Toward a Physically Based Gravity Wave Source Parameterization in a General Circulation Model

JADWIGA H. RICHTER AND FABRIZIO SASSI*

Climate and Global Dynamics Division, National Center for Atmospheric Research,[+] Boulder, Colorado

ROLANDO R. GARCIA

Atmospheric Chemistry Division, National Center for Atmospheric Research,[+] Boulder, Colorado

ABSTRACT

Middle atmospheric general circulation models (GCMs) must employ a parameterization for small-scale gravity waves (GWs). Such parameterizations typically make very simple assumptions about gravity wave sources, such as uniform distribution in space and time or an arbitrarily specified GW source function. The authors present a configuration of the Whole Atmosphere Community Climate Model (WACCM) that replaces the arbitrarily specified GW source spectrum with GW source parameterizations. For the nonorographic wave sources, a frontal system and convective GW source parameterization are used. These parameterizations link GW generation to tropospheric quantities calculated by the GCM and provide a model-consistent GW representation. With the new GW source parameterization, a reasonable middle atmospheric circulation can be obtained and the middle atmospheric circulation is better in several respects than that generated by a typical GW source specification. In particular, the interannual NH stratospheric variability is significantly improved as a result of the source-oriented GW parameterization. It is also shown that the addition of a parameterization to estimate mountain stress due to unresolved orography has a large effect on the frequency of stratospheric sudden warmings in the NH stratosphere by changing the propagation of stationary planetary waves into the polar vortex.

1. Introduction

Small-scale gravity waves (GWs) with horizontal wavelengths of tens up to several hundred kilometers are not resolved in general circulation models (GCMs) and hence need to be parameterized. In addition to the wide range of horizontal scales, gravity waves have vertical wavelengths ranging from a few to a few tens of kilometers and periods ranging from several minutes to several hours. Gravity waves are generated by a variety of sources including orography (e.g., Lilly and Kennedy 1973; Dörnbrack et al. 1999), convection (e.g., Dewan et al. 1998; Piani and Durran 2001), and geostrophic adjustment in regions of baroclinic instability (e.g., O'Sullivan and Dunkerton 1995; Zhang 2004). The nature of these wave sources is highly variable in space and time, implying similar variability for the generated gravity waves. Gravity waves propagating vertically, eventually dissipate, and deposit energy and momentum to the mean flow once they reach their critical levels or become unstable. This process occurs on very small scales and is not understood thoroughly [see the review by Fritts and Alexander (2003) and references therein]. The mean flow acceleration due to gravity waves increases with altitude due to the decreasing atmospheric density; as a result, in the mesosphere gravity waves are the dominant term in the global momentum budget. Gravity wave breaking also causes turbulence and mixes chemical constituents.

Gravity wave parameterizations are very important in GCMs; however, owing to the complex nature of gravity waves, representing them correctly in GCMs is quite a challenge. A gravity wave parameterization specifies

---

Corresponding author address: Jadwiga H. Richter, Climate and Global Dynamics Division, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307.
E-mail: jrichter@ucar.edu

gravity wave sources and then calculates effects of wave dissipation based on assumptions about wave dissipation. Most gravity wave parameterizations in the past have focused on representing the effects of wave dissipation and have developed various criteria for wave dissipation (e.g., Lindzen 1981; Warner and McIntyre 2001; Alexander and Dunkerton 1999). All of these parameterizations have made very simple assumptions about the wave sources. Typically, when one of these parameterizations is implemented in a GCM, it is assumed that nonstationary gravity wave sources are distributed uniformly in space and time or they specify more or less arbitrarily the variation of the sources with latitude and season (Garcia and Solomon 1985; Manzini and McFarlane 1998). The assumptions that most GCMs make about nonstationary GW sources are not necessarily realistic. Observations of gravity waves show that GW generation is very intermittent in space and time (Vincent and Reid 1983; Wu and Waters 1996) and can be linked to specific wave sources (Alexander and Pfister 1995; Alexander et al. 2000). Gravity waves generated by orography are primarily stationary and are better understood that nonstationary waves. The orographic GW source is usually treated by a separate parameterization (Palmer et al. 1996; McFarlane 1987).

Rind et al. (1988a) was the first to develop a GW parameterization that was linked to actual nonstationary wave sources. Rind et al. specified GW generation separately by convection and wind shear. At the time of the parameterization development, very little was known about the mechanisms of GW generation by those two sources, hence only theoretical assumptions linking GW properties to source properties were used. For convectively generated GWs, wave momentum flux was related to the convective mass flux generated by the model. Phase speeds of the waves were decided by the depth of convection: GWs generated by shallow convection had waves with phase speed of the background wind ± 10 m s^{-1}; GWs from deeper convection were launched with phase speeds of the background wind with ± 10, ± 20, and ± 40 m s^{-1}. For wind-shear-generated GWs, Rind et al. used a theoretical relationship between wave momentum flux and vertical wind shear above jet stream level. They assigned a single wavenumber and phase speed to the GWs generated in this fashion.

In the last decade, there have been several mesoscale modeling studies that have also shown how different and variable gravity waves can be from various wave sources. The wave source that has been investigated the most is convection. The modeling studies by Fovell et al. (1992), Alexander and Holton (1997), Lane et al. (2001), and Piani and Durran (2001) have shown that GW properties such as horizontal wavelength, frequency, and momentum flux phase speed spectra vary from simulation to simulation and are dependent on the properties of convection itself, as well as on the environmental wind and buoyancy frequency.

Beres et al. (2002) performed a sensitivity study of gravity wave properties to various profiles of tropospheric wind using a mesoscale model. They have found that the dominant vertical wavelength, and hence the dominant phase speed range, of convectively generated gravity waves is primarily determined by the vertical scale of the tropospheric heating. This finding is in agreement with previous theoretical findings by Salby and Garcia (1987). Beres et al. also found that the GW spectrum is strongly modified by the tropospheric wind and is most sensitive to wind shear in the upper troposphere. Using a combination of linear theory, mesoscale simulations of convection, and simplified heating simulations, Beres et al. (2004) and Beres (2004) have developed a method of estimating gravity wave properties based on the properties of convection and the environmental wind. A similar parameterization was developed by Song and Chun (2005). The Beres et al. (2004) convective source parameterization was implemented in the Whole Atmosphere Community Climate Model, version 2 (WACCM2) (Beres et al. 2005), replacing the arbitrarily specified source spectrum in the tropics. As a result of the parameterization, Beres et al. (2005) noted an improvement of the tropical stratospheric semiannual oscillation (SSAO) and the mesospheric semiannual oscillation (MSAO).

In recent years there have also been several efforts to better understand gravity wave generation from baroclinic waves and frontal systems. In an idealized study of a baroclinic wave life cycles, O'Sullivan and Dunkerton (1995) noted that gravity waves were generated from the jet stream, primarily in the exit region of upper tropospheric jet streaks. A similar study, but with much higher horizontal resolution, was performed by Zhang (2004) confirming these findings. Zhang related GW generation to balance adjustment (a generalization of geostrophic adjustment), a mechanism in which gravity waves are radiated in order to bring the flow back into a balanced state. Wang and Zhang (2007) performed a sensitivity study of GWs generated from baroclinic jet fronts. Wang and Zhang found that the dominant gravity wave frequency was linked to the large residual of the nonlinear balance equation; however, the relationships between the gravity wave and baroclinic system characteristics were not explored in enough depth to form a parameterization. A GW generation mechanism related to balance adjustment is generation by jet/front systems or frontogenesis (Griffiths and Reeder 1996; Reeder and Griffiths 1996). The distinction between wave generation by

balance adjustment and by frontogenesis is often hard to make as the two phenomena tend to be collocated. A comparison of the residual of the nonlinear balance equation at 300 mb with the Miller (1948) frontogenesis function at 600 mb in a MM5 run with 150-km resolution shows good agreement between the quantities (F. Zhang 2007, personal communication). Charron and Manzini (2002) implemented GW generation by frontal systems in a GCM using frontogenesis as a diagnostic for frontal systems. Charron and Manzini found several improvements to the stratospheric simulation as result of the added variability. Charron and Manzini retained the arbitrarily specified "background" spectrum in regions where frontogenesis was not present.

Aside from improving the simulated climate, parameterizations of individual gravity wave sources introduce more realism and model consistency into the representation of GWs in GCMs. This was the motivation behind the work of Rind et al. (1988a). Beres et al. (2005) showed that the monthly mean convective GW properties differ greatly between the tropics and extratropics and that they change with season. One of the most important reasons for connecting GW properties to wave sources is that, in a changing climate, parameterized GWs will also change and they will presumably change in a physically realistic manner. We present here results from a WACCM simulation without any arbitrary GW source spectra, utilizing the most up-to-date understanding of gravity wave generation by convection and fronts. This approach adds new physics to the GW parameterizations, provides GW representation consistent with the underlying climate, and causes GW source properties to change with changing climate. This paper is organized as follows: In section 2 we provide a brief model description and details of the new GW parameterization. In section 3 we present the resulting GW source spectra from individual wave sources and we show several features of the resulting model simulation. Section 4 contains a discussion and conclusions.

2. Model

a. General description

The GCM used in this study is the Whole Atmosphere Community Climate Model, version 3.5 (WACCM3.5). WACCM3.5 is based on the Community Atmosphere Model, version 3.5 (CAM3.5) with the vertical model domain extended to ~145 km. WACCM3.5 uses the finite-volume dynamical core of Lin (2004) with 66 vertical levels, with variable vertical resolution. The horizontal resolution for WACCM3.5 runs presented here is 1.9° × 2.5° (latitude × longitude). A previous version of the model, WACCM3, and the details of

model physics specific to WACCM3 were described in detail in Garcia et al. (2007) and will not be repeated here. Instead, we describe only the details of the model relevant to this study as well as changes in physical parameterizations from WACCM3 to WACCM3.5.

WACCM3.5 uses all of the CAM3.5 physical parameterizations. The major change between CAM3 and CAM3.5 has been the convection parameterization. CAM3 included the convection parameterization of Zhang and McFarlane (1995). This parameterization has been modified in CAM3.5 to include (i) a dilution approximation for the calculation of convective available potential energy (CAPE) (Neale et al. 2008) and (ii) convective momentum transport (Richter and Rasch 2008; Neale et al. 2008). There have been several changes made to the land model between CAM3 and CAM3.5. These are described in Oleson et al. (2008) and include modifications to surface datasets, parameterizations for canopy integration and interception, frozen soil, soil water availability, soil evaporation, surface and subsurface runoff, groundwater model, and introduction of a factor to stimulate nitrogen limitation on plant productivity.

Another significant difference between WACCM3.5 and CAM3.5 is the addition of surface stress due to unresolved orography or turbulent mountain stress. As in CAM3.5, WACCM3.5 uses the Holstlag and Boville (1993) nonlocal boundary layer diffusion to parameterize the turbulent and eddy transport in the planetary boundary layer (PBL). This nonlocal PBL scheme uses the surface fluxes to estimate the PBL height and nonlocal transport. In contrast to CAM3.5, however, WACCM3.5 uses an updated calculation of the surface fluxes that exploits the unresolved orographic variance to obtain an effective surface roughness. At the coarse resolution of GCMs the effects of topography are not fully captured, and the surface drag parameterization attempts to account for this. The concept of effective roughness length was first developed by Fiedler and Panofsky (1972). Fiedler and Panofsky defined roughness length for heterogeneous terrain as the roughness length that homogenous terrain would have to give the correct surface stress over a given area. The concept of effective roughness has been used in several Numerical Weather Prediction models (e.g., Wilson 2002; Webster et al. 2003).

A convenient and practical way to implement the effective roughness is to calculate an additional surface stress, expressed as

$$\tau = \rho C_d |\mathbf{V}| \mathbf{V}, \tag{1}$$

where ρ is the density, $\mathbf{V}$ is the horizontal wind vector in the bottom atmospheric layer, and C_d is the drag coefficient, calculated as

$$C_d = \frac{f(R_i)k^2}{\ln^2[(z + z_0)/z_0]}, \qquad (2)$$

where k is the von Kármán constant, z is the height of orography above the surface, and z_0 is an effective roughness length that represents unresolved topography and is estimated from the following:

$$z_0 = \min(0.1\sigma, 100). \qquad (3)$$

In the above, σ is the standard deviation of subgrid orography, expressed in meters. In (2), $f(R_i)$ is a function of the Richardson number ($R_i = gT_z/(T \, |\mathbf{V}_z|^2)$): $f(R_i) = 1$ if $R_i < 0$; $f(R_i) = 0$ if $R_i > 1$; and $f(R_i) = 1 - R_i$ if $0 < R_i < 1$.

