A New Gravity Wave Parameterization Including Three-Dimensional Propagation

Arata AMEMIYA and Kaoru SATO

Department of Earth and Planetary Science, University of Tokyo, Bunkyo, Japan

(Manuscript received 21 August 2015, in final form 28 January 2016)

Abstract

Recent studies suggest that the horizontal propagation of gravity waves (GWs) is important in the spatial distribution of gravity wave forcing (GWF), especially during winter in the high latitudes of the Southern Hemisphere (SH). However, most standard gravity wave parameterizations (GWPs) treat GW propagation simply in the vertical. In this study, a new orographic GWP that includes three-dimensional (3D) GW propagation is developed and its impact on large-scale dynamical fields is examined. Our GWP calculates the horizontal locations and changes in the 3D wavenumbers of GWs explicitly through vertical integration of the ray tracing equations. The GWF due to wave refraction, which occurs in inhomogeneous background fields even without wave dissipation, is also calculated. In addition, the computational cost of parallelization is greatly reduced by adopting a Taylor’s series approximation for the horizontal gradient of the background fields needed for the ray tracing calculation. Two numerical experiments are performed using the Model for Interdisciplinary Research of Climate (MIROC)-AGCM: one uses the new orographic GWP and the other uses a conventional GWP. In the experiment with the new GWP, the westward GWF is enhanced in the SH winter mesosphere above the core of the polar night jet. This enhancement results from the significant latitudinal propagation of the parameterized GWs toward the jet axis. The zonal wind is slightly stronger in the SH winter polar upper stratosphere, which is consistent with the differences in GWF caused mainly by refraction. However, the strength and seasonal evolution of the polar night jet is less affected by the different GWPs. This result may be because of the compensation by Eliassen–Palm flux divergence due to the resolved waves. These results suggest that 3D propagation in GWPs is potentially important for better representation of the momentum budget of the middle atmosphere in climate models.

Keywords gravity wave parameterization; climate model; middle atmosphere

1. Introduction

Gravity waves (GWs) play an essential role in the large-scale dynamics of the middle atmosphere, by re-distributing momentum through the wave–mean flow interaction, which is usually called GW forcing (GWF) (e.g., Fritts and Alexander 2003, Alexander et al. 2010). In the mesosphere, GWF drives strong global meridional flow from the winter pole to the summer pole and maintains a weak zonal wind layer near the mesopause (e.g., Plumb 2002). GWF also partly contributes to the formation of Brewer–Dobson circulation in the winter stratosphere (McLandress and Shepherd 2009; Butchart et al. 2010; Okamoto et al. 2011). In the tropical middle atmosphere, GWs are considered to be essential drivers of large-scale oscillations such as the Quasi-Biennial Oscillation and Semi-Annual Oscillation (Baldwin et al. 2001). In addition, GWF has been recently reported to contribute to the formation of a dynamically unstable field that is frequently observed in the mesosphere and to the generation of planetary-scale waves (Holton 1984; Norton and Thuburn 1996; Watanabe et al. 2009; Ern et al. 2011; Sato and Nomoto 2015).

Despite their importance, GWs cannot be suffi-
ciently resolved even in the current generation of climate models because GWs have small horizontal wavelengths. Thus, most climate models employ GW parameterizations (GWPs) to include the momentum deposition by GWs on resolved dynamical fields in the form of GWF. Currently, many varieties of GWP are used both for orographic and non-orographic GWs due to the growing demand for better representation of the middle atmosphere in climate projections (McLandress 1998; Kim et al. 2003). Recent developments include source specifications of non-orographic GWs originating from convection (Beres et al. 2005, Song and Chun 2005) and jet–front systems (Charron and Manzini 2002; Richter et al. 2010) based on physical processes and stochastic launching of GWs to imitate their intermittency (Eckermann 2011; Lott and Guez 2013; de la Camara and Lott 2015).

To reduce computational time, various approximations and simplifications are generally included in the GWPs. One common simplification is the columnar treatment of GWs, in which GW propagation is approximated as occurring purely in the vertical. In the rest of this paper, the GWP schemes using this approximation are referred to as “columnar” GWPs. The validity of this treatment, however, is not always assured. Recent studies based on high-resolution satellite observations (Preusse et al. 2009) and/or using a high-resolution general circulation model (GCM) with no GWPs (Sato et al. 2009, 2012) have shown that GWs propagate over a significant horizontal distance before reaching the upper stratosphere and mesosphere. Significant horizontal propagation occurs even for orographic GWs through the advection of the mean flow perpendicular to the wave-number vector and through the change in the wave-number vector due to refraction in spatially varying background fields. In addition, Geller et al. (2013) recently showed through a comparison of the absolute momentum fluxes among satellite observations, GW-permitting models and standard climate models with parameterized GWs, that the latitudinal distribution of parameterized GW fluxes are not realistic in several aspects. They suggested that the unrealistic distribution is partly due to an improper specification of non-orographic source momentum fluxes and partly due to the treatment of GW propagation in the GWP schemes.

The inclusion of three-dimensional (3D) propagation in spatially–varying background fields in parameterizations modifies GWF in two ways. First, the location where GWF is exerted may be horizontally distant from the GW sources. Second, the refraction due to the spatial gradient of the background fields causes GWF even without wave dissipation. The latter mechanism was first formulated by Bühler and McIntyre (2003) as the “remote recoil” effect associated with the balanced flow. Both of these GWF modifications are neglected in columnar GWPs.

