Dynamical control of the mesosphere by orographic and non-orographic gravity wave drag during the extended northern winters of 2006 and 2009

Article

Published Version

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Dynamical Control of the Mesosphere by Orographic and Nonorographic Gravity Wave Drag during the Extended Northern Winters of 2006 and 2009

CHARLES MCLANDRESS

Department of Physics, University of Toronto, Toronto, Ontario, Canada

JOHN F. SCINOCCA

Canadian Centre for Climate Modelling and Analysis, Victoria, British Columbia, Canada

THEODORE G. SHEPHERD

Department of Meteorology, University of Reading, Reading, United Kingdom

M. CATHERINE READER

University of Victoria, Victoria, British Columbia, Canada

GLORIA L. MANNEY

NorthWest Research Associates, and New Mexico Institute of Mining and Technology, Socorro, New Mexico

ABSTRACT

A version of the Canadian Middle Atmosphere Model (CMAM) that is nudged toward reanalysis data up to 1 hPa is used to examine the impacts of parameterized orographic and nonorographic gravity wave drag (OGWD and NGWD) on the zonal-mean circulation of the mesosphere during the extended northern winters of 2006 and 2009 when there were two large stratospheric sudden warmings. The simulations are compared to *Aura* Microwave Limb Sounder (MLS) observations of mesospheric temperature and carbon monoxide (CO) and derived zonal winds. The control simulation, which uses both OGWD and NGWD, is shown to be in good agreement with MLS. The impacts of OGWD and NGWD are assessed using simulations in which those sources of wave drag are removed. In the absence of OGWD the mesospheric zonal winds in the months preceding the warmings are too strong, causing increased mesospheric NGWD, which drives excessive downwelling, resulting in overly large lower-mesospheric values of CO prior to the warming. NGWD is found to be most important following the warmings when the underlying westerlies are too weak to allow much vertical propagation of the orographic gravity waves to the mesosphere. NGWD is primarily responsible for driving the circulation that results in the descent of CO from the thermosphere following the warmings. Zonal-mean mesospheric winds and temperatures in all simulations are shown to be strongly constrained by (i.e., slaved to) the stratosphere. Finally, it is demonstrated that the responses to OGWD and NGWD are nonadditive because of their dependence and influence on the background winds and temperatures.

1. Introduction

The northern winters of 2006 and 2009 were punctuated by two of the largest stratospheric sudden warmings

on record. Fortuitously, this was also a period when numerous research satellite instruments were in operation, providing an unprecedentedly detailed view of the atmosphere from the upper troposphere to the lower thermosphere (e.g., Manney et al. 2008a,b, 2009a,b; Randall et al. 2006, 2009). From those observations the following picture of the middle atmosphere emerged: In the months leading up to the warmings, which occurred in late January, the stratosphere was near its climatological

Corresponding author address: Charles McLandress, Department of Physics, University of Toronto, 60 St. George St., Toronto, ON M5S 1A7, Canada.
E-mail: charles@atmosp.physics.utoronto.ca

state, with a stratopause and zonal-mean westerly jet maximum located near 50 km. Following the warmings, the middle and upper stratosphere became extremely cold and undisturbed, and an elevated stratopause formed in the upper mesosphere and slowly descended to its climatological position by springtime. Chemical species such as nitrogen oxides ($NO_x = NO + NO_2$) and carbon monoxide (CO), which form in the upper atmosphere and are transported downward in winter, also underwent large changes, most notably strongly enhanced descent following the warmings.

These mesospheric observations have led to a concerted effort on the part of modeling groups to simulate and understand the dynamics and transport during and after these strong warming events. From the first comparisons of the mesospheric observations to standard assimilated meteorological analyses, it became apparent that models with lids higher than 80 km were needed in order to simulate the reformation of the stratopause at high altitudes after the warmings (Manney et al. 2008a).

In recent years, two high-lid data assimilation systems have been developed and used to study such warming events: the Canadian Middle Atmosphere Model (CMAM; Polavarapu et al. 2005) and the Navy Operational Global Atmospheric Prediction-Advanced Level Physics High Altitude (NOGAPS-ALPHA; Eckermann et al. 2009). One of the most remarkable findings to emerge was that the descent of the elevated stratopause following the warmings could be realistically reproduced without the assimilation of mesospheric temperature observations provided that the effects of gravity wave drag (GWD) were included, as was demonstrated by Ren et al. (2011) using the CMAM data assimilation system. Siskind et al. (2010) used a different approach by performing short-term forecasts initialized from an analysis of the NOGAPS-ALPHA system, but came to a similar conclusion. Both studies demonstrated that nonorographic GWD was responsible for the realistic mesospheric temperature response following the warmings, in support of the earlier suggestion by Ren et al. (2008) that the zonal-mean circulation in the mesosphere is slaved to the lower atmosphere through GWD.

Two recent studies using free-running (i.e., constrained only by prescribed sea surface temperatures and sea ice) versions of high-lid models have elucidated the causes of several key dynamical features of warming events like those in 2006 and 2009—in particular, the elevated stratopause and its subsequent descent. Using the Whole Atmosphere Community Climate Model (WACCM), Limpasuvan et al. (2012) argued that the formation of the elevated stratopause and its initial descent were due to mesospheric planetary wave drag.

Hitchcock and Shepherd (2013) used a time-dependent zonally averaged quasigeostrophic model driven by the zonal-mean torques from a free-running version of CMAM to diagnose the relative importance of the different types of wave drag and the radiative forcing, including their induced meridional circulations, which are a significant component of the mesospheric response to the sudden warming. This methodology allowed a complete quantitative attribution of the response to the different forcings. Their analysis showed that the mesospheric cooling immediately after the warming was due to a combination of radiative cooling and a transient circulation induced by the resolved wave drag in the stratosphere. As in the Limpasuvan et al. (2012) study, Hitchcock and Shepherd found that the formation of the elevated stratopause was initiated by the drag exerted by planetary waves propagating up from below. Both studies reconfirmed the findings of Siskind et al. (2010) and Ren et al. (2011) that the descent of the elevated stratopause during the recovery period was primarily driven by nonorographic GWD.

While free-running models are important for understanding the mechanisms responsible for the observed mesospheric response to warmings such as these, their model states are not sufficiently constrained to reproduce specific realizations of the atmospheric flow and thus cannot be compared quantitatively to observations of a particular event on a day-by-day basis. Since the winds and temperatures in the lower atmosphere have a strong influence on the propagation and absorption of gravity waves, free-running models likewise cannot be expected to reproduce the actual gravity wave fluxes entering the mesosphere for a particular event. It is only by constraining the winds and temperatures in the lower atmosphere (e.g., by using data assimilation) that the parameterized (and resolved) gravity wave fluxes could match the true fluxes. Doing so would then enable a quantitative evaluation of the mesospheric GWD in the model, which would be achieved by comparing the simulated large-scale winds and temperatures in the mesosphere to observations. Such a study has yet to be done. Moreover, none of the previous modeling studies examined the time period when the vortex was developing, focusing instead on the warmings and their aftermath, when the effects of nonorographic GWD were dominant. Thus, the extent to which orographic GWD controls the mesospheric circulation is unclear. It is also unknown whether the mesosphere is slaved to the lower atmosphere in late fall/early winter and, if it is, whether it is realized through orographic GWD, planetary wave drag, or a combination of the two.

Herein, we describe results from a study designed to investigate these issues using a version of CMAM in

which a simple relaxation procedure is employed to constrain the model to follow reanalysis data up to 1 hPa. Since the model includes interactive chemistry and online transport, we can examine the descent of chemical constituents like CO and NO_x. By validating our simulations against mesospheric observations of temperature and CO, as well as derived zonal winds, from the *Aura* Microwave Limb Sounder (MLS) we can quantify the role of parameterized GWD in driving the mesospheric circulation in the model and by inference in the real atmosphere. We assess the relative roles of orographic and nonorographic GWD using a set of sensitivity experiments in which the two types of drag are turned off, separately and together. This also enables us to assess the additivity of the responses to orographic and nonorographic GWD, which has hitherto not been done.

The advantage of our nudging approach is that the lower atmosphere (i.e., the portion below 1 hPa that is nudged) is to a large extent unaffected by the removal of one or another type of GWD, thus keeping the resolved waves and the momentum fluxes from the other GWD scheme largely unchanged throughout the lower atmosphere. This permits the separation of cause and effect in the mesosphere, which is not possible when the lower atmosphere is unconstrained, as in the case of the GWD sensitivity forecasts of Siskind et al. (2010). (Forecasts can only be run for short times before the stratospheric flow has departed so far from the observed state that the gravity wave fluxes entering the mesosphere are significantly changed.) Although the CMAM data assimilation system could be used here, the cost of performing the integrations would be prohibitive and there turns out to be little or no advantage for this type of study compared to our simple relaxational approach.

The paper is organized as follows. In section 2 we describe the data used. This includes a very brief discussion of the MLS data and a more in-depth discussion of CMAM and our simulations. We include here a fairly detailed description of the GWD parameterizations since they are crucial components of our study. In section 3 we present the results, starting with a description of the MLS observations and followed by a discussion of the control simulation for the "extended" winters (i.e., September–April) of 2006 and 2009. We then discuss the GWD sensitivity experiments, which enable us to understand the relative roles of orographic and nonorographic GWD in determining the zonal-mean dynamics and transport in the mesosphere. We end this section with a discussion of the additivity of the responses to orographic and nonorographic GWD. In section 4 we summarize our main results and discuss some implications of our findings.

2. Data

a. MLS observations

The MLS instrument on the *Aura* satellite has provided a nearly continuous set of measurements of temperature and trace gases in the middle atmosphere from August 2004 through the present. Here we use version 3.3 temperature and CO data (Livesey et al. 2011), as well as zonal winds derived from the MLS geopotential heights using a balanced wind formulation (Randel 1987) as described by Manney et al. (2008a). Daily average zonal means on a 2° latitude grid are used.

b. CMAM simulations

CMAM is a chemistry–climate model extending from the earth's surface to about 100 km. Detailed descriptions of the model and stratospheric chemistry scheme are given, respectively, in Scinocca et al. (2008) and de Grandpré et al. (2000) [with updates to the chemistry provided in Jonsson et al. (2004)]. The version of the model we use has a triangular spectral truncation of T47, corresponding to a 3.75° horizontal grid on which the physical parameterizations are evaluated. There are 71 levels in the vertical, with a resolution varying from several tens of meters in the lower troposphere to about 2.5 km in the mesosphere. To prevent wave reflections at the model lid a Rayleigh friction sponge is applied above about 85 km to the deviations of the horizontal winds from the zonal mean. Monthly and annually varying sea surface temperatures and sea ice distributions are prescribed using observations. The radiative forcings and chemical boundary conditions are the same as those used in SPARC CCMVal (2010).

