Secondary Generation of Gravity Waves Associated with the Breaking of Mountain Waves

TAKEHIKO SATOMURA AND KAORU SATO

Division of Geophysics, Graduate School of Science, Kyoto University, Kyoto, Japan

(Manuscript received 24 June 1997, in final form 21 January 1999)

ABSTRACT

The generation of small-scale gravity waves associated with the breaking of mountain waves in the stratosphere has been simulated within a fully compressible, nonhydrostatic, two-dimensional numerical model. The model includes the mesosphere and offers sufficiently high spatial resolution to characterize the breaking and generation of gravity waves in the stratosphere. The mean flow is initialized by CIRA86 at 40°N in February. A bell-shaped mountain with 30-km half-width and 1.5-km height is located in the bottom of the modeled domain.

The primary wave forced by the mountain propagates into the stratosphere with amplitude increasing with height and subsequently breaks in the lower stratosphere. After the primary wave breaking, significant wave activity is simulated in the stratosphere. These secondary waves are generated at both the upstream and the downstream edges of the breaking zone. Analysis is mostly focused on downstream small-scale waves. The horizontal and vertical wavelengths of the secondary gravity waves are 3–8 km and 3–20 km, respectively, and phase velocities are -1.5 to $+4$ m s^{-1}. The amplitudes of the vertical velocity component of the secondary waves are 0.1–0.2 m s^{-1} at altitudes of about 20 km. Theoretical consideration and model simulation suggest that the winter stratosphere can be a wave duct for small-scale gravity waves once generated there as in the present simulation, because zonal winds are minimized in the winter stratosphere between the height regions of the subtropical jet and the mesospheric jet. It is also suggested that both the convective instability and an instability related with normal-modes act separately in different areas to generate the secondary gravity waves.

1. Introduction

The amplitude of internal gravity waves propagating upward increases with height owing to the decreasing air density. The gravity waves break in high altitudes because statically unstable areas are locally produced, if no critical level is met before the wave reaches there. Today it is believed that the momentum deposition associated with the breaking of gravity waves is a significant forcing of the mean flow and causes the observed weak mean zonal winds near the mesopause (see, e.g., Holton 1982; Matsuno 1982) and the lower stratosphere (e.g., Tanaka and Yamanaka 1985; Palmer et al. 1986; McFarlane 1987). Further, gravity waves are also considered to play an important role in the quasi-biennial oscillation observed in the equatorial stratosphere [Dunkerton (1997a) for a review].

What, then, is the source of gravity waves in the middle atmosphere and how do they propagate from the source? From observational studies using aircraft (e.g., Fritts and Nastrom 1992) or mesosphere, stratosphere, and troposphere radar (e.g., Sato 1994), the dominant source of gravity waves is believed to be topography in midlatitudes. To understand topographic gravity waves, extensive field experiments have been completed over the Rocky Mountains (e.g., Lilly and Kennedy 1973), around the Alps (Kuettner 1986), and the Pyrenees (Bougeault et al. 1993). In general, however, observations by sondes, radars, or lidars are fixed in horizontal position and the observational tracks of aircraft offer sparse coverage in the vertical. Owing to these constraints it is difficult to study stationary waves such as mountain waves in detail by using the observational data only. Therefore, studies using numerical models and observational studies are complementary to each other.

While the detailed propagation processes of mountain waves from the ground to the middle atmosphere are not yet clearly understood, studies to introduce the effects of orographic gravity waves into the middle atmospheric general circulation model have been carried out since the 1980s. Lindzen (1981) hypothesized that the amplitudes of gravity waves would be ''saturated'' owing to local instability acting to prevent further growth of amplitudes with height. The obliterated wave momentum due to the gravity wave saturation is assumed to be deposited in background flow. Based on this assumption, the momentum deposition through the

Corresponding author address: Dr. T. Satomura, Division of Geophysics, Graduate School of Science, Kyoto University, Kitashirakawa Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan.
E-mail: satomura@kugi.kyoto-u.ac.jp

breaking of stationary waves has been parameterized in many global and regional numerical models. These attempts have succeeded in reducing the westerly bias of model results in the lower stratosphere and the troposphere (e.g., Palmer et al. 1986; McFarlane 1987; Iwasaki 1989). Recently, nonstationary gravity waves have also been considered important for precise modeling of the middle atmospheric circulation and new hypotheses and parameterizations have been proposed (e.g., Hines 1997a,b). These studies of gravity wave drag parameterizations have focused on momentum deposited by wave breaking; they ignored detailed processes of wave breaking and did not trace propagation of large-amplitude waves from the troposphere to the middle atmosphere for their practical purposes.

The breaking of gravity waves propagating from below has been studied by two-dimensional numerical models. Bacmeister and Schoeberl (1989) simulated the breaking of mountain waves in a deep (0–85 km) atmosphere using anelastic equations with constant basic horizontal wind and Brunt–Väisälä frequency. Their simulation indicated that a nonlinear wave–wave interaction produced large-amplitude gravity waves emitted downward from the breaking zone of the mountain wave. The downward gravity waves have almost the same horizontal wavelengths as the mountain wave. These downward gravity waves greatly reduced the momentum flux due to mountain waves below the overturning level. Walterscheid and Schubert (1990) studied the breaking processes of a long (300-km wavelength) monochromatic gravity wave in an isothermal middle atmosphere at rest. They used a compressible numerical model and concluded that the amplitude growth of the wave with height that is supposed to occur due to density decrease is substantially reduced because of early breakdown and conversion of kinetic energy into potential energy. Prusa et al. (1996) recently employed an anelastic numerical model and simulated the propagation and the breaking of gravity waves excited by an unsteady Gaussian obstacle located at an altitude of 15 km (the bottom of the model domain). It was shown that a monochromatic wave dominates the wave field at high altitudes because of the dispersion of gravity waves during vertical propagation and that the breaking levels lowered in time from $z \approx 100$ km to 60 km. They initialized the whole model domain by constant horizontal wind and Brunt–Väisälä frequency as did Bacmeister and Schoeberl (1989). Owing to the simplified basic state and the small obstacle located at a high altitude of 15 km, the gravity waves propagate from the stratosphere to the mesosphere without breaking. The processes of propagation and the breaking of mountain waves under realistic atmospheric conditions have not been examined in depth.

The purpose of the present work is to simulate breaking of mountain waves under realistic atmospheric conditions. It is expected that gravity waves are generated through the process of mountain wave breaking. The characteristics of the secondary gravity waves would be determined by the breaking processes and background flow. The excitation of internal gravity waves by instability of parallel flows was studied theoretically and also numerically (e.g., McIntyre and Weissman 1978; Fritts 1982; Sutherland and Peltier 1992; Sutherland et al. 1994; Dunkerton 1997b). The basic conditions of these studies are, however, confined to simple parallel flows. Dunkerton and Robins (1992) found a new mode of secondary instability ''radiating'' from an overturning region of primary monochromatic gravity waves. The amplitude of this radiating mode decreases so rapidly from the overturning zone that its vertical extent is about one vertical wavelength of the primary gravity wave or less.

To simulate propagation of mountain waves through a large vertical extent from the surface to the mesopause correctly, a fully compressible numerical model is employed. The horizontal and the vertical resolutions are maintained sufficiently high to resolve the breaking and associated processes.

