Interactions between Orographic Gravity Wave Drag and Forced Stationary Planetary Waves in the Winter Northern Hemisphere Middle Atmosphere

C. McLandress
Institute for Space and Terrestrial Science, North York, Ontario, Canada

N. A. McFarlane
Canadian Climate Centre, Downsview, Ontario, Canada

(Manuscript received 2 December 1991, in final form 5 October 1992)

ABSTRACT

A quasigeostrophic model is used to study the combined interaction among orographically generated gravity wave drag, forced planetary waves, and zonal mean flows in the Northern Hemisphere winter stratosphere and mesosphere. The localized gravity wave drag is shown to generate planetary waves in the mesosphere that, in turn, exert a substantial drag on the zonal mean flow via the Eliassen-Palm flux divergence. The amount of planetary wave drag is found to depend not only on the presence of the localized source of orographic gravity wave drag but also on the presence of upward-propagating planetary waves in the lower stratosphere.

The zonal mean wind field exhibits a split jet structure with the larger jet maximum situated in the upper stratosphere at 30°N. This feature is shown to arise from the presence of weak winds above the subtropical tropospheric jet maximum, which results in a region of low-level gravity wave breaking and reduced drag and larger winds above.

1. Introduction

The pioneering works of Murgatroyd and Singleton (1961) and Leovy (1964) revealed that a substantial sink of momentum is required in the mesosphere to balance the Coriolis torque that is associated with the zonal mean meridional circulation. It is now known that breaking, vertically propagating gravity waves provide this sink (Lindzen 1981; Matsuno 1982). The gravity wave drag (henceforth GWD) parameterization scheme proposed by Lindzen (1981) has been used successfully in simulating the zonally averaged circulation in the middle atmosphere (Holton 1983; Garcia and Solomon 1985, for example).

The study of Rind et al. (1988) was one of the more ambitious attempts to parameterize the effects of GWD on the general circulation in the entire region from the earth’s surface to the mesopause. This study demonstrated that a realistic simulation of the general circulation can be obtained by employing a combination of GWD sources associated with different physical processes such as topography, convection, and wind shear, with the source regions being localized geographically. This has been verified to some extent more recently by Hunt (1990), who used the Palmer et al. (1986) orographic GWD parameterization over land and the Holton (1983) scheme over oceans.

While these more recent studies have served to emphasize the importance of accounting for the effects of GWD throughout the lower and middle atmosphere, the interaction of the various drag mechanisms in general circulation models is complex and not easily analyzed. Moreover, these recent modeling efforts have not specifically addressed questions concerning the interaction between planetary waves and GWD, and consequently the role of GWD on the zonally asymmetric flow is less well understood. Hunt’s model, in fact, does not simulate orographically forced planetary wave disturbances in the stratosphere and mesosphere.

The excitation and transmission of gravity waves in the real atmosphere is clearly a spatially variable phenomenon. As noted by Dunkerton (1982), horizontal variations in the larger-scale tropospheric flow associated with the presence of planetary waves result in selective transmission of gravity waves into the middle atmosphere. This effect, combined with the effect of longitudinal variation of gravity wave excitation in the troposphere, gives rise to longitudinal variation of GWD in the middle atmosphere. Holton (1984), in fact, showed that introduction of a longitudinally varying source of orographically excited gravity waves leads to generation of planetary waves in the mesosphere as a result of the localized nature of the GWD.

© 1993 American Meteorological Society
The effects of nonzonal gravity wave breaking were also examined by Schoeberl and Strobel (1984), in the context of a two-dimensional model with a topographically generated GWD parameterization scheme to verify Dunkerton's conjecture of selective transmission. They noted that zonal variations in the breaking level may result in forcing of planetary waves. Furthermore, they showed that GWD could act to damp vertically propagating forced planetary waves in the mesosphere, provided that the waves are of relatively small amplitude and that the gravity wave flux emanating from the lower boundary is independent of longitude. Assuming that both of these conditions are met, which is most likely not the case in the atmosphere, they argued that the dissipated planetary waves would induce a secondary zonal circulation, which would act to decelerate the zonal winds. Miyahara (1985), using a linearized quasigeostrophic $\beta$-plane model, also concluded that planetary wave amplitudes in the mesosphere were reduced by GWD. This study, however, precluded the possibility of localized planetary wave generation by GWD.

In this paper we return to a simpler modeling approach to examine, more closely than hitherto, two specific aspects of the general circulation in the middle atmosphere, namely, 1) the effect of orographic GWD on forced planetary waves and 2) the combined effect of topographically generated GWD and forced planetary waves on the zonal circulation of the Northern Hemisphere winter mesosphere.

These questions are examined here using a quasilinear hemispheric model that has sufficient vertical extent to encompass the stratosphere and mesosphere. Orographic GWD is taken into account using a modified version of McFarlane's (1987) parameterization in which the gravity wave source at the surface is represented in terms of the topographic variance in mesoscales. In the present application the spatial distribution of the gravity wave source at the tropopause is obtained by allowing for filtering of gravity waves by the observed climatological synoptic-scale flow in the troposphere. A limited set of planetary waves, whose phases and amplitudes at the tropopause are also specified from observed climatology, is allowed to interact with the zonal mean flow and modulate and be affected by orographic GWD. Although simplified, this modeling approach does allow us to isolate some of the more important features of the combined action of orographic GWD and planetary waves.

2. Orographic gravity wave drag parameterization

The gravity wave drag parameterization used here is a modified version of McFarlane's (1987) formulation. In this section only a very brief overview of the scheme is presented—the reader is referred to that paper for more discussion of the parameterization. Details of the modifications used in our study are discussed in appendix A.

The momentum flux associated with orographically generated gravity waves is assumed to be predominantly vertically directed. Thus, the effects of GWD on the larger-scale flow are taken into account by including in the horizontal momentum equations a term representing the effect of the divergence of the vertical momentum flux associated with the gravity waves, as

$$\mathbf{j} = -\frac{1}{\rho} \frac{\partial \tau}{\partial z},$$

where the vertical momentum flux associated with mesoscale gravity waves is represented as

$$\tau = -\epsilon \rho N \bar{U} \bar{b}^2 \mathbf{n}.$$  \hspace{1cm} (2)

In Eq. (2), $\mathbf{n}$ is a unit vector in the direction of the mean flow at a reference level near the surface, $\bar{U}$ is the component of the local mean flow that is parallel to the reference level flow, and $\rho$ and $N$ are the local air density and Brunt–Väisälä frequency. The quantity $\bar{b}$ is the amplitude of the local streamline displacement associated with a typical gravity wave and the constant $\epsilon$ is the product of an efficiency factor and a representative horizontal wavenumber. The flux, $\tau$, is calculated using Eq. (20), which appears in appendix A.

At the model lower boundary (i.e., 100 mb) the gravity wave streamlining displacement is determined from tropospheric data by applying the above wave drag scheme to the five-year average of January NMC winds and temperatures for pressure levels at 700, 500, 300, 200, and 100 mb. Since the gravity wave generation occurs mainly over high terrain, the 700-mb level is chosen as the reference level. Consequently, with the constraint that the flux not exceed its local saturated value at the reference level, $\delta_{700\text{mb}}$ is evaluated using the subgrid-scale topographic variance field $S_D^2$, as

$$\delta_{700\text{mb}} = \min\{2S_D, F_C(U/N)_{700\text{mb}}\},$$  \hspace{1cm} (3)

where $F_C$ is the critical inverse Froude number and the standard deviation field, $S_D$, is shown in Fig. 1a. In the present study, $\epsilon$ is given the value of $8 \times 10^{-6}$ m$^{-1}$ and $F_C$ is set to 0.5. These are the values used by McFarlane (1987).