Orographic GWD (OGWD) and nonorographic GWD (NGWD) are parameterized using the schemes of Scinocca and McFarlane (2000) and Scinocca (2003), respectively. The OGWD scheme employs two vertically propagating zero-phase-speed gravity waves (GWs) to transport the horizontal momentum contained in all waves directed into the half-space to the left and right of the large-scale (i.e., resolved) horizontal velocity vector at the launch layer, which extends from the surface to the height of the subgrid topography. The orientation and magnitude of the momentum flux carried by these two waves is a function of the near-surface wind speed, its direction relative to the orientation of the subgrid topography, and the static stability in the launch layer. The two relevant tunable parameters in the OGWD scheme are the multiplicative-scale factor, $G(v) = 0.65$, and the inverse critical Froude number, $Fr_{crit} = 0.375$. In general, $G(v)$ scales the total vertical flux of horizontal momentum and Fr_{crit} determines the breaking height. However, Fr_{crit} also has influence over the total amount

of launch momentum flux because it is used to set the maximum amplitude of the parameterized waves at launch (i.e., waves are not permitted to exceed their breaking amplitude at launch). The above values of $G(v)$ and Fr_{crit} have been tuned for polar-ozone chemistry studies in CMAM since they produce reasonable zonal-mean zonal winds and polar temperatures in the winter lower stratosphere (Scinocca et al. 2008).

The NGWD scheme employs a spectrum of nonzero-phase-speed GWs propagating horizontally in the four cardinal directions. The launch level vertical wavenumber (m) energy spectral density is prescribed to be the so-called Desaubies spectrum, being independent of m for $m \ll m*$ and with an m^{-3} dependence for $m \gg m*$, where $m* = 2\pi/(1 \text{ km})$, and is identical for all azimuths and spatially uniform. The total nonorographic GW Eliassen–Palm flux in each azimuth at the launch level (~125 hPa) is the primary tuning parameter and is typically set to 4.243×10^{-4} Pa (as it is here) since it produces a reasonable seasonal evolution of the zonal-mean zonal winds and temperatures in the mesosphere in CMAM. Since the launch spectrum is isotropic, the net zonal and meridional Eliassen–Palm fluxes at the launch level are zero.

As the parameterized orographic and nonorographic GWs propagate upward they are subject to both critical level filtering and nonlinear saturation. The latter is treated in the NGWD scheme by not permitting the energy spectrum to exceed the observed m^{-3} form, while in the OGWD scheme a convective instability threshold is applied. In the uppermost model layer all remaining GW momentum flux is deposited in order to prevent any spurious downward influence (Shepherd and Shaw 2004). The parameter settings used for both GWD schemes are identical to those used in the CMAM contribution to SPARC CCMVal (2010).

Below 1 hPa the horizontal winds and temperatures are nudged (i.e., relaxed) to the 6-hourly horizontal winds and temperatures from the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Re-Analysis (ERA-Interim; Dee et al. 2011). A key aspect of the relaxational approach adopted here is that it is designed to primarily constrain the synoptic space and time scales in the model, which are well represented in the reanalysis data. This is accomplished by performing the relaxation in spectral space and applying it only to horizontal scales with $n \leq n_{max} = 21$, where n is the total wavenumber. The nudging tendency, which is applied to the vorticity, divergence, and thermodynamic equations, has the form $-(X - X_R)/\tau_0$, where $\tau_0 = 24$ h is the relaxational time scale and X and X_R are, respectively, the model and reanalysis spectral vorticity, divergence, or temperature coefficients. Linear interpolation is used to compute the reanalysis variables at intermediate times between adjacent 6-h intervals. The values of n_{max} and τ_0 were chosen so that the RMS differences between the nudged temperature fields and ERA-Interim are on average roughly similar to the RMS differences between pairs of different reanalysis datasets [i.e., ERA-Interim, 40-yr ECMWF Re-Analysis (ERA-40), National Centers for Environmental Prediction reanalysis (NCEP), and National Centers for Environmental Prediction—Department of Energy Reanalysis 2 (NCEP2)], where the latter differences provide an estimate of the uncertainty in the reanalysis products (Merryfield et al. 2013).

To demonstrate the impact of the nudging on the model temperatures, we present here results from several short simulations using three different nudging configurations. In the first the zonal means are nudged up to 1 hPa using the constant τ_0 and the deviations from the zonal means are nudged up to 1 hPa using a height-dependent relaxational time scale that is equal to τ_0 below 10 hPa and whose inverse is tapered to zero between 10 and 1 hPa. The tapering is used to minimize the possible distortion of vertically propagating (resolved) waves that might occur just above 1 hPa where the relaxational time scale jumps to infinity. We will refer to this as tapered nudging. The two other configurations are ones where the nudging is applied only up to 10 hPa and only up to 100 hPa, without any tapering.

The results of these three experiments are presented in Fig. 1, which shows polar-cap average (70°–90°N) temperatures in the stratosphere [10 hPa (32 km); top panel] and mesosphere [0.046 hPa (70 km); bottom panel] for the winter months of 2006. For each experiment an ensemble of three simulations was constructed using perturbed initial conditions. The colored curves denote the three CMAM ensembles, while the black dotted curves denote the MLS observations. At 10 hPa the agreement between the ensemble using tapered nudging (red) and MLS is very good, as expected since this is well within the nudging region where MLS and ERA-Interim are in good agreement (not shown). The agreement with MLS at 10 hPa worsens as the nudging height is lowered to 10 hPa (blue) and 100 hPa (green). At 70 km the ensemble using tapered nudging is still in good agreement with MLS, which is remarkable given that this is a full 20 km above the top of the nudging region. Moreover, the spread of the three ensemble members is very small, indicating that the polar-cap temperatures in the mesosphere are tightly constrained by (i.e., slaved to) the state of the lower atmosphere. As the nudging height is lowered to 10 and 100 hPa, the agreement with MLS at 70 km worsens and the inter-ensemble spread increases, with the 100-hPa nudging height yielding the poorest results.

(a) Temperature 10 hPa (32 km)
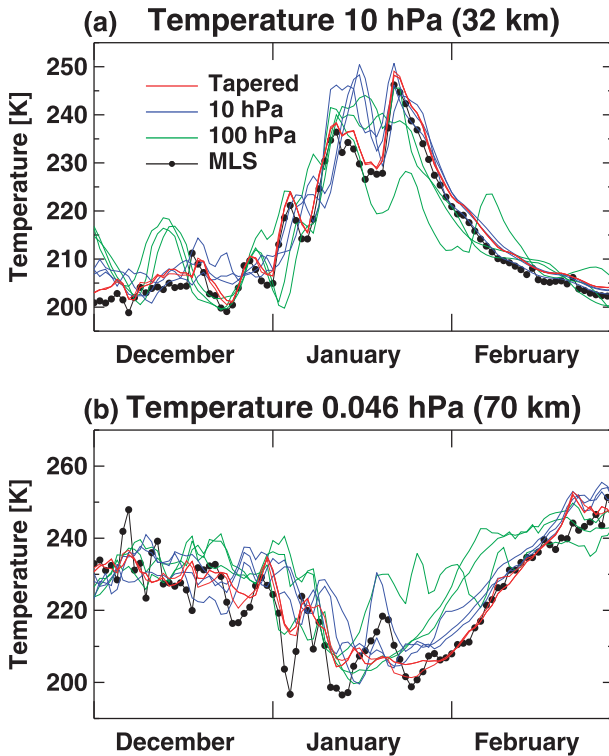

(b) Temperature 0.046 hPa (70 km)

FIG. 1. Polar-cap average temperatures at (a) 10 and (b) 0.046 hPa for the winter of 2006. Colored curves denote the following three ensembles of simulations (with each ensemble consisting of three members): tapered nudging (red), nudging only up to 10 hPa (blue), and nudging only up to 100 hPa (green). The black dots denote the MLS observations. Days range from 1 Dec 2005 to 28 Feb 2006.

The above analysis motivates the use of the tapered nudging configuration in the following four ensembles of simulations that form the backbone of this paper: 1) control, 2) without OGWD, 3) without NGWD, and 4) without GWD (i.e., with both OGWD and NGWD turned off). The latter three are otherwise identical to the control ensemble. The simulations extend from January 2004 to December 2010, but we present results for only the extended winters of 2006 and 2009. Each ensemble consists of three members, with the different members generated by perturbing the initial conditions in 2004. Unless stated otherwise, ensemble averages are shown. All plotted results are daily means, computed from instantaneous 6-hourly data. Log-pressure height (computed using a constant scale height of 7 km) is used for the vertical axes in all figures, as well as for specific altitudes mentioned in the text.

3. Results

a. MLS observations

The black dotted curves in Fig. 2 show MLS polar-cap temperatures at 70 km for the extended winters of 2006 and 2009. (The colored curves will be discussed in due course.) The two winters exhibit similar temporal behavior—namely, a moderate degree of day-to-day variability from mid-October until near the time of the major warmings (21 January 2006 and 24 January 2009), cooling during and immediately following the warmings (more pronounced in 2009), and a gradual and smooth warming and then cooling from February to April. The occurrence of a minor warming on 1 January 2006 appears to have resulted in the much less abrupt temperature drop at the time of the 2006 major warming.

Figures 3b, 3d, 4b, and 4d show time versus height sections of MLS polar-cap temperatures and the corresponding midlatitude average (40°–80°N) zonal winds for the two extended winters. As in Fig. 2, the results are qualitatively similar in the two years: Prior to the warmings, the stratopause is located near its climatological position at about 50 km, and the winds and temperatures exhibit the typical day-to-day variability of the late-fall/early-winter stratosphere. Immediately following the warmings, the stratosphere is cold and undisturbed, and a new stratopause forms in the upper mesosphere and steadily descends to its climatological position by springtime. The minor warming in early January 2006 is seen as the first patch of easterlies in Fig. 3d.

Figures 3f and 4f show polar-cap CO volume mixing ratios from MLS. The downward tilt of the contours in fall/early winter and late winter is indicative of the descent of CO from the thermosphere where it is produced by the photolysis of CO_2. Near or at the time of the major warmings, the signature of polar descent disappears because of the strong horizontal mixing that occurs as the vortex disappears, bringing in low values of CO from middle latitudes. Following the warmings, CO descends again until the final breakup of the polar vortex in springtime. An important point to note here is that the late-winter maximum of the MLS CO above about 60 km is substantially larger than the early-winter maximum (see also Fig. 8, black curves). The main difference in the temporal behavior of CO between the 2 years is that the signature of early-winter descent in 2009 is apparent right up until the major warming, while in 2006 it stops in early January at the time of the minor warming, and undergoes rapid day-to-day variations until the major warming.