2. Model description and experimental setup

The numerical model we use is based on the two-dimensional fully compressible model developed by Satomura (1989). While this model is able to generate terrain-following coordinates numerically, it incorporates terrain at the lower boundary through the use of a hybrid coordinates of the Cartesian coordinate and a conventional coordinate transformation given by Gal-Chen and Somerville (1975) as

$$\begin{cases} z^* = \dfrac{z_1(z - h)}{z_1 - h} & \text{for } z \le z_1 \\ z^* = z & \text{for } z > z_1, \end{cases} \tag{1}$$

where $h(x)$ is the lower boundary height of the integration domain and $z_1 = 15$ km. The compressible equations of motion, the continuity equation, and the thermodynamic equation are

$$\frac{d\mathbf{v}}{dt} + c_p(\Theta + \theta)\nabla\pi = \mathbf{k}g\frac{\theta}{\Theta} + \mathbf{F}_v, \tag{2}$$

$$\frac{d(\Pi + \pi)}{dt} + \frac{R_d(\Pi + \pi)}{c_v}\nabla \cdot \mathbf{v} = \frac{R_d(\Pi + \pi)}{c_v(\Theta + \theta)}F_\theta, \tag{3}$$

and

$$\frac{d\theta}{dt} = F_\theta, \tag{4}$$

where

$$\Pi(z) + \pi(x, z, t) = \left(\frac{p}{p_0}\right)^{R_d/c_p},$$

$$\frac{d\Pi(z)}{dz} = -\frac{g}{c_p\Theta(z)},$$

and $\mathbf{v} = (u, w)$. The right-hand sides of (2)–(4) are the contributions from turbulent mixing, for which we employ a conventional first-order closure formulation (Lilly 1962; Satomura 1989).

The model uses a second-order finite-difference approximation centered both in time and in space. To gain computational efficiency, both a time splitting method similar to Klemp and Wilhelmson (1978) and an implicit method in the vertical direction are applied to sound wave terms in the compressible equations. A fourth-order artificial viscous term is added to Eqs. (2) and (4) to damp small-scale computational modes.

The 512-km-wide integration domain is covered by grids with 500-m spacing in the x direction. In the vertical direction, the domain height is 97 km and is divided into 600 layers whose depth increases smoothly with height from 25 m near the ground to 500 m in the mesosphere. In the stratosphere and most of the middle and the upper troposphere, which are focused on in this study, the vertical grid interval is fixed at 100 m in order to keep sufficient resolution.

At the lateral boundaries, an open boundary condition (Miller and Thorpe 1981) based on Orlanski (1976) is employed. The upper boundary is a rigid wall and a thick ''sponge'' layer is put above a height of 50 km to avoid reflection of vertically propagating waves at the upper boundary. The lower boundary is a free-slip and thermally insulated rigid wall. A bell-shaped mountain

$$h(x) = \frac{a^2 h_0}{a^2 + (x - x_0)^2}, \tag{5}$$

where $h_0 = 1.5$ km and $a = 30$ km, is located at the center ($x_0 = 256$ km) of the lower boundary.

The initial wind and temperature fields are horizontally uniform. The vertical profiles of the initial wind and temperature are shown in Fig. 1, which are taken from the COSPAR International Reference Atmosphere 86 at 40°N in February and interpolated to the model grid points by cubic-spline functions. In this profile, the tropopause and the stratopause are located at altitudes of about 10 and 47 km, respectively. The wind has local maxima of 29 m s^{-1} at 13-km height corresponding to the subtropical jet and 68 m s^{-1} at 61-km height (the mesospheric jet). The vertical profile of the critical horizontal wavelength

$$\lambda_c \equiv \frac{2\pi}{l}, \tag{6}$$

where l is the Scorer parameter

$$l^2 = \frac{N^2}{u^2} - \frac{1}{u}\frac{d^2u}{dz^2}, \tag{7}$$

and N the Brunt–Väisälä frequency, is also shown in Fig. 1. The critical horizontal wavelength has a local maximum in the upper troposphere. The linear theory indicates that mountain waves whose horizontal wavelengths λ_x are shorter than λ_c are vertically evanescent (e.g., Smith 1977) and reflected downward. It should be noted that gravity waves with $\lambda_x < 10$ km generated by mountains cannot propagate upward into the stratosphere in the present basic state.

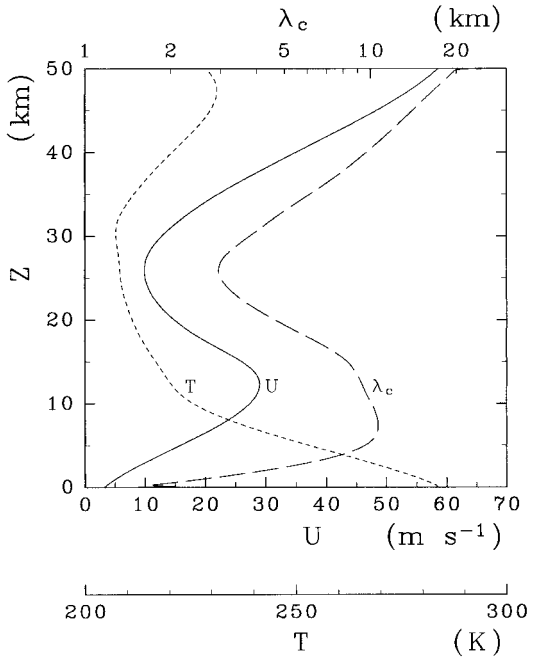

FIG. 1. Vertical profiles of initial horizontal wind (solid), temperature (dotted), and the critical wavelength (dashed) taken from CIRA86 at 40°N on February.

The simulation was run for 10 h and diagnostic field variables were saved at 5-min intervals between $t = 6$ h and 10 h for the analysis. A contour plot of horizontally averaged momentum flux, $\overline{u'w'}$, where u' and w' are wave components and the overbar indicates horizontal average, is shown in Fig. 2. This height–time diagram of averaged momentum flux indicates that the mountain waves were well established up to the stratosphere after 6-h integration and had reached a practically steady state at 10 h.

3. Results

Snapshots at 5 h, 7 h, 9 h, and 10 h around the mountain are shown in Fig. 3. Blue and red contours represent downward and upward vertical velocities, respectively, and the black lines are contours of potential temperature. The primary mountain wave propagates upward from the mountain to the stratosphere. The contours of potential temperature in Figs. 3b–d clearly show that the wave breaking occurs in the region of $x = 250$–300 km and $z = 21$–24 km. While the initial wind is weak at these heights (see Fig. 1), it is still more than 10 m s^{-1}; no critical level exists in the initial wind there. Thus it is probable that the increase of wave amplitude due to the decrease of air density with height results in this

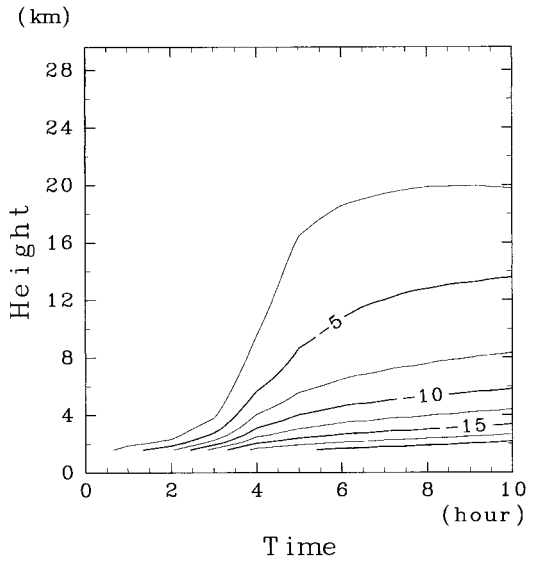

FIG. 2. Vertical momentum flux, $\overline{u'w'}$, as a function of z and t. The contour interval is 2.5×10^4 N m^{-1}.

wave breaking in the lower stratosphere. The group velocity of linear steady gravity waves in the lower stratosphere is more than several m s^{-1} for λ_x less than several tens of kilometers. Therefore, Fig. 3 shows that a part of the wave energy of the primary mountain wave is transmitted upward through the breaking zone.

Significant wave activity besides the mountain wave are found in two regions in Figs. 3c and 3d: downstream of the mountain throughout the troposphere and stratosphere, and over the mountain in the stratosphere. Waves over the mountain have longer wavelengths than the downstream waves. The phase lines of the waves over the mountain were nearly parallel to the mountain wave phase lines near the breaking zone, which is clear in close-up views (not shown). The constant phase lines of the downstream waves are not parallel to the mountain wave phases. The difference in phases and wavelengths suggests that the wave excitation mechanisms are different between the two regions. We will focus on the downstream short waves in the following sections, because they are easily separated from the primary mountain waves.

