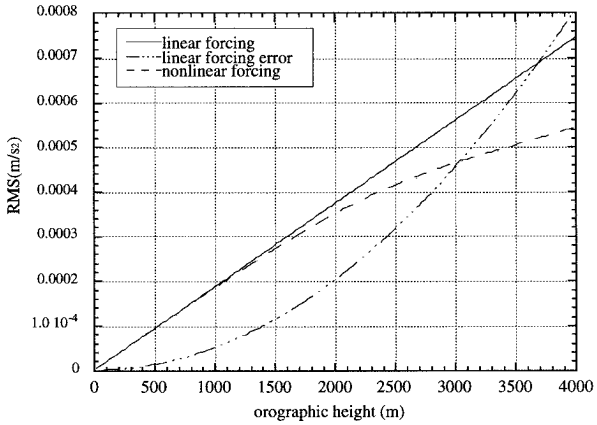

FIG. 3. A comparison of the magnitudes of the linear forcing approximation to mechanical forcing (solid) and the actual nonlinear mechanical forcing (dashed). The dash–dotted line represents the error of the linear approximation.

Since ψ^* and ψ_z^* both increase linearly with orographic height in the linear framework, (13b) shows that the forcing error increases as h^2 and that its value will be dependent upon the relative phases of ψ^*, ψ_z^*, and h. Once the linear solution is obtained for a single orographic height, the linear forcing, response, and error can be obtained for any height by simply scaling with h, h, and h^2, respectively. Since the linear forcing, response, and error only vary in magnitude (not shape) with varying h, the rms of the linear forcing, response, and error also scale as h, h, and h^2.

The rms of the linear forcing (13c), the true mechanical forcing ($N^2/f_o \mathbf{v}\cdot\nabla h$), and the linear error (13b) are plotted against orographic height in Fig. 3. (the parameter settings, orographic shape, and wind profile are the same as that used in Fig. 1). When compared to the forcing, how large can the error be in Fig. 3 before linear theory should be judged invalid? Also, at what mountain height is the difference between the linear approximation to the mechanical forcing and the true mechanical forcing large enough to invalidate the linear equations? Certainly this will vary from application to application.

One can estimate the extent to which linear theory will be valid by defining a critical mountain height as

$$\text{rms}[E_s(h)] = \zeta \text{rms}[L_s(h)]. \tag{21}$$

The critical height is defined as the orographic height at which the rms of the linear model error is equal to some fraction, ζ, of the linear forcing. The parameter, ζ, is somewhat arbitrary (but most likely between 0.1 and 0.5). Graphically, $\zeta = 1$ corresponds to the intersection point of the error and the linear forcing in Fig.

3. For this work we have chosen $\zeta = 0.33$, which is a reasonable choice, maybe even a little generous to linear theory. If another researcher feels that another choice of ζ is more appropriate, then all of our plots of h_{crit} scale linearly with ζ.

This analytical definition of h_{crit} is immune to any physical biases that one may have about the characteristics of the linear and nonlinear regimes. The strength of this formulation is demonstrated through the following example. By taking the linear equations (11) and (12) with no damping or Ekman pumping and with $[u]$ and $[\Theta_y]$ only a function of z, the solution has the form (Held 1995, personal communication)

$$\psi^* = \frac{[u]}{[u_z]}\left(\frac{\partial \psi^*}{\partial z} + \frac{N^2}{f}h\right). \tag{22}$$

The forcing error, E_s, for this case is identically zero since

$$E_s(x, y) = J\left(\psi^*, \frac{\partial \psi^*}{\partial z} + \frac{N^2}{f}h\right) = J\left(\psi^*, \frac{[u_z]}{[u]}\psi^*\right)$$
$$= \frac{[u_z]}{[u]}J(\psi^*, \psi^*) = 0, \tag{23}$$

and $h_{\text{crit}} \to \infty$. Thus, the linear solution is an exact solution regardless of the amplitude or shape of h, the vertical structure or amplitude of $[u]$, or the amplitude of the response.

One proposed criteria for determining if the linear equations are invalid has been the existence of closed isentropic surfaces at the lower boundary (see Valdes and Hoskins 1991; Cook and Held 1992). In the above example, closed isentropes could certainly exist yet the solution is still linear.

b. Critical height analysis

Given the importance of damping in the previous example, we begin by analyzing the effects of linear damping and Ekman pumping on the critical height. Figure 4 shows h_{crit} versus the Ekman pumping parameter, α, for the linear damping values of $\gamma = 0$ day^{-1}, $\gamma = 0.1$ day^{-1}, and $\gamma = 1$ day^{-1}. Consistent with the above analysis, when $\gamma = 0$ day^{-1} the critical height becomes very large as α goes to 0 because ψ^* and ψ_z^* are nearly in phase with the orography, thus producing an extremely small linear error.

As α increases, the amplitude of the response decreases while the phase of ψ^* and ψ_z^* shifts upstream of the orography. The phase differences between ψ^* and h and between ψ^* and ψ_z^* allow the error to become nonzero. Initially, the effect of eddy winds interacting with the orography due to the phase shift overwhelms the decreasing amplitude resulting in a reduction in h_{crit}. As α continues to increase above 400 m, the phase of the response does not shift while the amplitude continues to decrease. This results in a slight increase in the critical height. For very strong linear damping ($\gamma = 1$ day^{-1}), the phase of the response remains fixed at 90° upstream of the orography for all values of the Ekman pumping parameter. As α increases, the amplitude decreases and leads to a modest increase in the critical height. Critical heights for γ between 0 and 1 days^{-1} vary smoothly between these extremes. We choose $\gamma = 0.1$ day^{-1} and an Ekman pumping parameter of 150 m for the experiments described below.

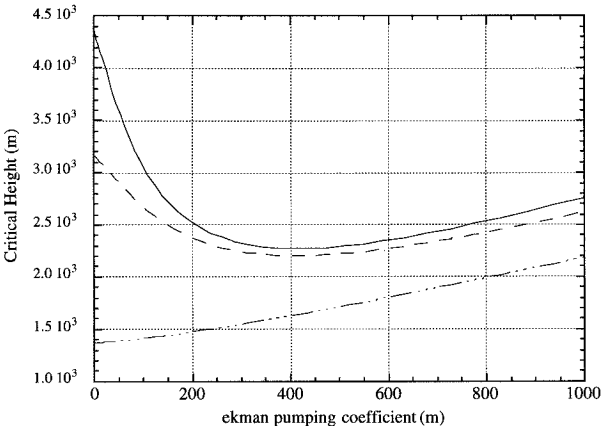

FIG. 4. The effect of dissipation on the mountain critical height. Critical height plotted against the Ekman pumping parameter for linear damping values of 0 day^{-1} shown by solid line, 0.1 day^{-1} shown by dashed line, and 1 day^{-1} shown by dashed–dotted line.

One of the most important factors controlling the amplitude and linearity of the response is the basic-state wind profile. Contours of h_{crit} are plotted for a wide range of surface winds and wind shears in Fig. 5. To generate this figure, meridionally uniform wind profiles were used with surface winds, u_s, ranging from 0.04 to 9 m s^{-1} and surface wind shears, du_s/dz, ranging from 0 m s^{-1}/H_s to 36 m s^{-1}/H_s. The orographic shape is circular with a 1500-km half-width.

In general, critical height increases monotonically with increasing surface wind shear. The critical height increases most dramatically at large values of surface wind shear but is relatively insensitive at smaller values of wind shear. The dashed line crudely separates the regions of parameter space that are sensitive to the surface wind from those that are not. A relevant feature of the basic state for determining whether zonal or meridional advection of potential temperature balances the mechanical forcing is the ratio of the surface wind to the surface wind shear, $H_u = u_s/(du_s/dz)$ (Held and Ting 1990). The dashed line in Fig. 5 denotes $H_u \approx 1$ km. Another relevant scale is the atmospheric scale height, H_s. For $H_u \approx H_s$ (to the right of the dashed line), the critical height is insensitive to u_s, while for $H_u \ll H_s$ (to the left of the dashed line) the critical height is sensitive to u_s.