While the synoptic structure of the two warming events was very different—the 2006 warming being characterized as a "vortex displacement" and the 2009 warming as a "vortex split"—the mesospheric response following these warmings is seen to be very similar. This is consistent with the finding of Hitchcock et al. (2013) that the morphology and time scale of the extended recovery of the stratosphere, which drives the mesospheric response, is insensitive to whether warmings occur as displacements
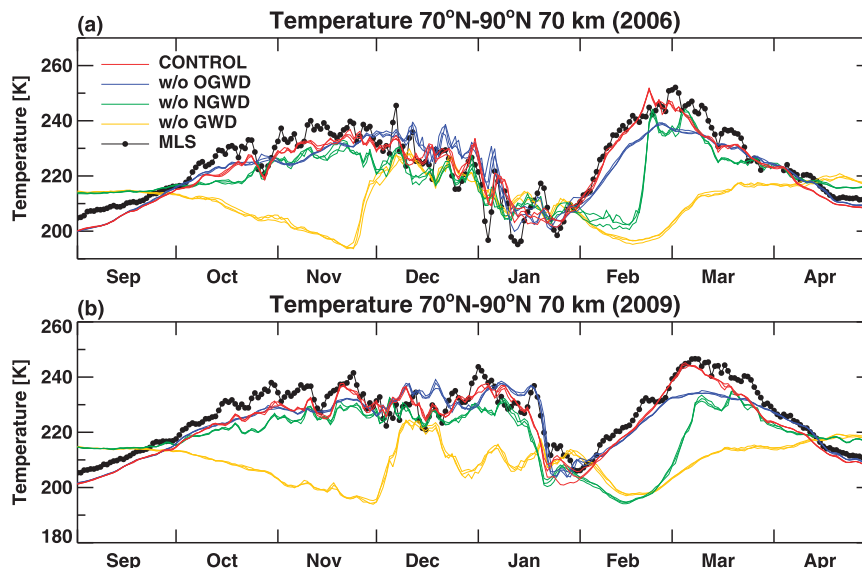
FIG. 2. Polar-cap average temperatures at 70 km for the extended winters of (a) 2006 and (b) 2009: control simulation (red) and the simulations without orographic GWD (blue), without nonorographic GWD (green), and without GWD (yellow). The black dotted curves denote the MLS observations. The three members of each CMAM ensemble are shown. Days range from 1 Sep 2005 (2008) to 30 Apr 2006 (2009).

or splits in the long-term observational record. It is also consistent with the fact that in the free-running CMAM, the mesospheric response to extended recoveries is the same for displacement and split events (Hitchcock and Shepherd 2013).

b. Control simulation

The red curves in Fig. 2 show the polar-cap temperatures at 70 km for the three members of the control simulation ensemble. As with Fig. 1, the agreement with MLS (black) is remarkably good, with differences less than about 10 K. Figures 3a, 3c, 4a, and 4c show the corresponding winds and temperatures for the ensemble average of the control simulation as a function of height and time. Good agreement with the MLS observations is seen throughout the entire height region above about 50 km where the model is not nudged. This supports the main conclusion of Ren et al. (2011) that it is not necessary to assimilate mesospheric temperature data to reproduce the observed mesospheric temperature response to the 2006 major SSW and the subsequent descent of the elevated stratopause. However, our results extend that conclusion to include the fall and early-winter months, which are not discussed by Ren et al. (2011). The good agreement between the control simulation and MLS is more clearly seen in Figs. 5a, 5b, 6a, and 6b, which show the corresponding differences with MLS for the extended winters of 2006 (left panels) and 2009 (right panels). Below about 85 km the absolute differences are less than about

5 K and 5 m s^{-1}. In the region above about 85 km the agreement is worse, but this is in the model sponge region where good agreement cannot be expected.

RMS differences between the control simulation and the MLS polar-cap temperatures and midlatitude zonal winds for the two extended winters combined are shown in Fig. 7 (red). As in Figs. 5 and 6, the differences are less than about 5 K and 5 m s^{-1} over most of the plotted domain. In addition, the three members of the control ensemble exhibit very little spread, indicating that the polar-cap temperatures and midlatitude average zonal winds are slaved to the state of the lower atmosphere over the full depth of the mesosphere for the entire extended winter periods. This lack of spread in the control simulation ensemble is also seen in Fig. 2, with the exception of the time period during the major warming in late January 2009 when the members differ from each other by up to about 10 K.

The polar-cap CO from the control simulation is also in good agreement with that from MLS (Figs. 3e,f and 4e,f). This is seen more clearly in Fig. 8, which shows results at 60 km for the two years. The control simulation CO (red) agrees to within about 10%–20% of the MLS observations (black) and exhibits the same seasonal variation, namely the two-peaked structure, with the smaller peak in early winter and the larger peak about 40 days after the major warmings. Qualitative agreement with the MLS CO observations in the mesosphere was found using an earlier version of CMAM (Jin et al. 2009).
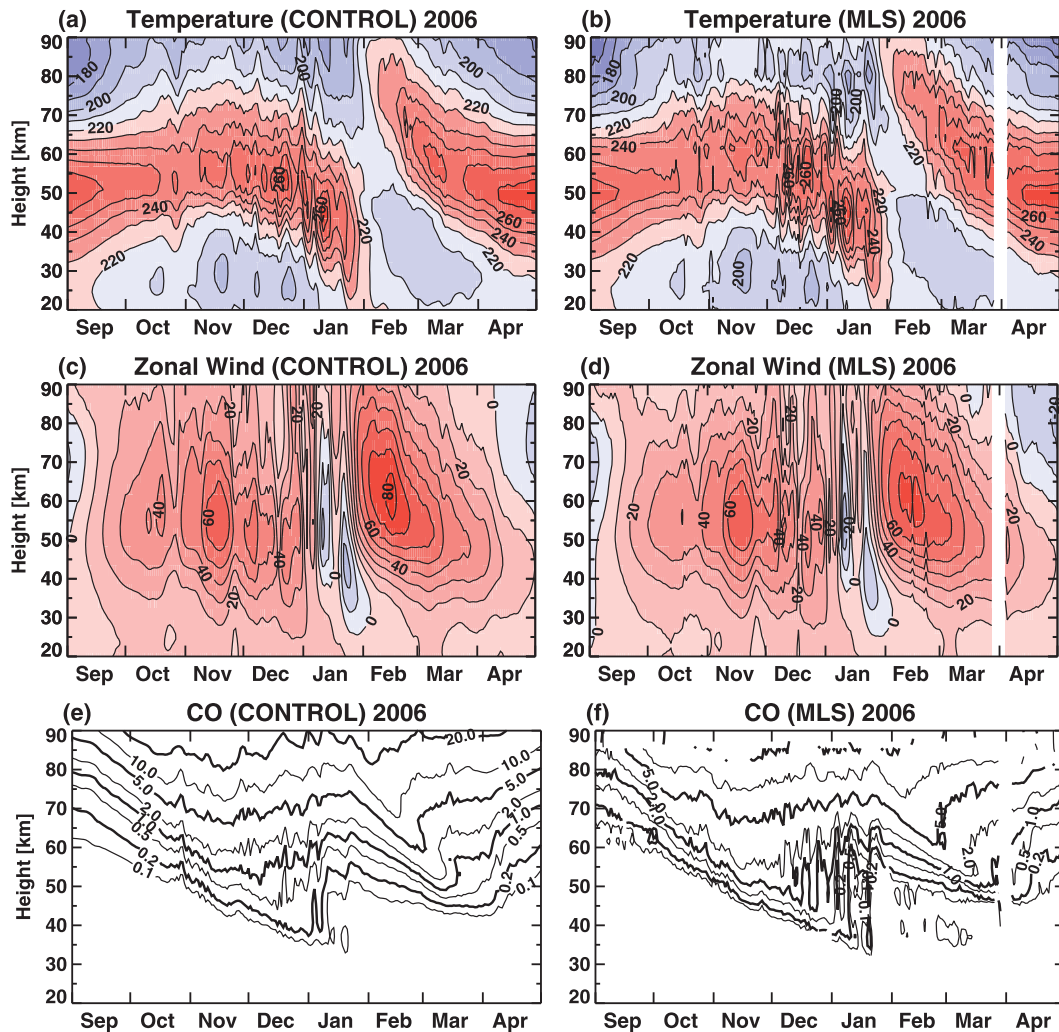
FIG. 3. (a),(b) Polar-cap average temperatures, (c),(d) zonal winds averaged from 40° to 80°N, and (e),(f) polar-cap average CO volume mixing ratios for the extended winter of 2006: (left) control simulation and (right) MLS observations. Contour intervals are 10 K and 10 m s^{-1} in (a),(b) and in (c),(d), respectively; CO has units of ppmv and a logarithmic contour interval. Temperatures less than 220 K and easterlies are shaded blue. Days range from 1 Sep 2005 to 30 Apr 2006. Missing MLS data not contoured. Above about 85 km there are few MLS CO data.

However, since that was a free-running version of the model it was not possible to make a quantitative comparison on a day-to-day basis as we have done here.

c. *Relative roles of orographic and nonorographic GWD*

In this section we examine the relative roles of OGWD and NGWD in producing the realistic zonal-mean mesospheric circulation found in the control simulation. Here we explain why the control simulation exhibits the same seasonal variation of mesospheric CO as seen in the MLS observations—namely, the double-peaked structure—with the late-winter peak being approximately twice as large as the early-winter peak.