The short waves seem to propagate both upward (x = 310–330 km, z = 20–40 km in Fig. 3c and x = 320–340 km, z = 20–40 km in Fig. 3d) into the stratosphere and downward (x = 310–320 km, z = 5–18 km in Figs. 3c and 3d) into the troposphere from the breaking zone. The waves propagating upward have shorter wavelengths than those propagating downward. There are other small-scale and small-amplitude waves around the breaking zone. We will examine these waves in detail below.

To analyze the structure of the short waves, a high-pass filter whose cutoff wavelength is 10 km was applied horizontally. A close-up view of filtered vertical velocity at t = 9 h near the tail of the breaking zone is shown in Fig. 4. We can find wave activity both above (region A) and below (region B) the breaking zone indicated by a border line between the regions A and B. The maximum amplitudes of the vertical velocities of the waves are 0.1–0.2 m s^{-1}. The waves in region A transmit energy upward and those in region B are evanescent as shown in section 3b. Waves in both regions A and B seem to originate from a thin sheet along their mutual border. We can also identify waves transmitting energy downward (region C) and upward (region D) from a small area around (x, z) ≈ (313 km, 20 km).

a. Phase velocity and wavelength

Figure 5 shows an x–t diagram of filtered vertical velocities at z = 27 km (regions A and D) and z = 15 km (regions B and C). It should be noted that waves have small but nonzero horizontal phase velocities. Further, waves in region A have a distinct negative phase velocity. The wave parameters and environmental variables are summarized in Table 1 where horizontal and vertical wavelengths, λ_x and λ_z, are estimated by assuming plane waves in each region. The mean values and variations of λ_x and λ_z given in Table 1 are calculated from 10 samples around the selected points for u and N (see below for point selection) in each region. The lengths of each sample are 40 km and 18 km for the estimation of horizontal and vertical wavelength, respectively. The vertical wavelength of linear gravity wave with horizontal wavelength of λ_x can be also calculated from the linear dispersion relation of internal gravity waves:

$$\lambda_{zcp} = 2\pi\left[\left(\frac{N}{u - c_p}\right)^2 - \left(\frac{2\pi}{\lambda_x}\right)^2 - \frac{1}{4H^2}\right]^{-1/2}, \quad (8)$$

where H is the density-scale height. We selected points for the estimation of u and N as listed in Table 1, where spatial variations of u and N in both x and z directions are rather small and neglected the effect of d^2u/dz^2. The density scale height H decreases slightly from 6.6 km at z = 15 km to 6.2 km at z = 27 km.

The theoretically estimated vertical wavelength, λ_{zcp}, corresponds well with the observed wavelength λ_z, indicating that waves in regions A, C, and D are likely internal gravity waves. These waves have similar horizontal wavelengths of between 3 and 8 km and horizontal phase velocities within −1.5 to +2 m s^{-1}, whereas vertical wavelengths vary from 3 km in region D to 30 km in C.

For waves in region B, however, λ_{zcp}^2 is negative and, thus, waves are evanescent (or external) in the vertical. In fact, the amplitudes of waves in region B clearly decrease before the waves reach z = 15 km (Fig. 4), although the altitude of source area of waves in region B is almost the same as that in region C.

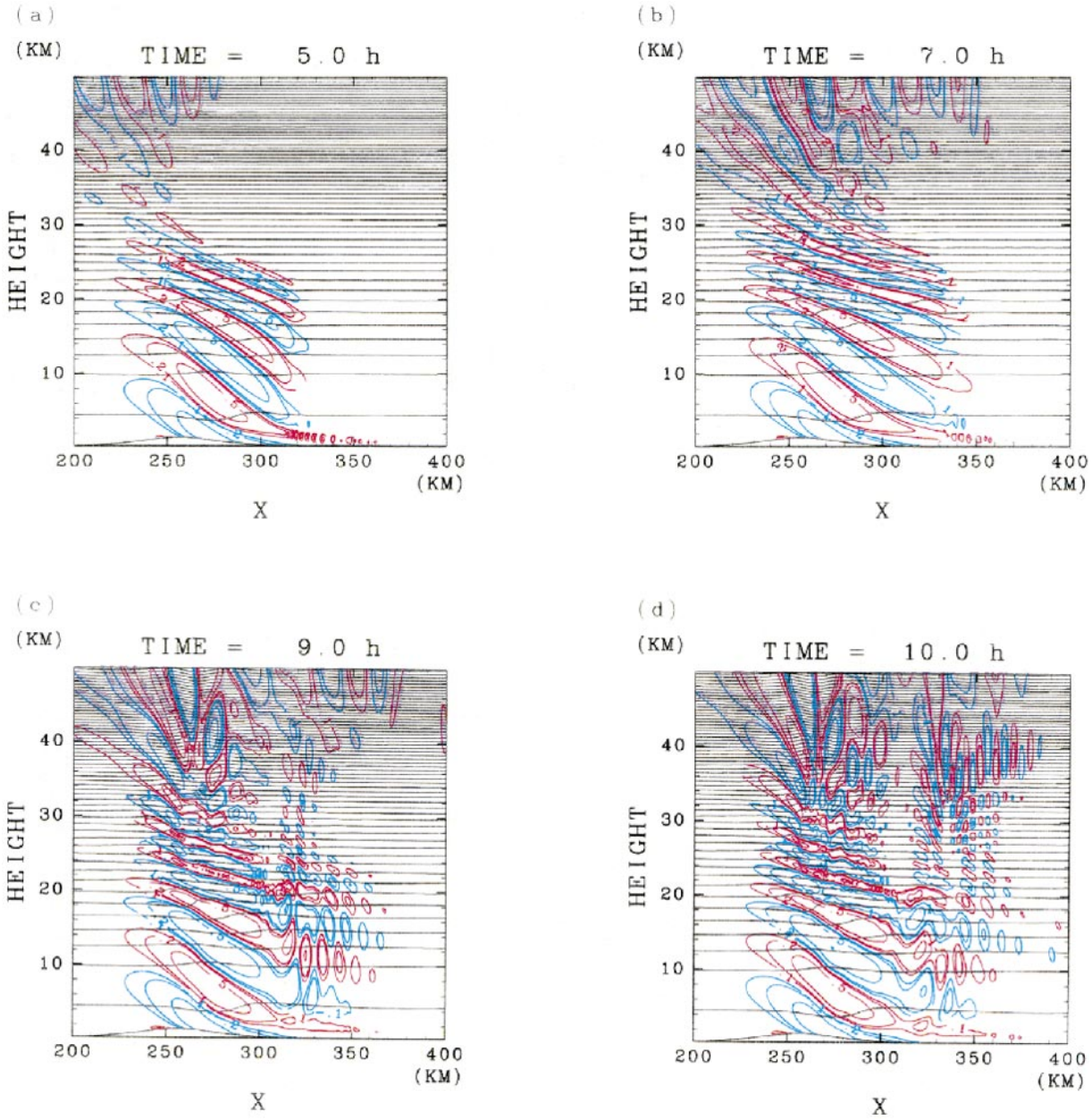

FIG. 3. Vertical velocity, w, in m s^{-1} and potential temperature at (a) 5 h, (b) 7 h, (c) 9 h, and (d) 10 h of simulation time. The blue and red lines are contours of negative and positive vertical velocity, respectively. The contour intervals of w are ± 0.1, ± 0.2, ± 0.5, and ± 1 m s^{-1}. The black lines are contours of potential temperature with an interval of 20 K. The mountain top is located at $x = 256$ km.

b. Momentum flux

The waves found in the regions A–D have the horizontal phase velocities less than the background horizontal wind velocity. If internal gravity waves have negative phase speed with respect to the mean flow as the waves in the four regions, the momentum flux $\rho u'w'$ should be negative for upward energy-propagating waves and positive for downward propagating waves, where dashed variables indicate wave components.