The stress calculated by (1) is used as an additional surface stress in the nonlocal PBL scheme to evaluate the PBL height and nonlocal transport [Eqs. (3.10)–(3.12) in Holstlag and Boville 1993]. The calculation of the additional surface stress is carried out only over land where the height of topography above sea level is nonzero in the model grid box.

b. Gravity wave parameterization

WACCM3.5 uses the orographic gravity wave parameterization based on McFarlane (1987). A tunable parameter in this parameterization is the efficiency with which gravity waves are launched. The orographic efficiency is set to 0.125 in WACCM3.5, consistent with that used by CAM3 and CAM3.5.

The nonorographic GW propagation parameterization is based on the formulation of Lindzen (1981). The spectrum is discrete: gravity waves are launched with phase speeds between -80 and 80 m s^{-1} at 2.5 m s^{-1} increments. In WACCM3.5 the arbitrarily specified source function is completely removed. Instead, GW launching occurs through two GW source parameterizations: convective and frontal. In our formulation, therefore, we assume that all nonorographic gravity waves are generated either by convection or frontal/baroclinic systems. We are not accounting for other, less common, sources of GWs such as shear instability. Shear instability tends to be collocated with frontal/baroclinic systems and the understanding of wave generation by this process, especially in the presence of frontal systems, is poor. For this reason, we choose to retain convection and and frontal/baroclinic systems as the only two nonorographic wave sources, acknowledging that this is an oversimplification. The main point of this study is a proof of concept that a reasonable middle atmospheric circulation can be obtained with the dominant GW source parameterizations; hence, we find our assumptions acceptable: other GW sources can be added in the future.

1) CONVECTIVE GW SOURCE PARAMETERIZATION

We use the Beres et al. (2004) parameterization to represent the convectively generated GW spectrum. The Beres parameterization determines GW properties, in particular the momentum flux phase speed spectrum, based on the convective heating properties. The convective heating depth is the most important aspect of the heating as it determines the dominant GW horizontal phase speed and the basic shape of the momentum flux spectrum in phase speed. Deep heating (≥ 10 km) tends to generate GWs with long vertical wavelengths and high horizontal phase speeds (20 to 30 m s^{-1}), whereas shallow heating (≤ 5 km) generates GWs with short vertical wavelengths and small horizontal phase speeds (10 to 15 m s^{-1}) (Beres 2004). The momentum flux phase speed spectrum is also highly dependent on the tropospheric wind, which creates asymmetry in the spectrum: this is taken into account in the parameterization. The magnitude of the momentum flux in the Beres et al. (2004) parameterization is primarily determined by the convective heating rate. The horizontal scale of the heating, as well as its frequency distribution, also affects the wave amplitude. Since these quantities are unknown in a GCM, following Beres et al. (2004) we assume a red-frequency distribution for wave periods between 10 and 120 min and that typical convective cells are 3 km wide. In a GCM, the convective heating rate is averaged over the model grid box. To arrive at a heating rate representative of heating cells, we assume that the convection takes up 5% of the grid box. Owing to these assumptions, the amplitude of GWs is the least certain aspect of the parameterization; however, the main features of the momentum flux phase speed spectrum are well constrained because they are based on the convective heating depth and mean wind, quantities that are represented well in a GCM.

The details of implementing the Beres (2004) parameterization were described in Beres et al. (2005), and here we provide only a brief summary. The convective GW source parameterization is activated whenever the deep convection parameterization of Zhang and McFarlane (1995) is active. The vertical heating profile from this parameterization is then used to determine the convective heating depth and heating rate. Together with the information about the mean wind in the heating region, the momentum flux phase speed spectrum at each model grid point is determined.

Our implementation of the convective source parameterization is almost identical to that presented in Beres et al. (2005) aside from the following differences: 1) in our current implementation GWs are launched everywhere convection is present and not only in the tropics and 2) GWs are launched at the top of convection instead of 100 mb.

2) FRONTAL GW SOURCE PARAMETERIZATION

Frontogenesis typically occurs in regions of strong wind field deformation and strong temperature gradient. There regions can be identified by the frontogenesis function (Miller 1948; Hoskins 1982), which we call F:

$$F \equiv \frac{1}{2}\frac{D|\nabla\theta|^2}{Dt} = -\left(\frac{1}{a\cos\phi}\frac{\partial\theta}{\partial\lambda}\right)^2\left(\frac{1}{a\cos\phi}\frac{\partial u}{\partial\lambda} - \frac{v\tan\phi}{a}\right) - \left(\frac{1}{a}\frac{\partial\theta}{\partial\phi}\right)^2\left(\frac{1}{a}\frac{\partial v}{\partial\phi}\right)$$
$$- \left(\frac{1}{a\cos\phi}\frac{\partial\theta}{\partial\lambda}\right)\left(\frac{1}{a}\frac{\partial\theta}{\partial\phi}\right)\times\left(\frac{1}{a\cos\phi}\frac{\partial v}{\partial\lambda} + \frac{1}{a}\frac{\partial u}{\partial\phi} + \frac{u\tan\phi}{a}\right). \tag{4}$$

In the above, u and v are the zonal and meridional components of the wind, respectively; θ is potential temperature; ϕ is latitude; λ is the longitude; and a is the radius of the earth. Large values of the frontogenesis function indicate regions of frontal activity. Frontogenesis can be calculated at all model levels; however, a single vertical level and a frontogenesis threshold need to be chosen for diagnosing fronts in a GCM. We choose 600 mb as the level at which frontogenesis is evaluated since this closely corresponds to the typical steering level of fronts. Charron and Manzini (2002) also used frontogenesis at 600 mb. The frontogenesis threshold for front identification is a tunable model parameter. Obviously, this value should correspond closely to the occurrence of fronts, but this is unknown precisely in a climate model. Charron and Manzini found a minimum threshold of 0.1 K^2 (100 km)$^{-2}$ h^{-1} as a reasonable value. Griffiths and Reeder (1996) suggested a value of 0.07 K^2 (100 km)$^{-2}$ h^{-1}. Here we use the value of 0.045 K^2 (100 km)$^{-2}$ h^{-1}. We find that a higher frontogenesis threshold does not produce enough gravity waves in the polar regions to drive a strong enough meridional circulation that can cool the mesopause to observed temperatures. The need to use a slightly lower frontogenesis threshold than employed by other authors suggests that we might be overestimating the GW generation by fronts and that other extratropical GW sources should be included in future simulations. Figure 1 shows a typical spatial distribution of GW generation by fronts in January. Most of the GW excitation occurs in the extratropical and polar NH, as expected. There are few regions of wave excitation in the SH.

At a given model time step, in all locations where the frontogenesis threshold is exceeded, GWs are launched from the source level of 600 mb. The momentum flux phase speed spectrum is Gaussian, cen-tered on the wind at source level with a half-width of 30 m s^{-1}. The amplitude of the waves is another tunable parameter of the scheme. Here we use a value of wave amplitude, $\tau_b = 1.5 \times 10^{-3}$ Pa. The value of τ_b, in large part, controls the height of the mesopause, which is now well constrained by observations. In reality, τ_b could vary according to frontal characteristics; however, the relationships between wave amplitude and frontal characteristics have not been established to date, hence we chose a constant value for τ_b.

In previous implementations of the gravity wave parameterization in WACCM and in other GCMs there was another tuning parameter, the efficiency parameter. This parameter accounted for intermittency in the wave sources that was not taken into account by the parameterizations. When GWs are launched from fronts, realistic intermittency already exists, and there is no longer a need for this parameter.

c. Simulation setup

All simulations presented here were carried out for 20 years and were forced with observed sea surface temperatures (SSTs) from 1985 to 2005. Because WACCM does not produce a internally generated quasi-biennial oscillation (QBO), we have relaxed the tropical winds to the observed QBO winds between 86 and 4 hPa with a relaxation time scale of 10 days between the latitudes 22°S and 22°N, decaying with a Gaussian distribution with a half-width of 10° centered at the equator. The relaxation procedure follows Balachandran and Rind (1995) and Matthes et al. (2004).

The simulations have interactive chemistry with specified realistic boundary conditions for the 1985–2005 period, as described in Garcia et al. (2007). The solar cycle also has observed variability. We include time variable boundary and solar cycle conditions to include natural variability in our simulations, which is of
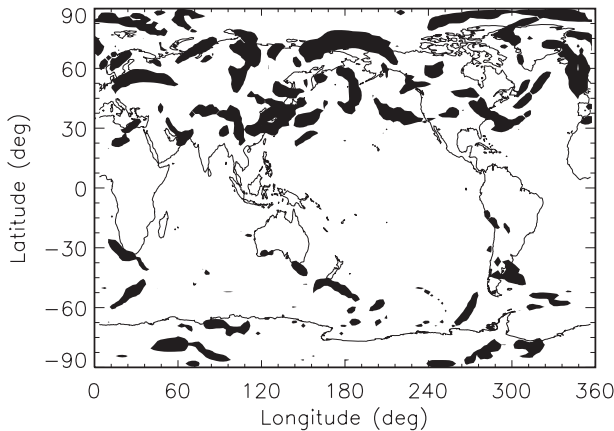
FIG. 1. Map of the regions that exceed the frontogenesis threshold on a selected day in January (depicted in black) in WACCM3.5.

particular importance for the simulation of stratospheric sudden warmings.

3. Results

a. Gravity wave sources

In this section we present the characteristics of gravity waves generated by the nonorographic source parameterizations. Figures 2 and 3 show the seasonal cycle of gravity wave momentum flux from convective and frontal-generated gravity waves. Convective GW momentum flux is confined mainly to the tropics (Fig. 2), whereas frontal GW momentum flux dominates in the extratropics (Fig. 3). Convective GW excitation does occur in the extratropics, especially during the passages of summertime fronts; however, on a monthly average the extratropical convective GW momentum flux is much smaller than that from persistent tropical convection. Convective GW momentum flux follows the seasonal cycle of convection: in NH winter (summer) it is confined primarily south (mostly north) of the equator. Since the north–south wind component is weak in the tropics, convective gravity waves are generated mainly in the eastward and westward directions by our parameterization.

Frontal GW momentum flux follows the seasonal cycle of fronts. Maximum frontal GW generation occurs in winter, with a minimum during the summer season (Fig. 3). The seasonal contrast is much less marked in the SH as compared to the NH. This is because strong baroclinicity is present in the SH in all seasons. Charron and Manzini (2002) also found this, which is an indication that the frontal source specification adds a degree of realism not present in arbitrary source specifications.

Frontal GW momentum flux is dominated by the westward component at 100 mb. This is due to the strong filtering of eastward propagating GWs by strong tropospheric westerlies between 600 and 100 mb. Since there is a strong north–south component to the 600-mb winds in the extratropics, the northward and southward GW momentum fluxes are also significant. They are of amplitudes comparable to the eastward frontal GW momentum flux. Eastward GW momentum flux is stronger than westward GW momentum flux in NH summer because strong tropospheric easterlies are present at that time, which create an asymmetry in the GW spectrum.

It would be useful to compare our estimates of 100-mb gravity wave momentum flux with observations. Unfortunately, there are no global observations of gravity wave momentum flux against which to verify our source parameterizations. Recently, the Stratospheric Processes and their Role in Climate (SPARC) project has created an initiative on the "Gravity Wave Momentum Budget for Global Circulation Studies." As a result of this project, observational constraints on GW momentum flux on a global scale may become available in the next few years [M. J. Alexander (project coordinator) 2008, personal communication]. However, at the moment there are no suitable observations that we can use to verify our model results.