Figures 5c–h and 6c–h show the differences between CMAM and MLS polar-cap temperatures and mid-latitude average zonal winds for the three sensitivity experiments. The overall differences far exceed those of the control simulation (top row). Consider first the simulation without OGWD (second row). The largest differences occur in late fall/early winter before the major warmings: the modeled lower mesosphere is up to 35 K too cold and the upper mesosphere up to 25 K too warm. (Note that the reasonably good agreement between this experiment and MLS seen in Fig. 2 is fortuitous since 70 km happens to be near the node in the temperature difference dipole seen in Figs. 5c and 5d.) Correspondingly, the modeled zonal winds are too
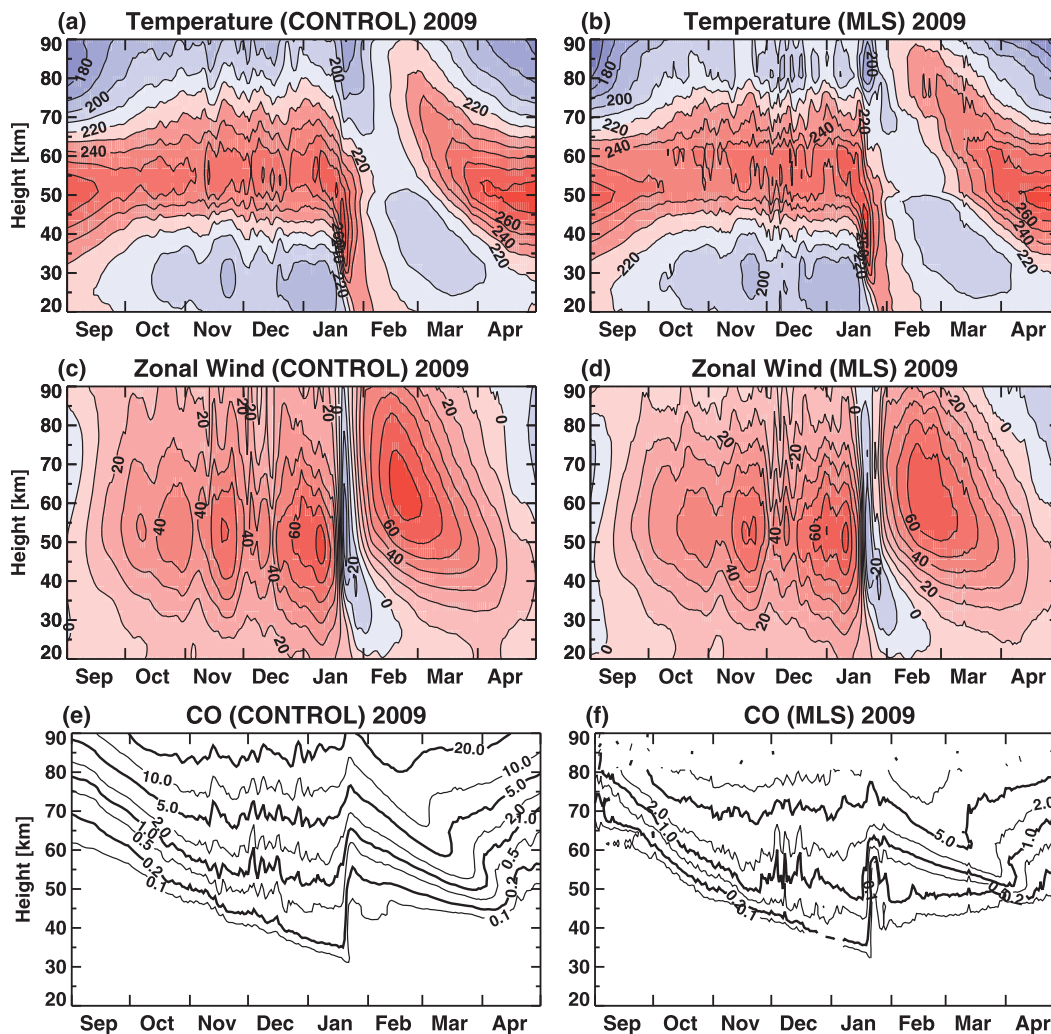
FIG. 4. As in Fig. 3, but for 2009, with days ranging from 1 Sep 2008 to 30 Apr 2009.

strong, with differences approaching about 50 m s^{-1} in the mesosphere. In 2009 the large wind and temperature differences extend right up to the start of the major warming, which is in contrast to 2006 when the differences start to weaken about the time of the minor warming in early January. In addition, the wind and temperature differences are about a factor of 2 smaller after the major warmings than before.

In the simulation without NGWD (Figs. 5e,f and 6e,f), the differences are also large but, in contrast to the simulation without OGWD, tend to be largest after the major warmings, not before. The removal of NGWD delays the descent of the stratopause (revealed by the MLS temperatures in Figs. 3b and 4b) until later in the winter when the lower-stratospheric westerlies have increased sufficiently to allow the upward propagation of the (parameterized) orographic GWs. There is

also a significant impact of NGWD at the beginning and end of the extended winters when the circulation is closer to summertime conditions. This is more clearly seen in early September when the removal of NGWD has resulted in negative zonal wind differences (i.e., easterlies) in the upper mesosphere (Figs. 6e and 6f).

In the simulation without GWD (Figs. 5g,h and 6g,h) the differences with respect to MLS are largest, with temperatures up to about 65 K too low and winds up to about 85 m s^{-1} too high. An elevated stratopause does not form, and the stratopause reformation at its climatological position near 50 km is delayed until March when the effects of solar heating start to become important. Although the wind and temperature difference fields exhibit the overall combined characteristics of the simulations without OGWD and without NGWD, it is
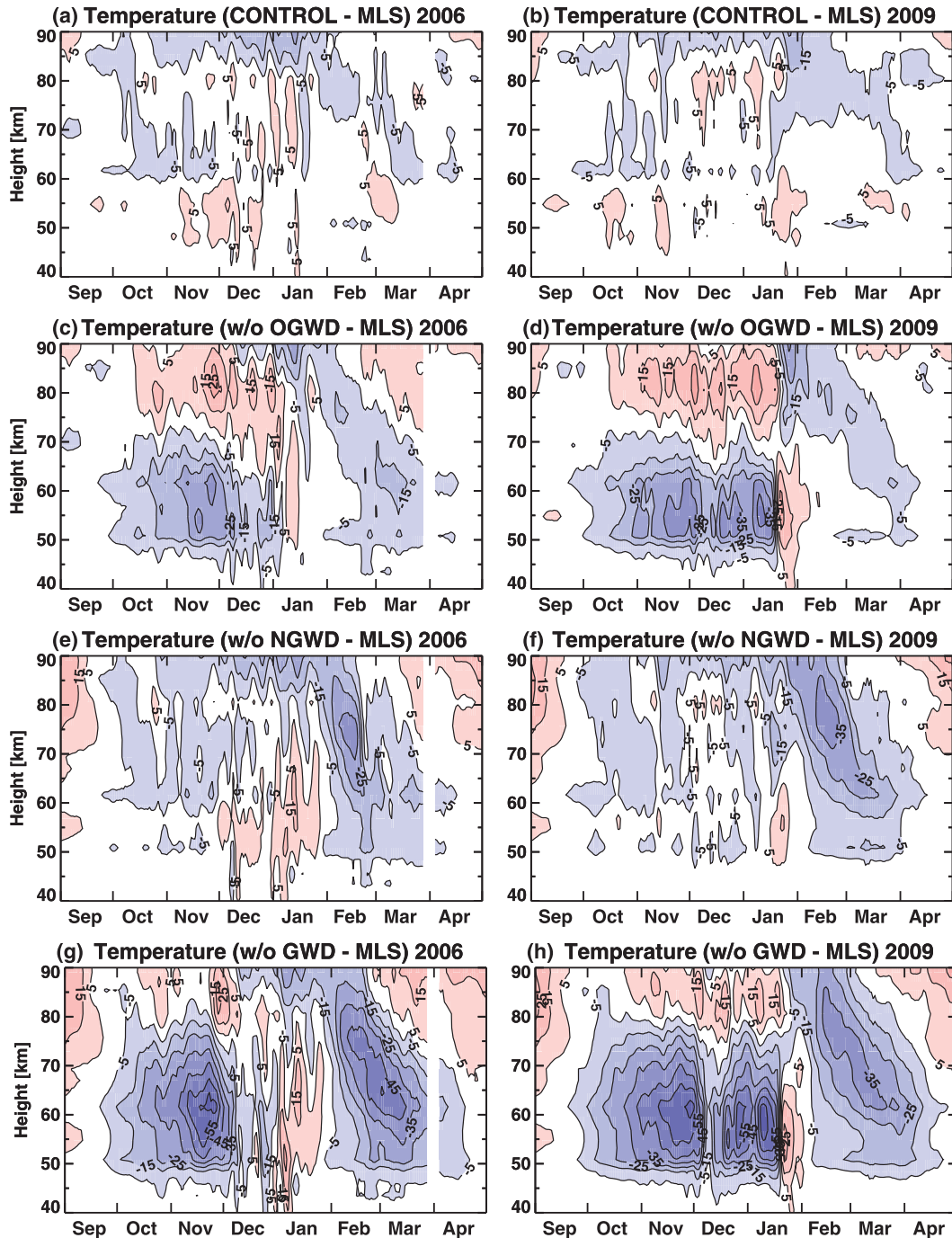
FIG. 5. Differences between MLS and CMAM polar-cap average temperatures for the extended winters of (left) 2006 and (right) 2009: (a),(b) control simulation and the simulations (c),(d) without orographic GWD, (e),(f) without nonorographic GWD, and (g),(h) without GWD. Contour interval is 10 K, with blue (red) shading for negative (positive) differences exceeding 5 K in magnitude. Data have been smoothed using a 3-day boxcar filter. Note that the height range differs from that in Figs. 3 and 4.

clear that at certain times and places (e.g., early winter between 50 and 70 km) the sum of the differences in the second and third rows of Figs. 5 and 6 do not equal those in the bottom row. This indicates that the responses to

OGWD and NGWD are not additive. This issue is investigated in section 3d.

The RMS wind and temperature differences (with respect to MLS) for the three sensitivity experiments are
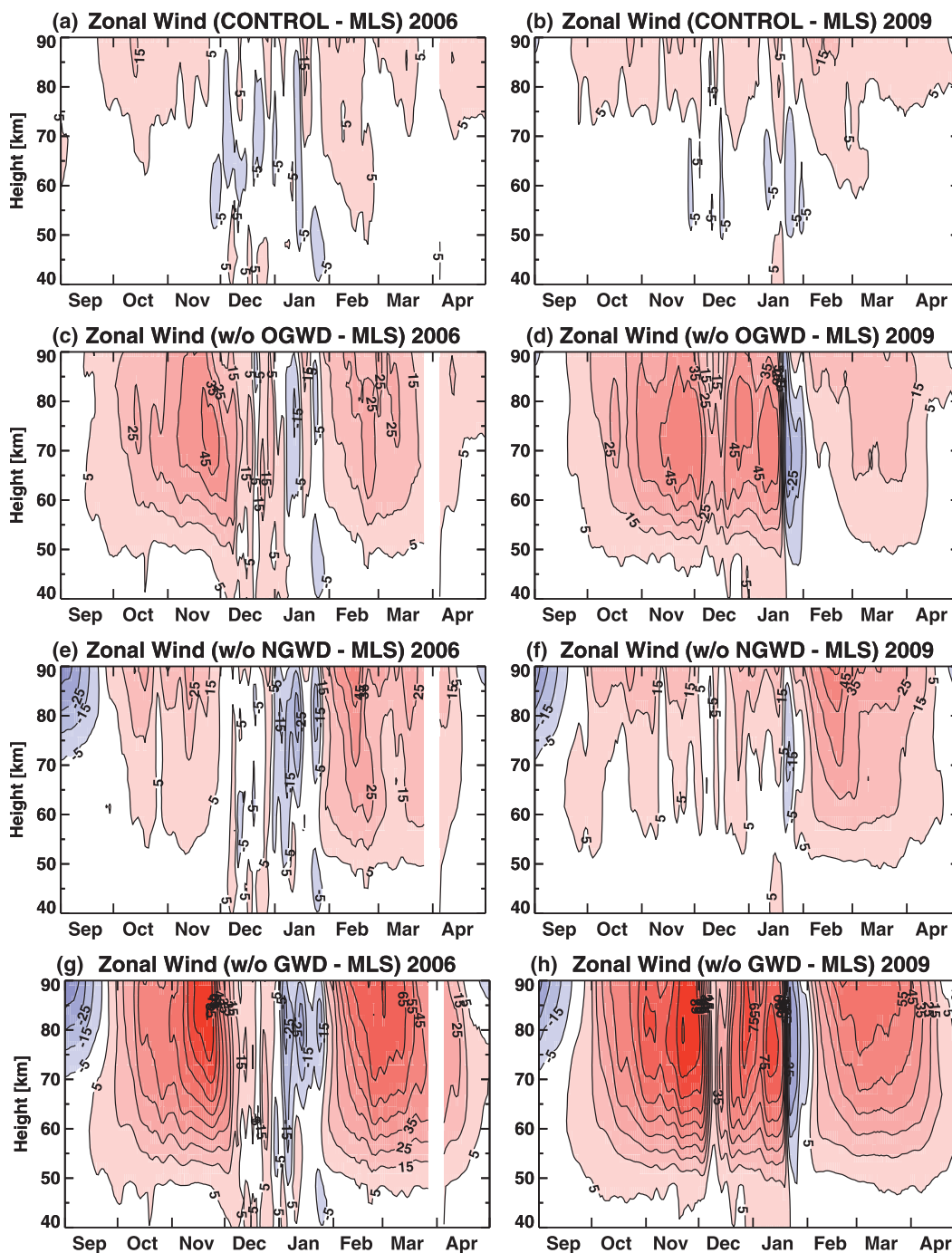
FIG. 6. As in Fig. 5, but for zonal winds averaged from 40° to 80°N. Contour interval is 10 m s^{-1}, with blue (red) shading for negative (positive) differences exceeding 5 m s^{-1} in magnitude.