Several studies have examined the effect of 3D propagation on momentum deposition using off-line ray–tracing calculation. For orographic GWs, Hasha et al. (2008) calculated the global-mean angular pseudomomentum flux with and without refraction and examined the difference. They concluded that a large difference is observed only for GWs near the inertial frequency and the overall impact of 3D propagation and refraction is not large. On the other hand, Durran (2009) suggested, based on a study by Chen et al. (2005), that horizontal and temporal variations in the large scale flow produced nontrivial modifications in the momentum flux of GWs. For non-orographic GWs, Kalisch et al. (2014) found large differences in the global distribution of the monthly mean GWF between a case assuming vertical propagation and a case with realistic oblique propagation. These results were obtained using the same source spectrum as satellite observations.

However, few studies, with the exception of one by Song and Chun (2008), have investigated the effect of the 3D propagation of GWs in climate models. Song and Chun developed a GWP that includes horizontal propagation for convectively generated GWs for the purpose of improving the representation of GWF in the equatorial middle atmosphere. Their scheme traces each GW packet with respect to time following its group velocity. Although their treatment improved the representation of time-dependent GWF, it had the drawback of high computational costs. In addition, only GWF due to wave dissipation is included, whereas GWF due to wave refraction is ignored in their parameterization.

For orographic GWs, the impact of 3D propagation on dynamical fields over long time scales has not been examined so far, even though it may cause large GWF around the core of the westerly jet in the winter stratosphere and mesosphere, as suggested by a high-resolution GCM simulation (Sato et al. 2009, 2012). The improvement of the latitudinal distribution of the parameterized GWF may have a significant impact on the seasonal evolution of the polar night jet in the Southern Hemisphere (SH) (McLandress et al. 2012).

In this study, we developed an orographic GWP scheme that considers the effects of 3D propagation of GWs. The scheme explicitly calculates GW horizontal
propagation and GWF due to both wave dissipation and refraction in spatially–varying background fields. An assumption for the background field is given in the scheme for effective parallelization. To examine these effects on simulated large-scale fields, we conducted a pair of numerical experiments using a general circulation model: one with the new GWP, and one with a conventional GWP. We focus particularly on the modified GWF around the polar night jet in the stratosphere and mesosphere of the winter SH.

This paper is organized as follows. Section 2 gives a theoretical description of GW propagation and interaction with a large-scale flow, followed by details of the new GWP scheme. In Section 3, we briefly describe the configuration of a general circulation model for the numerical experiments. In Section 4, the results of the experiments are shown, and the physical interpretation is discussed. Section 5 provides a summary and concluding remarks.

2. Methodology

In this section, we introduce the formulation of the new GWP scheme. First, the basis of the ray tracing calculation is described in Subsection 2.1, followed by the method of calculating the GWF in Subsection 2.2. In Subsection 2.3, the principal concept of our scheme is described. A method for effective parallelization is described in Subsection 2.4.

For simplicity, the formulation of the basic equations is described in Cartesian coordinates. The actual calculation of our GWP is performed based on equations in a spherical geometry, as demonstrated by Hasha et al. (2008) and described in their Appendix.

2.1 Ray–tracing equations

The temporal evolutions of the location and wavenumbers of GW packets are calculated using ray tracing equations. For inertia–gravity waves in the atmosphere with a density scale height, the dispersion relation is

\[ \frac{d_x}{dt} = c_{gx} = \bar{u} + \frac{k(N^2 - \tilde{\omega}^2)}{\tilde{\omega} \Delta} , \]

\[ \frac{d_y}{dt} = c_{gy} = \bar{v} + \frac{l(N^2 - \tilde{\omega}^2)}{\tilde{\omega} \Delta} , \]

\[ \frac{d_z}{dt} = c_{gz} = - \frac{m(\tilde{\omega}^2 - f^2)}{\tilde{\omega} \Delta} , \]

\[ \frac{d_k}{dt} = -k \frac{\partial N^2}{\partial x} - l \frac{\partial N^2}{\partial x} - \frac{1}{2\tilde{\omega} \Delta} \left( \frac{\partial N^2}{\partial x} (k^2 + l^2) - \frac{\partial \alpha^2}{\partial x} (\tilde{\omega}^2 - f^2) \right) , \]

\[ \frac{d_l}{dt} = -k \frac{\partial \bar{n}}{\partial y} - l \frac{\partial \bar{n}}{\partial y} - \frac{1}{2\tilde{\omega} \Delta} \left( \frac{\partial N^2}{\partial y} (k^2 + l^2) - \frac{\partial \alpha^2}{\partial y} (\tilde{\omega}^2 - f^2) \right) , \]

\[ \frac{d_m}{dt} = -k \frac{\partial \bar{n}}{\partial z} - l \frac{\partial \bar{n}}{\partial z} - \frac{1}{2\tilde{\omega} \Delta} \left( \frac{\partial N^2}{\partial z} (k^2 + l^2) - \frac{\partial \alpha^2}{\partial z} (\tilde{\omega}^2 - f^2) \right) , \]

where \( \Delta \equiv k^2 + l^2 + m^2 + \alpha^2 \) and \( d_g/\Delta \equiv \partial/\partial t + c_g \cdot \nabla \).