Momentum flux phase-speed spectra are a key component of the parameterization, as the GW wave phase speeds in large part determine at what level wave dissipation will occur. Figure 4 presents the momentum flux phase speed spectra at 100 mb for January and July for the convective GW source parameterization. This figure is very similar to Fig. 2 in Beres et al. (2005). Recall that the shape of the momentum flux phase speed spectra is determined by the Beres (2004) parameterization from the convective heating depth and the tropospheric wind at each model grid point. In the tropics, the momentum flux phase-speed spectra for convectively generated waves have two distinct peaks, one at positive phase speeds and one at negative phase speeds. The dominant GW phase speed and the breadth of the phase speed spectrum vary with latitude due to changes in the depth of the heating and the tropospheric wind. For example (see Fig. 4c), at 4°N the dominant phase speed of eastward (westward) propagating waves is 25 m s^{-1} (45 m s^{-1}), with the momentum flux for eastward propagating waves being 25% larger than for westward propagating waves. At 30°S, the spectrum is very different from that at 4°N: westward (eastward) propagating waves have a dominant phase speed near 20 m s^{-1} (of 70 m s^{-1}). The magnitude of momentum flux is about four times as large for westward propagating waves as it is for eastward
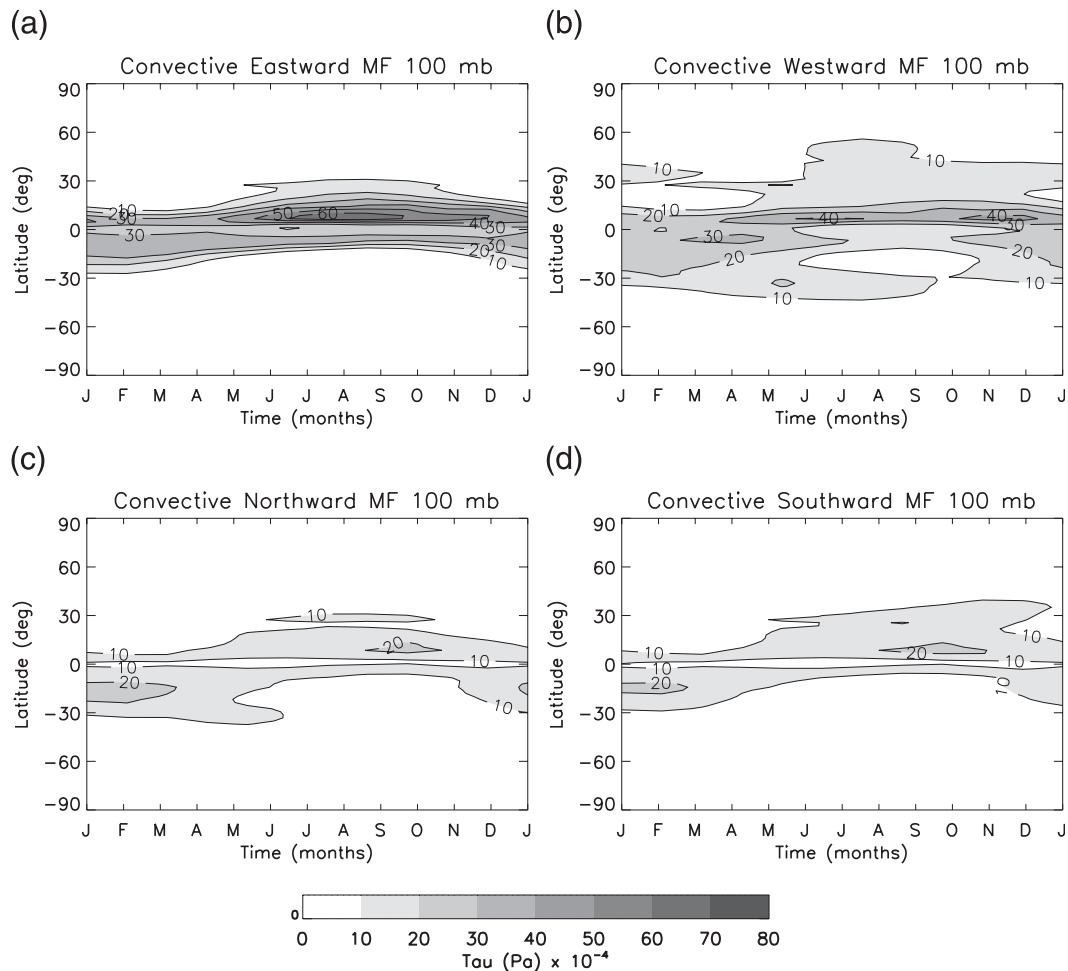
FIG. 2. Total (a) eastward, (b) westward, (c) northward, and (d) southward momentum flux at 100 mb from convectively generated gravity waves as a function of latitude and time of year (in Pa) in WACCM3.5.

propagating waves. In July (Fig. 4b), convectively generated waves become important outside of the tropics. The properties of the waves generated by extratropical convection are quite different from those generated in the tropics (Fig. 4d). At 60°S, the momentum flux phase speed spectrum has only one peak, at 30 m s^{-1}. This is due to the strong tropospheric wind shear, shown by the thick black and dashed lines in Fig. 4b. This is also seen in January near 40°N.

The momentum flux phase speed spectra for frontally generated waves are shown in Fig. 5. The overall momentum flux carried by frontally generated waves is lower than that carried by convectively generated waves by approximately 50%, especially in January. Note that the ratio of overall fontal to convective GW momentum flux is impacted by the tuning parameters in our parameterization: more observations are needed to verify the accuracy of this relationship. The frontally generated waves occur primarily in the extratropics. The

shape of the momentum flux phase speed spectra for these waves is primarily determined by our assumption of a Gaussian distribution at source level. Hence, at all latitudes the frontally generated gravity waves have one dominant phase speed. The dominant wave phase speed is, however, influenced by the wind at the source level of 600 mb. In regions of strong tropospheric shear (e.g., 40°N in January, shown in Fig. 5c), the dominant wave phase speeds are near, and sometimes even exceed, 50 m s^{-1}. In the polar regions, where tropospheric winds are close to zero, the dominant wave phase speeds are much smaller (Fig. 5d).

Figure 6 shows the drag on the mean flow imposed by the individual wave sources in WACCM3.5 averaged over December–February (DJF). The orographic GW drag (Fig. 6a) only plays a role in the winter lower stratosphere. Convective GW drag (Fig. 6b) is largest in the tropics, reaching 20 m s^{-1} day^{-1} near 100 km. The gravity wave drag from frontal systems is the largest

(a)
Frontal Eastward MF 100 mb

(b)
Frontal Westward MF 100 mb

(c)
Frontal Northward MF 100 mb

(d)
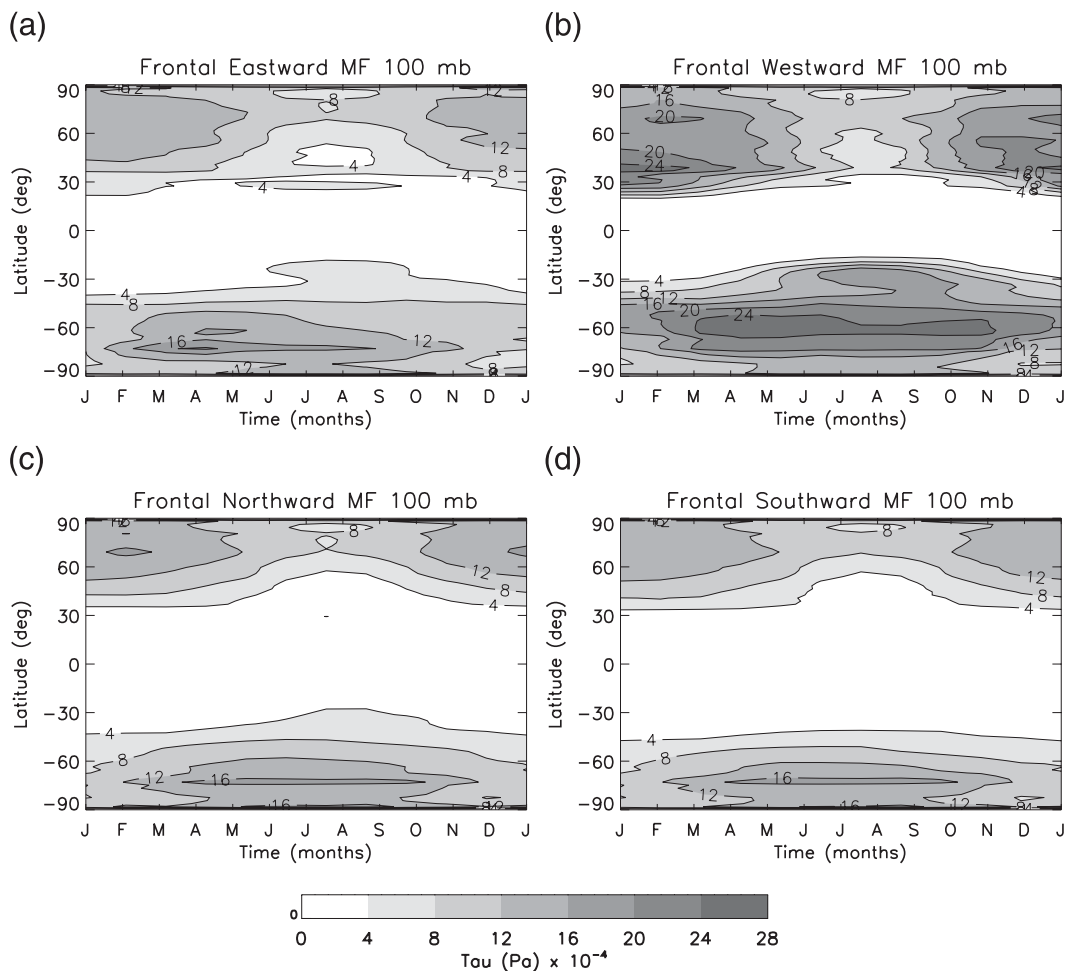### Frontal Southward MF 100 mb

Tau (Pa) × 10⁻⁴

FIG. 3. As in Fig. 2 but for frontally generated gravity waves. Note that the range of contours is smaller than in Fig. 2.

contributor to the total GW drag in the extratropics and is responsible for the cold summer mesopause and jet reversal at these altitudes. Eastward GW drag from fronts in the SH reaches over 100 m s^{-1} day^{-1} in a seasonal DJF mean, whereas westward GW drag in the NH reaches a little over 20 m s^{-1} day^{-1} in amplitude. The GW drag shown in Fig. 6 can be compared to the GW drag from individual wave sources presented in Rind et al. (1988a, their Figs. 5 and 6). Their orographic GW drag is approximately twice as large as in WACCM3.5, with a maximum of about 9 m s^{-1} day^{-1} near 60 km, 60°N. In the newer version of the Goddard Institute for Space Studies (GISS) Global Climate–Middle Atmosphere model, discussed in Rind et al. (2007), the orographic GW drag is reduced as compared to Rind et al. (1988a) and is more similar to that in WACCM3.5. Unlike in WACCM3.5, the convective GW drag in Rind et al. (1988a) has substantial amplitude in the extratropics, reaching amplitudes of 30–40 m s^{-1} day^{-1} both in the winter and summer hemisphere. In the tropics,

between 60 and 90 km, the convective GW drag in Rind et al. (1988a) is similar to that in WACCM3.5. The GW drag generated by wind shear in Rind et al. has very similar characteristics to that in WACCM3.5. It is contained to the extratropics with summer hemisphere amplitudes about four to five times larger than in the NH. Their GW drag amplitudes are also similar to those in WACCM3.5.

b. Middle atmospheric simulation

In this section we present the general characteristics of the middle atmosphere climate with the source-oriented gravity wave parameterization. We compare some of the general features of this simulation (called WACCM3.5 hereafter) to the simulation of the middle atmosphere with WACCM3 (Richter et al. 2008). The main point of this comparison is to assess how good the middle atmospheric climate is in the new simulation compared to the "best" simulation obtained with WACCM3, which did not contain individual wave source parameterizations.
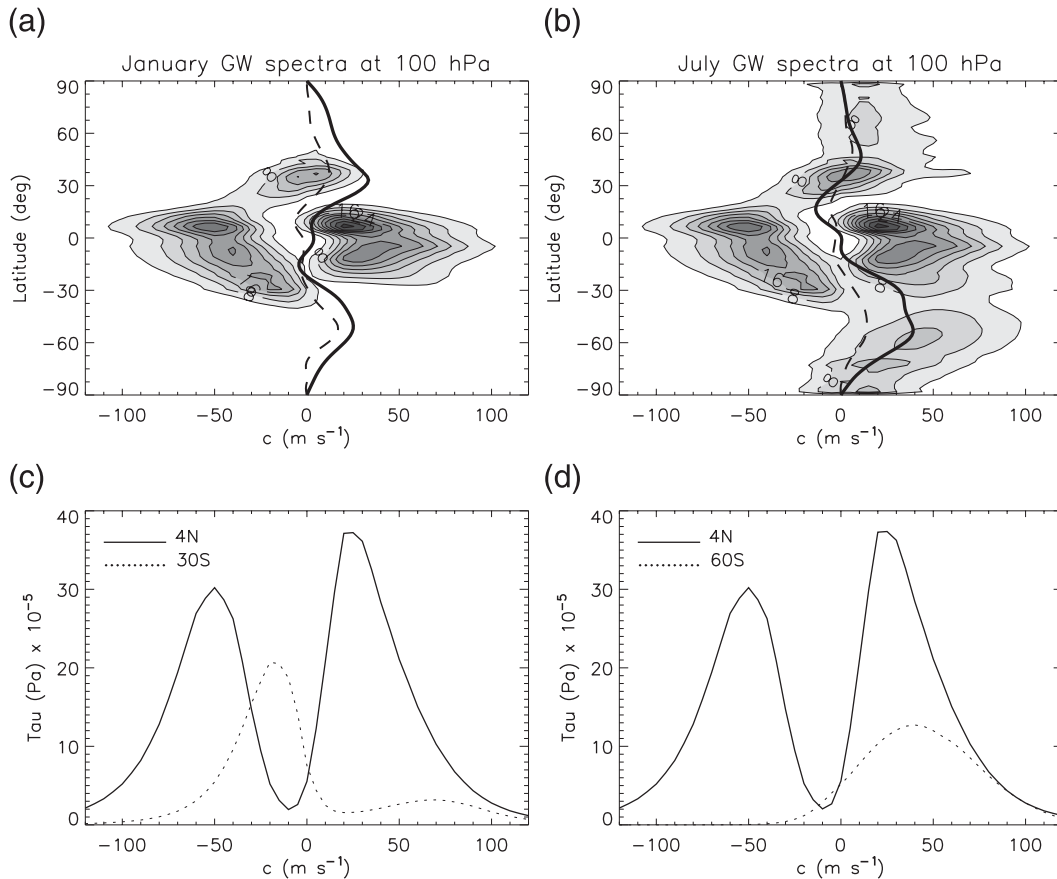
FIG. 4. Monthly averaged east–west momentum flux phase-speed spectra at 100 hPa derived from the convective gravity wave source parameterization for (a) January and (b) July in WACCM3.5. Contours are plotted in intervals of 2×10^{-5} Pa. The solid (dashed) line depicts the zonal mean wind at 100 (700) hPa. Cross sections through (a) at (c) 4°N (solid) and 30°S (dotted) and at (d) 4°N (solid) and 60°S (dotted).