shown in Fig. 7. The worst agreement is for the simulation without GWD (yellow), where the RMS differences peak at about 35 K and 50 m s^{-1}. As with the control ensemble, the spread between the different members of the sensitivity experiment ensembles is also very small. This is also seen in Fig. 2, which shows polar-cap temperatures

at 70 km for the two extended winters. Note that time series at other mesospheric heights (not shown) exhibit the same lack of spread. The lack of spread in the members of the ensemble without GWD is of particular interest because it demonstrates that the slaving of the mesosphere to the stratosphere is not only through GWD.
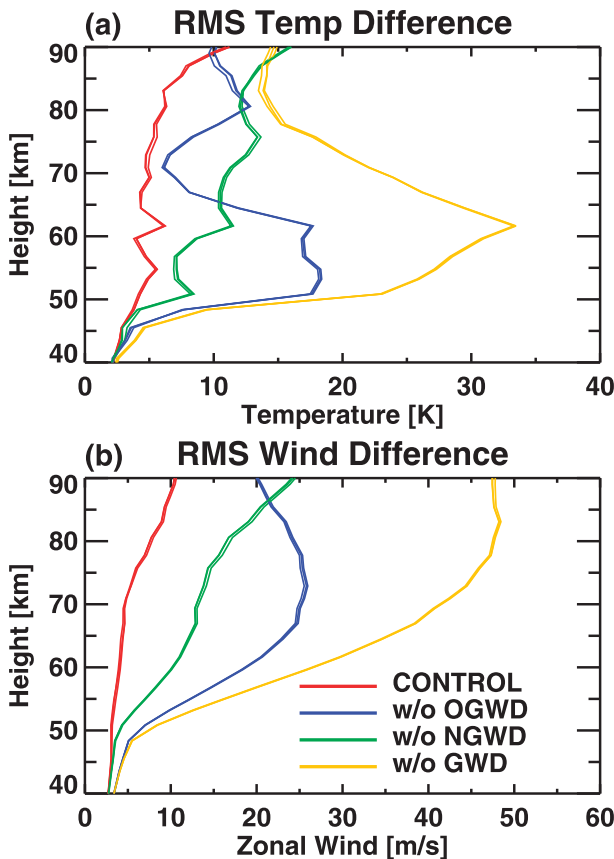
FIG. 7. RMS differences between MLS and CMAM for the combined extended winters of 2006 and 2009 (1 Sep–30 Apr) for the control simulation (red), and the simulations without orographic GWD (blue), without nonorographic GWD (green), and without GWD (yellow): (a) polar-cap average temperatures and (b) zonal winds averaged from 40° to 80°N. All three members of the four different ensembles are shown.
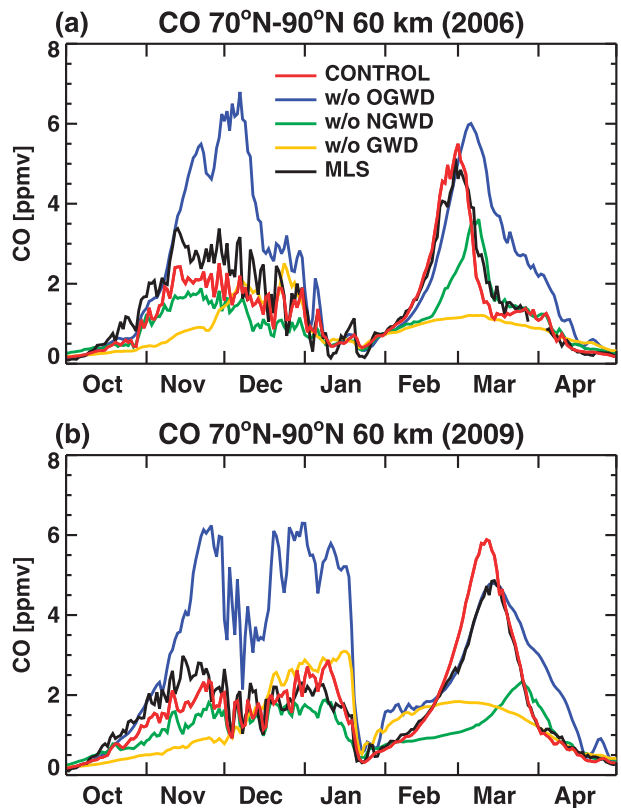

FIG. 8. Polar-cap average CO volume mixing ratios at 60 km for the extended winters of (a) 2006 and (b) 2009: control simulation (red) and the simulations without orographic GWD (blue), without nonorographic GWD (green), and without GWD (yellow). Black curves denote the MLS observations. Days range from 1 Oct 2005 (2008) to 30 Apr 2006 (2009).

In this case it is the resolved (planetary) wave drag that is constraining the mesospheric temperatures before the warmings. In the extended period after the warmings, planetary wave activity in the stratosphere is very weak, which is reflected in the simulations by extremely weak resolved wave drag below about 80 km (not shown). Thus, in the absence of GWD the mesospheric temperatures quickly relax toward radiative equilibrium, which explains the similarity of the three members of that ensemble in the period following the warmings. Figures 2 and 7 also demonstrate another important point—namely, that slaving to the stratosphere does not imply that the mesospheric circulation is realistic, only that its temporal variation is constrained. The realism of the mesospheric circulation depends on the realism of the slaving mechanisms.

Returning to the time series of polar-cap CO (Fig. 8), large differences are seen between the three sensitivity

experiments and MLS. In the simulation without OGWD (blue) the early-winter peak is too strong and the late-winter peak too prolonged. In the simulation without NGWD (green) the agreement with MLS is better than in the simulation without OGWD in early winter, but the late-winter peak is too weak, too brief, and is shifted to a later date, particularly in 2009. In the simulation without GWD (yellow) CO is far too weak in October and November and following the warmings.

In the next three figures we explain why the temporal behavior of mesospheric CO is so different for the three sensitivity experiments and why the agreement with MLS CO is best for the control simulation. We start by examining the residual vertical velocity w^*, which is largely controlling the concentration of CO over the winter pole. Downward control (Haynes et al. 1991) is then used to attribute the differences in w^* to the different types of wave drag.

Figure 9 shows polar-cap w^* for the control simulation and the three sensitivity experiments. In the upper mesosphere (i.e., above about 70 km) the control simulation
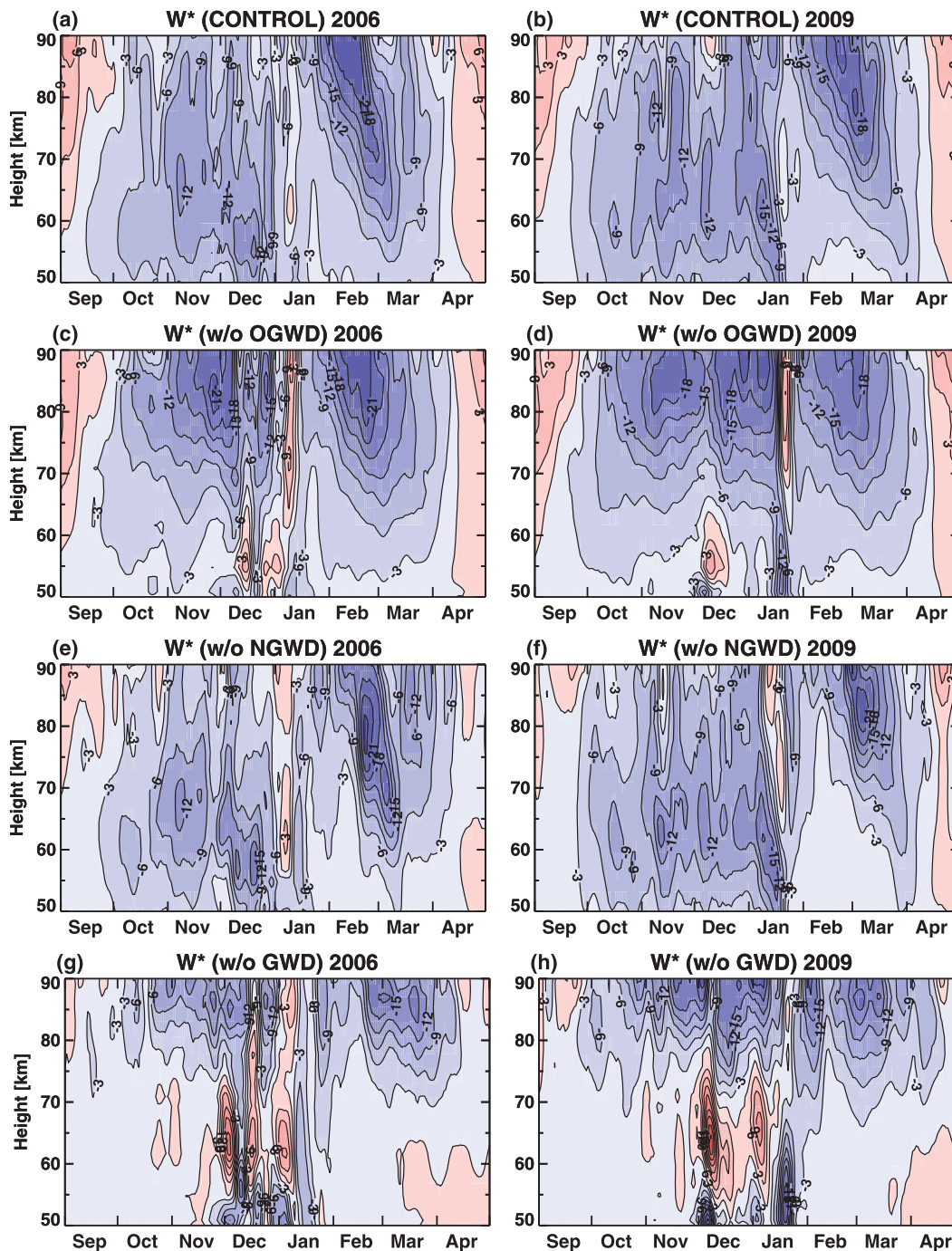
FIG. 9. Polar-cap average residual vertical velocities for the extended winters of (left) 2006 and (right) 2009: (a),(b) control simulation and the simulations (c),(d) without orographic GWD, (e),(f) without nonorographic GWD, and (g),(h) without GWD. Contour interval is 3 mm s^{-1}, with blue (red) shading for downwelling (upwelling). Data have been smoothed twice using a 5-day boxcar filter.