2.2 Calculation of wave amplitudes and GWF

In a standard ray–tracing model, the amplitude of each wave packet is obtained from the conservation law under several assumptions. For the propagation without dissipation, the “wave action,” which is a quantity with second-order perturbation amplification without dissipation, is conserved for the wave packet. The wave action density \( A \) is defined as

\[ A = \frac{\rho I_2}{2} \left( \bar{u}^2 + \bar{v}^2 + \bar{w}^2 + \frac{1}{N^2} \bar{b}^2 \right) \]

is the GW energy density that satisfies the following equation:

\[ \frac{\partial A}{\partial t} + \nabla \cdot (c_g A) = - \kappa A , \]

where \( \kappa \) is the dissipation rate, which is usually derived from the background turbulent viscosity and radiative relaxation coefficients. When the GW amplitudes become significantly large, nonlinear effects may lead GWs to break and/or partially dissipate. Such processes occur when GWs approach their critical level and/or reach very low density in the upper atmosphere. The partial dissipation of GWs is often
assumed in ray–tracing models based on linear instability theories (Marks and Eckermann 1995). In most GWPs, hypothetical wave saturation is considered to limit the wave amplitude (Lindzen 1981).

The wave forcings $\overline{X}$ and $\overline{Y}$ exerted on large-scale (resolved) momentum equations are expressed as the convergence of the pseudomomentum flux (e.g., Kinoshita et al. 2010):

$$\overline{X} = -\frac{1}{\rho_0} \nabla \cdot F_U, \quad (9)$$

$$\overline{Y} = -\frac{1}{\rho_0} \nabla \cdot F_V, \quad (10)$$

where $F_U$ and $F_V$ are the fluxes of zonal and meridional components of pseudomomentum and are defined as

$$F_U = kAc_g, \quad F_V = lAc_g. \quad (11)$$

### 2.3 New GWP including 3D propagation

Basically, most GWPs are designed to operate independently for each vertical column. This secures effective parallelization, similar to the parameterizations for other physical processes in GCMs. The following simplifications are needed in such columnar GWPs:

(a) Ignoring the background inhomogeneity and spherical geometry

(b) Ignoring the horizontal GW propagation

(c) Ignoring the time duration for propagation

Assumptions (a) and (b) are natural consequences of the columnar approach. With these two assumptions, it is clear from Eqs. (5) and (6) that we can take $x, y, k$ and $l$ as constants for each wave. The vertical wavenumber $m$ can change but is easily obtained from the dispersion relation (1) at each vertical level. Thus, the explicit integration of the ray–tracing equation is not needed for conventional GWPs.

Song and Chun (2008) made the first attempt to develop a GWP that explicitly addresses 3D propagation. Their scheme uses none of the simplifications above. Instead, they used explicit integration of the ray–tracing equations with respect to time to obtain 3D paths by taking into consideration the time duration needed for GW propagation in calculating GWF fields. Their scheme was developed to represent the time-dependent spatial distribution of forces by GWs originating from convection. However, its computational cost is not small.

By contrast, the new GWP scheme proposed in this study incorporates 3D propagation with a practical computational cost by keeping assumption (c). Although the time duration needed for GW propagation to the top of the model is generally longer than the single time step of GCMs, this treatment is reasonable at least in the region below the upper mesosphere where tidal motions are not dominant. The significant advantage of this treatment is that the time–integration of the ray–tracing equations can be transformed to the vertical integration, which makes the calculation quite effective.

The ground-based frequency $\omega$ is set as a constant in this study, although several previous studies have discussed possible changes in $\omega$ along the rays following the temporal tendencies of the background fields (e.g., Durran 2009; Senf and Achatz 2011). It is reasonable to neglect them in order to make it consistent with assumption (c) (See Section 5 for further discussion). For the orographic GWs addressed in this study, the assumption is simply that $\omega = 0$.

For each GW, the horizontal location and wave-numbers are calculated at each vertical level of the GCM by following discretized ray–tracing equations:

$$x(z_{k+\frac{1}{2}}) = x(z_{k-\frac{1}{2}})$$

$$+ \left( z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}} \right) \frac{1}{c_{gz}} \frac{\partial \omega}{\partial k} \left( z_{k-\frac{1}{2}} \right), \quad (12)$$

$$y(z_{k+\frac{1}{2}}) = y(z_{k-\frac{1}{2}})$$

$$+ \left( z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}} \right) \frac{1}{c_{gz}} \frac{\partial \omega}{\partial l} \left( z_{k-\frac{1}{2}} \right), \quad (13)$$

$$k(z_{k+\frac{1}{2}}) = k(z_{k-\frac{1}{2}})$$

$$- \left( z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}} \right) \frac{1}{c_{gz}} \frac{\partial \omega}{\partial x} \left( z_{k-\frac{1}{2}} \right), \quad (14)$$

$$l(z_{k+\frac{1}{2}}) = l(z_{k-\frac{1}{2}})$$

$$- \left( z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}} \right) \frac{1}{c_{gz}} \frac{\partial \omega}{\partial y} \left( z_{k-\frac{1}{2}} \right). \quad (15)$$

The calculation of the wave action density $A$ is performed following the conservation law (8) and the saturation criterion to express wave dissipation in a similar way to Marks and Eckermann (1995) (i.e., the
horizontal wind amplitude parallel to the horizontal wavenumber vector is not greater than the magnitude of the intrinsic phase speed). Equation (8) indicates that the total amount of wave action flux $c_gA$ is conserved for a steady non-dissipative wave packet. It should be noted that the wave action density $A$, which is obtained by removing higher wavenumber components, may be more appropriate for the ray calculation. The upper limit of $A$, which is obtained with limited spatial model resolution and computational resources (Hasha et al. 2008). Thus, we assume that the cross-section of the ray tube is constant, following previous studies. This assumption is equivalent to ignoring the divergence of the group velocity $c_g$ on a surface perpendicular to the direction of $c_g$. Under this assumption, the conservation law becomes

$$A|c_g| = \text{const.},$$

with which we can readily calculate $A$ along with ray trajectories. The upper limit of $A$, which is obtained by the saturation criterion, is given in order to describe the partial breaking of GWs.