The local value of the wind $U$ that appears in Eqs. (2) and (3) is chosen to be the one tangent to the 700-mb reference level velocity; that is,

$$U = \frac{\mathbf{V} \cdot \mathbf{V}_{700\text{mb}}}{|\mathbf{V}_{700\text{mb}}|}.$$  \hspace{1cm} (4)

The gravity wave momentum flux at 100 mb is shown in Fig. 1b. Comparison of Figs. 1a and 1b indicates that the largest fluxes occur not over the Himalayas where the roughest terrain is situated, but over eastern Asia where the subtropical jet is strongest. Furthermore, south of 15°N the flux almost completely vanishes. Examination of the dot product of the ve-
velocity vectors at 100 mb and 700 mb, shown in Fig. 1c, explains why this happens. The presence of the large shaded region south of 15°N indicates that the wind has turned sufficiently with height so that the angle between the 100-mb and reference level velocity vectors lies between 90 and 180 degrees, implying the presence of critical levels for orographic gravity waves in the troposphere. Such regions are also apparent in higher latitudes, notably over eastern Siberia. In such regions all the momentum flux is depleted in the region underlying the critical levels.

All GWD calculations in the quasigeostrophic dynamical model are performed on a three-dimensional grid that coincides, in the latitudinal and vertical directions, with the model grid but has 32 longitudinal points. The gravity wave momentum flux at the model lower boundary is obtained using the 100-mb streamline displacement and 700-mb wind field (to define $\mathbf{n}$) that were discussed above. Velocity tendencies due to GWD are computed locally on this grid using the finite-difference formulation of McFarlane (1987). The local wind given by Eq. (4) and the temperature fields are obtained by combining the contributions from the zonal mean and planetary wave components. The reference level, denoted by the subscript 0 in appendix A, is taken to be the 16-km level. A Fourier transform is used to represent the local GWD tendency field in terms of the zonal mean and the truncated set of planetary wave components of the model.

3. Model equations

a. Quasigeostrophic zonal mean and wave equations

The purpose of this study is to examine the interactions between orographic GWD and the stationary planetary wave and zonal flow structure of the winter Northern Hemisphere middle atmosphere. Topographically generated gravity waves are, at the tropopause level, largely confined to extratropical latitudes where the larger mountain chains such as the Himalayas and Rockies are found (Fig. 1). It is assumed here that on a seasonal time scale the large-scale flow in the extratropical stratosphere and mesosphere is sufficiently close to being in geostrophic balance that a quasigeostrophic model provides a simple but useful representation of the slowly varying part of the large-scale dynamics.

It can be argued (Boville 1987; Randel 1987) that a gradient balance assumption is preferable to the
The quasigeostrophic assumption in the middle atmosphere, although Marks (1989) has noted that a more serious objection to either of these assumptions is that the GWD terms in the momentum equations may be as large as the inertial terms. Despite this possibility, we have chosen to employ the quasigeostrophic assumption here mainly because the simplicity of the resulting model equations aids in the interpretation of various dynamical processes. The formulation of our model follows closely that of Matsuno (1971). The potential vorticity equations for the zonal flow and planetary waves are given as

$$\frac{\partial \tilde{q}}{\partial t} = -\tan \theta \frac{\partial}{\partial y} \left[ \cot \theta \left( \frac{\nabla \cdot \mathbf{F}}{\rho_0 \alpha \cos \theta} + \mathcal{F}_u + D_M - \gamma_z \tilde{U} \right) \right] - \frac{fR}{\rho_0 H} \frac{\partial}{\partial z} \left( \frac{\rho_0 \tilde{Q}}{N_0^2} \right)$$

(5)

$$\frac{\partial q'}{\partial t} + \tilde{U} \frac{\partial q'}{\partial x} - v' \frac{\partial \tilde{q}}{\partial y} = \left[ \frac{\partial}{\partial x} (\mathcal{F}_u') + \frac{f}{\cos \theta} \frac{\partial}{\partial \theta} \left( \mathcal{F}_u' \cos \theta \right) \right] + D' + R'$$

(6)

where

$$\tilde{q} = \frac{f}{\cos \theta} \frac{\partial}{\partial y} \left[ \frac{\cos \theta \frac{\partial \phi}{\partial y}}{f^2 + \gamma_z^2 \frac{\partial \phi}{\partial y}} + f \frac{\partial}{\partial z} \left( \frac{\rho_0 \frac{\partial \phi}{\partial z}}{N_0^2} \right) \right],$$

(7)

$$q' = \frac{f}{\cos \theta} \frac{\partial}{\partial y} \left[ \frac{\cos \theta \frac{\partial \phi'}{\partial y}}{f^2 + \gamma_z^2 \frac{\partial \phi'}{\partial y}} + \frac{1}{f} \frac{\partial^2 \phi'}{\partial x^2} + \frac{f}{\rho_0 \frac{\partial}{\partial z} \left( \frac{\rho_0 \frac{\partial \phi'}{\partial z}}{N_0^2} \right)} \right],$$

(8)

$$\tilde{U} = -\frac{f}{(f^2 + \gamma_z^2)} \frac{\partial \tilde{\phi}}{\partial y}.$$  

(9)

$$v' = f^{-1} \frac{\partial \phi'}{\partial x}, \ f = 2 \Omega \sin \theta, \ dx = a \cos \theta d \lambda, \ \text{and} \ dy = \alpha d \theta. \ \text{The vertical coordinate is defined as} \ z = -H \ \ln(p/p_0), \ \text{where the subscript} s \ \text{denotes the 1000-mb pressure level and} \ \rho_0 = \rho_s \ \exp(-z/H). \ \text{The background Brunt–Väisälä frequency,} \ N_0, \ \text{is obtained from the U.S. Standard Atmosphere reference temperature profile;} \ R \ \text{is the dry gas constant and} \ H \ \text{is the specified density scale height (i.e., 7 km).}$$

Equation (5) is obtained from the transformed Eulerian mean form of the zonal mean momentum and thermodynamic equations that are consistent with the quasigeostrophic approximation (Andrews et al. 1987, p. 268). The form of the Eliassen–Palm flux divergence, \( \mathbf{V} \cdot \mathbf{F} \), that is appropriate for this system can also be found in Andrews et al. (1987, p. 270).

The terms \( \mathcal{F}_u, D_M, \) and \( \tilde{Q} \) that appear in Eq. (5) represent the zonally averaged zonal component of the gravity wave flux divergence, the diffusive transport of zonal momentum, and the zonal mean diabatic heating, respectively. The exact form of these terms, excluding the diabatic heating, is given in appendix B.

The term \( \gamma_z \tilde{U} \) represents a Rayleigh friction term used to provide a zonal momentum sink in the tropics. The latitudinally dependent coefficient \( \gamma_z \) is given by

$$\gamma_z(\theta) = 10^{-5} \exp[-(\theta/15)^2],$$

(10)

where the latitude, \( \theta \), is evaluated in degrees. Note that apart from the latitudinal dependence of \( \gamma_z \), Eq. (9) is of the same form as used by Schoeberl and Strobel (1978a,b) and is obtained by assuming an approximate balance between the zonal mean components of the Coriolis and pressure gradient forces and the momentum sink, which is represented by the Rayleigh friction term. At extratropical latitudes where the Coriolis parameter, \( f \), is much larger than \( \gamma_z \), the zonal mean wind is very nearly geostrophic. Hence, the departures from geostrophy are significant only near the equator where the geostrophic approximation is of doubtful validity anyway.