Figure 6 shows the time series of the momentum flux $\rho u''w''$ starting from $t = 8$ h 30 min with a 30-min interval, where u'' and w'' are filtered velocities. It is clear that upward propagating waves in regions A and D have negative momentum fluxes and downward propagating waves in region C have positive momentum flux: the signs of the momentum fluxes are consistent with the internal gravity wave characteristics mentioned at the beginning of this section. The maximum absolute values of these fluxes are greater than 5×10^{-3} N m^{-2}.

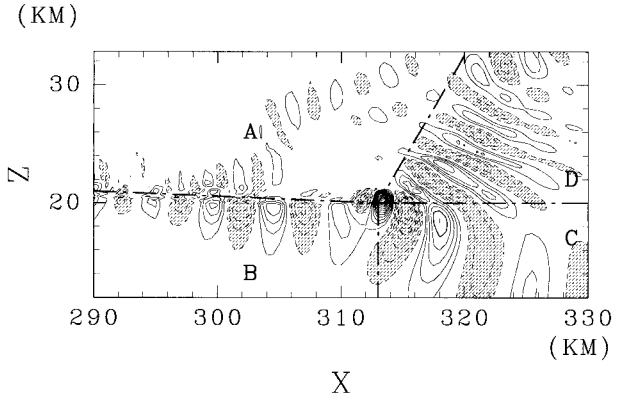

FIG. 4. Filtered vertical velocity at $t = 9$ h near the downstream tail of the breaking zone. The contour interval is 0.04 m s^{-1}. Dashed contours represent negative values. Thick dash–dotted lines separate regions A, B, C, and D.

Evanescent waves in region B, however, again show different characteristics. Here the momentum flux alternates between positive and negative signs with almost the same amplitude. The horizontally averaged momentum flux in region B is nearly equal to zero.

Upward and downward propagating waves are reflected and the sign of the momentum flux is changed at $z \approx 35$ km in region D and $z \approx 15$ km in region C, respectively. This reflection of waves will be discussed in the next section.

c. Spectral characteristics

Before we start discussion on the waves found around the breaking zone of the mountain wave in detail, it is worth confirming that waves in the regions A–D are not propagating from the troposphere but are generated in the stratosphere. The power spectrum of vertical velocity at $t = 10$ h and the curve of the upper height limit where wave packets started at the ground at $t = 0$ can reach at $t = 10$ h are shown in Fig. 7. The upper height limit is calculated by integrating the group velocity of each wavenumber estimated from the linear dispersion relation using the initial horizontal velocity and potential temperature profiles. Waves with horizontal wavelengths shorter than 10 km cannot propagate from the ground above $z = 15$ km where the critical wavelength is maximized (see Fig. 1). The upper limit of propagation for waves with longer wavelengths is lower because of their smaller vertical group velocities. The power spectra are calculated only for $z > 15$ km because the model coordinates are curvilinear below 15 km and hence the interpretation of the Fourier transform within such coordinates is difficult there. The line of the upper limit coincides well with the contour of 10^{-5} kg s^{-2}, particularly for $z > 20$ km in the wavelength range longer than about 10 km corresponding to the maximum critical wavelength at z = 15 km. Thus the line of upper limit could be used as an indicator of the upper-limit altitude reached by the primary mountain wave.

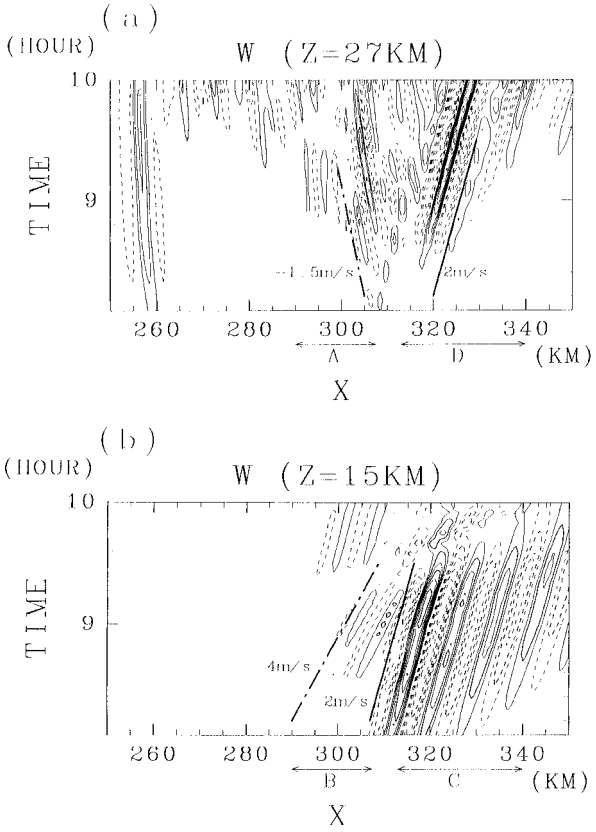

FIG. 5. Horizontal-time section diagram of filtered vertical velocity at (a) $z = 27$ km and (b) $z = 15$ km. The contour interval is (a) 0.02 m s^{-1} and (b) 0.01 m s^{-1}. Dashed contours represent negative values. Thick solid lines show horizontal phase velocity $c_p = 2$ m s^{-1}. Dash–dotted lines show (a) $c_p = -1.5$ m s^{-1} and (b) $c_p = 4$ m s^{-1}.

Substantial power penetrates into the ''inhibited'' area on the shortwave side ($\lambda_x < 10$ km and $z > 15$ km) in Fig. 7, corresponding to waves whose wavelengths are shorter than 10 km observed around the breaking zone. Therefore, the short waves around the breaking zone

TABLE 1. Wave parameters and environmental variables at 9 h. Here λ_x is horizontal wavelength, u is unfiltered horizontal velocity, λ_{zcp} is vertical wavelength calculated from the dispersion relation, — is external mode, c_p is horizontal phase velocity, and λ_z is vertical wavelength.

Region	λ_x (km)	c_p (m s^{-1})	λ_z (km)	λ_{zcp} (km)	u (m s^{-1})	N ($\times 10^{-2}$ s^{-1})
A	3.3–3.7	−1.5	2.3–9.0	2.8–7.6	5.3–9.3[a]	1.9–2.2
B	5.0–5.7	4	—	—	26–30[b]	2.1–2.3
C	5.5–8.5	2	12–33	8.0–41	17–26[c]	2.0–2.1
D	5.0–12	2	2.5–2.8	1.4–3.3	6.4–12[d]	1.9–2.2

[a] At (x, z) = (290–310 km, 24–25 km).
[b] At (x, z) = (310–330 km, 17–19 km).
[c] At (x, z) = (310–330 km, 17–19 km).
[d] At (x, z) = (290–310 km, 24–25 km).

are not those propagating from the bottom of the model, but are, in fact, those generated in the stratosphere. In this sense, it is appropriate to call these waves ''secondary gravity waves'' since they are generated by processes associated with the breaking of the primary mountain waves.

Note, also, that the thermal and wind structure of the winter stratosphere forms a wave duct for secondary gravity waves. As shown in a region of $x = 320$–350 km and $z = 25$–35 km in Figs. 6c and 6d, the energy of the secondary waves cannot propagate into the mesosphere but is reflected at $z = 35$–40 km since the wavelengths of secondary waves are less than 10 km and the critical wavelength λ_c surpasses 10 km once more in the upper stratosphere (Fig. 1). Owing to their nonzero phase velocities, the secondary gravity waves in regions C and D will propagate away downstream within the stratospheric duct from the breaking zone of the primary mountain wave. Thus, these secondary gravity waves will possibly disturb the stratospheric atmosphere even over the ocean or over flat surfaces.