When $H_u \approx H_s$ two regimes are identified. When $u_s < 1$ m s^{-1} the mechanical forcing is balanced by the Ekman pumping and linear damping. This portion of

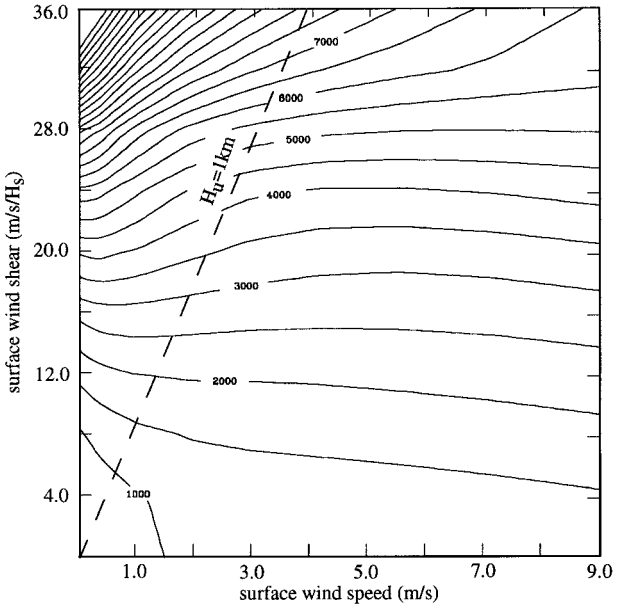

FIG. 5. Contours of critical height for a range of basic-state surface wind speeds and surface wind shears. The line $H_u \approx 1$ km separates the portion of the parameter space that is sensitive to the surface wind from the portion that is not.

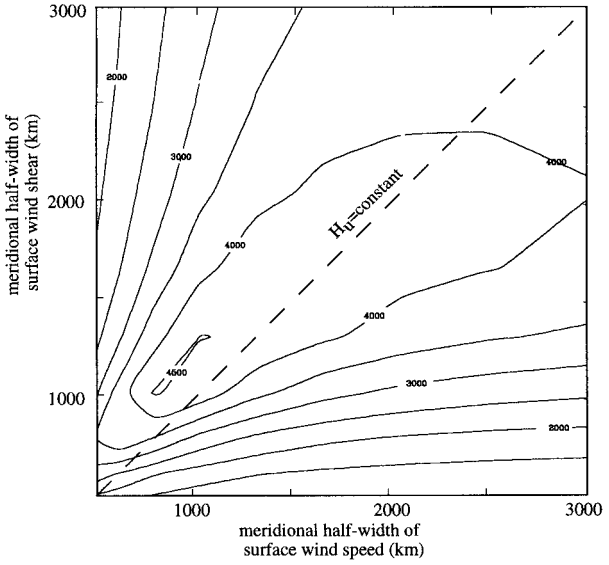

FIG. 6. The effect of the meridional profile of basic-state wind on the critical height. The half-widths of the surface wind and surface wind shear are the x and y axis, respectively. The maximum surface wind is 3 m s^{-1} and the maximum shear is 15 m s^{-1}/H. Along the line H_u = constant both the surface wind and surface wind shear have the same meridional shape.

the parameter space is quite small and is not of particular importance. When u_s increases above 1 m s^{-1} the mechanical forcing is balanced by the zonal advection of ψ_z^*. In this regime, the phase of the response remains fixed while ψ^* increases linearly with u_s. Since, in this case, the linear error (13b) is due mainly to the eddy wind interacting with the orography, the linear error grows linearly with u_s. The linear forcing (13c) also grows linearly with u_s, therefore h_{crit} is insensitive to u_s when $H_u \approx H_s$.

When $H_u \ll H_s$ the effects of Ekman pumping and linear damping are negligible and the forcing is balanced by the advection of the basic-state meridional temperature gradient with $\psi^* \approx [u_s/(du_s/dz)]h$ and in phase with the orography. As the surface wind increases, the amplitude of the response and the forcing increases linearly with u_s and the phase of the response shifts upstream. This change in phase causes the linear forcing error to increase faster than the linear forcing with increasing u_s resulting in a substantial decrease in critical height.

To analyze the effects of various meridional profiles, a maximum surface wind of 3 m s^{-1} and a maximum surface wind shear of 20 m s^{-1}/H_s were chosen to decay in the meridional direction as $-(y/w_{us})^2$ and $-(y/w_{uz})^2$, respectively, where w_{us} is the Guassian half-width of the surface wind profile and w_{uz} is the Guassian half-width of the shear. The half-widths, w_{us} and w_{uz}, measure the width of the surface westerlies and strong meridional temperature gradient, respectively. The meridional positions of the surface wind and shear maxima are coincident with the maximum orographic height.

Shown in Fig. 6 are the contours of the critical height for an array of w_{us} and w_{uz} values each ranging from 500 to 3000 km. When the meridional scales of the basic state are both much larger than the meridional scale of the orography, the wind profile varies insubstantially over the scale of the orography and h_{crit} approaches 4000 m. This is the same critical height obtained in Fig. 5 with $u_s = 3$ m s^{-1} and $du_s/dz = 20$ m s^{-1}/H_s. Choosing meridionally varying profiles cannot substantially increase the validity of linear theory since the critical heights in Fig. 6 are generally less than 4000 m. The maxima in h_{crit} lie along the $w_{us} = w_{uz}$ line. Along this line H_u is constant and the linear response is an exact solution (i.e., $h_{\text{crit}} \to \infty$) if no dissipation is present. When H_u varies on a meridional scale smaller than the orography, the manner in which the orographic forcing is balanced also varies meridionally. For example, if the surface wind is meridionally uniform ($w_{us} \to \infty$) and the surface wind shear decays in the merdional direction faster than the orography, two separate regions of linear balance develop. In the region of large shear the mechanical forcing is balanced by the meridional advection of potential temperature, while away from this region the forcing is balanced by zonal advection. Large errors in the linear prediction, which lead to small critical heights, develop due the incongruence of these two regions. Therefore, large meridional variations in H_u lead to substantial decreases in critical heights.

The importance of the tropopause strength in controlling the amplitude of the stationary wave response has recently been suggested (Lindzen 1994; Swanson and Pierrehumbert 1995). We analyzed the effects of

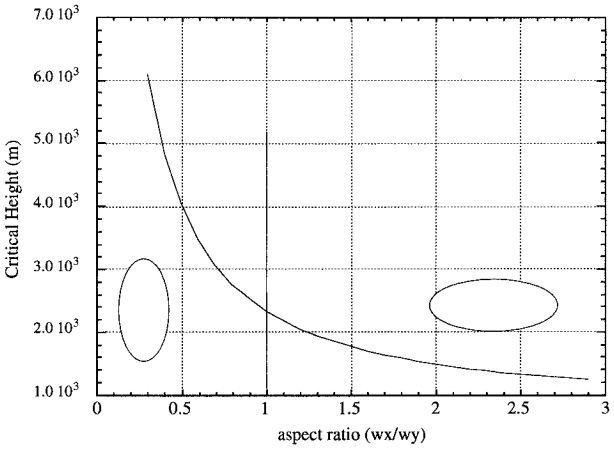

FIG. 7. Critical height plotted against the ratio of the orography's zonal half-width to meridional half-width. The mountain is circular when the ratio is 1.

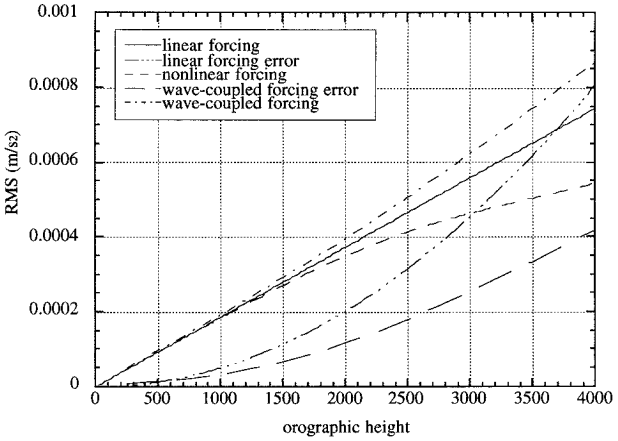

FIG. 8. An analysis of the accuracy of the wave-coupled boundary condition. The dark dash–dotted line denotes the wave-coupled mechanical forcing, and the dark dashed line represents the wave-coupled error. The three lighter lines are those shown in Fig. 3.

the tropopause strength by reducing the winds above the tropopause while holding the winds below the tropopause fixed. While increasing the upper-level potential vorticity gradient increases the stationary wave amplitude, it does not significantly alter the critical height in this dynamical framework.