The forcing by each GW is calculated as a vertical convergence of the pseudomomentum flux:

$$X(z_k) = -\frac{1}{\rho_0} \left( \frac{F_{uz}\left(z_{k+\frac{1}{2}}\right) - F_{uz}\left(z_{k-\frac{1}{2}}\right)}{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}} \right),$$

$$Y(z_k) = -\frac{1}{\rho_0} \left( \frac{F_{uz}\left(z_{k+\frac{1}{2}}\right) - F_{uz}\left(z_{k-\frac{1}{2}}\right)}{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}} \right),$$

where the vertical components of pseudomomentum fluxes $F_{uz}$ and $F_{uz}$ are obtained by $k$ and $A$ using (11). The location where these GWFs are exerted is determined by interpolation between $\left(x\left(z_{k+\frac{1}{2}}\right), y\left(z_{k+\frac{1}{2}}\right)\right)$ and $\left(x\left(z_{k-\frac{1}{2}}\right), y\left(z_{k-\frac{1}{2}}\right)\right)$, and the GWFs are simply distributed to four nearby grid points via linear weighting.

Several artificial conditions are given for the termination of the ray–tracing calculation. First, the lower limit for the vertical group velocity is given to avoid a large truncation error near the critical level. GWs with vertical propagation slower than this limit are totally dissipated, and the resultant wave forcing is given at that point. Second, GWs that reach the top of the model domain are artificially dissipated and removed as well. Third, GWs are removed without depositing momentum when internal reflection occurs where the squared vertical wavenumber is negative, as determined using the dispersion relation Eq. (1).

2.4 Approximation for background field variables

The 3D ray–tracing calculation requires parameters that describe the horizontal structure of the background field, which is essentially different from the columnar GWP schemes. This requires an additional message–passing interface (MPI) communication for ray calculation.

To minimize the cost of MPI communication, the background field variables and their first horizontal derivatives are approximated by the first-order truncation of a Taylor series expansion. This “quasi-column” approach only requires the information in one column and allows us to calculate ray paths in the area outside the column. For models employing a spectral dynamical core, the first horizontal derivatives of wind and temperature are easily obtained by an inverse spectral transformation. Similarly, a spatially smoothed background field is also easy to obtain by removing higher wavenumber components, which may be more appropriate for the ray calculation. This quasi-column treatment enables a significant reduction of MPI communication that is several times shorter than that needed for the calculation without the approximation.

The validity of this approximation, along with that of the vertical integration of ray–tracing equations, is confirmed by additional offline calculations. A detailed discussion is provided in Appendix.

3. Experimental setup

We implemented our orographic GWP scheme to a high-top version of the Model for Interdisciplinary Research on Climate (MIROC)-AGCM. The model has 80 vertical levels from the surface to the uppermost mesosphere (~ 0.006 hPa) and T42 horizontal resolution (See Watanabe et al. (2008) for details). In this study, we focused on the effect of including 3D propagation on the orographic GWP because orographic GWs are one of the most important GWs (e.g., Nastrom and Fritts 1992; Sato et al. 2009) and because their source specification is relatively less ambiguous compared with that for non-orographic GWs. Nevertheless, the basic parts of this new scheme are also applicable to non-orographic GWP.
The pseudomomentum fluxes and horizontal wave-numbers of two representative orographic GWs are given as the source for each grid on land, following the scheme by Scinocca and McFarlane (2000). The sign of the horizontal wavenumber vector, which is inherently chosen arbitrarily, is determined such that the intrinsic frequency is positive without losing generality. For example, zonal wavenumbers are negative in eastward winds at the surface.

Two experiments were performed to examine the impact of our new 3D GWP scheme. One experiment used the standard columnar orographic GWP, which we hereafter refer to as “CONTROL.” The other used the new scheme, which we call “OGW3D.” Note that the only differences between the two orographic GWPs are the presence or absence of 3D propagation and changes in the horizontal wavenumbers. All model settings other than those of the orographic GWPs are the same. For the non-orographic GWP, Hines’ scheme (Hines 1997) was used and included
prescribed isotropic and homogeneous sources at approximately 680 hPa, which are equivalent to 1.2 m s$^{-1}$ in horizontal wind standard deviation. The surface conditions, such as land albedo, sea surface temperature, and sea ice, and the mixing ratios of ozone and other minor constituents are prescribed as boundary conditions. These external data are the same as described in Watanabe et al. (2008). The monthly climatology for 1980–2000 is used to focus on dynamically induced variability. Simulations were performed for the two experiments over 50 model years.

4. Results

4.1 Difference in the mean field and GWF

Figures 1a and 1b show meridional cross-sections of the monthly mean climatology of the zonal mean zonal wind in July for OGW3D and CONTROL, respectively, and Fig. 1c shows their differences. The maximum zonal wind on the poleward side of the polar night jet is slightly stronger for OGW3D than for CONTROL. Figure 1d shows the difference in zonal mean temperature between the two experiments. The zonal mean temperature in the winter polar region is slightly colder in the lower stratosphere in OGW3D than in CONTROL. This difference is in the thermal wind balance with that in the zonal mean zonal wind.

