Case studies of the mesospheric response to recent minor, major, and extended stratospheric warmings

David. E. Siskind,1 Stephen. D. Eckermann,1 John P. McCormack,1 Larry Coy,1 Karl W. Hoppel,2 and Nancy L. Baker3

Received 26 February 2010; revised 6 July 2010; accepted 13 July 2010; published 23 October 2010.

[1] We have studied the mesospheric response to two recent stratospheric warmings by performing short-term forecasts at medium (1.5°) and high (0.5°) spatial resolution under different gravity wave drag (GWD) scenarios. We validated our models with our high-altitude analysis that extends from 0 to 90 km. For the minor warming of January 2008, reduced upper-level orographic GWD weakened the downward residual circulation and cooled the mesosphere. Parameterized nonorographic GWD increased the simulated mesospheric cooling. For the prolonged major warming of 2006, heavily attenuated orographic GWD led to pronounced cooling near 50 km. During the extended phase of this event, an unusually strong westerly polar vortex reformed in the lower mesosphere, which allowed westward propagating nonorographic gravity waves to reach the mesosphere and break, with net westward accelerations of over 50 m s⁻¹ d⁻¹. This, in turn, forced a strong residual circulation, yielding descent velocities over 2 cm s⁻¹ between 65°N and 85°N, consistent with previous reports of enhanced downward transport of trace constituents. The resulting adiabatic heating, as evidenced by the unusually vertically displaced stratopause at 80 km, is likely a direct consequence of this enhanced gravity wave driven descent. High-resolution simulations without parameterized GWD were closer to the analysis than medium-resolution simulations with parameterized orographic GWD only, but still did not fully simulate the mesospheric thermal response. Specifically, the 80 km temperature enhancement was still underestimated in these simulations. This suggests that higher spatial resolution is needed to adequately resolve extratropical gravity wave momentum fluxes.


1. Introduction

[2] One of the most dramatic meteorological phenomena in the middle atmosphere is the winter time sudden stratospheric warming (SSW). The largest SSWs can warm the polar winter stratosphere by as much as 60 K or more over a period of about a week [Labitzke, 1981]. Associated circulation changes occur in both the mesosphere and troposphere, though the potential impact of SSWs on Arctic weather remains controversial [Baldwin and Dunkerton, 2001; Mukougawa and Hirooka, 2004; Polvani and Waugh, 2004; Fletcher et al., 2007]. SSWs have also been reported to affect the winter thermosphere and ionosphere [Gregory and Manson, 1975; Chau et al., 2009]. The focus of this study is the mesosphere’s response to SSWs.

[3] Sudden stratospheric warmings occur when a large-amplitude planetary Rossby wave, forced in the troposphere, propagates into the stratosphere and interacts strongly with the local environment due to transience, dissipation or other nonlinear effects [Andrews et al., 1987; Haynes, 2005]. This wave mean flow interaction acts to decelerate the fast eastward polar night jet stream which defines the polar vortex. In some cases, the vortex is merely distorted or displaced from its usual polar-centric location. In other cases, the vortex is broken up completely.

[4] It has been known since the early work of Matsuno [1971] that this interaction between a planetary wave and the mean flow will also cause a mesospheric cooling to accompany a stratospheric warming [e.g., Andrews et al., 1987, Figures 6.7 and 6.8]. One area of uncertainty has been on the possible role of gravity waves in shaping this mesospheric response. Holton [1983] pointed out that the transition to stratospheric easterlies will act to reduce, or eliminate, the flux of tropospheric gravity wave momentum entering the mesosphere, leading to cooling via reduced
**Report Documentation Page**