In addition to the basic-state wind profile, the shape of the orography can modify the amplitude and linearity of the response. This is relevant to the earth's atmosphere since the major mountain ranges differ in shape. Figure 7 shows the critical height for a mountain of varying shape but fixed volume. The critical height is plotted against the ratio of the zonal mountain half-width to the meridional mountain half-width, w_x/w_y (19). The basic state is the same as shown in Fig. 2 and is meridionally uniform. Over a range of ratios from 0.3 to 3.0, h_{crit} decreases by a factor of 4 with the meridionally elongated mountains producing the largest critical heights. The amplitude of the surface response actually reduces by about 25% as the ratio is increased from 0.3 to 3.0. Also, the relative contribution of nonlinear temperature advection, $J(\psi^*, \psi_z^*)$, stays roughly constant at 35% of the total nonlinearity. When the mountain is zonally elongated, a decrease in h_{crit} results from the meridional wind interacting with the small meridional scale of the orography. The critical height varies in such a way that h_{crit}/w_y is approximately constant.

c. *Analysis of the hybrid wave–coupled lower boundary condition*

In order to incorporate the effects of perturbation winds interacting with orography, Chen and Trenberth (1988) incorporate a ''wave-coupled boundary condition'' into the linear balance equations on a sphere. In the quasigeostrophic framework, their wave-coupled lower boundary condition is expressed as

$$[u]\frac{\partial s^*}{\partial x} + v^*\frac{\partial [s]}{\partial y} + J\left(\psi^*, \frac{N^2}{f_o}h\right) = \frac{g}{f_o\Theta_0}Q^* - \alpha\frac{N^2}{g}\nabla^2\psi^* + D^*, \tag{25}$$

which differs from the standard linear equation (12) by the inclusion of the third term on the 1hs. The governing equation in the interior is unchanged from (11). By neglecting the nonlinearity in the interior and the nonlinear advection at the surface, the system of equations remains linear but the response is coupled to the orography at the lower boundary. Chen and Trenberth show that the eddy wind–orography interactions are often as large as the linear forcing in agreement with Saltzman and Irsch (1972).

Although the increased accuracy of the wave-coupled lower boundary condition often yields responses that are more accurate than those obtained using the classical lower boundary condition, by retaining only half of the second-order terms the boundary condition is inconsistent. Chen and Trenberth address this issue by showing that the nonlinear temperature advection at the surface (computed post facto) is 20%–50% of the retained second-order terms.

With the exact same winds and parameters as used in Fig. 3, the quasigeostrophic system [Eqs. (25) and (11)] is solved using the wave-coupled boundary condition. Plotted in Fig. 8 are $J([\psi] + \psi^*, (N^2/f_o)h)$ and $J(\psi^*, \psi_z^*)$, which are the wave-coupled mechanical forcing and error, respectively. Also shown in Fig. 8 are the results from Fig. 3 for comparison. The wave-coupled error is a factor of 2 smaller than the error from the classical linear boundary condition, but it still grows as $O(h^2)$. The resultant wave-coupled forcing is larger than the actual nonlinear forcing by about 60% at 4 km and leads to an overestimate in the response of about 30% when compared to the nonlinear response.

5. Nonlinear response to mechanical forcing

A critical height definition derived from linear theory may be able to predict when linear theory is invalid, but it gives no information about the structure of the nonlinear response. The fully nonlinear equations must be solved to ascertain that information. In this section we analyze the structure of the nonlinear response, compare it to the structure of the linear response, and explain the differences.

a. Response at the surface

The linear and nonlinear stationary wave solutions are generated in response to a 3-km mountain using C24 resolution. The wind profile is meridionally uniform with a surface wind of 3 m s^{-1} and surface wind shear of 15 m s^{-1}/H_s with a vertical profile similar in shape to that shown in Fig. 2. With this basic-state wind the critical height is approximately 3 km (Fig. 5), so a nonlinear response is expected.

Figures 9a and 9b show the linear and nonlinear eddy streamfunction responses at the surface, respectively. While the nonlinear response is slightly weaker than the linear response, the main difference is the structure. The axis connecting the high and low in the nonlinear solution is rotated 45° clockwise from the linear solution. The linear response is characterized by an east/west oriented dipole while the fully nonlinear response shows a southeast/northwest oriented dipole.

Figure 10a shows the full mechanical forcing ($d\Theta/dz$ $\mathbf{v}\cdot\nabla h$ as obtained from the nonlinear equations), which produces the response shown in Fig. 9b. The nonlinear forcing is rotated clockwise with rising motion more localized than the sinking motion. To attempt to explain this rotation and shed light on why the linear equations fail, each component of the nonlinear forcing is estimated using the linear response. Figures 10b and 10c show the traditional linear mechanical forcing ($d\Theta/dz$ $[u]\partial h/\partial x$) and the eddy-wind mechanical forcing ($d\Theta/dz$ $\mathbf{v}^*\cdot\nabla h$), and Fig. 10d shows the nonlinear temperature advection ($\mathbf{v}^*\cdot\nabla\Theta^*$). In Figs. 10b–d, the eddy terms, $\mathbf{v}^*$ and Θ^*, are obtained using the linear quasigeostrophic equations (11) and (12). A superposition of all three of these components results in a total forcing that is rotated clockwise with respect to the linear forcing alone (Fig. 10b) and is similar to the nonlinear forcing (Fig. 10a). The eddy wind–orography interaction is the dominant factor causing the rotation. Since the nonlinear temperature advection is out of phase with the eddy wind–orography interaction, the advection partially negates the tendency of the eddy wind–orography interaction to rotate the forcing and response.

It may be expected that the error involved in the wave-coupled lower boundary condition would look similar to Fig. 10d since this boundary condition accounts for the eddy wind–orography interaction but neglects the nonlinear potential temperature advection. The wave-coupled lower boundary condition includes forcing not present in the traditional lower boundary condition, so the wave-coupled response does not closely resemble the classical linear response. As a result, the wave-coupled error does not closely resemble the nonlinear potential advection error shown in Fig. 10d. In fact, the wave-coupled error is significantly larger than that ''predicted'' by Fig. 10d.

To further explain the nonlinear response, Fig. 11a shows contours of the entropy S (in units of K) and the total wind field, while Fig. 11b shows the total streamfunction. Closed isentropes of high entropy form on the equatorward mountain slope and thereby block the in-

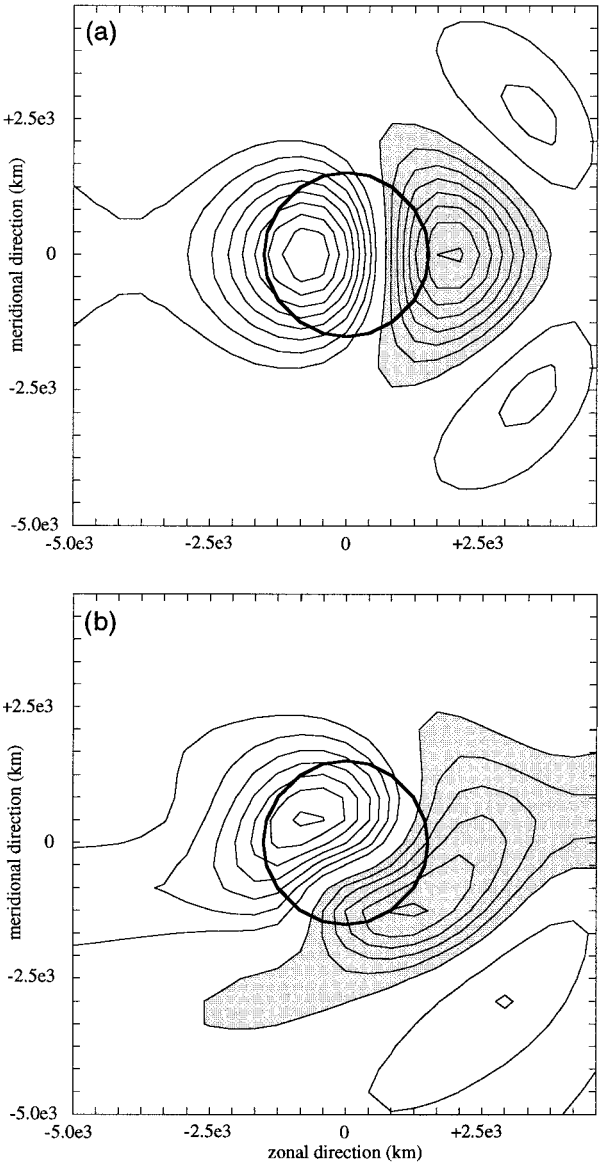

FIG. 9. The (a) linear and (b) nonlinear response to mechanical forcing at the surface in terms of eddy streamfunction. Half-width of the orography shown by dark circle. Contour interval is 5.0×10^5 m^2 s^{-1} with the zero contour not shown and negative values shaded.