Next, we compare the seasonal evolution of the polar night jet. Figures 2a and 2b show the time–height sections of the climatology of the zonal mean zonal wind averaged over 50–70°S for the two experiments. Figure 2c shows their difference. Figure 2d shows the difference in zonal mean temperature averaged over 60–90°S. Interestingly, the differences are quite small in terms of the strength of the polar night jet and its temporal evolution. The zonal mean zonal
wind in October is weaker in OGW3D by several m s\(^{-1}\), but it is within the range of interannual variability (approximately 10 m s\(^{-1}\)) of this season. The lower stratosphere is 1–2 K warmer in October in OGW3D than in CONTROL, but the statistical significance of this difference is also low. This statistical insignificance is primarily due to the large interannual variability in the strength of the polar night jet, which is likely related to the large variability in the planetary wave activity in the spring.

Next, we compare horizontal maps of the monthly mean climatology of the pseudomomentum flux fields of the parameterized orographic GWs for July at the 1 hPa surface (Figs. 3a, b). The difference in the horizontal distribution of the fluxes is clear. The GW pseudomomentum fluxes are broadly distributed primarily leeward of the orographic sources in OGW3D but are observed only above the orographic sources in CONTROL. This pattern is particularly clear in association with steep orography, such as the Antarctic Peninsula, where a large negative zonal GW flux is exerted leeward. This feature is consistent with satellite temperature observations (Preusse et al. 2002; Ern et al. 2011) and simulations by a high-resolution GCM model with no GWP (Sato et al. 2009, 2012). This implies that our GWP scheme successfully reproduced realistic GW horizontal propagation in OGW3D.

Figures 4a and 4c show the climatology of the zonal mean zonal component of GWF in the latitude–pressure cross section for July in the two experiments. A large negative zonal (i.e., westward) GWF is observed in the SH mesosphere in both experiments. In CONTROL, however, there is a distinct latitudinal gap spanning approximately 55–60°S, reflecting a gap in the orographic sources. By contrast, in OGW3D, the gap is obscured, and negative zonal forcing is enhanced at approximately 70°S in the mesosphere. In addition, there is a notable difference in the polar upper stratosphere: the negative zonal GWF is reduced more in OGW3D than in CONTROL.

The meridional component of GWF is shown in Figs. 4b and 4d for OGW3D and CONTROL, respectively. There are large differences, particularly in the stratosphere. In CONTROL (Fig. 4d), large meridional GWF is observed above the polar region, where GWs are launched with large meridional wavenumbers primarily from the steep border of the Antarctic continent. By contrast, in OGW3D (Fig. 4b), the meridional GWF in the stratosphere is enhanced. GWF is positive (negative) equatorward (poleward) of the jet axis latitude (~ 60°S) below 0.1 hPa and oppo-
Fig. 4. Climatology of zonal mean orographic GWF for (a, b) OGW3D and (c, d) CONTROL (a and c are zonal GWF and b and d are meridional GWF). The contour levels are 0.1, 0.3, 1.0 and 3.0 m s⁻¹ day⁻¹. The light shading indicates values smaller than −1.0 m s⁻¹ day⁻¹.

Figures 5a and 5b show the difference in the climatology of the zonal mean GWF fields for July between OGW3D and CONTROL. The main characteristic differences are summarized as the following three features: (i) intensification of the zonal deceleration at approximately 60°S in the mesosphere, (ii) reduction (and even reversal of the sign from negative to positive) of the zonal negative forcing in the polar stratosphere, and (iii) positive/negative meridional forcing in the middle/high latitudes in the stratosphere.

4.2 Impact of the refraction of GWs
The inclusion of 3D propagation can modify GWF through two processes: the change in propagation paths and in pseudomomentum. Note that the 3D wave mean–flow interaction can cause GWF due to refraction. This process does not need wave generation and/or dissipation (Bühler and McIntyre 2003, 2005). The existence of wave forcing due to refraction is clear by rewriting Eqs. (9) and (10) as in the following:
The first term on the right-hand side describes the forcing due to dissipation of GWs. The second term describes the forcing that is required to satisfy the momentum balance in case the horizontal wavenumber is modified in horizontally inhomogeneous mean fields. The second term is generally not zero but is neglected in conventional GWP schemes because horizontal wavenumbers are assumed to be constant.

Thus, it is interesting to examine how each term contributes to the differences in GWF between the two experiments. For this purpose, we performed a pair of offline calculations of GWF by applying the standard columnar GWP scheme and the 3D GWP scheme to the same background dynamical fields for July from the 50 years simulated by the OGW3D experiment and obtained the respective components of GWF on the right-hand sides of Eqs. (19) and (20) along with the GW parameters.

First, the results agreed well with those of the GWP-implemented simulations shown in Fig. 4 except for the uppermost mesosphere, where the tidal motions have large amplitudes. These results imply that the offline calculation successfully reproduced the GWF of the simulations (not shown in detail). The difference in the GWF components due to wave dissipation, represented by the first terms on the right-hand sides of Eqs. (19) and (20), are shown in Figs. 6a and 6d, respectively, whereas the difference due to wave refraction, represented by the second terms on the right-hand sides of Eqs. (19) and (20), are shown in Figs. 6b and 6e. Note that the GWFs shown in Figs. 6b and 6e are identical to those due to the refraction calculated in the new GWP because the GWF due to refraction is zero in the conventional columnar schemes. The sum of the first and second terms in Eqs. (19) and (20) are shown in Figs. 6c and 6f, respectively. The figures show that the offline calculation reproduced the essential features observed in the difference between the OGW3D and CONTROL experiments (Fig. 5), even though the columnar GWP scheme was applied to the background field of the OGW3D experiment for the offline calculation.