We compare the zonal mean wind with Unified Radio and Plasma Wave Experiment (URAP) observations, shown in Fig. 7, and compare temperature to the 40-yr ECMWF Re-Analysis (ERA-40) in the lower atmosphere and Sounding of the Atmosphere Using Broadband Emission Radiometry (SABER) data (Xu et al. 2007) in the mesosphere.

1) ZONAL MEAN WIND AND TEMPERATURE CLIMATOLOGY

Figure 8 shows the zonal mean wind and temperature averaged for DJF and June–August (JJA) for WACCM3 and WACCM3.5. The general features of the zonal mean wind in WACCM3.5, such as stratospheric jets and mesospheric wind reversals, look similar to WACCM3 and observations. Considering the substantial changes to the GW source specification, this is already a very encouraging result: we have taken out the arbitrarily specified GW source spectrum and replaced it with a physically based source representation, arriving

at a middle atmospheric simulation that resembles observations.

A major difference between the zonal winds in WACCM3 and WACCM3.5 as compared to observations is the extension of the tropospheric westerly jet into the SH lower stratosphere in DJF. This strong polar vortex is associated with cold lower stratospheric temperatures, which arise as a result of ozone depletion above Antarctica and persist too long. The problem is common in the SH summer in middle atmospheric GCMs (Garcia and Boville 1994; Eyring et al. (2006)). Comparison of Figs. 8a to 8c shows that the problem with cold polar temperatures has been reduced in WACCM3.5. The zero wind line still extends too far into the stratosphere; however, the amplitude of the vortex near the South Pole has been reduced from ~20 to less than 10 m s^{-1} in the lower stratosphere. To illustrate the improvement of the problem further, Fig. 9 shows the descent of the zero zonal mean wind line at 60°S during SH summer. This figure is similar to Fig. 2 of Eyring et al.
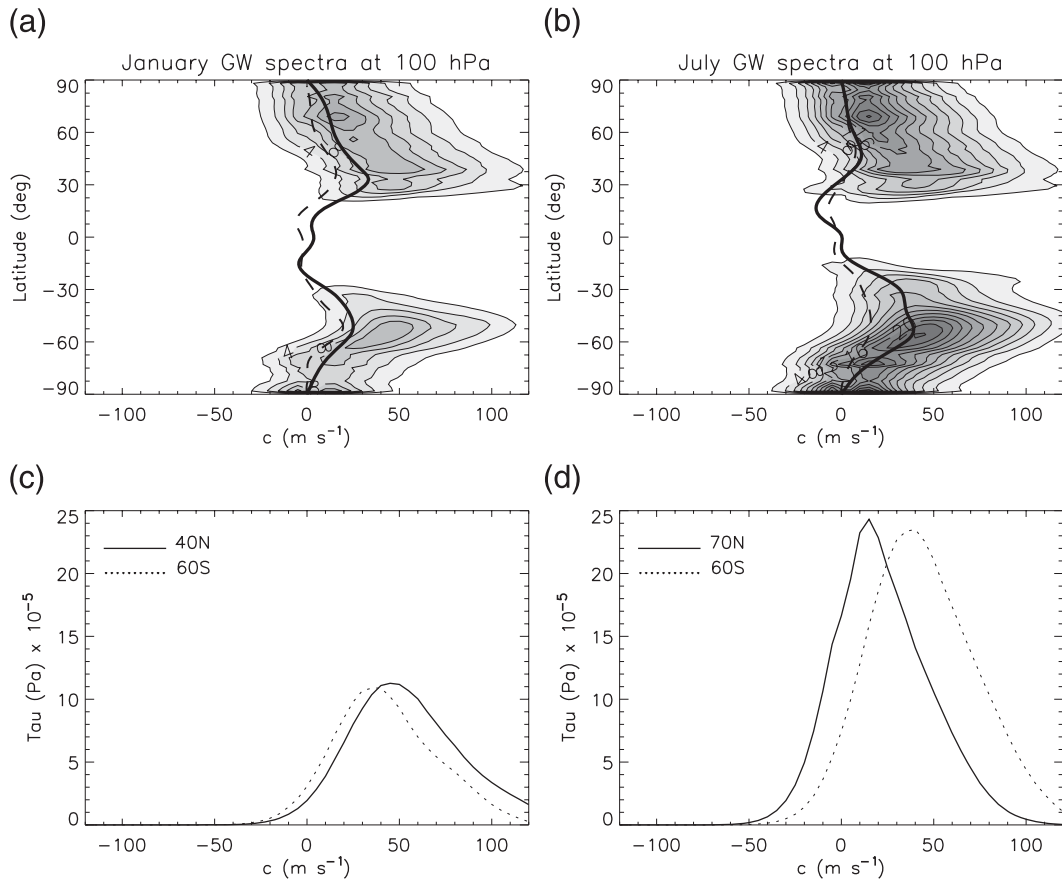
FIG. 5. As in Fig. 4 but derived from the frontal gravity wave source parameterization. The solid (dashed) line depicts the zonal mean wind at 100 (600) hPa. Cross sections through Fig. 5a at (c) 40°N (solid) and 60°S (dotted) and at (d) 70°N (solid) and 60°S (dotted).

(2006) and shows that the transition to easterlies is now much more comparable to observations, although a bias still persists. Comparison of December temperatures between 60° and 90°S (not shown) for WACCM3.5 and WACCM3 to ERA-40 shows that the cold bias has been reduced from −24 K in WACCM3 to −16 K in WACCM3.5 near 80 hPa. The improvement of the SH lower stratospheric winds in WACCM3.5, as compared to WACCM3, is most likely a direct result of altered GW drag. Comparison of planetary wave structures between the two simulations (not shown) shows that planetary wave amplitudes are very similar up to ∼300 mb, with largest differences above 100 mb, implying that the planetary wave generation in these two simulations is very similar. Above 300 mb, the zonal winds are already different between the two simulations, and it is likely that the changed planetary wave amplitude is largely a response to the mean flow change. In addition, the GW drag in the stratosphere (not shown) is very different between the two simulations, whereas the EP

flux divergence from resolved waves only shows small differences.

The SH stratosphere zonal-mean zonal wind in DJF is easterly. In WACCM3, the maximum winds were overestimated by 20 m s^{-1} compared to URAP, and this bias remains in WACCM3.5. However, the strong split in the jet near 60°S, present in WACCM3, is reduced in WACCM3.5. The NH stratosphere in DJF features an equatorward-tilted westerly jet. Averaged over DJF, the strength of the winds in WACCM3 was comparable to observations, and it is also comparable to observations in WACCM3.5. The shape of the NH stratospheric jet in DJF is more realistic in WACCM3.5, as the jet is narrower in the lower stratosphere.

The JJA zonal wind simulation in WACCM3.5 is similar to that in WACCM3 (Figs. 8b,d). In both simulations, the SH westerly jet is not tilted equatorward enough compared to URAP. In WACCM3, the strength of the winter jet was overestimated by ∼10 m s^{-1}. In WACCM3.5 this bias has increased to ∼30 m s^{-1}. This
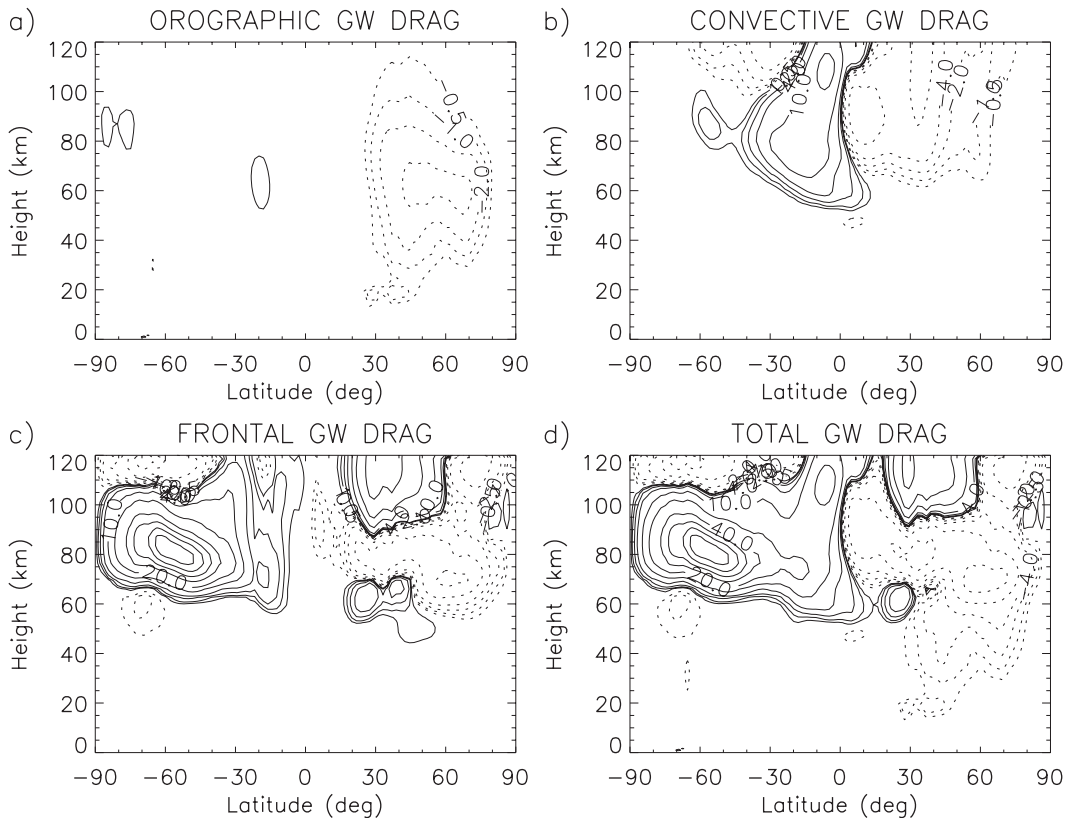
FIG. 6. (a) Orographic, (b) convective, (c) frontal, and (d) total GW drag averaged over DJF in WACCM3.5. Contour intervals are ±0.5, ±1, ±2, ±4, ±10, ±20, ±40, ±60, ±80, and ±100 m s^{-1} day^{-1}.

deterioration in the JJA winter jet is one of the very few degradations of the simulation in WACCM3.5 as compared to WACCM3. The NH stratospheric easterly jet in WACCM3.5 has a shape similar to that in WACCM3: compared to URAP, in both simulations the jet maximum is located too far equatorward. The jet maximum was 20 m s^{-1} too strong in WACCM3; in WACCM3.5, it is only 10 m s^{-1} too strong.