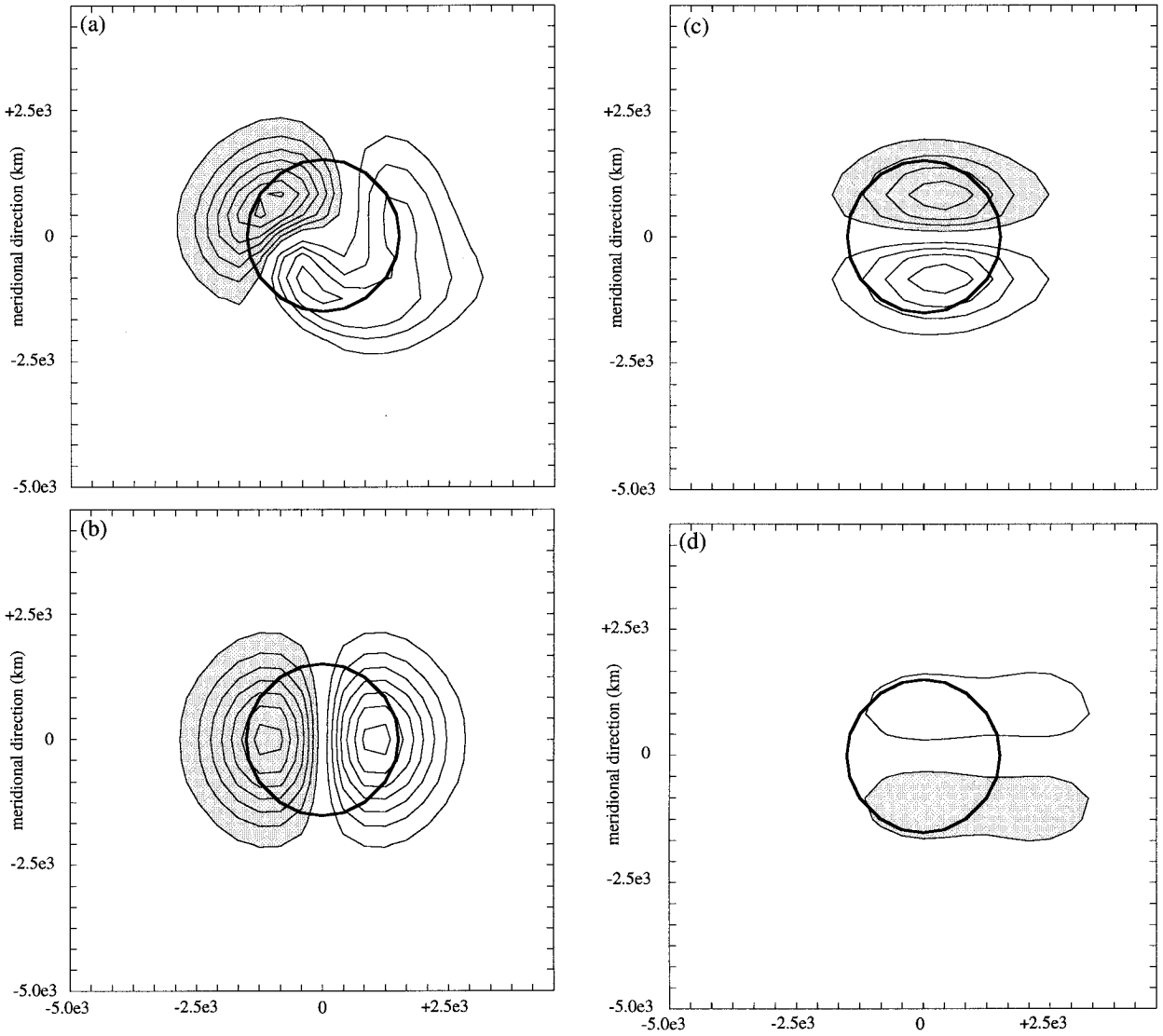

FIG. 10. (a) The mechanical forcing as obtained using the full nonlinear equations. (b) As in (a) but for the linear representation of that forcing. (c) The error produced as a result of linear theory neglecting the eddy wind/orography interaction. (d) The error produced by neglecting the nonlinear advection. Contour interval is 0.2 K day^{-1} with the zero contour not shown and negative values shaded.

cident flow. The isentropes cut off on the equatorward slope because the meridional derivative of $(\partial\Theta/\partial z)h$ is larger than the meridional derivative of the basic-state potential temperature at that location [see (10)]. Strong flow occurs over the poleward half of the orography. On the equatorward half of the mountain, the flow is directed toward the equator with near-zero zonal flow.

Without damping, the flow would be isentropic and lines of constant S would be streamlines. Ekman pumping acts as a source (sink) of S where the flow is anticyclonic (cyclonic) allowing cross isentropic flow toward higher (lower) entropy levels. Linear damping also generates cross-isentropic flow with air parcels in cold (warm) anomalies moving into higher (lower) entropy areas. Thus, dissipation processes are important in determining the degree of nonlinearity.

This analysis is generally consistent with that of Valdes and Hoskins (1991) and Cook and Held (1992). One difference is that Valdes and Hoskins describe the nonlinear response as a split flow (i.e., a doubly connected flow) with parcels near the surface moving poleward or equatorward around the mountain. Figures 11a and 11b suggest a slightly different interpretation, and the flow can be described as a large amplitude tilted wave rather than a split flow. In cases where nonlinearity is even more prominent, the wave continues to ''fold,'' but dissipation seems to inhibit the development of a cutoff circulation when only mechanical forcing is present.

b. Upper-level response

A convenient means of describing the mechanical forcing is as a dipole located at the surface where one

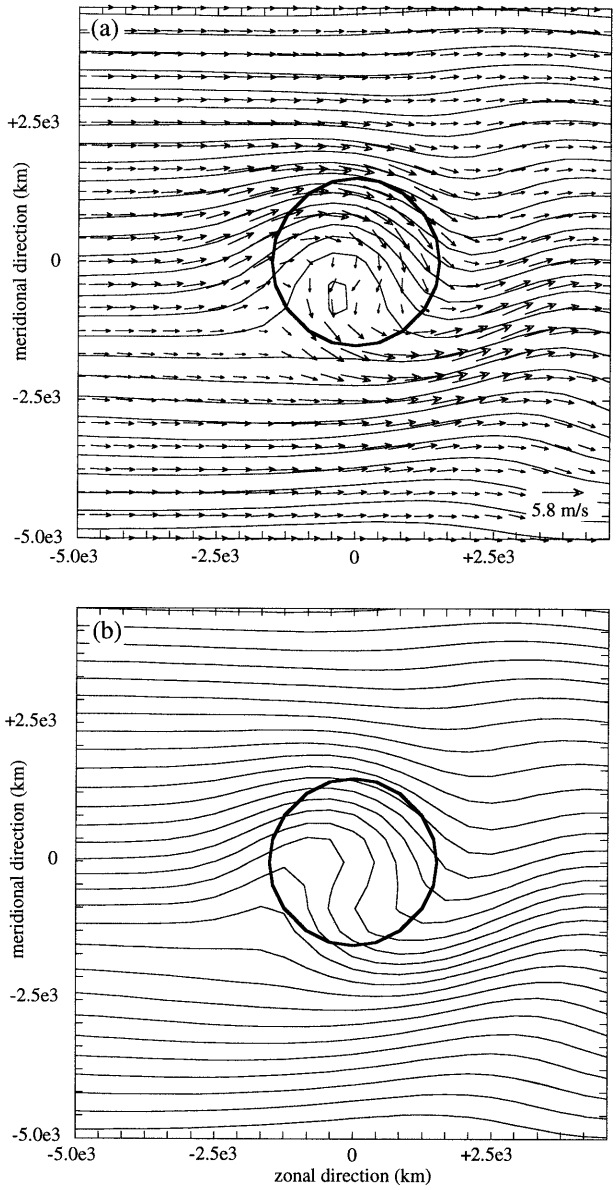

FIG. 11. The nonlinear response at the surface. (a) The surface entropy, S, along with the total wind vectors. (b) The streamlines of the flow. Contour interval is 3.0 K in (a) and 1.0×10^6 m^2 s^{-1} with the zero contour not shown and negative values shaded. The wind vector magnitude is shown in the lower corner of (a).