As in Matsuno (1971), the model equations are cast into semispectral form by expressing the longitudinal structure of the flow as a sum of the first five zonal harmonics. Both the planetary wave and zonal geopotential fields are specified at the model lower boundary by using five-year January (1982–86) mean 100-mb values from the National Meteorological Center (NMC) data. In order to prevent a large initial transient response, the planetary waves are switched on slowly with an e-folding time of 5 days. At the upper boundary \( \phi' \) and \( \partial / \partial t(\phi' / \partial z) \) are set to zero. At the pole both \( \phi' \) and \( \partial \phi' / \partial \theta \) vanish, while at the equator the boundary conditions \( \phi = 0 \) and \( \phi' = \phi^* \) are employed, where \( \phi^* \) is the geopotential height field corresponding to the radiative temperature \( T^* \) (see section 3b). The boundary condition at the equator is somewhat artificial and implies that the diabatic heating vanishes there. This choice for the equatorial boundary condition is motivated by the results of studies (Wehrbein and Leovy 1982; Shine 1987) of the radiative heating in the middle atmosphere. These studies indicate that, in the solstice period, the zonal mean temperature near the equator is fairly close to the radiatively determined value (see Fig. 5 of Shine 1987), which implies that the zonal mean diabatic heating is small there.

The model domain extends from the equator to the North Pole and from 16 km (i.e., 100 mb) to 128 km. The spectral form of Eqs. (5) and (6) is solved numerically on a 2-km by 5-deg grid. Spatial derivatives are discretized using centered finite differences, and the resulting system of equations is solved using the method devised by Lindzen and Kuo (1969).

Steady-state solutions are achieved within 100 days of simulated time. These are obtained by marching forward in time from a zonally symmetric initial state. A weak Robert filter (Robert 1966) is applied to both the wave and zonal geopotential in order to prevent time decoupling.
b. Diabatic heating formulation

The zonal diabatic heating is modeled as a simple Newtonian relaxation process, which drives the zonal temperature to a specified time-independent structure $T^*$. The heating term in Eq. (5) is therefore written as

$$\tilde{Q} = -\alpha(z)(\tilde{T} - \tilde{T}^*),$$

where $\tilde{T}^*$, shown in Fig. 2, is typical of a radiatively determined zonal mean temperature structure near the winter solstice. This field was constructed so as to resemble the radiatively determined temperature field shown in Fig. 6b of Shine (1987) below 80 km (the uppermost level shown in his paper). Wehrbein and Leovy (1982) show radiative equilibrium temperatures up to a height of 100 km. The structure of $\tilde{T}^*$ above 80 km has been devised using their Fig. 3 as guidance.

The height-dependent Newtonian cooling coefficient $\alpha$ is specified by Eq. (27) in appendix B. It is similar in structure to the “boxcar” profiles shown by Wehrbein and Leovy (1982) except that the maximum value corresponds to a relaxation time of 10 days, twice that of the Wehrbein and Leovy profile. This longer relaxation time was chosen empirically to give radiative heating rates of the same order of magnitude as found by Shine (1989). A justification for using a longer relaxation time here is that the reference temperature, $\tilde{T}^*$, is characteristic of a radiative equilibrium state rather than an observed state. Shine (1987) has shown that longer relaxation times are more appropriate in this case.

The diabatic heating, $R'$, for the planetary waves takes the form of Newtonian cooling with the same coefficient as is used for the zonal mean diabatic heating term.

4. Results

In this section the numerical results obtained using the quasi-linear model are presented. The discussion is divided into two main subsections that deal with the effects of orographic GWD on (i) the forced stationary planetary waves and on (ii) the zonally averaged circulation. The results corresponding to the basic model configuration described in sections 2 and 3 (i.e., with both orographic GWD and climatologically forced planetary waves) are subsequently referred to as case A. In order to elucidate various features of the flow structure, different experiments are performed by altering this basic model setup. Whenever these results are discussed, the changes to the model are duly noted.

a. GWD–planetary wave interaction

1) CASE A

The steady-state geopotential height field at 72 km for case A is shown in Fig. 3. The distortion associated with the large amplitude planetary waves, clearly evident by the presence of the broad trough over the eastern Pacific Ocean, is largely due to in situ forcing of planetary waves by the localized orographic GWD.
Fig. 4. (a) Zonal component of velocity (m s\(^{-1}\)) at 72 km for case A (contour interval of 10 m s\(^{-1}\)), (b) zonal component of velocity tendency due to GWD (m s\(^{-1}\)/day) at 72 km for case A (contour interval of 10 m s\(^{-1}\)/day is used and regions where GWD \(< -100\) m s\(^{-1}\) day are shaded), (c) meridional component of velocity tendency due to GWD (m s\(^{-1}\)/day) at 72 km for case A (contour interval of 2 m s\(^{-1}\)/day), and (d) zonal component of velocity tendency due to GWD (m s\(^{-1}\)/day) at 32 km for case A (contour interval of 0.5 m s\(^{-1}\)/day is used and regions where GWD \(< -3\) m s\(^{-1}\)/day are shaded).

The strong coupling between the planetary wave structure and the associated orographic GWD can be seen in Figs. 4a and 4b, which depict the zonal components of the total velocity and GWD (i.e., the zonal mean plus the first five planetary waves) at 72 km. A wind maximum of over 70 m s\(^{-1}\) is found just west of
North America with much weaker zonal winds of the order of 10 m s\(^{-1}\) occurring over central Asia. The maximum deceleration due to GWD is nearly 140 m s\(^{-1}\)/day and occurs over North America about 60 degrees to the east of the jet maximum. (Note that since the GWD depicted in this figure includes contributions from only the first five zonal harmonics, some nonzero values are seen over oceanic regions.) The meridional GWD component in Fig. 4c is much weaker as a consequence of the generally weaker meridional flow. There, two significant minima occur over North America and Greenland with decelerations of 20 and 30 m s\(^{-1}\)/day, respectively.

Comparison of Fig. 4b and Fig. 1b indicates that the largest GWD in the mesosphere is associated with the North American source region in spite of the fact that the gravity wave flux in subtropical latitudes at 100 mb is largest over the eastern portion of the Asian continent. At midstratospheric elevations, however, the Asian source region dominates the GWD field, as can be seen in Fig. 4d, with maximum decelerations of 7 m s\(^{-1}\)/day. Note that due to the smaller air density at higher elevations, the GWD maximum at 72 km exceeds the 32-km maximum by nearly a factor of 20.

The shift in the position of the maximum drag is attributed in part to the presence of planetary waves in the lower stratosphere. In Fig. 5 are shown the longitudinal and vertical structure of the zonal components of the velocity and GWD at 40°N. (Note that although the model top extends to 128 km, only the results in the region of interest, i.e., below 96 km, are plotted.) At 120°E (i.e., the position of the largest gravity wave flux at 100 mb shown in Fig. 1b) a region of wave breaking at 20 km, denoted by the −5 m s\(^{-1}\)/day closed contour, occurs. Comparing this to the corresponding wind field in Fig. 5a, it is seen that the presence of both large negative wind shear in this region, enhanced by the local wind maximum at 120°E, and large lower boundary gravity wave fluxes that are closer to saturation have resulted in the low-level gravity wave breaking. Consequently, at larger heights the GWD is weaker as a result of the greatly diminished vertical momentum flux.