The summer mesopause temperature and height are is an important climatological feature to represent correctly in models extending to the lower thermosphere. In WACCM3.5 the January summer mesopause is at 87 km log-pressure height and 128 K in January (143 K in a DJF average) (Fig. 10c). In July the summer mesopause is at 90.5 km log-pressure height with a temperature of 127 K (136 K in a JJA average) (Fig. 10d).
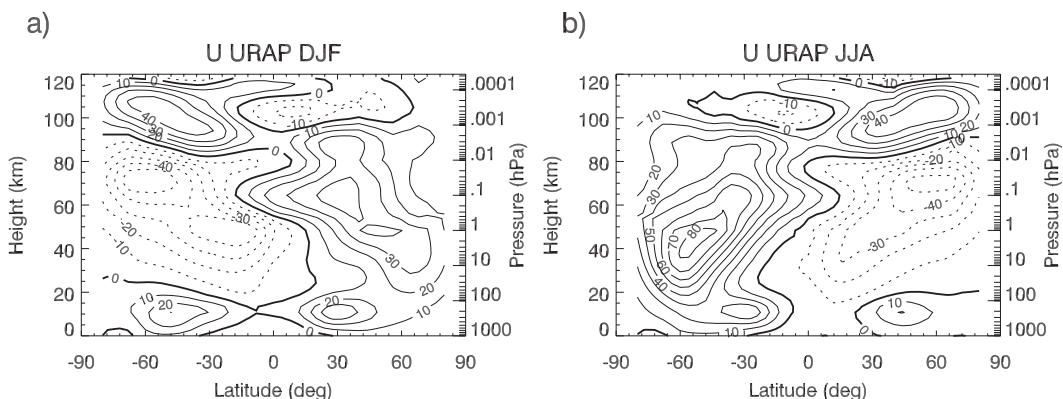

FIG. 7. Seasonally averaged URAP zonal wind (m s^{-1}) for (a) DJF and (b) JJA.
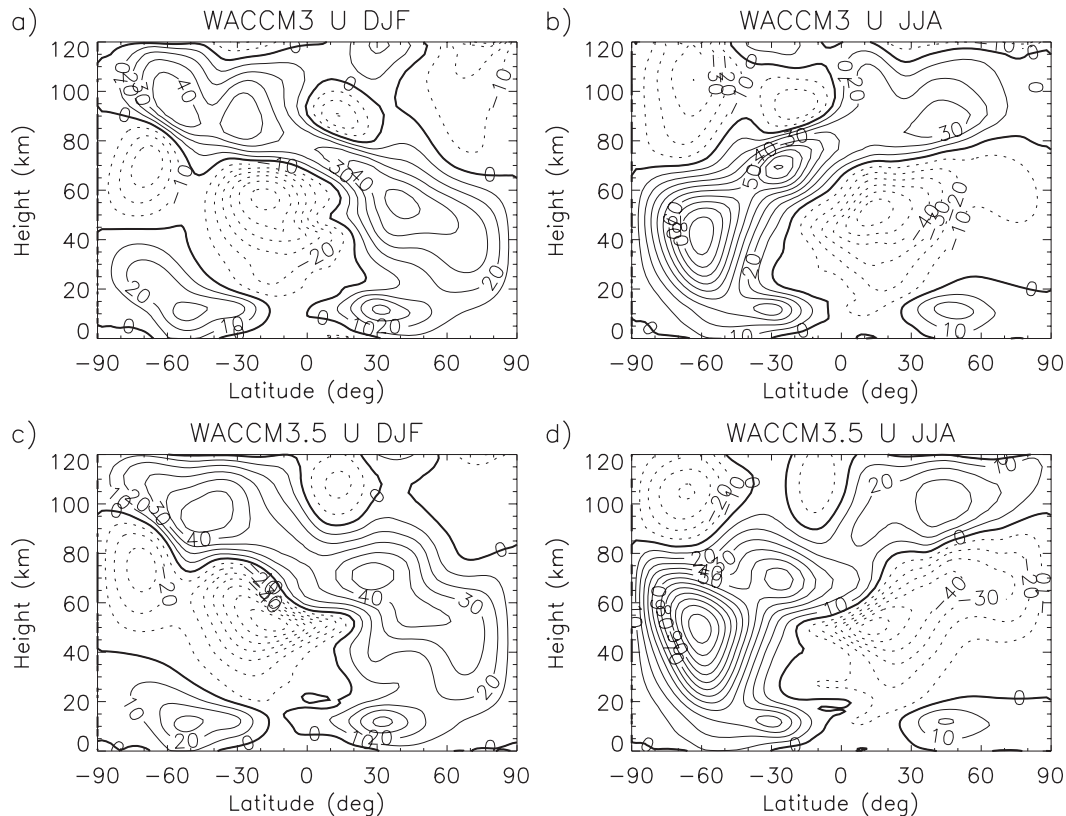
FIG. 8. Zonal mean wind (m s^{-1}) averaged over (left) DJF and (right) JJA for (top) WACCM3 and (bottom) WACCM3.5.

Compared to observations from SABER (Xu et al. 2007), the January summer mesopause temperatures are ~8 K too warm and are close to observations in July. For reasons that are at the moment not clear, WACCM3.5 does not reproduce the differences in mesopause structure between the Southern and Northern Hemisphere.

2) VARIABILITY

To assess variability in WACCM, we first look at the standard deviation of zonal mean temperature calculated from monthly mean output for the month of January. This is shown in Fig. 11 for ERA-40, WACCM3, and WACCM3.5. In ERA-40, the largest temperature variability occurs in the Northern Hemisphere stratosphere near 10 hPa, with standard deviations of 6 K near 100 hPa and maximum standard deviations exceeding 16 K between 10 and 1 hPa. It is worth noting that the maximum ERA-40 standard deviation between 10 and 1 hPa depends on the observation period chosen. The maximum standard deviation from 1980 to 2002 is only 10 K (6 K lower than between 1985 and 2002). Comparison of Figs. 11a to 11b shows that variability in the the NH stratosphere was underestimated in WACCM3. At

100 hPa the standard deviations of temperature were only about 3 K and only about 5 K between 10 and 1 hPa. In WACCM3.5 (Fig. 11c) the standard deviation of temperature is much closer to observations. The values reach 6 K near 100 hPa, similar to ERA-40, and exceed 12 K between 10 and 1 hPa.

Another good measure of the model's variability is the representation of sudden stratospheric warmings (SSWs). A sudden stratospheric warming is an abrupt warming of the polar stratosphere initiated by upward propagating planetary waves. Typically, GCMs have difficulties reproducing the observed SSW frequency (Charlton and Polvani 2007; Charlton et al. 2007). The reasons for this are not certain, but it is likely that the representation of the mean wind in the polar stratosphere has a large influence on SSW frequency.

To calculate the occurrence of SSWs in WACCM we use an algorithm based on the original World Meteorological Organization definition of SSWs. This algorithm is the same as described in Richter et al. (2008):

• *Major midwinter warming* is an event during which the temperature gradient between 60° and 90°N at 10 hPa is positive for at least five days and the zonal mean

FIG. 9. As in Fig. 8 but for zonal mean temperature (K).

wind at 60°N at 10 hPa is easterly during that time. If the temperature gradient becomes negative for less than three days and then becomes positive again, it is still considered a part of the same warming event.

• *Minor warming* is an event during which the temperature gradient between 60° and 90°N at 10 hPa is positive for at least five days, but there is no reversal of circulation at 10 hPa.

Table 1 summarizes the frequency of stratospheric warmings for ERA-40, WACCM3, and WACCM3.5. In WACCM3, the frequency of minor warmings was comparable to observations; however, this model was not able to produce very many major warmings. The major warming frequency in WACCM3 was only 0.1 averaged over both November–February (NDJF) and November–March (NDJFM). Note that the simulation presented in Richter et al. (2008) used climatological sea surface temperatures and had no prescribed or internally generated QBO; however, the lack of these factors is not likely to account for the low warming counts. WACCM3.5 shows a 40% increase in NDJF minor warmings compared to WACCM3. The new frequency of 1.0 is within 10% of the observed, and the NDJFM minor warming frequency is the same as observed. The

largest change from WACCM3 to WACCM3.5 occurs in the frequency of major stratospheric warmings. In NDJFM the frequency of major warmings is 0.6, the same as in ERA-40. In NDJF, the frequency of major warmings is 0.4, 4 times larger than in WACCM3 and close to the observed frequency of 0.5.

FIG. 10. Descent of the zero zonal-mean wind line at 60°S calculated from the climatological mean of monthly mean zonal winds. For ERA-40, data from 1985 through 2002 is used.
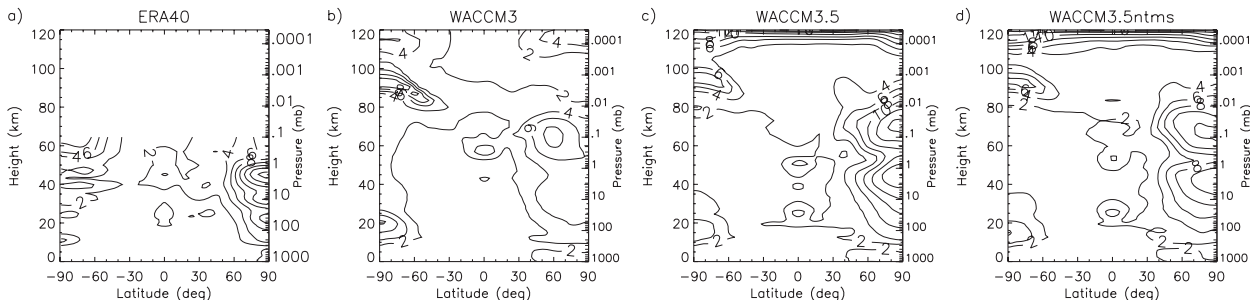
FIG. 11. January standard deviation of temperature (K) for (a) ERA-40, (b) WACCM3, (c) WACCM3.5, and (d) WACCM3.5ntms. Contour interval is 2 K.

The improvements in WACCM3.5 variability are associated with the improvement of the mean zonal wind during January–March. These are shown in Fig. 12. In WACCM3, the NH extratropical stratospheric winds were consistently too strong. In WACCM3.5 the NH stratospheric winds are very close to ERA-40 observations, especially in January and February. In March the 20 m s^{-1} contour still goes down to 20 hPa in WACCM3.5, whereas it only reaches 5 hPa in ERA-40. Note that Rind et al. (1988b) also noted a change in model variability and stratospheric warmings as a result of changes to the gravity wave parameterization.

The change in NH stratospheric variability and in the frequency of sudden stratospheric warmings is the largest improvement in WACCM3.5 as compared to WACCM3. One is tempted to conclude that these improvements are due only to the new representation of gravity waves in WACCM3.5. As we will demonstrate in the following subsection, a large part of the improvement in the representation of standard deviation of NH temperature in WACCM3.5 is due to the new gravity wave parameterization; however, the improvement in the frequency of sudden stratospheric warmings comes from the addition of turbulent mountain stress (TMS).

3) Cause of variability improvements

We have carefully examined all of the new additions to WACCM3.5 and have carried out sensitivity experiments with the GW tuning parameters in order to understand their role in simulating middle atmospheric climate. We did find that middle atmospheric climate is sensitive to the remaining GW tuning parameters. In

particular, the mesopause temperatures and height are dependent on the amplitude of frontally generated GWs (τ_b) and the frontogenesis function. We do not show the numerous simulations here but, in general, find that, if τ_b is set to a higher value, GWs break at a lower altitude, causing the mesopause height to be lower than observed (and vice versa). If the frontogenesis threshold is set to a lower value, there is not enough GW breaking in the mesosphere, causing the mesopause to be too high and its temperatures too warm.

Different settings of the remaining GW tuning parameters do impose changes of the NH stratospheric interannual variability. However, in all the simulations carried out with the source-oriented gravity wave parameterization, the standard deviation of NH stratospheric temperatures is much closer to observations than in WACCM3, implying that the more realistic variability in the GW sources, especially in the extratropics, improves the variability in the extratropical stratosphere. However, we find that the large improvement in the frequency of SSWs is not solely due to the source-oriented GW parameterization but, in large part, is due to the inclusion of turbulent mountain stress owing to surface roughness. We demonstrate this below by comparing the WACCM3.5 simulation with a WACCM3.5 simulation without TMS. We call the simulation without TMS WACCM3.5ntms and the only difference from WACCM3.5 in this simulation is the lack of TMS.

Figure 11d shows the temperature standard deviation during the month of January for WACCM3.5ntms. The values in the stratosphere are very similar to those in WACCM3.5. They differ only in the maximum values

TABLE 1. Frequency of occurrence of stratospheric sudden warmings: number of events per year.