lobe represents rising motion and the other represents sinking motion (see Figs. 10a,b). To test the sensitivity of the upper-level response to the orientation of the dipole, a set of linear model experiments was conducted in which a forcing dipole is rotated in 5 deg increments from an initial east–west orientation (as shown in Fig. 10b) to a north–south orientation. The results from these experiments (not shown) indicate that for this basic state the wave train will rotate with the forcing up to about 50 degrees rotation. At rotation angles above 50 degrees, the stationary wave response consists of two wave trains, one propagating equatorward and one propagating poleward.

This generation of two wave trains is consistent with the theory of Rossby wave propagation. The rotation of the response with the forcing must cease before the dipole is oriented north–south because Rossby-like waves cannot propagate parallel to the potential vorticity gradient. When the forcing is a north–south oriented dipole, the $y = 0$ line is the axis of asymmetry. Since all other relevant quantities are zonally symmetric, the stationary wave response is asymmetric about the y axis taking the form of two wave trains: one propagating poleward, the other equatorward. The north–south dipole forcing was discussed by Valdes and Hoskins (1991) and our numerical findings are in agreement except that the stationary wave energy is not reduced when the dipole is rotated as they hypothesized.

The nonlinear model results indicate that regardless of the orographic height, the mechanical forcing rotates clockwise only to about 50 to 60 degrees and no farther. Therefore, the upper-level stationary wave response generally has the form of a single, equatorward propagating wave train. The characteristics of the upper-level wave train can be easily modified by altering the meridional structure of the basic state.

Figures 12a and 12b show the eddy streamfunction and horizontal components of a 3D wave flux vector (Plumb 1985) at the $z = 1.2H_s$ level from the linear and nonlinear models, respectively. For the linear system, the meridional potential vorticity gradient and the mechanical forcing are symmetric about $y = 0$, so the linear response is meridionally symmetric with wave energy dispersing equally toward pole and equator. The nonlinear response shows a meridional bias with the majority of the wave activity propagating equatorward. This is consistent rotation of dipole forcing discussed above. The nonlinear response is weaker than the linear response by about 40% when comparing momentum fluxes and 10% when comparing the eddy streamfunction amplitude.

Modification of the interior flow also contributes to the change in propagation direction. In the nonlinear case the waves are not propagating on a zonally uniform potential vorticity gradient, but rather on an atmospheric flow which has zonal variations in the potential vorticity gradient. The stationary wave number, $K_{\text{stat}} = (q_y/u)^{0.5}$, acts as the refractive index with waves curving toward higher values of K_{stat} (Hoskins and Ambrizzi 1993). While the initial orientation of the stationary wave train is due mainly to the orientation of the forcing, the curving of the wave train along its path is due to variations in the refractive index.

Nonetheless, the effects of nonlinearity are mainly felt at the surface. Figure 13 shows the rms of the nonlinear advection terms (13a) plotted against orographic height. While the nonlinear terms reach approximately 25% of the linear terms, there is a separation; the two linear terms balance each other and the two nonlinear

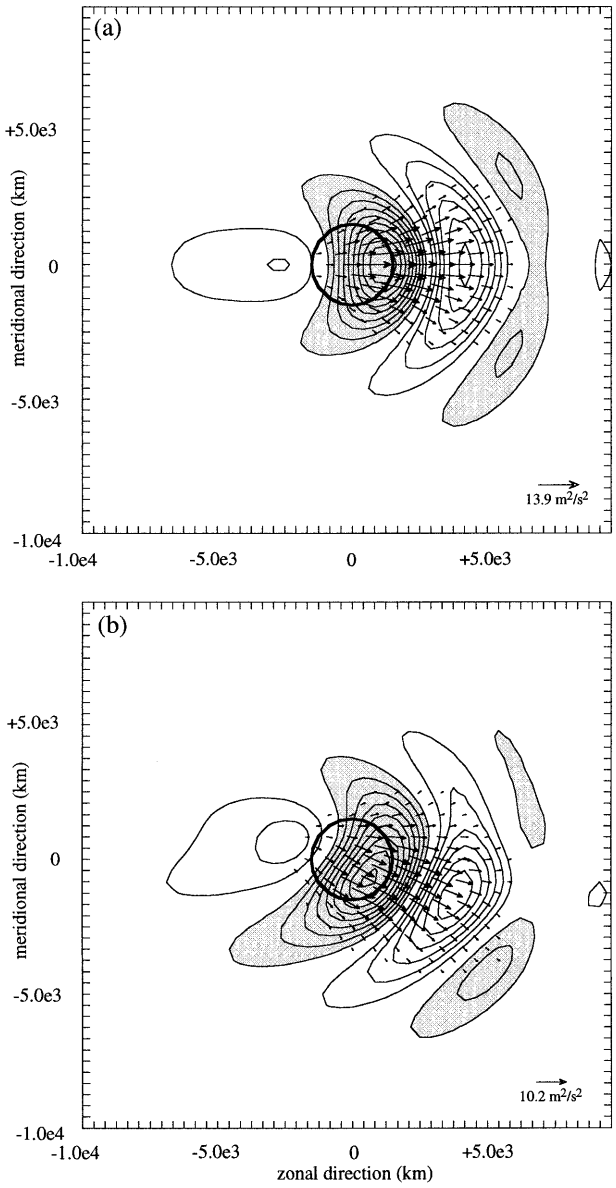

FIG. 12. The (a) linear upper-level response and (b) the nonlinear upper-level response shown in terms of eddy streamfunction (contours) and horizontal wave fluxes (vectors). Contour interval is 1.0 $\times$ 10^6 m^2 s^{-1} with the zero contour not shown and negative values shaded. The vector magnitudes are shown in the lower corner of each figure part.

terms balance each other. This can be understood through Rossby wave propagation. The Rossby wave phase relation of the eddy wind field is such that the geostrophic portion of the eddy wind field is nondivergent (the quasigeostrophic equations use only the geostrophic flow). Therefore, if the wave energy is nearly conserved (no wave breaking, diabatic heating, or large dissipation), $\nabla \cdot \mathbf{v}^* q^* \approx 0$ leads to $\mathbf{v}^* \cdot \nabla q^* \approx 0$ and the nonlinear terms identically cancel.

The nonlinear results presented here are typical of the nonlinear response to mechanical forcing. While changing the basic-state winds or dissipation parameters changes the extent to which nonlinearity is important, we find that the structure of the nonlinear response is similar regardless of the winds or parameters.

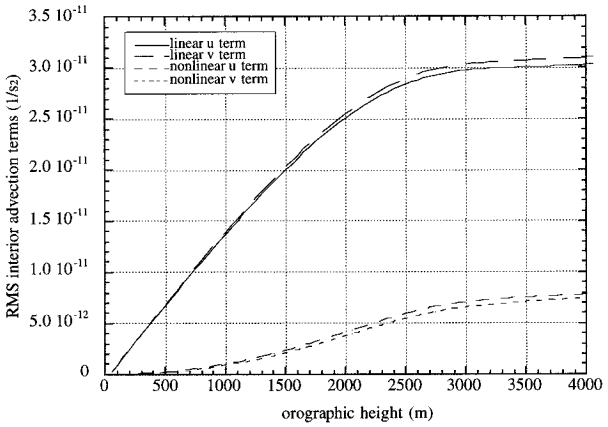

FIG. 13. Magnitude of the linear advection terms in the interior (upper lines) and the magnitude of the nonlinear terms in the interior (lower lines) plotted against orographic height.