It is clear that the GWF due to refraction has large magnitudes comparable to those due to dissipation in
the stratosphere for both zonal and meridional components. In the zonal direction, GWF is largely eastward in the polar stratosphere and mesosphere and amounts to 0.3–1 m s⁻¹ day⁻¹. In the meridional direction, GWF is strongly northward and southward to the north and south of the polar night jet, respectively. These spatial distributions resemble the major differences observed in the stratosphere in Figs. 5a and 5b (as well as in Figs. 6c, f), indicating the difference mainly originates from GW refraction.

To see more direct evidence of the refraction, a statistical analysis was conducted on the wave parameters obtained from the offline calculations. Because GWP schemes used in the present study are based on the ray-tracing equations, the ray paths, wavenumbers, and amplitudes of respective GWs can be traced. Figure 7a shows a scatter diagram of the GW latitude at the source level (horizontal axis) versus that
at 1 hPa (vertical axis) for all GWs in July over fifty years. Note that waves that do not experience latitudinal propagation should stay on the diagonal. The distribution is broad in the diagram, indicating that a considerable number of the waves propagate latitudinally until they reach 1 hPa. Additionally, a relatively small number of waves are present at approximately 60°S at the source level, yet a considerable number of waves are present near that latitude at 1 hPa. This pattern indicates that a significant number of waves originate to the north and south of 60°S and contribute to GWF near 60°S above 1 hPa, where a clear gap in GWF was observed in CONTROL (Fig. 4).

Figure 7b is a scatter diagram of the zonal wavenumber at 100 hPa versus that at 1 hPa for all GWs launched from the equator to 60°S, and (d) for waves launched from 60°S to 90°S.

Fig. 7. Scatter diagrams of the parameterized GWs’ (a) latitude at the surface vs. that at 1 hPa, (b) zonal wavenumbers at 100 hPa vs. those at 1 hPa, and (c) meridional wavenumbers at 100 hPa vs. those at 1 hPa, for waves launched from the equator to 60°S, and (d) for waves launched from 60°S to 90°S.

in July over 50 years. The majority of waves plot on the left half of the diagram. This distribution indicates that most waves reaching 1 hPa have negative zonal wavenumbers at 100 hPa. This is a natural consequence of wind filtering by the strong westerly wind that prevails in the winter stratosphere. Another important feature is that a large number of waves on the right half of Fig. 7b plot above the diagonal, whereas a large number of waves on the left half plot below the diagonal. This pattern implies that the absolute value of the zonal wavenumber tends to increase through propagation into the stratosphere, which is consistent with the findings of Hasha et al. (2008). The reason for this tendency will be discussed later.

Figures 7c and 7d show the scatter diagrams for the
meridional wavenumbers for 60–90°S and 0–60°S, respectively. These latitude regions are shown separately because the waves cause meridional GWFs with different signs (see Fig. 8b). The difference in the refraction feature between the two latitude ranges is clear. Most meridional wavenumbers decrease with height at 0–60°S but increase with height at 60–90°S. This feature is consistent with the meridional propagation tendencies of GWs seen in Fig. 7a.

The distinct meridional refraction is attributable to strong latitudinal shear of the background zonal wind around the polar night jet. The ray–tracing equations are approximately expressed as

\[
\frac{d}{dt}\frac{\partial}{\partial l} = -k \frac{\partial u}{\partial x} - l \frac{\partial \bar{v}}{\partial y},
\]

(21)

\[
\frac{d}{dt}\frac{\partial}{\partial l} = -k \frac{\partial \bar{u}}{\partial y} - l \frac{\partial \bar{v}}{\partial y},
\]

(22)

Thus, a large positive \(\partial u/\partial y\) to the south of the jet leads to an increase in \(l\), whereas a large negative \(\partial u/\partial y\) to the north causes a decrease in \(l\) (note that \(k < 0\) for most waves; see Fig. 7b). In addition, this means that \(\partial l/\partial z > 0\) \((\partial l/\partial z < 0)\) to the south (north) of the jet; thus, southward (northward) GWF was produced due to the meridional refraction to the south (north) of the jet (see Eq. 20). This influence is consistent with the features observed below 0.1 hPa in Fig. 6d where the polar night jet is situated.

By contrast, the refraction in the zonal direction is not simple because the zonal structure of the background wind field is generally modulated by quasi-steady or transient planetary waves and synoptic-scale waves. In addition, orographic GWF is generally not zonally uniform but localized and typically has a zonally asymmetric structure, reflecting the source distribution. A robust tendency of increasing absolute values of the zonal wavenumber was observed in the offline calculation by Hasha et al. (2008). However, our results show a larger contribution of the refraction effect to the upward angular-momentum transport and may be consistent with the indication of Durran (2009). This zonal refraction is important in the momentum budget in the middle atmosphere and in the evolution of the horizontal structure of the background field, but the study of these details is beyond the scope of this paper.

4.3 Modification of the resolved wave forcings

As described above, the zonal mean zonal wind and temperature do not differ significantly in terms of the meridional cross section and seasonal variation between the CONTROL and OGW3D experiments for the SH winter and spring, regardless of the significant differences in GWF. Thus, we examined the possible reasons for this using an analysis based on the Transformed Eulerian Mean (TEM) formulation (Andrews et al. 1987).

From the zonal momentum equation, the difference between the two experiments, denoted by \(\Delta(*)\), is described as

\[
\frac{\partial (\Delta \bar{u})}{\partial t} - f \Delta \bar{v} \bar{u} \approx \frac{1}{\rho_0 a \cos \phi} \Delta (\nabla \cdot F) + \Delta \bar{X},
\]

(23)

where \(\bar{X}\) is the zonal mean eastward parameterized GWF calculated by Eq. (11) and the EP flux divergence, \(\nabla \cdot F\), is the zonal forcing by resolved disturbances. Importantly, the responses to a given \(\Delta \bar{X}\) can occur not only for the mean fields \(\Delta \bar{u}\) and \(\Delta \bar{v}\) but also for \(\nabla \cdot F\).