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

<table>
<thead>
<tr>
<th>1. REPORT DATE</th>
<th>2. REPORT TYPE</th>
<th>3. DATES COVERED</th>
</tr>
</thead>
<tbody>
<tr>
<td>06 JUN 2010</td>
<td></td>
<td>00-00-2010 to 00-00-2010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. TITLE AND SUBTITLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case studies of the mesospheric response to recent minor, major, and extended stratospheric warmings</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6. AUTHOR(S)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naval Research Laboratory, Space Science Division, 4555 Overlook Ave SW, Washington, DC, 20375</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>8. PERFORMING ORGANIZATION REPORT NUMBER</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>10. SPONSOR/MONITOR’S ACRONYM(S)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>11. SPONSOR/MONITOR’S REPORT NUMBER(S)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>12. DISTRIBUTION/AVAILABILITY STATEMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approved for public release; distribution unlimited</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>13. SUPPLEMENTARY NOTES</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>14. ABSTRACT</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>15. SUBJECT TERMS</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>16. SECURITY CLASSIFICATION OF:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. REPORT</td>
</tr>
<tr>
<td>unclassified</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>17. LIMITATION OF ABSTRACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same as Report (SAR)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>18. NUMBER OF PAGES</th>
<th>19a. NAME OF RESPONSIBLE PERSON</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>

Standard Form 298 (Rev. 8-98)
Prepared by ANSI Std Z39-18
wave-induced diabatic descent [Garcia and Boville, 1994]. Later work emphasized the possibility of large zonal asymmetries in gravity wave transmission [Dunkerton and Butchart, 1984] and the possibility that the strong zonal asymmetries in mesospheric gravity wave flux during SSWs may be important for planetary wave evolution [McLandress and McFarlane, 1993; Miyahara, 1985]. For example, Smith [1996, 1997] presented evidence from UARS data that longitudinally asymmetric gravity wave fluxes can force secondary planetary waves in the upper mesosphere and lower thermosphere, a result later seen in models [e.g., Liu and Roble, 2002; Smith, 2003]. Observations of mesospheric gravity waves during SSWs show both large zonal asymmetries and large variations among SSW events [e.g., Dowdy et al., 2007], as well as an overall reduction in lower mesospheric gravity wave flux [e.g., Wang and Alexander, 2009; Wright et al., 2010]. In general, however, the relative impacts of interactions among gravity waves and planetary waves in driving the mesospheric response to SSWs remain highly uncertain.

[5] Indeed, only within the last decade have sufficient global measurements of mesospheric temperatures been available to aid modeling studies of these complex multiscale mesospheric responses to SSWs. Coy et al. [2005] studied the minor SSW of August 2002 using the forecast model component of the Advanced-Level Physics High-Altitude (ALPHA) prototype of the Navy Operational Global Atmospheric Prediction System (NOGAPS) and compared their modeled mesospheric response to temperatures from the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) instrument on board the NASA/TIMED satellite. Good agreement with the observed temperature changes reported by SABER was evident up to about 70 km, with both NOGAPS-ALPHA and SABER showing a shallow (~10 km) lower mesospheric cooling layer and similar wave 1 temperature structures. They found that this shallow lower mesospheric cooling could be understood as a straightforward quadrupole circulation response to the impulsive momentum forcing from planetary waves. One limitation of this study was the use of simple Rayleigh friction instead of a comprehensive middle atmospheric gravity wave drag (GWD) parameterization.

[6] By contrast, similarly motivated experiments with the Canadian Middle Atmosphere Model (CMAM) by Ren et al. [2008] applied parameterized orographic and nonorographic gravity wave drag at all altitudes. They modeled the major SSW of September 2002 and found that the absorption of westward propagating gravity waves in the stratosphere resulted in a large eastward gravity wave–drag anomaly in CMAM. However, unlike Coy et al. [2005], they were unable to validate their simulated mesospheric temperature response since SABER was not observing high southern latitudes in September 2002. To better understand the role of gravity waves in the mesosphere’s response to SSWs, a global model parameterizing both orographic and nonorographic gravity wave drag is needed, but incorporating validating observations that extend from the troposphere through the mesosphere. We attempt such a study here using the NOGAPS-ALPHA forecast assimilation system (see section 2).