4. Discussion

a. Generation mechanisms

In our model, gravity waves with different wavelengths and phase velocities and vertically evanescent disturbances were generated in different areas. To look into the generation mechanisms of each region defined in Fig. 4, greatly magnified views of the Brunt–Väisälä frequency, horizontal velocity, and $\log_2$Ri, where Ri is the Richardson number, near the downstream tail of the breaking zone at 9 h, are shown in Fig. 8 with filtered vertical velocity. At 9 h, the mountain waves have reached steady development for several hours as shown in Fig. 2.

It is clear that gravity waves in region A originate in the middle or upper part of the wave-breaking zone of the mountain wave where the vertical stratification is unstable (Fig. 8a). Strong convection is expected to prevail under such unstable conditions. Given an unstable layer of depth 1 km, the linear theory of convective instability between two rigid stress-free plates predicts that the horizontal wavelength of the fastest growing mode is 2.8 km (e.g., Chandrasekhar 1961); this value is similar to that of the gravity waves observed in region A. The horizontally averaged heat flux associated with the filtered components is clearly positive from $z = 20.1$ km to 21.2 km corresponding to the breaking zone (not shown).

Convective motions are considered to generate gravity waves by both mechanical oscillator and obstacle-related mechanisms (e.g., Clark et al. 1986; Fovell et al. 1992). Mechanical oscillators would generate gravity waves that propagate to right and left, whereas gravity waves generated by the obstacle effect propagate with the same horizontal phase velocity as the convective movement. The phase velocity of the gravity waves in region A (-1.5 m s^{-1}, see Fig. 5a) is almost the same as the horizontal wind velocity in the overturning zone (≈ -1 m s^{-1}, see Fig. 8b). Considering that the horizontal wind speed gives a good estimation of the advection speed of the embedded convection, the obstacle effect may be the generation mechanism of the secondary waves in region A.

The generation of secondary gravity waves in regions B–D may be explained in terms of the instability studied by Sutherland and Peltier (1992, referred to as SP92 hereafter) and Dunkerton (1997b). SP92 found unstable but propagating modes by a linear stability analysis of horizontally uniform flow. They used a vertical profile of horizontal velocity and Brunt–Väisälä frequency extracted from one wavelength of Long's solution (Long 1953) above a critical position in the lee of the obstacle.

The origin of disturbances in region B is located in a stable zone just beneath the breaking point. This stable zone overlaps a strong shear zone (Fig. 8), which is 2 km thick at $x = 305$ km and in which the horizontal wind decreases from 30 m s^{-1} at $z = 19$ km to a slowly reversed wind at $z = 21$ km. The wavenumber of the temporally fastest growing mode (6.546 in dimensionless units, see SP92) corresponds to about 7.5 km if the thickness and strength of shear are 2 km and 1.5×10^{-2} s^{-1}, respectively, as in the strong shear zone in region B. The most unstable mode analyzed by SP92 has a maximum amplitude just below the neutral buoyancy layer. Therefore the characteristics of the structure and wavelength of the most unstable mode analyzed by SP92 also agrees well with disturbances in region B.

The phase velocity of the most unstable mode of SP92 is about 0.2 m s^{-1}, while the phase velocity of waves in region B is 4 m s^{-1}. It should be noticed that the levels of the minimum horizontal velocity and the neutral stratification coincide in SP92 whereas the primary mountain wave in our simulation developed into a nonlinear state much further than the condition of SP92. In fact, looking into Figs. 8a and 8b, the neutral stratification is formed 500 m below the level of the minimum horizontal velocity at 9 h. If we assume that the phase velocity of the most unstable mode is almost stationary relative to the horizontal velocity at the level of neutral stratification, it then becomes about 3 m s^{-1}. This value corresponds well with the phase velocity of the disturbances in region B.

In region B, the local critical wavelength[1] $\lambda_c = 6$–7.5 km at $z = 18$–19 km owing to the stronger horizontal wind in region B. The horizontal wavelength of waves is $\lambda_x \approx 5.4$ km in region B. Therefore the propagation condition $\lambda_x > \lambda_c$ is not satisfied and waves are vertically evanescent in region B.

[1] Here λ_c is calculated by Eqs. (6) and (7), which are not exact for waves with nonzero phase velocities. The error is, however, only about 10% because the background wind is strong ($\approx$30 m s^{-1}).

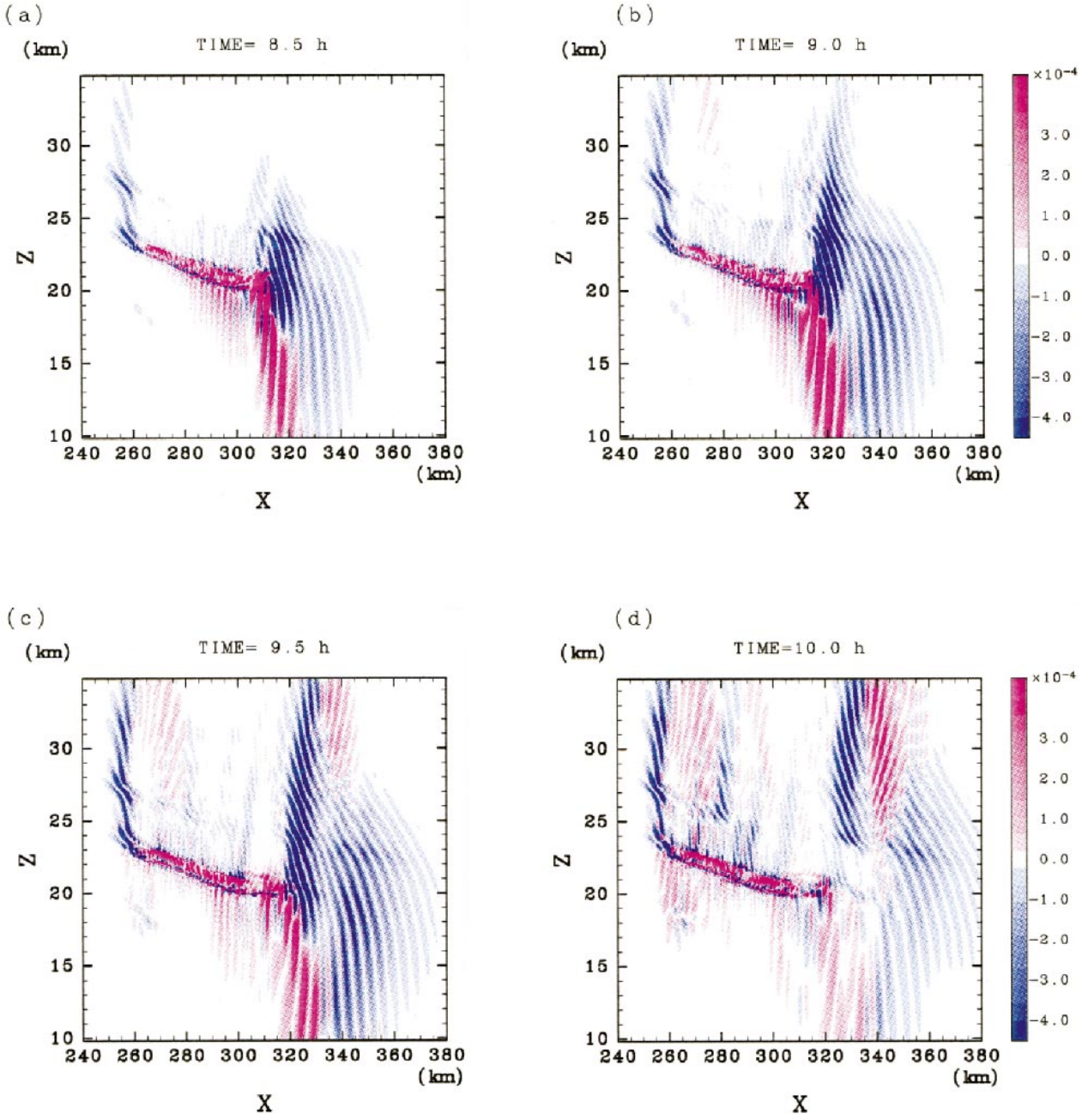

FIG. 6. Horizontal-height section of momentum flux $\rho u'' w''$ (N m^{-2}) at (a) 8 h 30 min, (b) 9 h, (c) 9 h 30 min, and (d) 10 h. The momentum flux is calculated using filtered wind velocities.