Figure 8a is identical to Fig. 5a and shows the differences in the zonal mean zonal component of the orographic GWF between the two experiments for July. Figures 8b and 8c show the difference in \(\nabla \cdot F\) and in the non-orographic GWF, respectively. The large negative difference observed in the zonal GWF in the mesosphere above the polar night jet (i.e., 0.5–0.05 hPa at 60–80°S and 0.1–0.01 hPa at approximately 50°S) is largely compensated by a positive difference in \(\nabla \cdot F\) of the same order of magnitude. This may be due to a compensation mechanism between parameterized and resolved waves, as suggested by recent studies (McLandress et al. 2012; Cohen et al. 2013, 2014; Scheffler and Pulido 2015). The difference in non-orographic GWF (Fig. 8c) is relatively large in the upper mesosphere. Because we used the Hines scheme for non-orographic GWF with the same configuration for both experiments, the difference is attributable to the background fields. The large positive difference in the non-orographic GWF at high latitudes above 0.1 hPa appears to partly cancel the negative difference in the orographic GWF, similar to the resolved wave forcing, although the statistical significance is not large.

Consequently, the total difference, as shown in Fig. 8d, is not large above 0.3 hPa because of this cancellation. This is consistent with the insignificant difference in the high-latitude winter mesosphere between the two experiments. The total difference in the polar upper stratosphere is also not significant and is likely due to partial cancellation between wave forcings. Note, however, that there is a significant positive difference in zonal wind in the polar upper strato-
Fig. 8. The differences in (a) the zonal component of orographic GWF, (b) EP flux divergence, (c) the zonal component of non-orographic GWF, and (d) total parameterized and resolved wave forcing between OGW3D and CONTROL based on monthly climatology for July. The contour levels are 0.1, 0.3, 1.0, 3.0, and 10.0 m s$^{-1}$ day$^{-1}$. The light and dark shading indicates absolute values greater than 1 and 10 m s$^{-1}$ day$^{-1}$, respectively.
sphere (Fig. 1c), which is consistent with the difference in orographic GWF (Fig. 8a).

5. Summary and concluding remarks

In this study, a new orographic GWP scheme that includes 3D propagation has been developed. By performing two experiments using the MIROC-AGCM, one with the new GWP (OGW3D) and one with a conventional GWP (CONTROL), the impact on the large-scale dynamical fields was examined, with a focus on the field in July, when a strong stratospheric polar night jet is present in the SH.

The new GWP scheme has the following characteristics:
1. The change in the horizontal location and wave-numbers of individual GWs are calculated explicitly using discretized ray–tracing equations.
2. The integration of the ray–tracing equation is performed with respect to height instead of time, assuming the propagation time duration is negligible. The calculation is also approximated using a Taylor series expansion for background fields, which reduces MPI communication in parallel computing.
3. The calculation of wave forcing is performed according to generalized 3D wave–mean interaction theory, which includes GWF arising from refraction.

The implementation of the 3D propagation in orographic GWP resulted in significant differences in the distribution of GWF during the SH winter. There are three major differences. First, the negative zonal GWF in the high latitude mesosphere above the polar night jet is significant in OGW3D, whereas a gap exists in the GWF in CONTROL because of a lack of orographic sources. Second, a negative zonal GWP in the polar stratosphere is largely diminished in OGW3D. This attenuation is mainly due to cancellation by positive zonal GWF arising from refraction. Third, meridional GWF in the stratosphere is significantly enhanced and exhibits distinct dipole structure: positive (northward) for 0–60°S and negative (southward) for 60–90°S. The third feature is attributable to the refraction-generated GWF associated with strong shear in the background wind.

Regardless of the significant differences in GWF, the strength of the polar night jet and its seasonal evolution are not substantially different. The TEM analysis revealed that the difference in zonal GWF is largely compensated by the EP flux divergence of resolved waves as well as by non-orographic GWF. Thus, the implementation of 3D propagation in GWP affects the wave forcing not only by direct modulation of GWF but also by indirect modulation of the resolved waves. Consequently, the overall impact on the zonal mean field was not as large as might be expected from the significant differences in GWF. However, slight differences are observed. The polar night jet in the SH was slightly weaker in October in OGW3D than in CONTROL. The zonal mean zonal wind in the polar upper stratosphere of the SH is stronger in OGW3D than in CONTROL, consistent with the difference in zonal GWF.

This study focused primarily on the changes in orographic GWF due to the inclusion of 3D propagation and the associated impacts on the large-scale fields present during the SH winter. It would also be interesting to examine the importance of 3D GW propagation in the Northern Hemisphere, which features short time-scale phenomena, such as sudden stratospheric warmings. The impact of 3D propagation on the interaction of GWF with transient planetary waves, as suggested in the present study, may be more important in the dynamics of the Northern Hemisphere.

Our scheme is also applicable to non-orographic GWPs, with a proper specification of discretized spectra of GWs. However, quantitative assessment of momentum transport by non-orographic GWP is still challenging because the observations needed for source specification are not sufficient. In addition, several aspects that were neglected in this study might be important in some regions. In particular, in the upper mesosphere and above areas where tidal waves are dominant, the assumption of instantaneous propagation and/or constant frequency may be inappropriate. Careful treatment in such regions is important.