Although the presence of a well-defined wave 1 component in the 72-km geopotential (Fig. 3) is expected as a response to lower boundary forcing, the orientation of the broad trough over the Pacific Ocean relative to the broad region of GWD over North America (Fig. 4b) is a manifestation of the response to the spatial variation of GWD. In the context of the quasilinear model used here, this response is associated with the spatial distribution of large-scale vorticity sources and sinks associated with the localized structure of the GWD.

The quasi-linear form of the steady state potential vorticity budget [see Eq. (6)] requires that the eddy component of the curl of the GWD force be balanced by the combined effects of advection of eddy potential vorticity by the mean zonal flow (i.e., \(\bar{U}\partial q'/\partial x\)), meridional advection of the zonal mean planetary vorticity by the eddy component of the flow (i.e., \(v'\partial q'/\partial y\)), and dissipation of eddy potential vorticity due to the combined effects of thermal damping, momentum diffusion, and Rayleigh friction [i.e., \(R', D'\) in Eq. (6)]. Figure 6 shows these four terms at the 72-km level projected onto the first five zonal harmonics. Comparison of the panels in this figure reveals that the marked potential vorticity source/sink dipole associated with the broad maximum of GWD over North America is balanced predominantly by the \(\bar{U}\partial q'/\partial x\) term. The contributions due to meridional advection and dissipation are generally smaller and are mainly important in subtropical regions, and near the pole in the case of the advection term. This spatial distribution of the terms in the eddy potential vorticity budget is different from that obtained by Holton (1984), who
found the local balance to be predominantly between meridional advection of planetary vorticity and the curl of the GWD force.

The degree to which the planetary waves can directly alter the zonal mean flow field is determined by the Eliassen–Palm flux divergence (henceforth EPFD).

FIG. 6. Planetary wave potential vorticity tendencies (s⁻²) at 72 km for case A due to (a) GWD, (b) zonal advection of relative potential vorticity, (c) meridional advection of planetary potential vorticity, and (d) dissipation (i.e., Rayleigh friction, diffusion, and Newtonian cooling). Contour interval of 1 s⁻² is used in all panels.
This quantity along with Eliassen–Palm flux vectors is shown in Fig. 7. In order to evaluate the relative contribution of EPFD to the net zonal mean momentum forcing, the zonally averaged GWD is presented in Fig. 8. (The discussion of the zonal GWD term itself is left to the section on GWD–zonal flow interaction.) The substantial size of the EPFD deceleration in middle latitudes throughout the entire mesosphere indeed indicates the important role played by planetary wave driving of the zonal wind in our model. This is an interesting result that differs from Holton’s (1984) finding that GWD-generated planetary waves produced only weak EPFD. The differences between ours and Holton’s results are discussed in more detail in a following subsection.

2) Case B (No GWD on Planetary Waves)

The quasi-linear nature of the model used here means that the only nonlinear processes affecting the planetary waves are the interaction with the mean zonal flow and the nonlinearity resulting from use of local values of the wind in the GWD formulation. The direct contribution of GWD to the planetary wave structure is through the bracketed terms on the right-hand side of Eq. (6) representing the curl of the GWD force. The effect of these terms is revealed in the second experiment, denoted case B, in which GWD is computed in the same localized manner as before but is permitted to act only on the zonal mean flow [i.e., the GWD curl term in Eq. (6) is arbitrarily suppressed].

The planetary wave structure for the first two zonal harmonics for case B, and for comparison purposes for case A, is presented in Figs. 9a–d. In addition, the geopotential differences for the two cases (i.e., case A minus B) are shown in Figs. 9e and 9f. An examination of these figures reveals that planetary wave amplitudes in the mesosphere are drastically reduced in case B. That this is due solely to the presence of localized GWD and not to possible reduced planetary wave vertical propagation resulting from the different refractive properties of the zonal mean winds was verified in an.
additional experiment (results not shown) in which the case B experiment was modified by linearizing the dynamical equations about the case A zonal wind. The resulting wave structure also exhibited greatly reduced amplitudes and so confirms the hypothesis that it is the GWD that causes the in situ generation of planetary waves in the mesosphere. Comparison of Figs. 9a and 9b indicates that the vertical phase variation of wave 1 is also substantially reduced in the mesosphere, indicating reduced vertical propagation by removal of the direct forcing by localized GWD.

Another consequence of the suppression of the direct GWD forcing of the planetary waves is that the zonal mean deceleration associated with EPFD is substantially reduced, as can be seen in Fig. 10a in which is shown the zonal forcing due to EPFD and GWD at 60°N for both cases A and B. For case B the EPFD forcing is only about 1 or 2 m s⁻¹/day in contrast to the value of 20 m s⁻¹/day for case A. The net forcing due to the combined effects of EPFD and GWD, however, shown in Fig. 10b, is nearly identical. Thus, as the EPFD forcing is reduced by suppressing the GWD forcing of the planetary waves, the zonal GWD increases to compensate for this reduction. This effect is found at all latitudes, as can be seen in Fig. 11, which shows the difference in the net zonal forcing (i.e.,

![Figure 8](image1.png)

**Fig. 8.** Zonal mean wind deceleration due to GWD (m s⁻¹/day) for case A. Below 40 km a contour interval of 0.5 m s⁻¹/day is used.

![Figure 9](image2.png)

**Fig. 9.** Planetary wave geopotential amplitude (dm) (solid) and phase (deg) (dashed) for (a) $m = 1$ for case B (i.e., no GWD on planetary waves), (b) $m = 1$ for case A (i.e., with GWD on planetary waves), (c) $m = 2$ for case B, (d) $m = 2$ for case A,
Fig. 9. (Continued) (c) $m = 1$ difference (case A–case B), and (f) $m = 2$ difference. Note that the phase $\alpha$ is defined by $A \cos(m\lambda + \alpha)$, so that an increase of phase with height means a westward tilt of the ridges.
GWD, EPFD, and also the remaining dissipation terms) for the two cases. Despite the fact that large differences occur in the individual terms, the net difference is small. The corresponding zonal wind profiles at 60°N for the two cases are shown in Fig. 12. The winds for case B are about 10 m s⁻¹ stronger, as would be expected from the larger values of GWD.

3) Cases C and D (Modified Lower Boundary Forcing)

The occurrence of large EPFD forcing in our results and not in Holton’s (1984) is found to be directly related to the presence of planetary waves propagating upward from the troposphere, an effect ignored by Holton. To demonstrate this, results from two additional experiments are shown here. In the first of these (case C) the lower boundary planetary wave forcing is switched off so that planetary wave generation occurs only as a result of the presence of localized GWD. Aside from the changes to the lower boundary planetary wave and gravity wave flux structures, case C is identical to the case A configuration described earlier. (The case C scenario is comparable to Holton’s except that he allowed for only a single planetary wave and used a much simpler horizontal lower boundary gravity wave flux structure.) In the second experiment (case D) the lower boundary planetary wave forcing is modified by shifting the ridge of wave one 90 degrees to the west, keeping the remaining waves at their climatologically prescribed values. Since both of these experiments involve modifying the planetary wave lower boundary forcing, the resulting gravity wave vertical momentum flux at 16 km would differ as a result of the changes in the local wind speeds. In order to avoid this and so permit a direct comparison of the different results, the lower boundary gravity wave streamline displacement was chosen so that the momentum flux at 16 km equaled that obtained in case A.