The source area of waves in region C seems to be $(x, z) \approx (312$ km, 21 km) at 9 h, near the downstream end of the breaking zone. The thickness of the vertical shear zone at $x = 312$ km is about 3 km (from $u = 28$ m s^{-1} at $z = 18$ km to $u = 1$ m s^{-1} at $z = 21$ km). The Brunt–Väisälä frequency decreases to almost zero at $z = 21$ km where $u = 1$ m s^{-1} (Figs. 8a and 8b). The most unstable mode analyzed by SP92 for this background state has about a 10-km wavelength and about 1 m s^{-1} phase velocity. These values agree well with the parameters of the gravity waves in region C and the instability possibly causes the gravity waves in this region. The horizontal length of the source area, however, extends over about one wavelength only. Thus, it may be difficult to apply the analysis of SP92, which assumed a horizontally uniform wind and temperature.

It is possible that the gravity wave generation in region C is controlled by spatial structure of the unstable

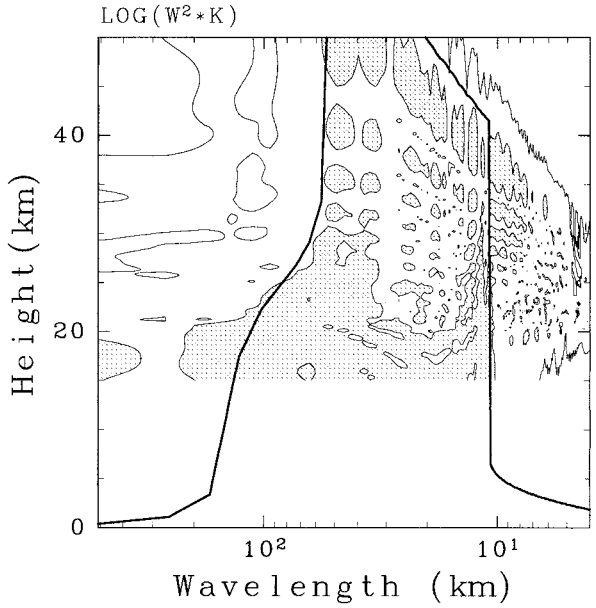

FIG. 7. Power spectrum of vertical velocity weighted by air density at $t = 10$ h. The contours are 10^{-3}, 10^{-5}, 10^{-7} kg s^{-2}. The shaded area represents values $>10^{-5}$ kg s^{-2}. The thick solid curve indicates the upper limit of the height where wave packets starting from the ground at $t = 0$ h can reach at 10 h.

region. Scinocca and Peltier (1993) observed a large-amplitude stationary disturbance similar to the most unstable mode of SP92 in their nonlinear numerical simulation of downslope windstorms. The wavelength of the disturbance agrees with the most unstable mode of SP92 and the source area of the disturbance is located at the downstream edge of the breaking zone: the characteristics of the large-amplitude stationary disturbance observed by Scinocca and Peltier (1993) are very similar to those of the waves in region C.

Figure 9 shows close-up views of x–t diagrams of filtered vertical velocities near the downstream end of the breaking zone. The left edge of the wave packet in region C (Fig. 9b), which is considered to be its source area, moves downstream from $(x, t) \approx$ (302 km, 7.4 h) to $(x, t) \approx$ (310 km, 8.4 h). This movement corresponds closely with the stretch of the mechanically unstable (Ri < 0.25) layer in both their locations and their speeds. As seen clearly in Fig. 9b, the phase velocity of gravity waves coincides with the stretching speed of the unstable layer. Since this unstable layer is associated with the primary wave breaking, the correspondence between the wave characteristics and the stretch of the unstable layer suggests that the gravity waves in region C are generated by a similar mechanism as shown by Scinocca and Peltier (1993), that is, through a nonlinear process at the edge of the breaking zone.

The local critical wavelength λ_c in region C is 5.5–7.0 km. The horizontal wavelength of the secondary wave $\lambda_x \approx 7.6$ km in region C. Therefore the propagation condition $\lambda_x > \lambda_c$ is satisfied there and gravity waves can propagate far from the source area.

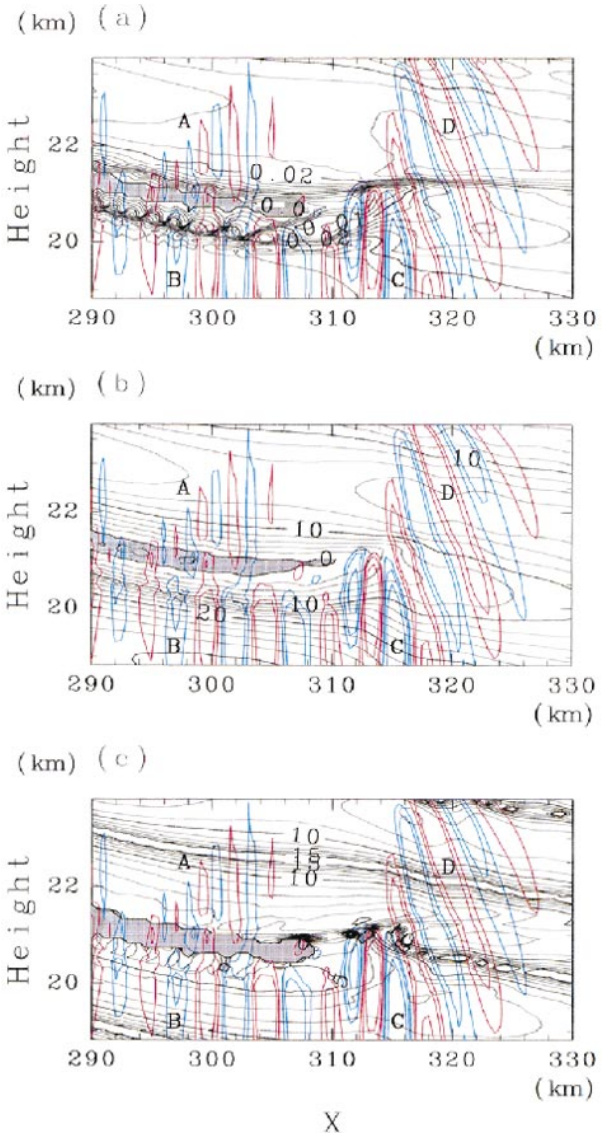

FIG. 8. Close-up of (a) Brunt–Väisälä frequency, N; (b) horizontal velocity, u; and (c) $\log_2$Ri, where Ri is the Richardson number, at $t = 9$ h near the tail of the breaking zone. The contour interval in (a) is 0.0025 s^{-1}, (b) 2 m s^{-1}, and (c) 1. A contour line of $u = -1$ m s^{-1} is added in (b). Shaded areas indicate (a) $N < 0$, (b) $u < 0$, and (c) Ri < 0.25. Filtered vertical velocity is also plotted in blue (negative) and red (positive) contours for ± 0.05 m s^{-1}, ± 0.1 m s^{-1}, and ± 0.2 m s^{-1}. Regions are indicated by letters A–D.

Gravity waves in region D appear from $(x, z) \approx$ (314 km, 21 km) at $t = 9$ h, which is also near the downstream edge of the breaking zone. Judged from the fact that the horizontal wavelength and phase velocity of gravity waves in region D are very similar to those in region C, it is inferred that gravity waves in region D are generated by the same generation mechanism in region C.