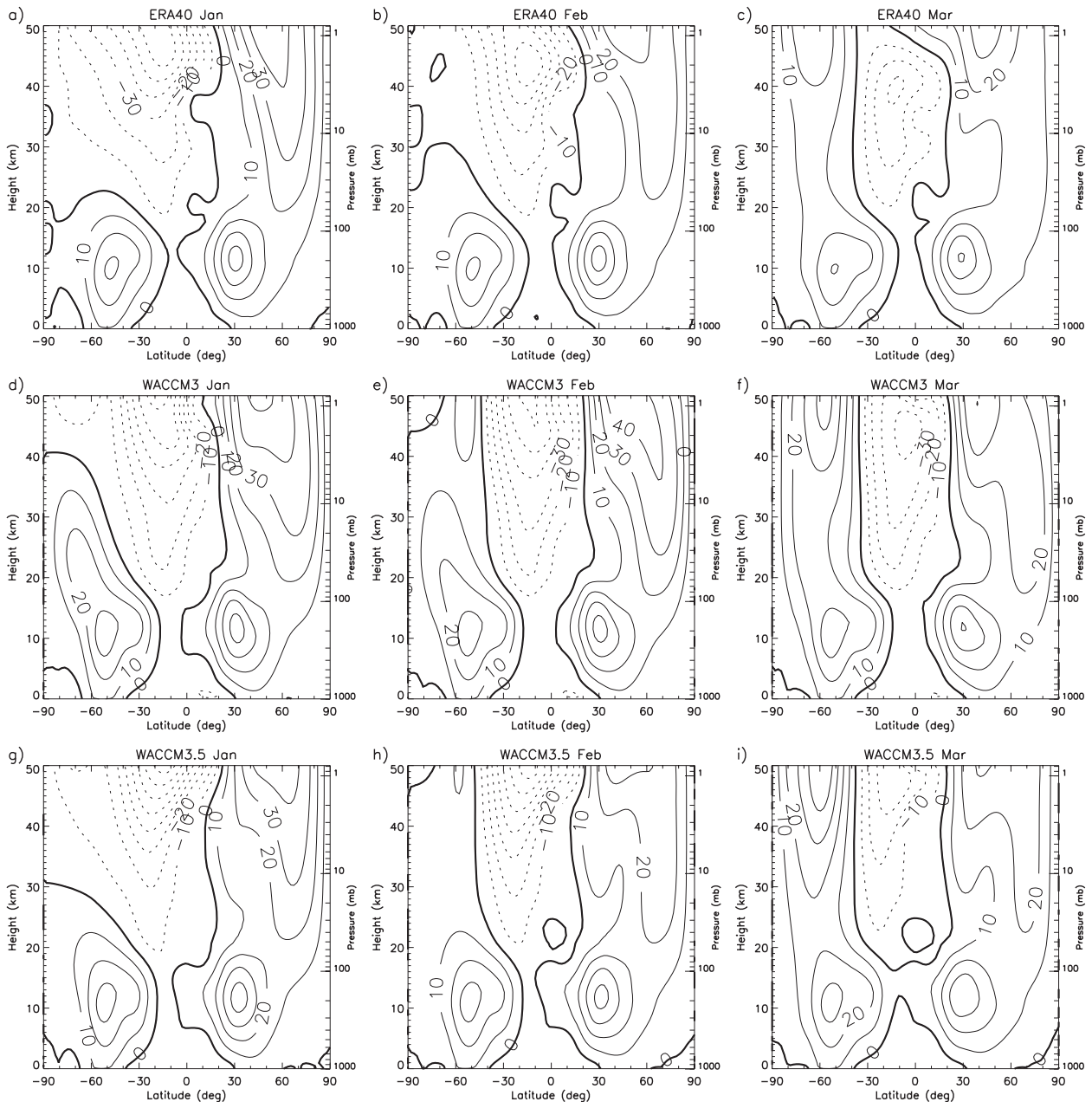| Warming type	ERA-40	WACCM3	WACCM3.5	WACCM3.5ntms
Major midwinter (NDJF)	0.5	0.1	0.4	0.1
Major midwinter (NDJFM)	0.6	0.1	0.6	0.25
Minor (NDJF)	0.9	0.7	1.0	0.4
Minor (NDJFM)	1.4	1.4	1.4	0.65

FIG. 12. (left) January, (middle) February, and (right) March zonal mean winds for (top) ERA-40, (middle) WACCM3, and (bottom) WACCM3.5. Contour interval is 10 m s^{-1}.

over the polar cap between 10 and 1 hPa, where they are 2–4 K larger in WACCM3.5 than in WACCM3.5ntms. The statistics of stratospheric sudden warmings for WACCM3.5ntms are shown in the last column of Table 1. We see that there is a striking change in SSW frequency between WACCM3.5 and WACCM3.5ntms. The frequency of both minor and major warmings in WACCM3.5ntms is much reduced as compared to WACCM3.5. There are only two major midwin-

ter warmings in NDJF, a fourfold decrease from WACCM3.5; the reduction in NDJFM is more than half. The frequency of minor warmings in WACCM3.5ntms is only half of that in WACCM3.5. Since there are no differences in GW tuning between WACCM3.5 and WACCM3.5ntms, this suggests that the frequency of stratospheric sudden warmings is largely controlled by the resolved waves rather than parameterized gravity waves.
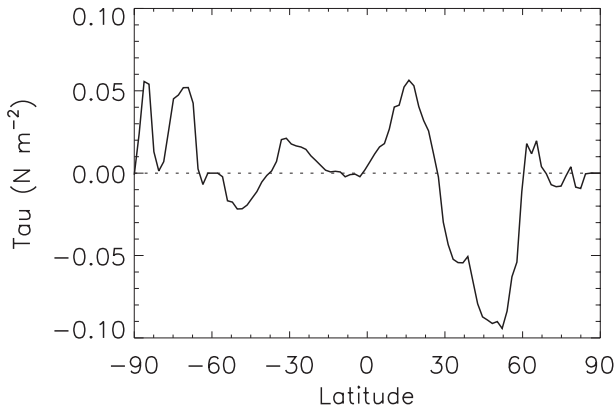
FIG. 13. Zonally averaged turbulent mountain stress (N m^{-2}) in WACCM3.5ntms for December, January, and February.

FIG. 14. DJF zonal wind difference: WACCM3.5 − WACCM3.5ntms. Contours are $\pm(1, 2, 3, 4, 5, 10, 15, 20, 25)$ m s^{-1}. Light and dark shading represent regions with Student's t-test values at the 95% and 99% levels, respectively.

TMS in WACCM3.5 adds a surface drag term; hence, the circulation in the troposphere is changed as compared to WACCM3.5ntms. Figure 13 shows the zonal component of TMS averaged zonally for DJF. The largest drag on the mean flow occurs between 30° and 60°N. Figure 14 shows the zonal wind difference between WACCM3.5 and WACCM3.5ntms for the same time period. Corresponding to the large values of TMS, the surface winds in WACCM3.5 are reduced between 30° and 60°N. In a zonal average this change is by 2 m s^{-1} near 40°N; however, these changes are much larger at individual longitudes corresponding to mountain regions. Changes to the surface momentum budget also cause changes to the upper tropospheric circulation in the extratropics. There is a statistically significant region of increased westerlies, in WACCM3.5, centered on 40°N between 300 and 50 hPa. This wind change is primarily due to changes in orographic gravity wave drag. Figure 15 shows that the
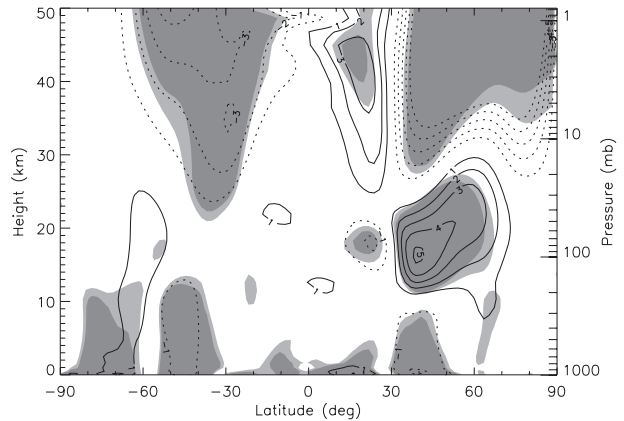
gravity wave drag near 40°N (between 300 and 50 hPa) is up to 3 m s^{-1} day^{-1} stronger in WACCM3.5ntms than in WACCM3.5, almost a threefold increase. This change in orographic GW drag leads to weaker westerlies in the lower stratosphere in WACCM3.5ntms. The orographic gravity wave generation is very much dependent on the surface winds and, since those have increased by the removal of TMS in WACCM3.5ntms, the orographic gravity wave generation also has increased, and so has the forcing from gravity waves on the mean flow. The differences in the zonal wind in the NH troposphere between WACCM3.5 and WACCM3.5ntms apparently are not due to the EP flux divergence from resolved waves, as we will illustrate below.

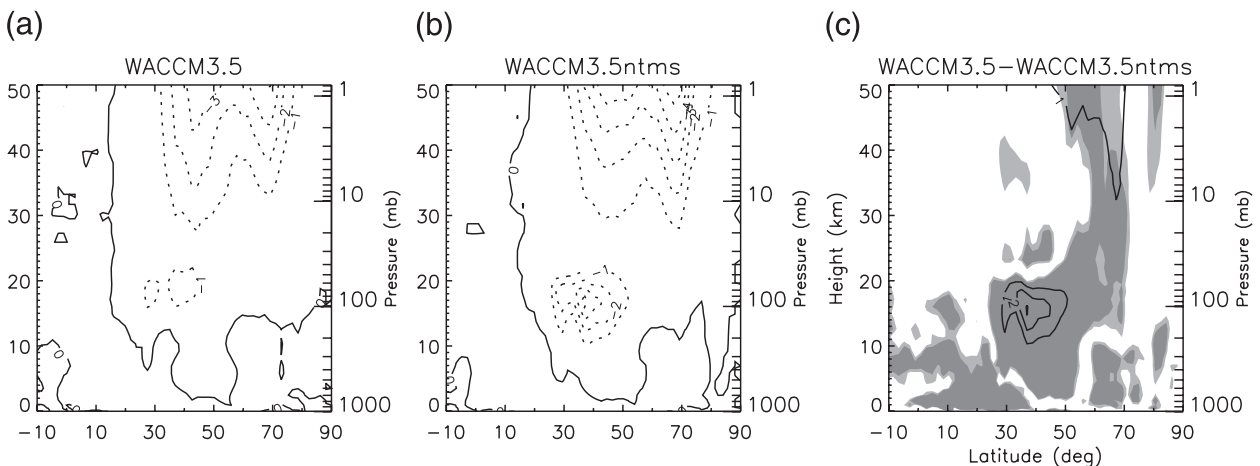In the NH stratosphere the winds in WACCM3.5 are much weaker, as compared to WACCM3.5ntms,

FIG. 15. DJF averaged gravity wave drag for (a) WACCM3.5, (b) WACCM3.5ntms, and (c) WACCM3.5 − WACCM3.5ntms. Contours interval is 1.0 m s^{-1} day^{-1}. Light (dark) shading in (c) represent regions with Student's t-test values at the 95% (99%) levels.
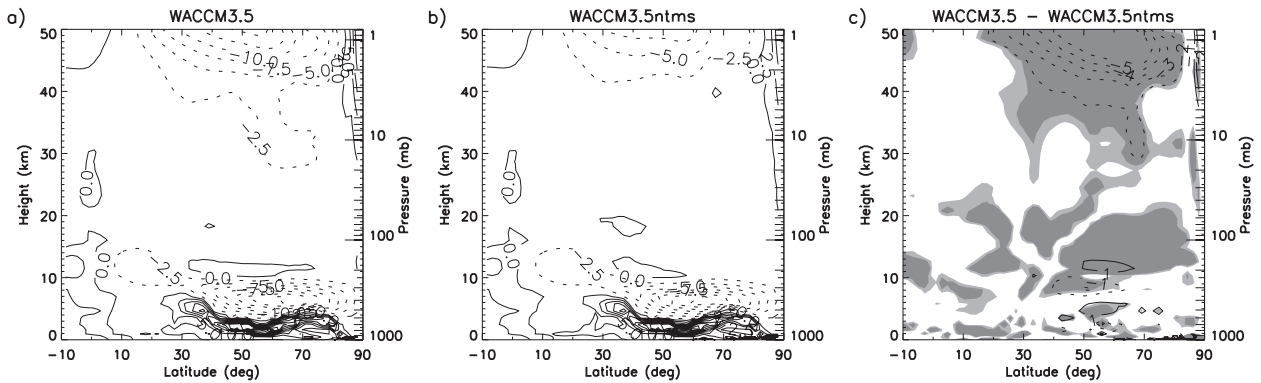
FIG. 16. As in Fig. 15 but of EP flux divergence. Contours interval is 2.5 m s^{-1} day^{-1} in (a) and (b) and in (c) $\pm(1, 2, 3, 4, 5, 10, 15, 20)$ m s^{-1} day^{-1}.

especially above 10 mb where differences can exceed 15 m s^{-1} in a seasonal average. As we show below, these changes are mainly related to stronger upward propagation and stronger momentum deposition by planetary waves in this region.