(Figs. 9a,b) exhibits weak downwelling in late fall/early winter and strong downwelling in late winter following the warmings, thus explaining the weak early-winter and strong late-winter peaks of CO in Fig. 8. In the simulation without OGWD (Figs. 9c,d) the late-fall/early-winter downwelling in the upper mesosphere is a factor of 2 stronger than in the control, while in the region below it is weaker. This is more clearly seen in Fig. 10a, which shows
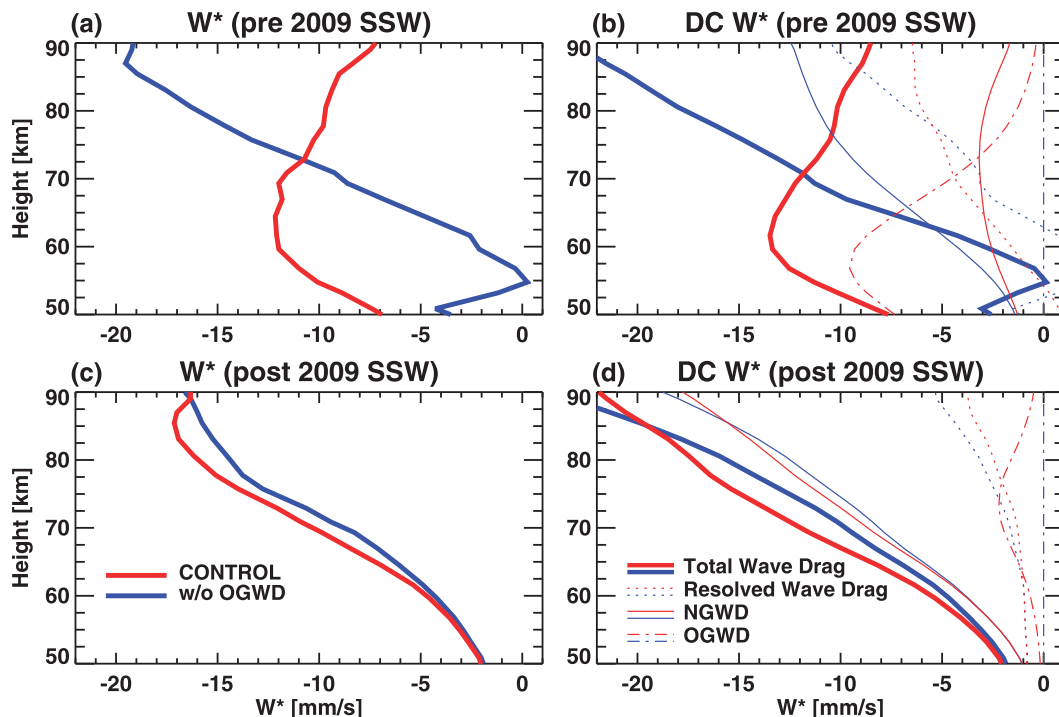
FIG. 10. Vertical profiles of polar-cap average residual vertical velocity averaged over 40-day periods (a) before and (c) after the 2009 major SSW for the control simulation (red) and simulation without OGWD (blue). (b),(d) Corresponding downward control estimates of $w*$ computed using the total wave drag (thick solid), only nonorographic GWD (thin solid), only orographic GWD (dashed–dotted), and only resolved wave drag (dotted). The averaging periods start 50 days before and 10 days after the date of the actual warming (24 Jan). Total wave drag is the sum of the resolved wave drag and orographic and nonorographic GWD.

vertical profiles of $w*$ for a 40-day period prior to the 2009 warming. It is the strong downwelling in the upper mesosphere that is causing the unrealistically large early-winter peak in CO in the simulation without OGWD (Fig. 8). The anomalous upper-mesospheric downwelling and lower-mesospheric upwelling that occurs when OGWD is switched off also results in adiabatic warming and cooling, respectively, which explains the vertical dipole structure in the temperature differences in early winter (Figs. 5c,d). After the warmings the polar-cap $w*$ is similar in strength in the two simulations (Figs. 9a–d and 10c), but the tongue of downwelling is somewhat wider (in time) in the simulation without OGWD, thus explaining why the late-winter peak of CO is wider than in the control. In the simulation without NGWD (Figs. 9e,f) the early-winter downwelling is similar to that of the control, but the tongue of downwelling following the warmings is narrower and in the 2009 winter does not descend as far down, thus explaining the weak late-winter peak of CO seen in Fig. 8. In the simulation without GWD (Figs. 9g,h) the downwelling in the lower mesosphere is very weak except in December and January, whereas in the upper mesosphere there is

strong downwelling from October to mid-April. Since there is no parameterized GWD in this simulation, $w*$ is largely being driven by resolved wave drag.

Figures 10b,d show the downward control (DC) estimates of the polar-cap $w*$ ($\hat{w}$) computed using the different types of wave drag and averaged over the two 40-day periods. The good overall agreement between $\hat{w}$ computed using the total wave drag (thick curves) and $w*$ (Figs. 10a,c) demonstrates the validity of the DC calculation, except in the model sponge. Note that the magnitude of $\hat{w}$ is somewhat larger than $w*$, as expected, since $w*$ does not have time to fully adjust to the steady-state conditions assumed in the downward control. An examination of the contributions from the individual forcings (thin curves) reveals that NGWD (solid) is largely responsible for the strong downwelling in the upper mesosphere prior to the warming (Fig. 10b) in the simulation without OGWD (blue curve), with a secondary contribution coming from resolved wave drag (dotted). The latter is an indirect response to the removal of OGWD, resulting from the increased westerlies above 50 km (Fig. 6d) that allow planetary waves to propagate to higher elevations where they dissipate, as inferred from

large negative values of resolved wave drag above about 85 km in early winter (results not shown).[1] The late-fall/early-winter increase in NGWD in the upper mesosphere in the simulation without OGWD is a direct consequence of the increase in the strength of the mesospheric westerlies seen in Fig. 6d, which has raised the region of saturation of the westward-propagating nonorographic GWs (not shown). This indicates that there is a strong interplay between OGWD and NGWD in late fall/early winter, mediated by their impact on the background winds and temperatures.

In the 40-day period after the warming, downwelling is driven primarily by NGWD (Fig. 10d), with OGWD becoming more important later on as westerlies strengthen in the lower stratosphere (not shown). In comparison to the period before the warming, the downwelling driven by OGWD after the warming is much weaker and occurs higher up. Qualitatively similar results are found for the winter of 2006.

d. Additivity of responses to orographic and nonorographic GWD

While the wind and temperature differences (with respect to MLS) for the simulation without GWD (Figs. 5g,h and 6g,h) appear to be qualitatively similar to the sum of the differences of the simulations without OGWD and without NGWD, there is clear evidence that the responses to OGWD and NGWD are nonadditive. To quantify this we introduce a diagnostic, which we shall refer to as the GW additivity anomaly ΔX, defined by

$$\Delta X = \Delta X_{\text{GWD}} - (\Delta X_{\text{OGWD}} + \Delta X_{\text{NGWD}}), \quad (1)$$

where X represents a quantity like temperature or zonal wind, ΔX_{GWD} is the response of X to GWD (i.e., X for the control minus X for the simulation without GWD), ΔX_{OGWD} is the response to OGWD (i.e., control minus simulation without OGWD), and ΔX_{NGWD} is the response to NGWD (i.e., control minus simulation without NGWD). If the response is additive, then $\Delta X = 0$.

Figure 11 shows the GW additivity anomalies for midlatitude zonal wind ΔU (Fig. 11a) and polar-cap temperature ΔT (Fig. 11b) for the winter of 2009. Qualitatively similar results are found for 2006. There are two
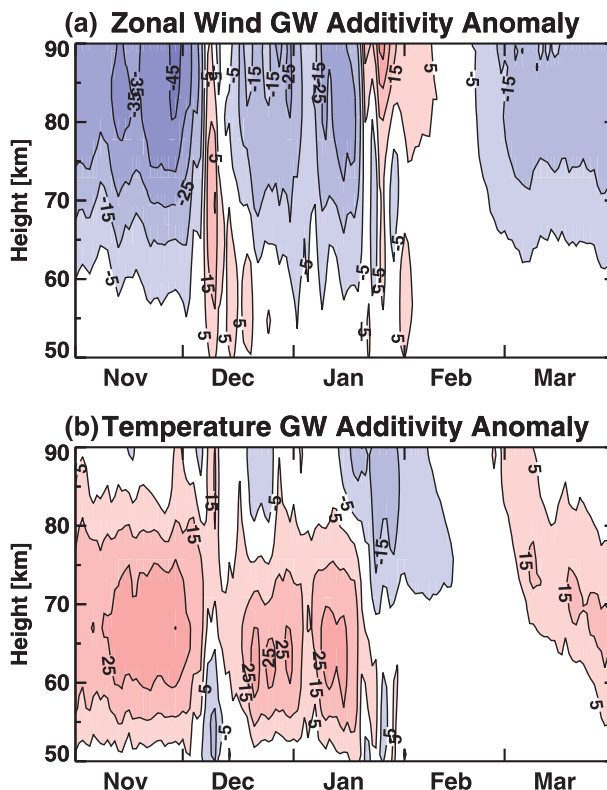
[1] While it is possible that in situ generation of planetary waves could be occurring in the upper mesosphere as a result of baroclinic instability, for example, the sign of the resolved wave drag generated by such waves would be opposite in sign (i.e., positive) to that found in the simulation without OGWD. Consequently, the dominant contribution to the resolved wave drag in the upper mesosphere in this simulation appears to be from planetary waves propagating up from below.