It is worth noting that the leftmost wave in region D appears from almost the same location $x \approx 312$ km

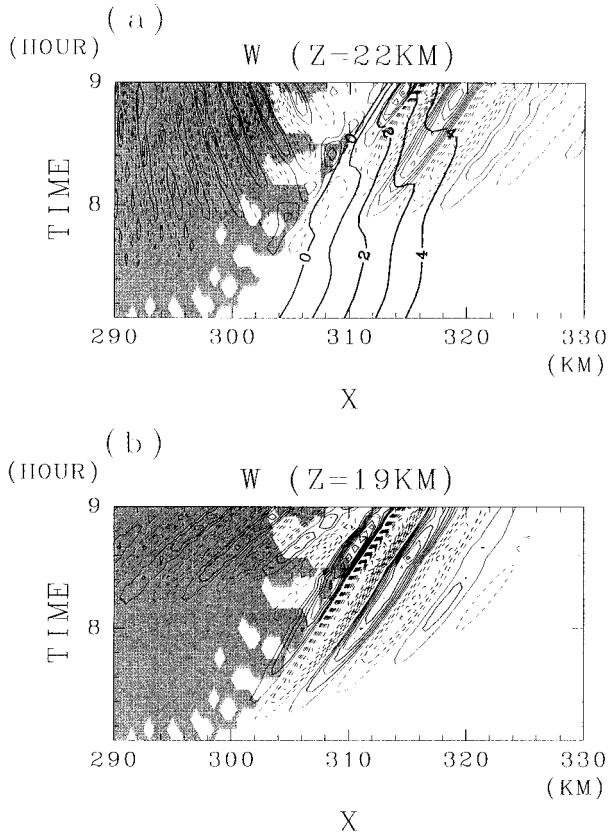

FIG. 9. Horizontal-time section diagram of filtered vertical velocity (thin lines) at (a) $z = 22$ km and (b) $z = 19$ km. The contour interval is (a) 0.02 m s^{-1} and (b) 0.04 m s^{-1}. Dashed contours represent negative values. Areas where the minimum Ri between $z = 19$ km and 22 km is less than 0.25 are shaded. Thick solid contours in (a) denote averaged intrinsic phase velocities $\overline{u} - c_p = 0$, 1, 2, 3, and 4 m s^{-1}.

for $t = 8$–8.5 h (Fig. 9a), whereas gravity waves in region C appear along the edge of the breaking zone propagating downstream (Fig. 9b). The waves in region D are observed within an area where the intrinsic phase velocity is rather larger than the horizontal wind. Thick solid curves in Fig. 9a show contours of averaged intrinsic phase velocitiy $u - c_p$ in the height range of 19–22 km:

$$\overline{u} - c_p = \int_{z=19\,\mathrm{km}}^{z=22\,\mathrm{km}} (u - c_p)\, dz,$$

where $c_p = 2$ m s^{-1} as in Table 1. The value of $\overline{u} - c_p$ is smaller than 1 m s^{-1} for $x < 310$ km. Although the WKB approach is not applicable to the region where the wind varies with similar scale of gravity waves, it is natural to consider that the existence of the area where the intrinsic phase velocity becomes nearly equal to zero prevents secondary gravity waves from propagating freely though it to region D. Therefore it is deduced that the difference in the timing of appearance between gravity waves in regions C and D be due to the difference of the background wind field.

b. Reflection of generated waves

If the critical wavelength λ_c increases slowly in the direction of the wave propagation, linear theory predicts that plain stationary gravity waves are reflected at a height where $\lambda_c = \lambda_x$. Assuming that gravity waves are almost stationary in regions C and D, the heights where $\lambda_c = \lambda_x$ are about 31 and 16 km, respectively (Fig. 1 and Table 1). In fact the gravity waves are reflected at these altitudes, as shown in Fig. 6. Departure from zero phase velocity of about 2 m s^{-1} modifies the reflection altitudes only slightly to $z \approx 35$ km for region C and $z \approx 15$ km for region D. The wave reflection in the upper stratosphere is also observed in the power spectra (Fig. 7) where strong and weak powers in w appear alternately in the vertical. The winter stratosphere acts as a wave duct for small-scale gravity waves and such trapped gravity waves can be generated in association with breaking of large-scale gravity waves propagating from the lower troposphere as shown in this study.

c. Observational evidence

The secondary gravity waves simulated in this study have short horizontal wavelengths $\lambda_x = 3$–8 km, long vertical wavelengths $\lambda_z = 3$–20 km, and small phase velocities $c_p = -1.5$ to $+4$ m s^{-1}. Among the observational studies reporting strong vertical wind fluctuations in the middle atmosphere, Sato and Hirota (1988) first succeeded in determining directly horizontal structures of gravity waves with the MU radar (Shigaraki, Japan; 35°N, 136°E) by using its multibeam capability to observe the winter lower stratosphere. The determined wave parameters were $\lambda_x < 5$–30 km, $\lambda_z \approx 5$ km, and $c_p = 1$–5 m s^{-1} in the height range of 12–22 km; these parameters are very similar to the secondary waves found in this study.

A subtropical westerly jet ($\approx$80 m s^{-1}) was located at an altitude of around 12 km during the observation period of Sato and Hirota (1988). The small-scale gravity waves as observed by them cannot propagate from the ground through the jet to the lower stratosphere. In the winter, a strong monsoon wind (west-north-westerly wind) blows over Japan and mountain waves develop frequently. On the windward side of the MU radar in winter, there are mountains of about 1 km in height. Therefore, it is reasonable to consider that the gravity waves detected by Sato and Hirota (1988) were the secondary gravity waves generated in association with breaking of mountain waves as observed in our simulation. Sato (1990) showed a high correlation between the activity of vertical wind disturbances observed by the MU radar in winter and the horizontal wind at 90 kPa ($\approx$1 km) at Yonago, which is about 250 km west-northwest of the radar site and on the upstream side of

the mountain. This fact also suggests that mountain waves control the small-scale wave activity in the downstream area.

5. Conclusions

Using a two-dimensional nonhydrostatic model, we have simulated the generation of new secondary gravity waves associated with the breaking of primary mountain waves in the lower stratosphere. The wave parameters of the secondary waves are 3–8 and 2–20 km for horizontal and vertical wavelengths, respectively, and -1.5 to $+4$ m s^{-1} for phase velocities, which well satisfy the theoretical dispersion relation of internal gravity waves. The amplitudes of the vertical velocities of the secondary waves are 0.1–0.2 m s^{-1}. These secondary gravity waves propagate horizontally as well as vertically from the primary wave-breaking area because of nonzero phase velocities. These waves are reflected vertically where vertical wavelengths become infinite and are trapped in the stratosphere, because the vertical structure of the environmental wind and temperature in the winter lower stratosphere works as a duct for gravity waves with such short horizontal wavelengths. The horizontal wavelengths of the secondary gravity waves studied in this paper are much shorter than the primary mountain wave ($\approx$180 km). Hence, these waves are clearly different from the downward propagating waves found by Bacmeister and Shoeberl (1989), whose horizontal wavelength was almost the same as that of the primary mountain wave.

Comparing the wave parameters with unstable modes in linear and nonlinear analyses made in the previous studies, three kinds of generation mechanisms are suggested for the secondary internal gravity waves: 1) an obstacle effect due to convective motions in the shear flow near the breaking zone (region A); 2) an unstable normal mode of slowly growing breaking mountain wave (SP92; Scinocca and Peltier 1993) (region B); 3) a mode growing nonlinearly at the edge of the breaking mountain wave (regions C and D). Mountain wave simulations over narrower mountains (a = 10 km and 5 km) have also been performed (not shown). Secondary gravity waves similar to those reported in this paper are also generated in these cases. This fact suggests that the generation mechanisms of small-scale gravity waves as discussed in this study act over a wide range of breaking mountain wave parameters.