Figure 16 shows the EP flux divergence for WACCM3.5, WACCM3.5, and their difference averaged over DJF. The EP fluxes shown are calculated from 20 years of daily data. We have also calculated the EP fluxes from monthly mean data to isolate the stationary wave component (not shown). Figures 16a and 16b show that there is large region of resolved wave generation between 30° and 70°N in the lowermost troposphere, indicated by positive EP flux divergence values. Immediately above that region, there is already some momentum deposition to the mean flow from the resolved waves. Figure 16c shows that there are not many regions of significant differences between the resolved waves in WACCM3.5 and WACCM3.5ntms in the generation region and up to about 7 km. This suggests that the generation of resolved planetary waves is not affected by the turbulent mountain stress. Significant differences in resolved waves start appearing above 200 hPa: in WACCM3.5 the momentum deposition from planetary waves is smaller than in WACCM3.5ntms, indicated by more positive EP flux divergence. Above 30 mb the situation is reversed: there is much stronger momentum deposition in WACCM3.5, related to the reduced zonal mean wind values. Comparison of EP flux divergence derived from daily data to that derived from monthly mean data (not shown) shows that most of the momentum deposition above 200 hPa is due to stationary planetary waves in both simulations considered here.

The changes in EP flux divergence shown in Fig. 16 are consistent with the planetary wave structure at various altitudes shown in Fig. 17. At 500 and 200 hPa the amplitude and phase of the stationary planetary waves are very similar between WACCM3.5 and WACCM3.5ntms. However, at 100 hPa, above the region where there is more momentum deposition in WACCM3.5ntms, the stationary wave amplitude is much stronger in WACCM3.5 as compared to WACCM3.5ntms. Hence, in WACCM3.5 planetary waves propagating into the stratosphere are stronger, leading to more planetary wave activity reaching the polar vortex and initiating more SSWs. This is further illustrated in Fig. 18, which shows EP flux vector differences between WACCM3.5 and WACCM3.5ntms. Compared to WACCM3.5ntms, WACCM3.5 exhibits stronger poleward planetary wave propagation at latitudes north of 40°N near 15 km (Fig. 18c). The increased propagation of planetary waves into the stratospheric polar vortex causes more frequent initiation of SSWs.

It is interesting to know what causes more poleward planetary wave propagation. This question cannot be answered with certainty because interactions among planetary waves, gravity waves, and the mean flow involve feedbacks. In particular, wave forcing affects the background winds, while the winds affect the propagation of the waves, so establishing clear causal links is difficult. One possibility is that gravity waves directly affect the planetary waves and hence alter their direction and amplitude. Another possibility is the response of planetary waves to the changed mean flow as a result of changed GW drag. Figure 18 also shows the squared refractive index in WACCM3.5 and WACCM3.5ntms. The refractive index squared, n_s^2, is defined as

$$n_s^2 \equiv \frac{\overline{q}_\phi}{a\overline{u}} - \frac{s^2}{a^2 \cos^2\phi} - \frac{f^2}{4N^2H^2}, \qquad (5)$$

following Eq. (5.3.7) of Andrews et al. (1987) for quasigeostrophic waves. Figure 18a shows that in WACCM3.5 there is a minimum in refractive index centered on 50°N,
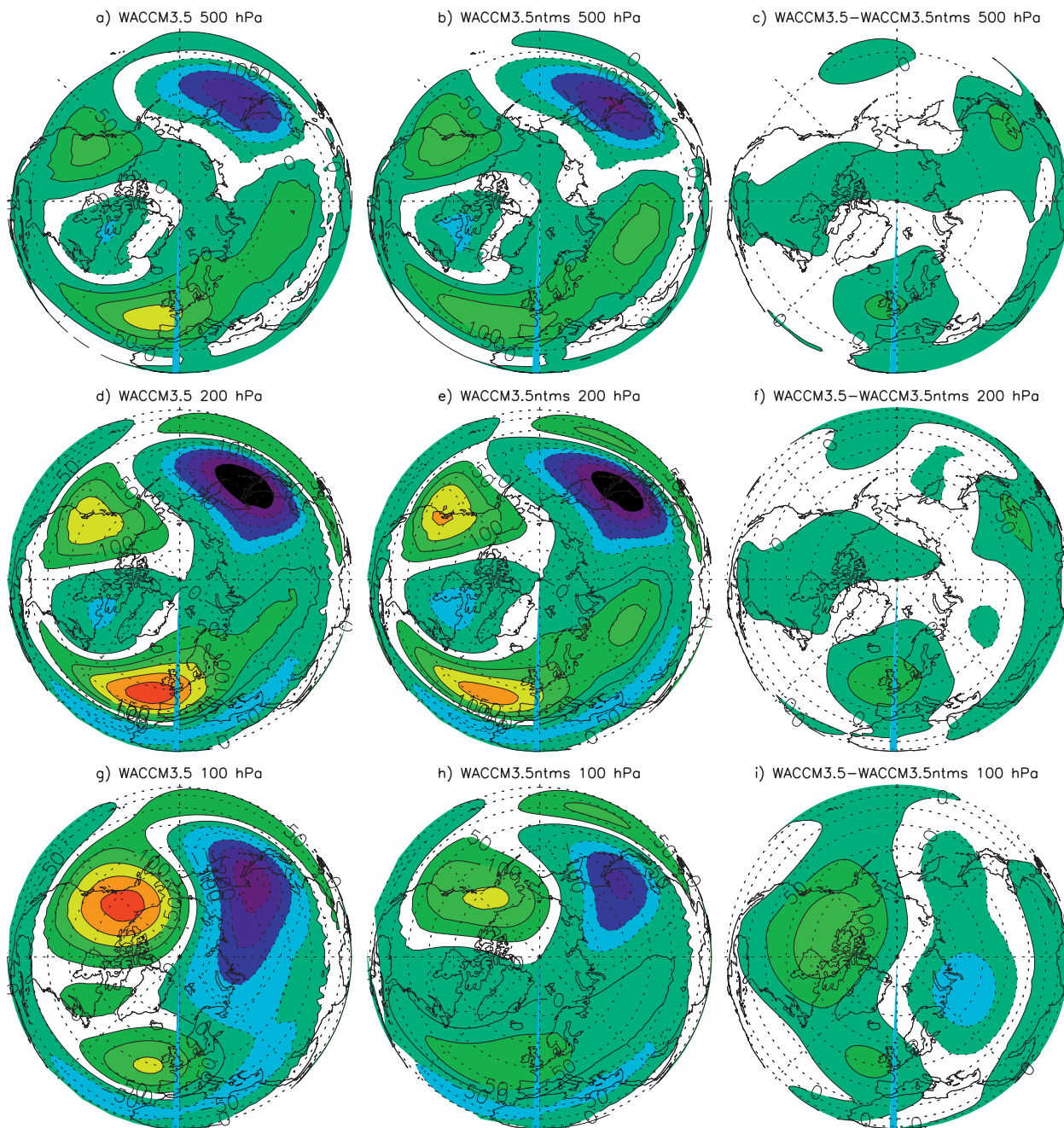
FIG. 17. DJF averaged geopotential height deviations from a longitudinal mean for (left) WACCM3.5, (middle) WACCM3.5ntms, and (right) their difference. The fields are plotted at (top) 500 hPa, (middle) 200 hPa, and (bottom) 100 hPa. Contour interval is 50 m in all panels.

12 km. According to Matsuno (1970), planetary waves will tend to be refracted away from regions of refractive index minimum. Hence, in WACCM3.5 equatorward propagation of planetary waves could be inhibited by the refractive index minimum, causing planetary wave activity to be diverted toward the polar regions. In WACCM3.5ntms (Fig. 18b) the refractive index mini-

mum is centered on 40°N, 15 km. Hence, compared to WACCM3.5, planetary waves can propagate farther equatorward, and fewer waves can propagate into the polar stratosphere. Note that, since the refractive index is defined for quasigeostrophic flow, it may not fully describe the changes between the two simulations in planetary wave structure. However, since the refractive
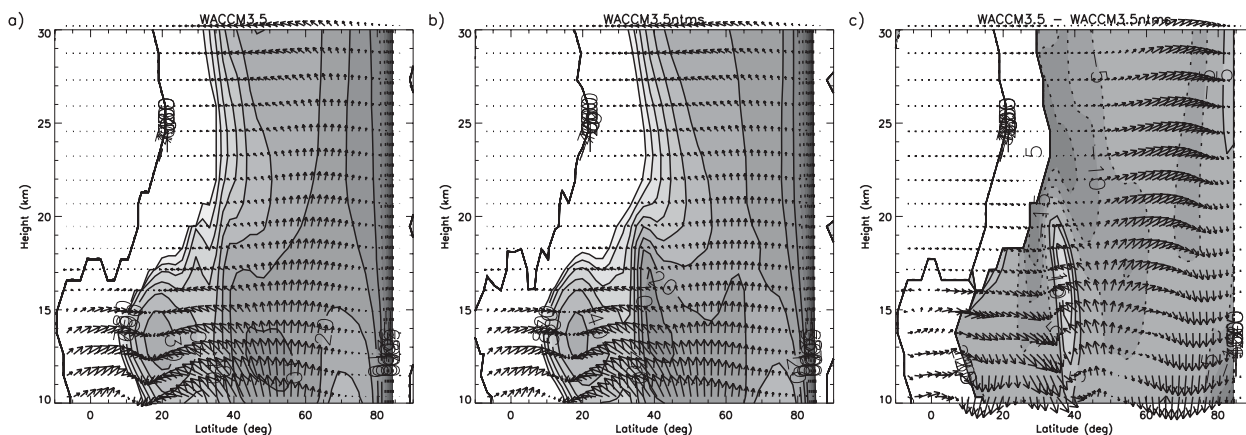
FIG. 18. January refractive index for wavenumber 1 ($a^2 n_1^2$) for (a) WACCM3.5, (b) WACCM3.5ntms, and (c) WACCM3.5 − WACCM3.5ntms. Contour interval is 10 in (a) and (b) and 5 in (c); arrows depict EP flux vectors.

index is one of the few diagnostics available to examine the response of planetary waves due to the mean flow changes, we find our analysis useful.

In short, TMS leads to altered orographic GW generation and altered tropospheric circulation. The altered GW drag and/or tropospheric circulation cause changes in the upward propagation of stationary planetary waves, which in turn affect the wave propagation into the polar vortex and the initiation of SSWs.

4. Discussion and conclusions

We have shown the effects of including a source-oriented GW parameterization in the WACCM middle atmospheric GCM. In our GW source parameterization we have included two dominant nonstationary wave sources: convection and frontal systems, in addition to the stationary orographically generated gravity waves. With the use of only those sources, the mean zonal wind and temperature structure of the middle atmosphere (in WACCM3.5) is as good as, and in several important respects better than, the simulation of the middle atmosphere with an arbitrarily specified parameterization (in WACCM3). To achieve this, we have lowered the frontogenesis threshold below values suggested by other authors (see section 2). This suggests that we may be overestimating the gravity wave generation from frontal systems and that another extratropical GW source, such as shear generation, may need to be considered in future studies.

Many of the features of the mean middle atmospheric circulation, such as the structure of stratospheric and mesospheric jets, are similar between WACCM3 and WACCM3.5. This suggests that to first order the specification of the GW sources does not alter the general feature of the middle atmospheric climatology; how-

ever, it affects the details. The cold polar stratosphere bias is improved in WACCM3.5, as well as the NH winter lower stratospheric winds. Planetary wave structure is always altered as a result of GW drag or zonal wind changes. It is very difficult to say how much of the improvement in WACCM3.5 is due to gravity waves versus planetary waves, given that the two continuously interact. Nevertheless, as we have described earlier, in the SH DJF the largest changes in the lower stratosphere are in GW drag, and EP flux divergence from planetary waves is similar between WACCM3 and WACCM3.5. Hence, this improvement is likely due directly to changed GW drag. However, as shown in the previous section, in the NH the altered planetary wave structure has a large effect on the model's variability.