Figures 13a and 13b show the vertical profiles of planetary wave 1 geopotential amplitude and phase at 60°N for both cases A and C. For case C the maximum geopotential amplitude occurs in the upper mesosphere in the region of wave breaking and the phase variation in the vertical is more constant in the stratosphere and lower mesosphere than it is in case A. Note that the phase in Fig. 13b is defined so that a decrease with height indicates a westward tilt with height of the ridge, in contrast to the definition used in Fig. 9.

Comparing the EPFD forcing terms at 60°N, shown in Fig. 13c, indicates that for case C the decelerations have been greatly reduced in the stratosphere and mesosphere from over 20 m s⁻¹/day (for case A) to under
These are shown in Fig. 13d. As in the case where GWD forcing of planetary waves was suppressed (case B), the net zonal mean momentum forcing is nearly the same regardless of the position or the amplitude of the planetary waves at the lower boundary. At the present time we are unable to explain this behavior and are inclined to think that it may be fortuitous. It should be kept in mind that in all three cases the lower boundary gravity wave vertical momentum flux is kept the same.

4) DISCUSSION

The results presented here are, to our knowledge, the first to demonstrate that both planetary wave amplitudes and EPFD are enhanced due to the effects of localized GWD. Holton (1984), while showing that planetary waves are generated by zonal asymmetries in the GWD force, also concludes that such waves do not have a significant associated EPFD. In our results, the large EPFD term is attributed to the presence of planetary waves forced at the tropopause in addition to the localized GWD.

We have also found that, provided the lower boundary gravity wave momentum flux is the same, the net drag on the zonal mean wind remains quite insensitive to changes in the planetary wave forcing.

5 m s\(^{-1}\)/day in agreement with the reduced wave amplitude and phase tilt of Figs. 13a and 13b. This result is in accordance with the findings of Holton (1984).

The planetary wave one structure and EPFD term for the second experiment (case D) are also shown in Figs. 13a–c. Shifting the position of the wave at the lower boundary has reduced both the amplitude and vertical phase tilt. Furthermore, as can be seen from Fig. 13c, the resulting EPFD in the mesosphere has been reduced by over a factor of 2 from the corresponding case A values.

This last experiment is of interest because it indicates that the mean zonal drag exerted by the planetary waves depends on the relative position of the GWD-generated mesospheric planetary wave and the upward-propagating stratospheric wave forced at the lower boundary. Since the position of the GWD-forced wave is fixed geographically to the gravity wave source regions in the troposphere, shifting the lower boundary wave 1 phase by 90 degrees westward results in a vertical phase shift considerably smaller than when the climatological value is used. This interference mechanism was noted in Schoeberl and Strobel (1984) as a means of altering the planetary wave drag on the mean zonal flow.

Last, we present the sum of the EPFD and mean zonal GWD at 60°N for the three cases A, C, and D.
Fig. 13. (a) Geopotential amplitude (m) for $m = 1$ at 60°N for case A (solid), case C (i.e., no lower boundary planetary wave forcing) (long dashes), and case D (i.e., wave 1 ridge at lower boundary shifted 90 degrees westward) (short dashes); (b) same as panel (a) but for geopotential phase (deg) (note phase is defined so that a decrease in phase implies a westward tilt with height); (c) same as panel (a) but for EPFD; and (d) same as panel (a) but for sum of GWD and EPFD.
Our findings differ from Miyahara’s (1985) and Schoeberl and Strobel’s (1984) results, in which they conclude that stationary planetary wave amplitudes in the mesosphere are suppressed by breaking gravity waves. They both assume, however, that the lower boundary gravity wave source has no horizontal variation, hence removing a mechanism that is directly responsible for planetary wave generation in our model.

As indicated above, all the results presented here are for steady-state solutions to the model equations. In a steady state, nonvanishing of the EPFD due to the planetary waves must be due to processes that act to enhance or deplete planetary wave activity. This is a well-known consequence of the more general form of the Eliassen–Palm theorem derived by Andrews and McIntyre (1976) (cf. Andrews et al. 1987, section 3.6). Since we employ time-averaged lower boundary planetary wave forcing, transience due to temporal variation in tropospheric planetary wave activity plays no role. Consequently, the smallness of the EPFD in both the case B and C experiments is a reflection of the fact that the combined effects of momentum diffusion and Newtonian damping are relatively weak in the dynamical model.

It must again be remarked that the quasi-linear nature of the dynamical model precludes any effects due to interactions among the planetary waves. Since the inclusion of these terms would violate the model’s conservation properties, this effect cannot be properly examined here. Some indication of the importance of wave–wave interactions can be inferred, however, by examining the planetary wave amplitudes in Figs. 9b and 9d. The predominance of wave 1 at stratospheric and mesospheric heights (for example, at 60°N and 70 km, the wave 1 amplitude is over twice as large as that of wave 2 and over three times as large at 50 km) would seem to imply relatively weak interaction between these waves, and so little effect on wave 1. Wave 2, on the other hand, could be more significantly modified by interaction of wave 1 with itself.

b. GWD–zonal mean flow interaction

It has been known for some time (e.g., Lindzen 1981; Holton 1983) that GWD provides the zonal mean momentum sink that balances the large Coriolis torque in the mesosphere. In fact, when our model was integrated without any GWD (results not shown), zonal winds of over 400 m s⁻¹ occurred in the mesosphere despite the presence of large amplitude–forced planetary waves.

1) CASE A

The zonally averaged wind and temperature fields obtained with GWD present (case A) are presented in Fig. 14. The temperature field in the polar mesosphere is in excess of 50 K warmer than the “radiative-equilibrium” forcing field seen in Fig. 2. The reversal of

Fig. 14. (a) Zonal mean wind (m s⁻¹) and (b) zonal mean temperature (K) for case A.
the equator to pole temperature gradient in this region is associated with closing off of the mean zonal flow. These realistic features of the zonal mean circulation are also found in previous studies: for example, in the simulations of Holton (1983) and Garcia and Solomon (1985). The split zonal jet structure, however, is not found in those studies but is apparent, though to a lesser extent than shown in Fig. 14a, in the observed climatological mean zonal flow for January (Andrews et al. 1987; Marks 1989).

The net diabatic heating field, shown in Fig. 15, exhibits maximum cooling at the polar stratopause and has a magnitude that is only slightly larger (1–2 K/day) than that obtained by Shine (1989), who used the Cooperative Institute for Research in the Atmosphere (CIRA) temperature data for the period near solstice. Also shown in Fig. 15 is the residual mean velocity vector that arises in response to the drag terms in the zonal momentum equation. In the mesosphere and upper stratosphere the flow is poleward at the equator and in the tropics, attaining a maximum of approximately 4 m s⁻¹ at a height of about 70 km, with mass continuity and the lateral boundary conditions dictating downward motion at the pole.

For a steady state the residual mean meridional component of this circulation, $\vec{V}^*$, is maintained by the zonal momentum sources and sinks. The role of each of these is illustrated in Fig. 16, which depicts the individual contributions to $\vec{V}^*$ at several levels in the vertical. (Note that the range of the ordinate differs for each of the panels.) This figure shows that the residual mean circulation in middle latitudes is maintained primarily by the combined action of GWG and planetary wave EPFD. The increase in the effect of GWG with height up to the 80-km level is evident. The vertical variation of the contribution from EPFD is more complicated. It is positive and dominant in high latitudes in the lower stratosphere (20 km) but negative in middle latitudes. It is positive and of significant magnitude in middle latitudes throughout the mesosphere. The negative values in polar regions of the mesosphere are counteracted by the contribution from momentum diffusion. In subtropical latitudes the role of the equatorial Rayleigh friction sponge in maintaining a cross-equatorial flow is also seen.