FIG. 11. GW additivity anomalies for (a) zonal winds averaged from 40° to 80°N (ΔU) and (b) polar-cap temperature (ΔT) for the winter of 2009; see text for definitions of ΔU and ΔT. Nonadditivity occurs where $\Delta U \neq 0$ or $\Delta T \neq 0$. Contour intervals are 10 m s^{-1} and 10 K, with blue (red) shading for negative (positive) values exceeding 5 m s^{-1} and 5 K in magnitude.

periods when the responses are nonadditive: the winter months leading up to the major warming in late January, and the month or so following the warming. In general, ΔT is positive (peak values of about 25 K near 65 km) and ΔU is negative (peak values of about -45 m s^{-1} near 85 km), with the sign difference and vertical offset of the maxima being consistent with thermal wind balance. The strength of the nonadditivity can be expressed as the ratio $R_X = \Delta X_{\text{GWD}}/(\Delta X_{\text{OGWD}} + \Delta X_{\text{NGWD}})$. For the midlatitude zonal wind R_U averaged from November to January increases from about 1.2 at 65 km to about 1.8 at 85 km, indicating that the zonal wind response to OGWD and NGWD is strongly nonadditive in the upper mesosphere at this time. Additivity is only apparent during the month following the major warming (as indicated by the white area in Fig. 11). However, such additivity is specious as it occurs during a period when both the mesospheric OGWD and lower-mesospheric resolved wave drag are extremely weak.

Positive ΔT in Fig. 11b means that the temperature response to GWD (ΔT_{GWD}) exceeds that of the sum of the separate responses ($\Delta T_{\text{OGWD}} + \Delta T_{\text{NGWD}}$). To help
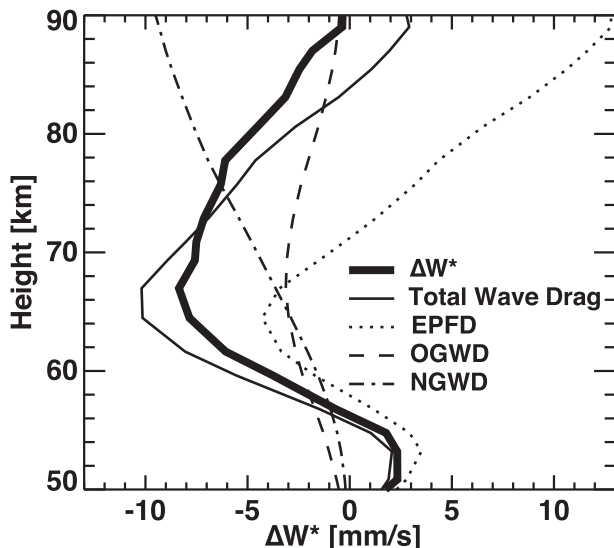
FIG. 12. GW additivity anomaly for polar-cap residual vertical velocity Δw^* (thick solid) averaged from 1 Nov 2008 to 31 Jan 2009, $\Delta \hat{w}$ (thin solid), and the decomposition of $\Delta \hat{w}$ into its components: resolved wave drag (EPFD; dots), orographic GWD (dashed), and nonorographic GWD (dashed dots). Nonadditivity occurs where $\Delta w^* \neq 0$ and $\Delta \hat{w} \neq 0$. See text for definitions of Δw^* and $\Delta \hat{w}$.

understand why this happens, we resort again to downward control to diagnose the contributions from the different types of wave drag. Since the nonadditivity is strongest prior to the major warming (Fig. 11), we focus on that period and time average the results to further simplify the analysis. Figure 12 shows the GW additivity anomaly associated with the polar-cap residual vertical velocity Δw^* (thick curve) averaged from 1 November to 31 January. Above about 55 km Δw^* is negative, with a peak value at about 65 km. The negative Δw^* causes adiabatic warming, thus explaining the positive ΔT in Fig. 11.

The causes of nonzero Δw^* can be understood by constructing the DC estimate of Δw^* from (1); that is,

$$\Delta \hat{w} = \Delta \hat{w}_{\text{GWD}} - (\Delta \hat{w}_{\text{OGWD}} + \Delta \hat{w}_{\text{NGWD}}), \qquad (2)$$

where $\hat{w} = \hat{w}^{(E)} + \hat{w}^{(N)} + \hat{w}^{(O)}$, with the bracketed superscripts E, N, and O denoting the respective contributions from resolved wave drag (denoted EPFD for Eliassen–Palm flux divergence), OGWD, and NGWD to the DC estimate of w^*. Using (2) and the equation for $\hat{w}$ yields the DC contributions to $\Delta \hat{w}$ from the three types of wave drag, which are given by

$$\Delta \hat{w}^{(E)} = \Delta \hat{w}_{\text{GWD}}^{(E)} - (\Delta \hat{w}_{\text{OGWD}}^{(E)} + \Delta \hat{w}_{\text{NGWD}}^{(E)}), \quad (3)$$

$$\Delta \hat{w}^{(O)} = -\Delta \hat{w}_{\text{NGWD}}^{(O)}, \qquad\qquad\qquad (4)$$

$$\Delta \hat{w}^{(N)} = -\Delta \hat{w}_{\text{OGWD}}^{(N)}. \qquad\qquad\qquad (5)$$

The simpler forms of (4) and (5), as compared to (3), follow from the fact that two of the terms in (2) cancel because the DC estimate of the OGWD or NGWD contribution to $\hat{w}$ is zero in the simulations without OGWD or NGWD and without both types of GWD.

The results of the decompositions (3)–(5) are given by the thin curves in Fig. 12, where the dotted, dashed, and dashed–dotted curves denote, respectively, the EPFD, OGWD, and NGWD contributions to $\hat{w}$. To demonstrate the validity of this decomposition, the DC estimate $\Delta \hat{w}$ is also plotted (thin solid curve), and is seen to be in good qualitative agreement with Δw^* (thick solid curve). Above about 70 km the dominant contributions are from NGWD and EPFD, with the two being of opposite sign, which explains the diminishing magnitude of Δw^* with height in that region. The maximum negative values of Δw^* near 65 km arise from the nearly equal and negative contributions from all three wave drag components, while the positive values of Δw^* near 55 km are primarily from EPFD.

The physical mechanisms responsible for the contributions of the three types of wave drag to the nonadditivity documented in Fig. 12 can be understood by examining the changes in wave drag between the different simulations. Here we discuss only NGWD and EPFD since overall they provide the dominant contributions to $\Delta \hat{w}$ in the upper mesosphere shown in Fig. 12. From (5), the NGWD contribution to $\Delta \hat{w}$ is simply the negative of the response of $\hat{w}^{(N)}$ to OGWD (i.e., $\hat{w}^{(N)}$ from the simulation without OGWD minus $\hat{w}^{(N)}$ from the control simulation). Figure 10b shows $\hat{w}^{(N)}$ for those two simulations (thin solid blue and red curves) for a slightly different time period than used in Fig. 12. The difference between those two curves is negative, which yields the negative contribution from NGWD to $\Delta \hat{w}$ seen in Fig. 12. The physical mechanism for this change in NGWD was discussed in the previous section—namely, the increase in westward NGWD in the upper mesosphere that occurs in response to the stronger westerlies in the mesosphere when OGWD is turned off.

The EPFD contribution to nonadditivity in Fig. 12 is more complicated since, as can be seen from (3), it must be derived from all four simulations. Consequently, it will not be analyzed in detail as for NGWD. The physical mechanism responsible for its contribution can be understood simply as resulting from the strong dependence of resolved wave drag on the zonal-mean winds and temperatures, which in turn depend strongly upon the OGWD and NGWD.

The fact that the zonal wind and temperature responses to OGWD and NGWD are nonadditive is not entirely surprising given that both the parameterized and resolved wave drag are strongly dependent upon the background

winds. The nonadditive responses simply indicate that the different types of wave drag interact strongly with each other through their influence on the background state.

4. Summary and discussion

A version of the Canadian Middle Atmosphere Model that is nudged toward reanalysis data up to 1 hPa is used to examine the relative roles of orographic and non-orographic gravity wave drag in determining the zonal-mean circulation of the mesosphere during the extended northern winters of 2006 and 2009. These years are chosen because of the occurrence of two of the largest stratospheric sudden warmings on record, which strongly coupled the lower and upper atmosphere through vertically propagating planetary-scale Rossby waves and small-scale gravity waves. By examining the extended period from September to April we are able to study not only the long recovery period following the warmings, which has been the focus of previous studies, but also the fall/early-winter period leading up to the warmings when the middle atmosphere was much closer to its climatological state.

Four sets of simulations are performed. The first is a control simulation in which both the OGWD and NGWD parameterizations are active. The three others are sensitivity experiments in which the GWD parameterizations are turned off, separately and together. Each set of simulations consists of three members initialized on 1 January 2004 using perturbed initial conditions. The use of ensembles of three enables us to assess the degree to which the mesospheric circulation is slaved to the state of the lower atmosphere.

We validate our simulations using satellite observations of temperature, zonal wind (derived from geopotential height), and CO from the *Aura* MLS instrument. The control simulation is shown to be in remarkably good agreement with MLS, with RMS temperature and wind differences of less than 5 K and 5 m s^{-1} above the top of the nudging region between 50 and 85 km. The simulated CO is also shown to be in good (to within 10%–20% at 60 km) agreement with MLS, and exhibits the same temporal variation over the course of the two extended winters—namely, a two-peaked structure, with a small late-fall/early-winter peak and a larger late-winter peak a month or so after the major warmings.

The sensitivity experiments reveal that the relative roles of OGWD and NGWD are very different between fall/early winter and late winter. In the months leading up to the warmings, OGWD is shown to have a larger overall impact than NGWD. During this time period the simulation without OGWD exhibits excessively strong

(with respect to MLS) zonal-mean zonal winds, a large amplitude temperature difference dipole in the vertical, and overly large values of CO above about 60 km. NGWD, on the other hand, is shown to have a more pronounced effect in the month after the warmings as the vortex is recovering, which is in agreement with the findings of Ren et al. (2011). In the simulation without NGWD the descent of the elevated stratopause is delayed until the lower-level zonal winds have increased to the point where the parameterized orographic GWs are able to reach the mesosphere, in agreement with the findings of Hitchcock and Shepherd (2013). The presence of OGWD in late winter (a month or so after the warmings) also drives the realistic descent of CO at that time. In the simulation without NGWD there is much weaker descent following the warming and consequently a much weaker late-winter peak of CO in the lower mesosphere. In the simulation without any GWD, the late-fall/early-winter winds in the mesosphere are extremely strong and the stratopause does not descend from high altitudes after the warming but reforms later than observed, in early spring, at its climatological height as a result of increased solar heating.