In the source-oriented GW parameterization presented here, there are still tuning parameters and uncertainties in source specification. These are mainly related to the frontal/baroclinic wave generation. In particular, relationships between GWs generated from frontal systems and the properties of those systems are needed; the amplitude and momentum flux phase-speed spectrum of these waves are not well known. Another uncertainty in the frontal GW source parameterization is related to the diagnostic of the GW generation region. At the moment, we use the frontogenesis function following Hoskins (1982) and Charron and Manzini (2002); in the future we plan to use the nonlinear balance equation diagnostic of Zhang (2004). Global observations of gravity waves would also be helpful in constraining gravity wave source parameterizations; in particular, observations of gravity wave phase speed spectra are needed. Even if there was not much improvement in the middle atmospheric climate as a result of including individual GW source parameterizations, including realistic source specifications is arguably superior to a standard

approach because it is model consistent and allows for middle atmospheric circulation changes as a result of changes in tropospheric sources.

We have found that the NH stratospheric interannual variability and the frequency of SSWs are much closer to observations in WACCM3.5 than in WACCM3. The improvement in interannual variability can be attributed to the source-oriented GW parameterization, which provides more realistic GW source variability. The improvement in SSW frequency is primarily due to the added turbulent mountain stress near the surface that, through interactions with the mean wind and orographic gravity waves, alters the propagation of stationary planetary waves into the polar vortex. This suggests that both GWs and planetary waves must be represented correctly in GCMs in order to properly represent stratospheric variability.

The WACCM3.5 model configuration is being used for the recent Chemistry-Climate Model Validation Activity (CCMVAL) for SPARC. As part of this activity, simulations are being carried out with WACCM3.5 from the early 1950s into the twenty-first century. These simulations will allow for the examination of GW source changes with changing climate.

REFERENCES

Alexander, M. J., and L. Pfister, 1995: Gravity wave momentum flux in the lower stratosphere over convection. *Geophys. Res. Lett.,* **22,** 2029–2032.

——, and J. Holton, 1997: A model study of zonal forcing in the equatorial stratosphere by convectively induced gravity waves. *J. Atmos. Sci.,* **54,** 408–419.

——, and T. J. Dunkerton, 1999: A spectral parameterization of mean-flow forcing due to breaking gravity waves. *J. Atmos. Sci.,* **56,** 4167–4182.

——, J. H. Beres, and L. Pfister, 2000: Tropical stratospheric gravity wave activity and relationship to clouds. *J. Geophys. Res.,* **105,** 22 299–22 309.

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

Balachandran, N., and D. Rind, 1995: Modeling the effects of UV variability and the QBO on the troposphere–stratosphere system. Part I: The middle atmosphere. *J. Climate,* **8,** 2058–2079.

Beres, J. H., 2004: Gravity wave generation by a three-dimensional thermal forcing. *J. Atmos. Sci.,* **61,** 1805–1815.

——, M. J. Alexander, and J. R. Holton, 2002: Effects of tropospheric wind shear on the spectrum of convectively generated gravity waves. *J. Atmos. Sci.,* **59,** 1805–1824.

——, ——, and ——, 2004: A method of specifying the gravity wave spectrum above convection based on latent heating properties and background wind. *J. Atmos. Sci.,* **61,** 324–337.

——, R. R. Garcia, B. A. Boville, and F. Sassi, 2005: Implementation of a gravity wave source spectrum parameterization dependent on the properties of convection in the Whole Atmosphere Community Climate Model (WACCM). *J. Geophys. Res.,* **110,** D10108, doi:10.1029/2004JD005504.

Charlton, A. J., and L. M. Polvani, 2007: A new look at stratospheric sudden warmings. Part I. Climatology and modeling benchmarks. *J. Climate,* **20,** 449–469.

——, and Coauthors, 2007: A new look at stratospheric sudden warmings. Part II: Evaluation of numerical model simulations. *J. Climate,* **20,** 470–488.

Charron, M., and E. Manzini, 2002: Gravity waves from fronts: Parameterization and middle atmosphere response in a general circulation model. *J. Atmos. Sci.,* **59,** 923–941.

Dewan, E. M., and Coauthors, 1998: MSX satellite observations of thunderstorm-generated gravity waves in mid-wave infrared images of the upper stratosphere. *Geophys. Res. Lett.,* **25,** 939–942.

Dörnbrack, A., M. Leutbecher, R. Kivi, and E. Kyrö, 1999: Mountain-wave-induced record low stratospheric temperatures above northern Scandinavia. *Tellus,* **51A,** 951–963.

Eyring, V., and Coauthors, 2006: Assessment of temperature, trace species, and ozone in chemistry-climate model simulations of the recent past. *J. Geophys. Res.,* **111,** D22308, doi:10.1029/2006JD007327.

Fiedler, F., and H. A. Panofsky, 1972: The geostrophic drag coefficient and the "effective" roughness length. *Quart. J. Roy. Meteor. Soc.,* **98,** 213–220.

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

Fritts, D., and M. J. Alexander, 2003: Gravity wave dynamics and effects in the middle atmosphere. *Rev. Geophys.,* **41,** 1003, doi:10.1029/2001RG000106.

Garcia, R. R., and S. Solomon, 1985: The effect of breaking gravity waves on the dynamics and chemical composition of the mesosphere and lower thermosphere. *J. Geophys. Res.,* **90** (D2), 3850–3868.

——, and B. A. Boville, 1994: "Downward control" of the mean meridional circulation and temperature distribution of the polar winter stratosphere. *J. Atmos. Sci.,* **51,** 2238–2245.

——, D. R. Marsh, D. E. Kinnison, B. A. Boville, and F. Sassi, 2007: Simulation of secular trends in the middle atmosphere, 1950–2003. *J. Geophys. Res.,* **112,** D09301, doi:10.1029/2006JD007485.

Griffiths, M., and M. J. Reeder, 1996: Stratospheric inertia-gravity waves generated in a numerical model of frontogenesis. I: Model solutions. *Quart. J. Roy. Meteor. Soc.,* **122,** 1153–1174.

Holstlag, A. M., and B. A. Boville, 1993: Local versus nonlocal boundary layer diffusion in a global climate model. *J. Climate,* **6,** 1825–1842.

Hoskins, B. J., 1982: The mathematical theory of frontogenesis. *Annu. Rev. Fluid Mech.,* **14,** 131–151.

Lane, T. P., M. J. Reeder, and T. L. Clark, 2001: Numerical modeling of gravity wave generation by deep tropical convection. *J. Atmos. Sci.,* **58,** 1249–1274.

Lilly, D. K., 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.

Lin, S.-J., 2004: A "vertically Lagrangian" finite-volume dynamical core for global models. *Mon. Wea. Rev.,* **132,** 2293–2307.

Lindzen, R. S., 1981: Turbulence and stress owing to gravity wave and tidal breakdown. *J. Geophys. Res.,* **86** (C10), 9707–9714.

Manzini, E., and N. A. McFarlane, 1998: The effect of varying the source spectrum of a gravity wave parameterization in a middle atmosphere general circulation model. *J. Geophys. Res.,* **103,** 31 523–31 539.

Matsuno, T., 1970: Vertical propagation of stationary planetary waves in the winter Northern Hemisphere. *J. Atmos. Sci.,* **27,** 871–883.

Matthes, K., U. Langematz, L. Gray, K. Kodera, and K. Labitzke, 2004: Improved 11-year solar signal in the Freie Universität Berlin Climate Middle Atmosphere Model (FUB-CMAM). *J. Geophys. Res.,* **109,** D06101, doi:10.1029/2003JD004012.

McFarlane, N. A., 1987: The effect of orographically excited wave drag on the general circulation of the lower stratosphere and troposphere. *J. Atmos. Sci.,* **44,** 1775–1800.

Miller, J. E., 1948: On the concept of frontogenesis. *J. Meteor.,* **5,** 169–171.

Neale, R. B., J. H. Richter, and M. Jochum, 2008: The impact of convection on ENSO: From a delayed oscillator to a series of events. *J. Climate,* **21,** 5904–5924.

Oleson, K. W., and Coauthors, 2008: Improvements to the Community Land Model and their impact on the hydrological cycle. *J. Geophys. Res.,* **113,** G01021, doi:10.1029/2007JG000563.

O'Sullivan, D., and T. J. Dunkerton, 1995: Generation of inertia–gravity waves in a simulated life cycle of baroclinic instability. *J. Atmos. Sci.,* **52,** 3695–3716.

Palmer, T. N., J. Shutts, and R. Swinbank, 1996: Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parameterization. *Quart. J. Roy. Meteor. Soc.,* **112,** 1001–1039.

Piani, C., and D. Durran, 2001: A numerical study of stratospheric gravity waves triggered by squall lines observed during the TOGA COARE and COPT-81 experiments. *J. Atmos. Sci.,* **58,** 3702–3723.

Reeder, M. J., and M. Griffiths, 1996: Stratospheric inertia–gravity waves generated in a numerical model of frontogenesis. II: Wave sources, generation mechanisms, and momentum fluxes. *Quart. J. Roy. Meteor. Soc.,* **122,** 1175–1195.

Richter, J. H., and P. J. Rasch, 2008: Effects of convective momentum transport on the atmospheric circulation in the Community Atmosphere Model, version 3. *J. Climate,* **21,** 1487–1499.

——, F. Sassi, R. R. Garcia, K. Matthes, and C. Fischer, 2008: Dynamics of the middle atmosphere as simulated by the Whole Atmosphere Community Climate Model, version 3 (WACCM). *J. Geophys. Res.,* **113,** D08101, doi:10.1029/2007JD009269.

Rind, D., R. Suozzo, N. K. Balachandran, A. Lacis, and G. Russel, 1988a: The GISS global climate global atmosphere model: Part I: Model structure and climatology. *J. Atmos. Sci.,* **45,** 329–370.

——, ——, and ——, 1988b: The GISS global climate global atmosphere model: Part II: Model variability due to interactions between planetary waves, the mean circulation, and gravity wave drag. *J. Atmos. Sci.,* **45,** 371–386.

——, J. Lerner, J. Jonas, and C. McLinden, 2007: The effects of resolution and model physics on tracer transports in the NASA Goddard Institute for Space Studies general circulation models. *J. Geophys. Res.,* **112,** D09315, doi:10.1029/2006JD007476.

Salby, M. L., and R. R. Garcia, 1987: Transient response to localized episodic heating in the tropics. Part I: Excitation and short-time near-field behavior. *J. Atmos. Sci.,* **44,** 458–498.

Song, I.-S., and H.-Y. Chun, 2005: Momentum flux spectrum of convectively forced internal gravity waves and its application to gravity wave drag parameterization. Part I: Theory. *J. Atmos. Sci.,* **62,** 107–124.

Vincent, R. A., and I. M. Reid, 1983: HF Doppler measurements of mesospheric gravity wave momentum fluxes. *J. Atmos. Sci.,* **40,** 1321–1333.

Wang, S., and F. Zhang, 2007: Sensitivity of mesoscale gravity waves to the baroclinicity of jet-front systems. *Mon. Wea. Rev.,* **135,** 670–688.

Warner, C. D., and M. E. McIntyre, 2001: An ultrasimple spectral parameterization for nonorographic gravity waves. *J. Atmos. Sci.,* **58,** 1837–1857.

Webster, S., A. R. Brown, D. R. Cameron, and P. Jones, 2003: Improvements to the representation of orography in the Met Office unified model. *Quart. J. Roy. Meteor. Soc.,* **129,** 1989–2010.

Wilson, J. D., 2002: Representing drag on unresolved terrain as a distributed momentum sink. *J. Atmos. Sci.,* **59,** 1629–1637.

Wu, D. L., and J. W. Waters, 1996: Satellite observations of atmospheric variances: A possible indication of gravity waves. *Geophys. Res. Lett.,* **23,** 3631–3634.

Xu, J., H.-L. Liu, W. Yuan, A. K. Smith, R. G. Roble, C. J. Mertens, J. M. Russell III, and M. G. Mlynczak, 2007: Mesopause structure from Thermosphere, Ionosphere, Mesosphere, Energetics, and Dynamics (TIMED)/Sounding of the Atmosphere Using Broadband Emission Radiometry (SABER) observations. *J. Geophys. Res.,* **112,** D09102, doi:10.1029/2006JD007711.

Zhang, F., 2004: Generation of mesoscale gravity waves in upper-tropospheric jet–front systems. *J. Atmos. Sci.,* **61,** 440–457.

Zhang, G. J., and N. A. McFarlane, 1995: Role of convective scale momentum transport in climate simulation. *J. Geophys. Res.,* **100** (D1), 1417–1426.