The vertical and latitudinal structure of the zonally averaged GWG, shown previously (see Fig. 8), exhibits maximum decelerations of about 30 m s⁻¹/day near the 68-km level. The sharp decrease of GWG poleward of 70°N and between 30° and 15°N is in part a result of the inhibiting of gravity wave breaking by strong westerly flow in the underlying regions. It noteworthy that the magnitude of the zonally averaged GWG found here is in reasonable agreement with the estimates of the small-scale drag for December and January obtained by Marks (1989) using the CIRA data. Maximum values of the combined drag due to EPFD and GWG are 25–35 m s⁻¹/day in midlatitudes at 68 km, in agreement with the estimate of the total drag force obtained by Shine (1989) as diagnosed from the residual mean circulation determined from the CIRA data. In order to highlight the effects of gravity wave breaking in the lower stratosphere, a 0.5 m s⁻¹/day contour interval is used below 40 km. Decelerations of 2 m s⁻¹/day occur above the tropospheric jet maximum at 35°N in reasonable agreement with the results of McFarlane (1987) (see his Fig. 11).

The presence of GWG at lower elevations is also demonstrated in Fig. 17, which shows the zonally averaged gravity wave momentum flux at five successive levels. Note that the range of the ordinate is different for each of the panels in this figure. In the region above the subtropical tropospheric jet where the vertical wind shear is negative, a large part of the flux is absorbed between the model lower boundary and 32 km. In contrast, in more northerly latitudes, where the zonal winds increase with height, the flux is only slightly reduced. Within the remainder of the stratosphere (32–48 km) GWG is mostly confined to the middle latitude region between 35°N and 70°N even though the gravity wave momentum flux does not vanish outside this region. The lateral extent of GWG expands again in the mesosphere above the levels of maximum zonal winds.

2) CASE E (SPLIT JET STRUCTURE)

The split jet structure seen in Fig. 14a is an interesting feature of our simulation that has not been observed.
in other middle-atmosphere models that employ parameterized GWD (Holton 1983; or Garcia and Solomon 1985, for example). As will now be demonstrated, it is the latitudinal curvature of the zonal momentum sources and sinks, and in particular the GWD term, that dictates the departure of the zonal wind from the radiative-equilibrium value. It will furthermore be shown that the double jet structure can be directly attributed to the presence of the lower-stratospheric wind minimum situated just above the subtropical tropospheric jet.

The quasigeostrophic nature of our model requires, in a steady state, that the Coriolis torque balance the momentum sources and sinks. Thus, the residual mean meridional velocity is given by $\frac{\bar{V}^*}{f} = -\overline{F}/f$, where $\overline{F}$ denotes the sum of the momentum sources and sinks. Substituting this into the continuity equation, integrating vertically from $z$ to $\infty$, and then using the steady thermodynamic equation leads to

$$\bar{T} - \bar{T}^* = \frac{\beta(z)H}{R \cos \theta} \frac{\partial}{\partial \theta} \left[ \cot \theta \int_{z}^{\infty} \rho_0 \bar{F} \, dz \right],$$  

(12)
where $\beta(z) = N^2(z)/(\rho_0 \alpha)$. Differentiating this result with respect to latitude and employing the geostrophic zonal wind equation (9) finally yields

$$\frac{\partial \bar{U}}{\partial z} - \frac{\partial \bar{U}^*}{\partial z} = -\frac{\beta(z)}{a} \frac{f}{f^2 + \gamma_z}$$

$$\times \frac{\partial}{\partial \theta} \left[ \frac{1}{\cos \theta} \frac{\partial}{\partial \theta} \left( \cot \theta \int_z^\infty \rho_0 F dz' \right) \right], \quad (13)$$

where $\bar{U}^*$ is the zonal wind that is in thermal wind balance with the radiative-equilibrium temperature field $\bar{T}^*$. This simple but revealing result indicates that the departure of the vertical wind shear from its radiative-equilibrium value is determined solely by the latitudinal gradient of the quantity $\int_z^\infty \rho_0 F dz'$. [Note that in the GWD case where $F = \rho_0 \partial \bar{r}/\partial z$ this integrated quantity simply equals $\bar{r}$ provided that $\bar{T}(\infty) = 0$.]

To explain the split jet structure in our model simulation, in light of the above analysis, the following set
of experiments was performed. In all cases, no planetary waves were forced at the lower boundary nor in the interior by GWD, and the 16-km gravity wave momentum flux obtained from case A was used to generate the lower boundary gravity wave streamline displacement. In addition to GWD, a zonal momentum sink term of the form

$$\Delta F^* = -ke^{-(z-25)/15}e^{-(\theta-20)/15} (\vec{U} - \vec{U}_0), \quad (14)$$

where $\theta$ and $z$ are evaluated in kelvin and kilometers, respectively, was included to drive the zonal wind in the region above the subtropical jet to the prescribed value $U_0$. Thus, by varying the magnitude of $K$ or the value of $U_0$, the lower-level zonal wind in the tropics can be altered. Note that because of its Gaussian shape, the effect of $\Delta F^*$ is restricted mainly to the region centered around 20°N and 25 km.

In Fig. 18a is shown the latitudinal structure of the zonal wind at 40 km for five different cases. For comparison purposes, the wind obtained without any GWD present (but with the equatorial sponge and diffusion present) is plotted. In the case with GWD and with $K = 0$ or $U_0 = 5$ m s$^{-1}$ the double jet structure at this level is clearly evident. As $K$ or $U_0$ is increased, the subtropical jet maximum is progressively reduced and is all but absent when $K = 5 \times 10^{-6}$ s$^{-1}$ and $U_0$ is taken to be the 16-km zonal mean value.

The vertical profiles of the zonal wind at 30°N for all the experiments are depicted in Fig. 18b. In the case with GWD and with $K = 5 \times 10^{-7}$ s$^{-1}$ and $U_0 = 5$ m s$^{-1}$ a jet of about 90 m s$^{-1}$ occurs and is situated at 45 km. As $K$ or $U_0$ is increased, the zonal winds in the upper stratosphere decrease, and the height of the jet maximum is subsequently lowered.

The reduction in the stratospheric jet maximum by increasing the low-level winds is related to the upwelling gravity wave vertical momentum flux shown in Fig. 18c. In the case where $K = 0$ or $U_0 = 5$ m s$^{-1}$, the flux is reduced by an order of magnitude by 20 km and then continues upward at a constant value up to about 50 km where wave breaking again occurs. As $K$ is increased, the flux drop becomes considerably smoother and is spread out over a much broader region, resulting in a more continuous region of wave breaking.

The latitudinal structures of the quantity $\int_0^z \rho_0 F^* dz'$ for the GWD component (i.e., using $F^* = \vec{F}_0$) and for the total forcing (i.e., using $F^* = \vec{F}_0 + \vec{D}_M - \gamma_z \vec{U}$) are shown in Figs. 18d and 18e, respectively. Note that the curves in Fig. 18d are nearly identical to the actual gravity wave vertical momentum flux—the slight differences being due to the use of the local density rather than the basic-state density in the actual GWD calculation. As $K$ is reduced (and the double jet structure enhanced), the latitudinal gradient of the flux between 15° and 35°N is increased. A similar feature can be seen in the case A results shown in Fig. 17, which depicts the large depletion of gravity wave momentum flux in subtropical latitudes between 16 and 32 km below a region of little flux depletion (i.e., no breaking) between the 32- and 48-km levels.