6. Factors controlling the amplitude of nonlinear response

a. Effects of surface wind strength

Figures 14a and 14b show the rms of the eddy streamfunction versus orographic height at the surface and at $z = 1.2H_s$, respectively, for different surface wind speeds. The surface wind is varied from 1 to 7 m s^{-1} in 1.5 m s^{-1} increments with a 10 m s^{-1}/H_s surface wind shear. For orographic heights below about 1 km, the amplitude of the response is linear with height for all surface wind speeds and increasing the surface wind speed increases the amplitude of the response. These relations hold for larger orographic heights, but the amplitude of the nonlinear response ''saturates'' and becomes less sensitive to changes in orographic height.

The circles indicate the amplitude and orographic height at which the rms of the nonlinear response deviates from the rms of the linear prediction by a factor of 1/3. When the surface wind speed is increased, not only does the amplitude of the response increase, but linear theory remains valid for larger orography. For surface wind speeds between 1.0 to 2.5 m s^{-1}, the linear response deviates from the nonlinear response by one-third at orographic heights around 2 km. At larger surface wind speeds of 5.0 to 7.0 m s^{-1}, the linear prediction overestimates the response by one-third for orographic heights of 3 km. The linear theory based critical height defined in (23) captures this trend by predicting an increase of h_{crit} of 1/2 km between $u_s = 1$ m s^{-1} and $u_s = 7$ m s^{-1} with $du_s/dz = 10$ m s^{-1}/H_s (see Fig. 5).

The amplitude of the stationary response at upper levels is larger than the response at the surface because

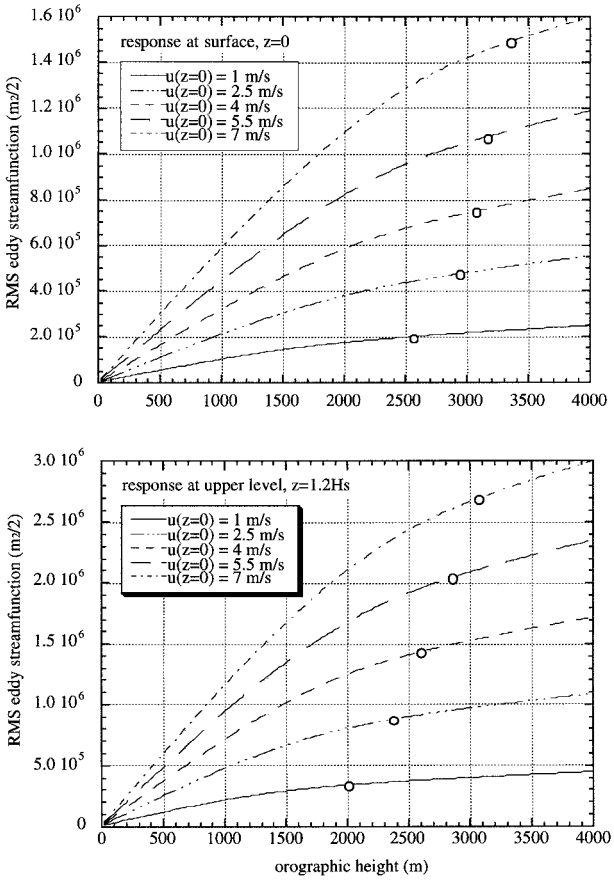

FIG. 14. Amplitude of the nonlinear response (a) at the surface and (b) at $z = 1.2H_s$, vs orographic height. The wind shear is constant at 10 m s^{-1} H_s and each line denotes a different surface wind speed.

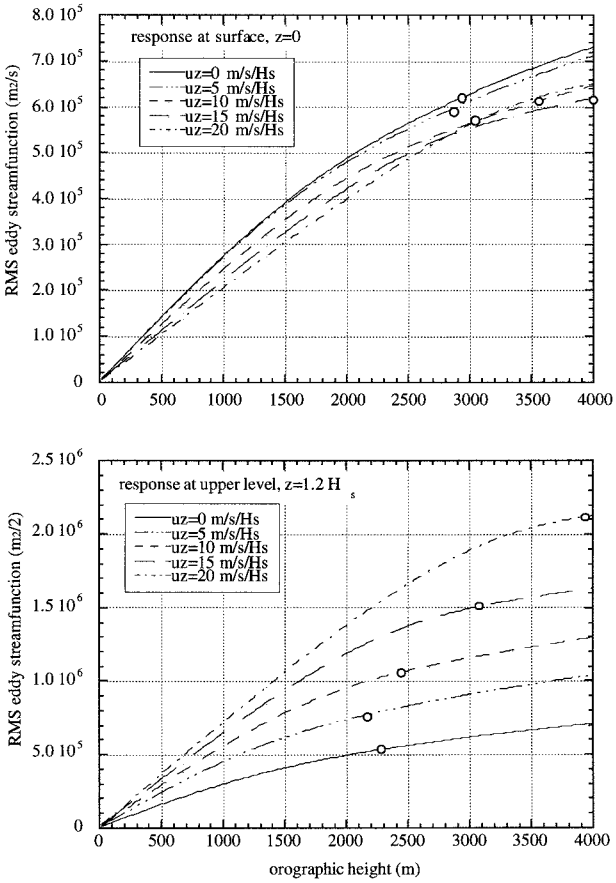

FIG. 15. Amplitude of the nonlinear response (a) at the surface and (b) at $z = 1.2H_s$ vs orographic height. The surface wind is constant at 3 m s^{-1} and each line denotes a different surface wind shear.

the external mode dominates the response and the vertical structure of that mode depends on the vertical wind profile (Held et al. 1985). Since the zonal wind speed increases vertically, the amplitude of the external mode increases with height. The surface and the upper-level responses have the same characteristics in the transition from the linear to the nonlinear regime.

Valdes and Hoskins (1991), using a linear primitive equation model, found an increase in the amplitude of the response with surface wind speed, in agreement with our results. However, in their nonlinear primitive equation model, they found that the upper-level response is relatively insensitive to the magnitude of the surface wind while our results show that the amplitude increases with increasing surface wind strength. This discrepancy is most likely due to the difference in how the lower boundary condition is applied (as discussed in section 2). The quasigeostrophic lower boundary condition is implemented on the $z = 0$ surface (2) and the surface wind represents in some sense the vertically averaged wind that interacts with the orography. In the primitive equation system, the lower boundary condition is implemented on the $z = h$ surface, so the strength of the wind that interacts with the orography is both a function of the wind on the $z = 0$ surface and the zonal wind shear. Valdes and Hoskins analyze the nonlinear response using a wintertime basic-state wind field and alter the surface wind by multiplying by factors from 0 to 1.5. Since the shear was strong (maximum of $\sim$35 m s$^{-1}/H_s$) and the maximum surface wind was about 4 m s^{-1}, the changes in the surface wind made relatively small changes to the wind which interacted with the orography resulting in small changes in the response.

b. Effects of zonal wind shear strength

Figures 15a and 15b show how the amplitudes of the surface and upper-level responses depend on wind shear. The surface wind is held fixed at 3 m s^{-1} and the shear is varied from 0 m s$^{-1}/H_s$ to 20 m s$^{-1}/H_s$ in increments of 5 m s$^{-1}/H_s$. A linear regime at smaller orographic heights is evident. Within this linear regime, dramatic increases in the shear reduce the amplitude of the response at the surface only slightly (Fig. 15a). Since the amplitude of the external mode depends on the upper-

level winds, the amplitude of the upper-level response is more sensitive to changes in the wind shear.

The circles on each figure denote the orographic height at which the linear response deviates from the nonlinear response by one-third. For small values of zonal wind shear (0 and 5 m s^{-1}/H_s), increasing the shear has little effect on the critical height, in agreement with Fig. 5. For larger values of wind shear (10–20 m s^{-1}/H_s), an increase in shear substantially increases the maximum orographic height at which linear theory is valid. This too is consistent with Fig. 5, which shows an increase in h_{crit} of about 1500 m from $du_s/dz = 10$ m s^{-1}/H_s to $du_s/dz = 20$ m s^{-1}/H_s.