Recent numerical simulations indicate that disturbances with three-dimensional structures appear in well-developed K–H billows (e.g., Fritts et al. 1996). Since the mechanisms of the secondary wave generation simulated in this paper are not the K–H instability, it is uncertain whether the three-dimensionality is essential to generate the secondary gravity waves associated with mountain wave breaking in the real atmosphere. However, it is necessary to confirm our results by three-dimensional models in the future. Time-dependent environmental flow is possibly of more relevance to mountain wave generation (e.g., Lott and Teitelbaum 1993) and consequently to the generation mechanisms of the secondary gravity waves. Thus, studies incorporating the time evolution of the background flow should be also made in order to fully understand the dynamical effects of mountains on the atmosphere.

Acknowledgments. We acknowledge valuable discussions with Dr. David C. Fritts. The numerical experiment was done on the KDK system at Radio Atmospheric Science Center and on the Data Processing Center at Kyoto University. This research was supported by the Center for Climate System Research, University of Tokyo; a Grant-in-Aid for Scientific Research B-2-08458146 of the Ministry of Education, Science and Culture, Japan (TS); and the Nissan Science Foundation (KS). Figures were drawn by the GFD-DENNOU Library.

REFERENCES

Bacmeister, J. T., and M. R. Schoeberl, 1989: Breakdown of vertically propagating two-dimensional gravity waves forced by orography. *J. Atmos. Sci.,* **46,** 2109–2134.

Bougeault, P., and Coauthors, 1993: The atmospheric momentum budget over a major mountain range: First results of the PYREX field program. *Ann. Geophys.,* **11,** 395–418.

Chandrasekhar, S., 1961: *Hydrodynamic and Hydromagnetic Stability.* Clarendon Press, 652 pp.

Clark, T. L., T. Hauf, and J. P. Kuettner, 1986: Convectively forced internal gravity waves: Results from two-dimensional numerical experiments. *Quart. J. Roy. Meteor. Soc.,* **112,** 899–925.

Dunkerton, T. J., 1997a: The role of gravity waves in the quasi-biennial oscillation. *J. Geophys. Res.,* **102,** 26 053–26 076.

——, 1997b: Shear instability of internal inertia-gravity waves. *J. Atmos. Sci.,* **54,** 1628–1641.

——, and R. E. Robins, 1992: Radiating and nonradiating modes of secondary instability in a gravity-wave critical layer. *J. Atmos. Sci.,* **49,** 2546–2559.

Fovell, R., D. Durran, and J. R. Holton, 1992: Numerical simulations of convectively generated stratospheric gravity waves. *J. Atmos. Sci.,* **49,** 1427–1442.

Fritts, D. C., 1982: Shear excitation of atmospheric gravity waves. *J. Atmos. Sci.,* **39,** 1936–1952.

——, and G. D. Nastrom, 1992: Sources of mesoscale variability of gravity waves. Part II: Frontal, convective, and jet stream excitation. *J. Atmos. Sci.,* **49,** 111–147.

——, T. L. Palmer, Ø. Andreassen, and I. Lie, 1996: Evolution and breakdown of Kelvin–Helmholtz billows in stratified compressible flows. Part I: Comparison of two- and three-dimensional flows. *J. Atmos. Sci.,* **53,** 3173–3191.

Gal-Chen, T., and R. C. J. Somerville, 1975: On the use of a coordinate transformation for the solution of the Navier–Stokes equations. *J. Comput. Phys.,* **17,** 209–228.

Hines, C. O., 1997a: Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere. Part 1: Basic formulation *J. Atmos. Solar-Terr. Phys.,* **59,** 371–386.

——, 1997b: Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere. Part 2: Broad and quasi monochromatic spectra, and implementation. *J. Atmos. Solar-Terr. Phys.,* **59,** 387–400.

Holton, J. R., 1982: The role of gravity wave induced drag and diffusion in the momentum budget of the mesopshere. *J. Atmos. Sci.,* **39,** 791–799.

Iwasaki, T., 1989: A parameterization scheme of orographic gravity

wave drag with two different vertical partitionings. *J. Meteor. Soc. Japan,* **67,** 11–27.
Klemp, J. B., and R. B. Wilhelmson, 1978: The simulation of three-dimensional convective storm dynamics. *J. Atmos. Sci.,* **35,** 1070–1096.
Kuettner, J. P., 1986: *The Aim and Conduct of ALPEX*. Vol. 27, GARP Publ. Ser., WMO, 3–14 pp.
Lilly, D. K., 1962: On the numerical simulation of buoyant convection. *Tellus,* **14,** 148–172.
——, and P. J. Kennedy, 1973: Observations of a stationary mountain wave and its associated momentum flux and energy dissipation. *J. Atmos. Sci.,* **30,** 1135–1152.
Lindzen, R. S., 1981: Turbulence and stress owing to gravity wave and tidal breakdown. *J. Geophys. Res.,* **86,** 9707–9714.
Long, R. R., 1953: Some aspects of the flow of stratified fluids. Part I: A theoretical investigation. *Tellus,* **5,** 42–58.
Lott, F., and H. Teitelbaum, 1993: Linear unsteady mountain waves. *Tellus,* **451,** 201–220.
Matsuno, T., 1982: A quasi one-dimensional model of the middle atmosphere circulation interacting with internal gravity waves. *J. Meteor. Soc. Japan,* **60,** 215–226.
McFarlane, N. A., 1987: The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere. *J. Atmos. Sci.,* **44,** 1775–1800.
McIntyre, M. E., and M. A. Weissman, 1978: On radiating instabilities and resonant overreflection. *J. Atmos. Sci.,* **35,** 1190–1196.
Miller, M. J., and A. J. Thorpe, 1981: Radiation conditions for the lateral boundaries of limited-area numerical models. *Quart. J. Roy. Meteor. Soc.,* **107,** 615–628.
Orlanski, I., 1976: A simple boundary condition for unbounded hyperbolic flows. *J. Comput. Phys.,* **21,** 251–269.
Palmer, T. N., G. J. Shutts, and R. Swinbank, 1986: Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parametrization. *Quart. J. Roy. Meteor. Soc.,* **112,** 1001–1039.
Prusa, J. M., P. K. Smolarkiewicz, and R. R. Garcia, 1996: Propagation and breaking at high altitudes of gravity waves excited by tropospheric forcing. *J. Atmos. Sci.,* **53,** 2186–2216.
Sato, K., 1990: Vertical wind disturbances in the troposphere and lower stratosphere observed by the MU radar. *J. Atmos. Sci.,* **47,** 2803–2817.
——, 1994: A statistical study of the structure, saturation and sources of inertio-gravity waves in the lower stratosphere observed with the MU radar. *J. Atmos. Terr. Phys.,* **56,** 755–774.
——, and I. Hirota, 1988: Small-scale gravity waves in the lower stratosphere revealed by the MU radar multi-beam observation. *J. Meteor. Soc. Japan,* **66,** 987–999.
Satomura, T., 1989: Compressible flow simulations on numerically generated grids. *J. Meteor. Soc. Japan,* **67,** 473–482.
Scinocca, J. F., and W. R. Peltier, 1993: The instability of Long's stationary solution and the evolution toward severe downslope windstorm flow. Part I: Nested grid numerical simulations. *J. Atmos. Sci.,* **50,** 2245–2263.
Smith, R. B., 1977: The influence of mountains on the atmosphere. *Advances in Geophsyics,* Vol. 21, Academic Press, 87–230.
Sutherland, B. R., and W. R. Peltier, 1992: The stability of stratified jets. *Geophys. Astrophys. Fluid Dyn.,* **66,** 101–131.
——, C. P. Caulfield, and W. R. Peltier, 1994: Internal gravity wave generation and hydrodynamic instability. *J. Atmos. Sci.,* **51,** 3261–3280.
Tanaka, H., and M. D. Yamanaka, 1985: Atmospheric circulation in the lower stratosphere induced by the mesoscale mountain wave breakdown. *J. Meteor. Soc. Japan,* **63,** 1047–1054.
Walterscheid, R. L., and G. Schubert, 1990: Nonlinear evolution of an upward propagating gravity wave: Overturning, convection, transience and turbulence. *J. Atmos. Sci.,* **47,** 101–125.