The role of the equatorial Rayleigh friction sponge and the diffusion terms in smoothing out the latitudinal gradients of $\int_0^z \rho_0 F^* dz'$ is clearly seen in Fig. 18e. It is the latitudinal structure of this quantity that determines the actual zonal wind-shear difference $\Delta S$, where $\Delta S = \vec{U}_z - \vec{U}_z^*$. In Fig. 18f is plotted $\Delta S$ for the five experiments. In the case without GWD, $\Delta S$ is very nearly zero, with the slight departures from zero attributed to the presence of the equatorial sponge and diffusion. With GWD present and when the lower-level winds are reduced by using small $K$ or $U_0$, $\Delta S$ is positive in the 20–45-km region and corresponds to the large winds seen in Fig. 18b. As the lower-stratospheric wind minimum is increased, $\Delta S$ exhibits a corresponding decrease.

6. Summary and concluding remarks

In this paper a quasigeostrophic quasi-linear dynamical model is used to study the interactions among orographically generated gravity wave drag, stationary planetary waves, and the zonal mean flow in the wintertime Northern Hemisphere stratosphere and mesosphere. The dynamical model enables planetary waves, forced at the tropopause using climatological data, to interact nonlinearly with the mean zonal flow. The planetary waves not only are affected by the presence of the nonzonal GWD but also affect the zonally averaged GWD through their influence on the local wind field.

The orographic GWD scheme is a modification of that used in the Canadian Climate Centre general circulation model (McFarlane 1987). The wave momentum flux at the lower boundary (100 mb) is specified as that which is consistent with the parameterization scheme when applied to the climatological flow in the troposphere (between 700 mb and 100 mb). Diabatic heating in the winter hemisphere is modeled as a simple Newtonian cooling process in which the temperature is relaxed toward a specified zonally symmetric reference value typical of radiative equilibrium in the middle atmosphere near the equinox and the model equations are marched to a steady state.

The introduction of longitudinal variation in the GWD force gives rise to significant enhancement of planetary wave amplitudes in the mesosphere. A substantial portion of the drag acting on the zonal mean flow is attributed to the Eliassen–Palm flux divergence associated with this enhanced planetary wave activity. The size of the planetary wave drag is determined to a large extent by the presence of planetary waves forced at the lower boundary. In the case where no boundary forcing is used or in the case where the climatological position of the wave 1 at the boundary is altered, sub-
Fig. 18. Case E (i.e., split jet structure) results: (a) zonal mean wind (m s$^{-1}$) at 40 km for varying lower stratospheric momentum forcing given by Eq. (14) — $K = 0$ (solid), $K = 5 \times 10^{-7}$ s$^{-1}$, and $U_0 = \overline{U}$ (16 km) (long dashes), $K = 5 \times 10^{-7}$ s$^{-1}$ and $U_0 = 5$ m s$^{-1}$ (two short dashes and one medium dash), and no GWD with $K = 0$ (three short dashes and one long dash); (b) same as for panel (a) but at 30°N; (c) gravity wave vertical momentum flux (Pa) at 30°N (refer to panel a for curve identification); (d) $\int_{z_2}^{z_1} \rho_0 F_z dz$ at 40 km (i.e., gravity wave vertical momentum flux);

stantial reduction of the Eliassen-Palm flux divergence occurred. Another interesting feature is that the total zonal momentum drag (i.e., EPFD and GWD) remains nearly constant even though the individual terms varied considerably as the planetary wave lower boundary forcing was altered, or if the direct GWD forcing of the planetary waves was suppressed. As indicated above, this result may be fortuitous.

Concerning the zonal mean circulation, a split zonal jet was found with the wind maximum being located at more southerly latitudes than previous studies. This structure was attributed to the presence of gravity wave breaking in the lower stratosphere just above the tropospheric jet. The resulting large depletion of the vertical momentum flux at higher levels caused very steep latitudinal gradients in the GWD that tended to drive up the zonal wind speeds in the subtropics.

The reliability of these conclusions may of course be questioned in view of the simplifications employed both in the large-scale dynamical model and the gravity wave drag parameterization. First, the use of time-averaged climatological values for the boundary forcing precludes the possibility of planetary wave transience playing a large role in determining the zonal mean circulation. This effect is to be studied in a forthcoming study.

The second point concerns the GWD parameterization. Modeling studies such as that of Rind et al. (1988) strongly suggest that the effect of wave drag associated with gravity waves emanating from sources other than topographic variations is also of considerable importance in the large-scale dynamics of the mesosphere. Due to reduced topography in the Southern Hemisphere, orographic GWD alone cannot reproduce the observed wintertime flow. Other gravity wave sources, such as shear generation, are required. Rind et al. (1988) showed that in the Northern Hemisphere winter mesosphere, too, shear-generated GWD is important. The need for other GWD sources is particularly true in the summer hemisphere, where the predominantly easterly mean zonal flow structure inhibits upward propagation of orographic gravity waves. The substantial drag force (cf. Marks 1989) that is required to explain the observed departure from radiative equilibrium in the summer hemisphere may well be associated with breaking of gravity waves that have arrived in the mesosphere after undergoing substantial horizontal as well as vertical propagation.
Fig. 18. (Continued) (e) $\int_0^{40} \rho F dz'$ at 40 km where $F$ denotes sum of GWD, equatorial sponge, and diffusion; and (f) difference between vertical wind shear of zonal wind and radiative-equilibrium wind at 30°N given by Eq. (13).
As the problem of identifying sources and parameterizing, in a spatially varying manner, the effects of such waves is very poorly understood, we have opted to ignore them here and confine attention to the winter hemisphere where orographic gravity waves are important. Since we have been interested in the processes involved in the interactions between localized GWD and the large-scale circulation, the exact nature of the GWD-generating mechanisms is not crucial.

The finding that planetary wave amplitudes are enhanced in the mesosphere due to longitudinal variation of gravity wave drag raises a question concerning the extent to which this effect may be modulated by nonlinear interactions among planetary waves, neglected here under the quasi-linear assumption. This question cannot be addressed adequately in the context of the quasigeostrophic model formulation used here, because such nonlinear interactions are not easily included in a manner that ensures maintenance of potential vorticity conservation constraints (Matsuno 1984). Beyond this, however, the available observations of wind and temperature in the upper mesosphere region are not, to our knowledge, sufficiently numerous to confirm or deny the ubiquity of planetary waves in that region. The recent work of Miyahara et al. (1991) sheds some light on this question by suggesting that diabatic heating and acceleration rates that are consistent with radar wind observations in the mesopause region are not entirely explicable in terms of vertical propagation of gravity waves, and may, in fact, be indicative of the presence of planetary waves in that region.

Acknowledgments. The authors would like to express their thanks to Dr. D. Rind, Dr. T. G. Shepherd, and an anonymous reviewer for helpful comments and criticisms. The first author would like to thank the Atmospheric Environment Service for providing the financial support that made this work possible.

APPENDIX A

Details of Orographic GWD Parameterization

A disadvantage of the McFarlane (1987) parameterization is that \( \tau \) changes discontinuously at a breaking level, denoted \( z_B \). Such behavior may be unrealistic in part because the wave response is not likely to be monochromatic. This possibility affects other aspects of the parameterization as well. For example, in regions where the orography is of a rolling nature rather than a series of ridges, an azimuthal spectrum of waves may be excited. Those with orientations at large angles to the reference level flow, though having much smaller associated values of Reynolds stress, will break at lower levels. The turbulence associated with saturation of these waves may lead to some damping of unsaturated waves. Moreover, nonlinear interactions among the waves may lead to excitation of waves with smaller vertical scales that are more prone to breaking (Lindzen and Forbes 1983).