A closer examination of the simulations reveals that there is a strong interplay between the OGWD and NGWD in late fall/early winter, with the strength of the upper-mesospheric NGWD being controlled by the strength of the upper-stratospheric/lower-mesospheric westerlies, which are in turn controlled by the OGWD. In the simulation without OGWD anomalously strong downwelling occurs in the upper mesosphere, which explains the overly large values of lower-mesospheric CO and the anomalously high temperatures in the upper mesosphere in this simulation. This occurs because of an increase in westward NGWD in the upper mesosphere in response to the increased westerlies in the lower mesosphere brought about by the removal of OGWD.

The lack of spread in the zonal-mean temperature and zonal wind time series between the different members of each ensemble demonstrates just how strongly the zonal-mean circulation of the mesosphere is slaved to the lower atmosphere. Moreover, in the late fall/early winter of 2006 and 2009 this occurs through a combination of OGWD and resolved wave drag, while in late winter following the major warmings it is through NGWD. An important point to note is that the slaving of the mesosphere to the stratosphere does not imply that the mesospheric circulation is realistic, only that its temporal variation is constrained. The realism of the mesospheric circulation depends on the realism of the slaving mechanisms. This is demonstrated in the sensitivity experiments in which the different types of GWD are turned off—the members of each of those ensembles exhibit

little spread but are substantially different from the observations (Fig. 2).

We also examine the additivity of the responses to OGWD and NGWD, and find that they are strongly nonadditive in the upper mesosphere before the major warmings when OGWD, NGWD, and resolved wave drag are simultaneously active. That the responses are nonadditive is not too surprising given that the breaking criterion used in the GWD parameterizations has a strong dependence on the background winds. Moreover, turning off one or both of the GWD terms alters the zonal-mean winds, thus changing the propagation and breaking of planetary waves. Nonadditivity also has the practical implication that the OGWD and NGWD schemes cannot be tuned separately.

We turn now to a brief discussion of some possible implications of our study. The first is that the good agreement between the control simulation and the MLS observations over a wide range of NH winter conditions provides a measure of confidence in the GWD parameterizations employed in CMAM, at least in terms of their zonal-mean effects. (In the case of OGWD it must be remembered that the local values can far exceed that of the zonal mean, so our study does not address the realism of the local values of OGWD.) The two-wave OGWD scheme used here was shown by Scinocca and McFarlane (2000) to transport 30%–50% more momentum flux into the middle atmosphere in winter than the simpler single-wave McFarlane (1987) scheme. Our results suggest that this increase in middle-atmosphere OGWD in the Scinocca–McFarlane scheme is realistic and that the use of two waves to represent the full spectrum of orographic GWs is an improvement over simpler one-wave OGWD schemes.

Strictly speaking, the good agreement between the control simulation and the MLS observations only applies to the total GWD, and there could in principle be compensating errors between the OGWD and NGWD schemes. However, as seen in Figs. 5 and 6, the agreement holds both before the major warmings, when OGWD is dominant, and after the major warmings, when the weak lower-stratospheric winds filter most of the OGWD, leaving the NGWD dominant at mesospheric heights. The different zonal wind conditions expose the mesosphere to different relative influences of OGWD and NGWD, suggesting that the two GWD components must each be fairly accurate.

A second implication of our study concerns data assimilation. Ren et al. (2011) showed that it was not necessary to assimilate mesospheric measurements to reproduce the zonal-mean mesospheric temperature response to the 2006 major SSW and the prolonged descent of the elevated stratopause following the warming.

Our study supports that finding, and extends it to include the entire period from September to April for the winters of 2006 and 2009. The benefit of assimilating mesospheric data would therefore seem to lie in reproducing zonal asymmetries in the mesosphere, which are not investigated here, or the mesospheric state during summertime, which is prone to in situ instabilities (e.g., Plumb 1983) and thus is not expected to be as strongly slaved to the stratosphere.

A final remark is that our modeling approach could be used to better constrain the source parameters of NGWD parameterizations. While this suggestion has been made before in the context of data assimilation (e.g., Ren et al. 2011), nudging to reanalysis data makes it possible to perform a suite of simulations using a variety of different GWD parameter settings, which would be prohibitively expensive with a data assimilation system.

REFERENCES

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.,* **137,** 553–597, doi:10.1002/qj.828.

de Grandpré, J., S. R. Beagley, V. I. Fomichev, E. Griffioen, J. C. McConnell, A. S. Medvedev, and T. G. Shepherd, 2000: Ozone climatology using interactive chemistry: Results from the Canadian Middle Atmosphere Model. *J. Geophys. Res.,* **105** (D21), 26 475–26 491.

Eckermann, S. D., and Coauthors, 2009: High-altitude data assimilation system experiments for the northern summer mesosphere season of 2007. *J. Atmos. Sol.-Terr. Phys.,* **71,** 531–551.

Haynes, P. H., C. J. Marks, M. E. McIntyre, T. G. Shepherd, and K. P. Shine, 1991: On the "downward control" of extratropical diabatic circulations by eddy-induced mean zonal forces. *J. Atmos. Sci.,* **48,** 651–678.

Hitchcock, P., and T. G. Shepherd, 2013: Zonal-mean dynamics of extended recoveries from stratospheric sudden warmings. *J. Atmos. Sci.,* **70,** 688–707.

——, ——, and G. L. Manney, 2013: Statistical characterization of Arctic polar-night jet oscillation events. *J. Climate,* **26,** 2096–2116.

Jin, J. J., and Coauthors, 2009: Comparison of CMAM simulations of carbon monoxide (CO), nitrous oxide (N_2O), and methane (CH_4) with observations from Odin/SMR, ACE-FTS, and Aura/MLS. *Atmos. Chem. Phys.,* **9,** 3233–3252, doi:10.5194/acp-9-3233-2009.

Jonsson, A. I., J. de Grandpré, V. I. Fomichev, J. C. McConnell, and S. R. Beagley, 2004: Doubled CO_2-induced cooling in the middle atmosphere: Photochemical analysis of the ozone

radiative feedback. *J. Geophys. Res.,* **109,** D24103, doi:10.1029/2004JD005093.

Limpasuvan, V., J. H. Richter, Y. J. Orsolini, F. Stordal, and O.-K. Kvissel, 2012: The roles of planetary and gravity waves during a major stratospheric sudden warming as characterized in WACCM. *J. Atmos. Sol.-Terr. Phys.,* **78–79,** 84–98, doi:10.1016/j.jastp.2011.03.004.

Livesey, N. J., and Coauthors, 2011: Earth Observing System (EOS) Aura Microwave Limb Sounder (MLS) Version 3.3 Level 2 data quality and description document. Jet Propulsion Laboratory Tech. Rep. JPL D-33509, 156 pp. [Available online at http://mls.jpl.nasa.gov/data/v3-3_data_quality_document.pdf.]

Manney, G. L., and Coauthors, 2008a: The evolution of the stratopause during the 2006 major warming: Satellite data and assimilated meteorological analyses. *J. Geophys. Res.,* **113,** D11115, doi:10.1029/2007JD009097.

——, and Coauthors, 2008b: The high Arctic in extreme winters: Vortex, temperature, and MLS and ACE-FTS trace gas evolution. *Atmos. Chem. Phys.,* **8,** 505–522, doi:10.5194/acp-8-505-2008.

——, and Coauthors, 2009a: Satellite observations and modeling of transport in the upper troposphere through the lower mesosphere during the 2006 major stratospheric sudden warming. *Atmos. Chem. Phys.,* **9,** 4775–4795, doi:10.5194/acp-9-4775-2009.

——, and Coauthors, 2009b: Aura Microwave Limb Sounder observations of dynamics and transport during the record-breaking 2009 Arctic stratospheric major warming. *Geophys. Res. Lett.,* **36,** L12815, doi:10.1029/2009GL038586.

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

Merryfield, W. J., and Coauthors, 2013: The Canadian Seasonal to Interannual Prediction System. Part I: Models and initialization. *Mon. Wea. Rev.,* in press.

Plumb, R. A., 1983: Baroclinic instability of the summer mesosphere: a mechanism for the quasi-two-day wave? *J. Atmos. Sci.,* **40,** 262–270.

Polavarapu, S., S. Ren, Y. Rochon, D. Sankey, N. Ek, J. Koshyk, and D. Tarasick, 2005: Data assimilation with the Canadian Middle Atmosphere Model. *Atmos.–Ocean,* **43,** 77–100.

Randall, C. E., V. L. Harvey, C. S. Singleton, P. F. Bernath, C. D. Boone, and J. U. Kozyra, 2006: Enhanced NO_x in 2006 linked to strong upper stratospheric Arctic vortex. *Geophys. Res. Lett.,* **33,** L18811, doi:10.1029/2006GL027160.

——, ——, D. E. Siskind, J. France, P. F. Bernath, C. D. Boone, and K. A. Walker, 2009: NO_x descent in the Arctic middle atmosphere in early 2009. *Geophys. Res. Lett.,* **36,** L18811, doi:10.1029/2009GL039706.

Randel, W. J., 1987: The evaluation of winds from geopotential height data in the stratosphere. *J. Atmos. Sci.,* **44,** 3097–3120.

Ren, S., S. Polavarapu, and T. G. Shepherd, 2008: Vertical propagation of information in a middle atmosphere data assimilation system by gravity-wave drag feedbacks. *Geophys. Res. Lett.,* **35,** L06804, doi:10.1029/2007GL032699.

——, ——, S. R. Beagley, Y. Nezlin, and Y. J. Rochon, 2011: The impact of gravity wave drag on mesospheric analyses of the 2006 stratospheric major warming. *J. Geophys. Res.,* **116,** D19116, doi:10.1029/2011JD015943.

Scinocca, J. F., 2003: An accurate spectral nonorographic gravity wave drag parameterization for general circulation models. *J. Atmos. Sci.,* **60,** 667–682.

——, and N. A. McFarlane, 2000: The parametrization of drag induced by stratified flow over anisotropic orography. *Quart. J. Roy. Meteor. Soc.,* **126,** 2353–2393.

——, ——, M. Lazare, J. Li, and D. Plummer, 2008: The CCCma third generation AGCM and its extension into the middle atmosphere. *Atmos. Chem. Phys.,* **8,** 7055–7074, doi:10.5194/acp-8-7055-2008.

Shepherd, T. G., and T. A. Shaw, 2004: The angular momentum constraint on climate sensitivity and downward influence in the middle atmosphere. *J. Atmos. Sci.,* **61,** 2899–2908.

Siskind, D. E., S. D. Eckermann, J. P. McCormack, L. Coy, K. W. Hoppel, and N. L. Baker, 2010: Case studies of the mesospheric response to recent minor, major, and extended stratospheric warmings. *J. Geophys. Res.,* **115,** D00N03, doi:10.1029/2010JD014114.

SPARC CCMVal, 2010: SPARC Report on the evaluation of chemistry-climate models. SPARC Rep. 5, WCRP-132, WMO/TD-No. 1526, 434 pp. [Available online at http://www.atmosp.physics.utoronto.ca/SPARC/ccmval_final/index.php.]