The impact of refraction for non-orographic GWs is expected to be more significant than for orographic GWs, as non-orographic GWs generally have a broader range of horizontal wavelengths and phase speeds. Thus, the proper treatment of 3D propagation for both orographic and non-orographic GWs should be an important step toward the improvement of climate models.

Acknowledgments

We are grateful to Shingo Watanabe for providing the source code for MIROC-AGCM and informative instruction. We also thank Eliza Manzini, Marco Giorgetta, and Toru Nozawa for their insightful discussions. The numerical calculations were performed by the supercomputer system FX10 at the Information Technology Center of the University of Tokyo.
The figures were prepared using the GFD-DENNOU library. This work was supported by the Program for Leading Graduate Schools and by JSPS Grant-in-Aid for Scientific Research (25247075), Japan.

Appendix: Validity of ray–tracing calculations in GWP

In our GWP scheme, the ray–tracing calculation is approximated due to practical requirements. First, the integration of the ray–tracing equation is performed with respect to vertical model levels using a simple forward Euler scheme. The truncation error by this treatment may be large where vertical propagation of GWs is slow and/or the vertical model resolution is not sufficient. Second, the background wind and static stability are approximated using a first-order truncation of horizontal Taylor series expansion for effective parallelization. The truncation error depends on the horizontal distance from the wave origin. We performed several test calculations to confirm the validity of these approximations for the GWP scheme.

We prepared a standard ray–tracing scheme, which is designed to accurately calculate ray paths and wavenumbers for given initial parameters. The ray–tracing equations are integrated with respect to time, with a sufficiently small time step. The third-order Runge–Kutta scheme is used for the integration of first-order ODEs. The calculation of the amplitude
and GWF is the same as the GWP scheme. The cutoff vertical group velocity, which provides the lower limit for waves to be allowed to propagate, is set as 0.01 m s\(^{-1}\). The results calculated by this approach provide ideal references for comparison.

Given the same background and initial wave parameters, calculations are performed using three different schemes. The first (Scheme A) involves the standard ray–tracing scheme. The second (Scheme B) uses the new GWP but without the Taylor expansion approximation. The third (Scheme C) uses the new GWP with the Taylor expansion approximation. Through comparison of the three results, we separately examined the validity of the vertical integration and that of using the Taylor series expansion.

We compared ray paths under several idealized situations. As background fields, we assumed constant static stability (\(N = 0.016 \text{ s}^{-1}\)) and a zonal wind jet modified by planetary waves with a zonal wave number of 2 (See the contours of the zonal wind structure in Fig. A1). These assumptions mimic the polar vortex with large-scale wind perturbation. GWs are launched from the surface at the latitudes of all grid centers and have initial wavenumbers of \(k \simeq 2 \times 10^{-5} \text{ (m}^{-1})\) (which corresponds to a horizontal wavelength of 300 km), \(l = 0\), and a ground-based frequency of zero. Fig. A1 shows the results for GWs launched near 80°S, where the latitudinal propagation was the largest (i.e., the most severe case of the Taylor series expansion approximation). The vertical profiles of the

Fig. A2. A comparison of the three offline calculations of (a)–(c) the zonal mean zonal component of GWF, and (d)–(f) the zonal pseudomomentum flux at the 1 hPa surface. The calculations are performed using the same GW sources and background data, i.e., 6-h output from OGW3D for July in a particular year, and different schemes: (a, d) standard time-integration of the ray tracing equation, (b, e) our GWP without the approximation of the background using Taylor expansion, and (c, f) our GWP with the approximation. For (a)–(c), the contours and shading are the same as in Fig. 4. For (d)–(f), the contour levels are ±0.1, 0.3, 1.0, 3.0, 10.0, and 30.0 mPa. The light, intermediate, and dark shadings are used for absolute values larger than 0.1, 1.0, and 10.0 mPa, respectively.
positions of GW and GW parameters along the ray paths in the three schemes are shown with different line types. GWs propagate as much as 30° in longitude and 6° in latitude and break at approximately 1 hPa. Regardless of such a large horizontal propagation, the three schemes produced only minor differences in terms of both wavenumber and location.

Additionally, the GWFs of the offline calculations using a realistic background field are compared. Each calculation is performed by each of the three schemes described above using the same background data, i.e., the 6-h output from OGW3D for July in a particular year.

Figures A2a, A2b and A2c show the meridional cross section of the zonal mean zonal component of orographic GWF, calculated using Schemes A, B, and C, respectively. The essential spatial pattern and amplitude of the GWF in Fig. A2c are similar to those in Fig. 4a, except that the large negative GWF in the upper polar stratosphere is not present in Fig. A2c. This may reflect the fact that the amplitude of the positive GWF due to refraction (Fig. 6b) is largely dependent on planetary wave activity, which has large interannual variability. In the comparison of Figs. A2a, A2b and A2c, the overall spatial GWF patterns agree well with each other. The difference between Figs. A2a and A2b lies mainly in the level of the maximum amplitude of the GWF in the upper mesosphere. This difference indicates that the vertical integration resulted in a slightly higher estimate of the breaking levels. The difference between Figs. A2b and A2c is primarily in the upper polar stratosphere, where the GWF due to refraction is dominant. However, these differences are smaller than the difference between OGW3D and CONTROL (Fig. 5). Figures A2d, A2e and A2f show the maps of the zonal component of the pseudomomentum flux at 1 hPa, as calculated using Schemes A, B, and C, respectively. They agree well with each other in terms of the essential spatial distribution features due to the three-dimensional propagation discussed in Section 4.1.

References


Song, I.-S., and H.-Y. Chun, 2005: Momentum flux spec-