These results show that the amplitude of the response is not predictive of the validity of linear theory. In fact, the largest critical heights are associated with the largest responses (Fig. 14a). At the surface the phase of the response is actually more important than the amplitude in determining the degree of linearity. The amplitude of the upper-level response is determined by both the amplitude of the surface response and the vertical structure of the external mode. Therefore, large amplitude upper-level responses may be due to large upper-level zonal winds, and not associated with large surface forcing.

7. Conclusions

A stationary wave model has been developed to investigate how nonlinearity effects the atmosphere's response to large-scale, mechanical forcing by orography. The model is fully nonlinear, so nonlinear terms (advective) that are potentially important in the interior and at the surface are included as well as ''nonlinearity'' in the lower-boundary forcing that develops when eddy winds interact with orography. The model is used in linear and nonlinear modes to investigate the validity of approximations to mechanical forcing of stationary waves by orography, the dependence of the boundary between linear and nonlinear response regimes on various parameters, and the dynamics of the nonlinear response. Forcing of stationary waves by diabatic heating associated with orography is not considered here, but it is addressed in a companion paper (Ringler and Cook 1997, manuscript submitted to *J. Atmos. Sci.*).

The model is based on quasigeostrophic dynamics on the midlatitude β plane. This system retains much of the pertinent nonlinearity while remaining relatively simple to understand in depth. When the model is used in its nonlinear mode, the governing equations are iterated until a balanced nonlinear solution is obtained. Since the model retains up to 24 waves in both the zonal and meridional directions, the planetary-scale stationary waves are well resolved. The model domain is sufficiently large to eliminate contamination of the response by boundary effects and resonance.

A critical height separating linear and nonlinear response regimes is defined as the mountain height at which the magnitude of the terms that are discarded during linearization are comparable to the terms that are retained. We find that a number of factors, not just the mountain height, control whether the response is linear or nonlinear. A stronger meridional temperature gradient increases the critical height; for example, with a surface wind of 4.0 m s^{-1} the critical height increases from 1 km to 6 km when the zonal wind shear increases from 0.0 m s^{-1}/H_s to 28.0 m s^{-1}/H_s. Changes in the surface wind speed have opposite effects, depending upon the shear. When a weak (strong) vertical shear is present, increasing the surface wind increases (decreases) the critical height. When the wind shear is 24.0 m s^{-1}/H_s, increasing the surface wind from near zero to 9.0 m s^{-1} decreases the critical height from 6 km to 4 km.

Our results indicate that, in order to evaluate the degree to which the response to a given orography in a given basic state will be captured accurately by linear dynamics, the most important factor to consider is the relative phase of the surface response and the orography. Thus, the amplitude of the response and the extent to which linear theory is valid are not well correlated in general.

The stationary wave model employs linear damping at the surface, and we find that this form of dissipation plays a crucial role in determining the critical height. For small values of linear damping, large critical heights are obtained for Ekman pumping parameter values near zero or larger than 1000 m. A minimum in the critical height is obtained for an Ekman pumping value of 400 m. The location of the minima (Fig. 4) depends on the basic-state winds. Since linear stationary wave models often employ large damping at the surface to control the amplitude of resonant responses, the amplitude and phasing of the surface response may be distorted.

In addition to comparing the linear and fully nonlinear boundary conditions, comparisons are made to the hybrid wave-coupled lower boundary condition developed by Chen and Trenberth (1988). Our analysis suggests that the wave-coupled boundary condition reduces the error between a factor of 2 to 3 over the linear boundary condition. However, since the wave-coupled boundary condition accounts for interactions between the response and the orography while neglecting nonlinear advection, the accuracy of the wave-coupled boundary condition is variable, depending on the relative importance of eddy wind–orography interaction to nonlinear advection.

The surface response to orographic forcing in the nonlinear regime is characterized by a high and low located to the northwest and southeast of the mountain center, respectively (with the Coriolis parameter > 0). (The linear response consists of a high due west of the elevation maximum and a low to the east.) The different orientation of the nonlinear response with respect to the linear response is due to eddy wind interactions with orography. However, nonlinear advection of potential temperature also plays an important role and weakens the reorientation. For larger mountain heights, the amplitude of both the nonlinear mechanical forcing and the

response saturate, increasing much less sharply than linear theory predicts.

Observed seasonal variations in surface wind speed and meridional temperature gradients are more than sufficient to produce large changes in the amplitude and structure of the mechanically forced response to mountains of the size of the Rockies and Himalayas. Compared to the wintertime response, the summertime response should be more nonlinear and weaker at upper levels due to the smaller meridional temperature gradient. The magnitude of the upper-level response is also sensitive to the wind shear, so a portion of the year-to-year variability in stationary wave amplitude and structure could be due to variations in the upper-level winds.

In the context of mechanically forced stationary waves, nonlinearity is most important at the surface in the vicinity of the orography. Including nonlinearity in the interior provides a zonally asymmetric wind that can alter propagation characteristics. However, while the amplitudes of the advective terms are comparable to the linear terms in the interior, the nonlinear terms tend to cancel.

We have explored the dependence of the boundary between the linear and nonlinear responses to mechanical orographic forcing on how the forcing is represented and on the basic-state characteristics. Excluded here is any consideration of the role of the time-dependent flow in determining the stationary response. We have also neglected the effects of the presence of mountains on the diabatic heating field; this should be considered as an integral part of the atmosphere's response to orography, and these thermal effects are considered within this modeling framework in Ringler and Cook (1997, manuscript submitted to *J. Atmos. Sci.*).

Acknowledgments. We would like to thank Drs. S. Colucci, P. Gierasch, and L. Trefethen for reading an earlier draft of this paper. Also, we appreciated the helpful discussions with Dr. I. Held and the comments by two reviewers. Some of the results presented were completed using computing resources at the Cornell Theory Center. This paper is a part of TDR's dissertation and was supported by the NASA Graduate Student Researchers Program Grant NGT-51147.

APPENDIX A

Partial Differential Equation Operators

In order to expand the governing equations given in (4) and (6) about an asymmetric flow, define

$$\psi^{i+1}(x, y, z) \equiv \psi^{i}(x, y, z) + \delta(x, y, z), \quad \text{(A1)}$$

where it is formally assumed that

$$|\delta|/|\psi| \ll 1. \quad \text{(A2)}$$

In (A1), the total solution at the iteration $i + 1$ is composed of a known flow field found from the previous iteration (superscript i) and a correction to the flow field, δ. The procedure is to substitute the expansion (A1) into (4) and (6) and truncate all terms $O(\delta^2)$. All terms which remain are then $O(\delta^0)$ or $O(\delta^1)$. The terms which are $O(\delta^1)$ are gathered on the lhs to form the operator, H, and all terms which are $O(\delta^0)$ are gathered on the rhs to form the forcing vector, F, written as

$$H\delta = F. \quad \text{(A3)}$$

Let $\epsilon = f_o^2/N^2$, then the operator, H, has the form

$$\begin{aligned}
&H(0 < z < z_{\text{top}}) \\
&= +\frac{\partial\psi^i}{\partial x}\left[\frac{\partial^3}{\partial x^2\partial y} + \frac{\partial^3}{\partial y^3} + \left(\frac{\partial\varepsilon}{\partial z} - \frac{\varepsilon}{H_s}\right)\frac{\partial^2}{\partial y\partial z} + \varepsilon\frac{\partial^3}{\partial y\partial z^2}\right] - \frac{\partial\psi^i}{\partial y}\left[\frac{\partial^3}{\partial x^3} + \frac{\partial^3}{\partial x\partial y^2} + \left(\frac{\partial\varepsilon}{\partial z} - \frac{\varepsilon}{H_s}\right)\frac{\partial^2}{\partial x\partial z} + \varepsilon\frac{\partial^3}{\partial x\partial z^2}\right] \\
&\quad + \left[\beta + \frac{\partial^3\psi^i}{\partial x^2\partial y} + \frac{\partial^3\psi^i}{\partial y^3} + \left(\frac{\partial\varepsilon}{\partial z} - \frac{\varepsilon}{H_s}\right)\frac{\partial^2\psi^i}{\partial y\partial z} + \varepsilon\frac{\partial^3\psi^i}{\partial y\partial z^2}\right]\frac{\partial}{\partial x} + \left[\frac{\partial^3\psi^i}{\partial x^3} + \frac{\partial^3\psi^i}{\partial x\partial y^{23}} + \left(\frac{\partial\varepsilon}{\partial z} - \frac{\varepsilon}{H_s}\right)\frac{\partial^2\psi^i}{\partial x\partial z} + \varepsilon\frac{\partial^3\psi^i}{\partial x\partial z^2}\right]\frac{\partial}{\partial y} \\
&\quad + \mu\left(\frac{\partial^4}{\partial x^4} + 2\frac{\partial^4}{\partial x^2\partial y^2} + \frac{\partial^4}{\partial y^4}\right) + \gamma\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \left(\frac{\partial\varepsilon}{\partial z} - \frac{\varepsilon}{H_s}\right)\frac{\partial}{\partial z} + \varepsilon\frac{\partial^2}{\partial z^2}\right)
\end{aligned}$$