Holton (1982, 1983) accounts for processes leading to wave damping below the breaking level by assuming that the associated wave drag decays exponentially below the breaking level. Here a different method is used, as follows.

The displacement amplitude required for vertical invariance of the Reynolds stress associated with the gravity waves is easily seen to be

\[
\delta_l = \delta_0 \left( \frac{\rho_0 N_0 U_0}{\rho NU} \right)^{1/2},
\]

where the subscripted quantities are evaluated at the reference level and \( \delta_0 \), the reference level streamline displacement, is specified. We now define a local inverse Froude number, \( F \), as

\[
F = \frac{N \delta_l}{U},
\]

and consider a situation in which \( F_0 < F_C \) and \( F \) initially decreases with height, attaining a minimum value, \( F_m \), at \( z_m \). Above that point \( F \) increases monotonically with height such that \( F > F_C \) for \( z > z_B \). In McFarlane (1987), wave breaking occurred at the level where \( F \) exceeded its critical value, \( F_C \), which results in a discontinuous change in the flux.

Regions where \( F \) increases (decreases) with height are those in which the presence of a gravity wave causes isentropes to become increasingly (decreasingly) vertically steepened (McFarlane 1987), the breaking level being that at which the isentropes are thought to be sufficiently steep that convective overturning is very effective at limiting the amplitude of the wave. If damping of the wave below this level is associated with breaking of waves of shorter vertical wavelengths, it seems reasonable to assume that such a process is likely to be most effective in regions where \( F \) is increasing with height. This notion is employed here in a simple way by taking \( \delta \) to be of the form

\[
\delta = \delta_l \exp \left( -\int_0^z D(z')dz' \right),
\]

where

\[
D(z) = \begin{cases} 
0, & \text{if } z \leq z_m \\
(\partial F/\partial z)/(F_C - F_m + F), & \text{if } z > z_m,
\end{cases}
\]

which leads to

\[
|\tau| = \begin{cases} 
|\tau_0|, & \text{if } z \leq z_m \\
|\tau_0| F_C^2/(F_C - F_m + F)^2, & \text{if } z > z_m,
\end{cases}
\]

so that \( D(z) \) vanishes and the vertical momentum flux is independent of height in the region where \( F(z) \) decreases in the vertical direction.
This formulation can be generalized to circumstances in which there are multiple regions of wave damping associated with nonmonotonic vertical variations of \( F \). Let \( F_p(j) \) be the \( j \)th maximum of \( F \) located at \( z_p(j) \) and let \( F_m(j) \) be the minimum located just below \( z_p(j) \) at \( z_m(j) \). Then, in the region \( z_p(j - 1) \leq z < z_p(j) \),

\[
|\tau| = \begin{cases} 
|\tau_0| F_C^2/(\Gamma_j + F)^2 & \text{if } z_m(j) \leq z < z_p(j) \\
|\tau_0| F_C^2/(\Gamma_j + F_m(j))^2 & \text{if } z_p(j - 1) \leq z < z_m(j),
\end{cases}
\]

where

\[
\Gamma_j = F_C + \sum_{i=1}^{j-1} F_p(i) - \sum_{i=1}^{j} F_m(i).
\]

This formulation yields a smooth vertical variation of the wave drag while ensuring that the saturated wave limit is approached when \( F \gg F_C \). As a result of this, the gravity wave breaks below \( F_C \). An estimate of the height at which breaking first occurs for realistic winter mean flow conditions can be determined from Fig. 2 of McFarlane (1987). With this new formulation, breaking will begin at the first local minimum of \( F \), which occurs at approximately 300 mb rather than at 100 mb where \( F \) exceeds 1.

**APPENDIX B**

**Specification of Model Dissipation Processes**

Holton (1983) and Garcia and Solomon (1985) allowed the diffusivity implied by Lindzen's (1981) GWD formulation to act on both the mean velocity and temperature fields. The propriety of this procedure is debatable, especially in regard to diffusion of heat (cf. Chao and Schoeberl 1984). Here the GWD diffusivity formulation is not used for either heat or momentum. Rather, a fixed height-dependent diffusivity is used to prevent the development of excessive vertical gradients in the horizontal momentum field. There is no explicit vertical diffusion of heat.

As a means of controlling the magnitude of horizontal gradients in the mean zonal velocity, a biharmonic angular velocity diffusion term, formulated in the manner of Holton and Wehrbein (1980), is employed. With the inclusion of these diffusion processes, the momentum source term, \( \bar{D}_M \), in Eq. (5) is written as

\[
\bar{D}_M = -\frac{K_H}{a^3 \cos^2 \theta} \frac{\partial^2 \omega}{\partial \theta^2} + \frac{1}{\rho_0} \frac{\partial}{\partial z} \left[ \rho_0 K_T \frac{\partial \bar{U}}{\partial z} \right],
\]

where \( K_H = 5 \times 10^{16} \text{ m}^4 \text{s}^{-1} \) and

\[
K_T(z) = 50 \left[ 1 + \tanh \left( \frac{z - 60}{20} \right) \right] \tanh \left( \frac{z - 16}{5} \right) \times \tanh \left( \frac{128 - z}{5} \right),
\]

where the height, \( z \), is evaluated in kilometers.

For the planetary waves, both vertical and horizontal diffusion of potential vorticity are employed. This is in addition to Rayleigh damping of momentum in the tropics. A similar damping region is imposed adjacent to the upper boundary of the model. This upper Rayleigh friction sponge layer is employed to prevent unwanted reflection of planetary waves at the upper boundary. The term \( D' \) in Eq. (6), which embodies all of these processes, is written as

\[
D' = \frac{1}{\rho_0} \frac{\partial}{\partial z} \left[ \rho_0 K_T(\xi(\phi')) + K_W \nabla^2 q' - \xi(\gamma_w \phi') \right],
\]

where \( K_w = 10^6 \text{ m}^2 \text{s}^{-1} \) and the Rayleigh friction coefficient \( \gamma_w \) is specified as

\[
\gamma_w = \gamma_z + 10^{-7}[1 + 100 e^{-[(z-128)/20]^2}],
\]

where \( z \) is in kilometers. In Eq. (22), \( \xi \) and \( \nabla^2 \) designate the planetary wave vorticity and 2D spherical Laplacian operators, respectively.

Last, the Newtonian cooling term, \( R' \), in Eq. (6) is given by

\[
R' = -\frac{(2\alpha a)^2}{\rho_0} \frac{\partial}{\partial z} \left[ \frac{\alpha_0 \partial \phi'}{N_0^2 \partial z} \right],
\]

with the Newtonian cooling coefficient, which is also used in Eq. (11), specified by

\[
\alpha(z) = \alpha_0 \left( 1 - e^{-[(z-128)/30]^2} \right) \left[ 1 - 0.8 e^{-[(z-28)/25]^2} \right] \times \left( 1 - 0.3 e^{-[(65-z)/10]^2} \right),
\]

where \( \alpha_0 = 1.25 \times 10^{-6} \text{ s}^{-1} \) and \( z \) is evaluated in kilometers.

**REFERENCES**


Garcia, R. R., and S. Solomon, 1985: The effect of breaking gravity waves on the dynamics and chemical composition of the me-