and

$$\begin{aligned}
H(z = 0) = &+\frac{\partial\psi^i}{\partial x}\frac{\partial^2}{\partial y\partial z} - \frac{\partial\psi^i}{\partial y}\frac{\partial^2}{\partial x\partial z} + \frac{\partial}{\partial y}\left(\frac{\partial\psi^i}{\partial z} + \frac{N^2}{f_o}h\right)\frac{\partial}{\partial x} - \frac{\partial}{\partial x}\left(\frac{\partial\psi^i}{\partial z} + \frac{N^2}{f_o}h\right)\frac{\partial}{\partial y} + \frac{\alpha N^2}{f_o}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) \\
&+ \mu\left(\frac{\partial^4}{\partial x^4} + \frac{\partial^4}{\partial x^2\partial y^2} + \frac{\partial^4}{\partial y^4}\right) + \gamma\frac{\partial}{\partial z},
\end{aligned}$$

The forcing vector, F, contains all $O(\delta^o)$ terms and is given as

$$F(0 < z < z_{\text{top}})$$

$$= +\frac{\partial \psi^i}{\partial x}\left[\beta + \frac{\partial^3 \psi^i}{\partial x^2 \partial y} + \frac{\partial^3 \psi^i}{\partial y^3} + \left(\frac{\partial \varepsilon}{\partial z} - \frac{\varepsilon}{H_s}\right)\frac{\partial^2 \psi^i}{\partial y \partial z} + \varepsilon \frac{\partial^3 \psi^i}{\partial y \partial z^2}\right] - \frac{\partial \psi^i}{\partial y}\left[\frac{\partial^3 \psi^i}{\partial x^3} + \frac{\partial^3 \psi^i}{\partial x \partial y^2} + \left(\frac{\partial \varepsilon}{\partial z} - \frac{\varepsilon}{H_s}\right)\frac{\partial^2 \psi^i}{\partial x \partial z} + \varepsilon \frac{\partial^3 \psi^i}{\partial x \partial z^2}\right]$$

$$- \mu\left(\frac{\partial^4 \psi^i}{\partial x^4} + 2\frac{\partial^4 \psi^i}{\partial x^2 \partial y^2} + \frac{\partial^4 \psi^i}{\partial y^4}\right) + \frac{g f_o}{\rho \Theta}\frac{\partial}{\partial z}\left(\frac{\rho Q}{N^2}\right) - \gamma\left(\frac{\partial^2 \psi^i}{\partial x^2} + \frac{\partial^2 \psi^i}{\partial y^2} + \left(\frac{\partial \varepsilon}{\partial z} - \frac{\varepsilon}{H_s}\right)\frac{\partial \psi^i}{\partial z} + \varepsilon \frac{\partial^2 \psi^i}{\partial z^2}\right)$$

and

$$H(z = 0) = -\frac{\partial \psi^i}{\partial x}\frac{\partial^2 \psi^i}{\partial y \partial z} + \frac{\partial \psi^i}{\partial y}\frac{\partial^2 \psi^i}{\partial x \partial z} - \frac{\partial \psi^i}{\partial x}\frac{\partial}{\partial y}\left(\frac{\partial \psi^i}{\partial z} + \frac{N^2}{f_o}h\right) + \frac{\partial \psi^i}{\partial y}\frac{\partial}{\partial x}\left(\frac{\partial \psi^i}{\partial z} + \frac{N^2}{f_o}h\right)$$

$$- \frac{\alpha N^2}{f_o}\left(\frac{\partial^2 \psi^i}{\partial x^2} + \frac{\partial^2 \psi^i}{\partial y^2}\right) - \mu\left(\frac{\partial^4 \psi^i}{\partial x^4} + \frac{\partial^4 \psi^i}{\partial x^2 \partial y^2} + \frac{\partial^4 \psi^i}{\partial y^4}\right) - \gamma\frac{\partial \psi^i}{\partial z}.$$

The analytical form of these operators is then discretized via a Fourier series expansion in the x and y. The equations are finite differenced in the vertical. The resulting matrix operator, $\mathbf{H}$, is a block tridiagonal and is solved by an L–U decomposition (Golub and van Loan 1993).

REFERENCES

Ashe, S., 1979: A nonlinear model of the time-averaged axially asymmetric flow induced by topography and diabatic heating. *J. Atmos. Sci.*, **36,** 109–126.

Chen, S. C., and K. E. Trenberth, 1988: Orographically forced planetary waves in the Northern Hemisphere winter: Steady state model with wave coupled lower boundary formulation. *J. Atmos. Sci.*, **45,** 657–681.

Cook, K. H., and I. M. Held, 1988: Stationary waves of the Ice Age climate. *J. Climate*, **1,** 807–819.

——, and ——, 1992: The time-mean response of the atmosphere to large-scale atmospheric flow. *J. Atmos. Sci.*, **49,** 525–539.

da Silva A. M., 1989: Comments on ''Orographically forced planetary waves in the Northern Hemisphere winter: Steady state model with wave-coupled lower boundary formulation.'' *J. Atmos. Sci.*, **46,** 2101–2103.

Golub, G. H., and G. F. van Loan, 1993: *Matrix Computations*. The Johns Hopkins University Press, 170–171.

Haupt, S. E., J. C. McWilliams, and J. T. Tribbia, 1993: Modons in shear flow. *J. Atmos. Sci.*, **50,** 1181–1198.

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

——, and M. Ting, 1990: Orographic versus thermal forcing of stationary waves: The importance of the mean low-level wind. *J. Atmos. Sci.*, **47,** 495–500.

——, R. L. Panetta, and R. T. Pierrehumbert, 1985: Stationary external Rossby waves in vertical shear. *J. Atmos. Sci.*, **42,** 865–883.

Hoskins, B. J., and T. Ambrizzi, 1993: Rossby wave propagation on a realistic longitudinally varying flow. *J. Atmos. Sci.*, **50,** 1661–1671.

Lenters, J. D., K. H. Cook, and T. D. Ringler, 1995: Comments on ''On the influence of the Andes on the general circulation of the Southern Hemispere.'' *J. Climate*, **8,** 2113–2115.

Lindzen, R. S., 1994: The Eady problem for a basic state with zero pv gradient but $\beta \neq 0$. *J. Atmos. Sci.*, **51,** 3221–3226.

Lorenz, E. N., 1960: Energy and numerical weather prediction. *Tellus*, **12,** 364–373.

Plumb, R. A., 1985: On the three-dimensional propagation of stationary waves. *J. Atmos. Sci.*, **42,** 217–229.

Ting, M., and I. M. Held, 1990: The stationary response to a tropical SST anomaly in an idealized GCM. *J. Atmos. Sci.*, **47,** 2546–2566.

Trenberth, K. E., and S. C. Chen, 1988: Planetary waves kinematically forced by Himalayan orography. *J. Atmos. Sci.*, **45,** 2934–2948.

Tung, K. T., 1983: On the nonlinear versus linearized lower boundary conditions for topographically forced stationary long waves. *Mon. Wea. Rev.*, **111,** 60–66.

Saltzman, B., and F. Irsch, 1972: Note on the theory of topographically forced planetary waves in the atmosphere. *Mon. Wea. Rev.*, **100,** 441–444.

Swanson, K., and R. T. Pierrehumbert, 1995: Stationary Rossby waves in a homogenized potential vorticity atmosphere. *J. Atmos. Sci.*, **52,** 990–994.

Valdes, P. J., and B. J. Hoskins, 1991: Nonlinear orographically forced planetary waves. *J. Atmos. Sci.*, **48,** 2089–2106.