[7] A second motivation for the present paper is observations of recent SSWs that significantly alter the simple classic stratospheric warming/mesospheric cooling paradigm. First in 2004, then again in 2006 and 2009, prolonged stratospheric warmings have dramatically perturbed the vertical temperature structure of the extratropical winter middle atmosphere [Manney et al., 2005, 2008; Hoppel et al., 2008]. Unusual features include dramatic cooling at 50 km and the reformation of the stratopause feature near 80 km. Temperature increases of up to 70 K were observed at these altitudes. Many years ago, Labitzke [1972] did suggest prolonged cooling at 50 km and hinted at a rewarming in the mesosphere during the extended phase of a SSW; however, the data didn’t exist at that time to observe the dramatic zonal mean warming at 80 km reported by the above authors. Thus the 80 km stratopause can be considered a new phenomenon to atmospheric science.

[8] This high-altitude warm layer has been linked to interesting transport phenomena [Siskind et al., 2007; Manney et al., 2009a] including large downward intrusions of nitric oxide (NO) and carbon monoxide (CO) [Randall et al., 2006, 2009; Manney et al., 2009a] and an unusual brightening in the OH Meinel band airglow Winick et al., 2009] presumably due to enhanced downwelling of atomic oxygen. The NO descent is of further interest because it can react with stratospheric ozone. Previously enhanced fluxes of NO into the stratosphere were thought to be solely due to enhanced solar or geomagnetic activity [Siskind et al., 2000; Randall et al., 2007]. But the 2006 and 2009 NO intrusions came at times of low solar–terrestrial activity and therefore must be mostly due to the unusual meteorology of the middle atmosphere [Randall et al., 2006]. Thus these prolonged SSWs offer a new mechanism of coupling between the upper and middle atmosphere.

[9] Using NOGAPS-ALPHA, this time with parameterized middle atmospheric orographic gravity wave drag instead of Rayleigh friction, Siskind et al. [2007] showed that the highly disturbed lowermost stratosphere in 2006 significantly increased the critical level removal of parameterized orographic gravity waves which normally deposit significant amounts of momentum in the 50–60 km region. This suggestion has since received observational support from Wang and Alexander [2009] and Wright et al. [2010]. The unusually low temperatures at the altitude of the conventional undisturbed polar winter stratopause were linked to this reduced orographic gravity wave drag and associated weakening of the descending branch of the mesospheric residual circulation which normally warms the winter polar stratopause [Hitchman et al., 1989]. But while Siskind et al.’s [2007] simulations provided a hint of a high-altitude warming it was much less than observed. They pointed out that the omission of parameterized nonorographic gravity wave drag from their simulations did not allow them to evaluate the cause of the high-altitude layer, Hauchecorne et al. [2007] specifically noted that enhanced high-altitude penetration of westward propagating gravity waves would be favored by the reformed lower mesospheric vortex and suggested that the breaking of these waves would drive a circulation that would heat the upper mesosphere. Thus the second key objective of the present study is to evaluate these suggestions using NOGAPS-ALPHA experiments that incorporate parameterized orographic and nonorographic gravity wave drag at all altitudes.

[10] Our approach is to examine and model two recent SSWs: the minor SSW of January 2008 [e.g., Chau et al.,
which is particularly relevant to the unusual event of 2006, and by implication, 2004 and 2009. The extended warming will be seen to differ noticeably from the other more standard cases in the inferred role that nonorographic GWD plays in driving key features of the observed mesospheric response.

Section 2 describes the version of NOGAPS-ALPHA that we use in this study. Section 3 investigates the two targeted SSWs using NOGAPS-ALPHA observational analyses and model simulations, decomposing the response and sensitivities in terms of the component zonal momentum forcing and related circulation diagnostics. Section 4 discusses the calculated vertical transport velocities from the individual cases in order to present a more unified picture of the mesospheric response into, for example, major or minor warmings, or Canadian warmings or final warmings [e.g., Charlton and Polvani, 2007]. However, it is not obvious to what extent these are useful identifiers for the mesosphere. Here we will build upon the basic “major” and “minor” terminology by introducing the concept of the so-called “extended warming” which is particularly relevant to the unusual event of 2006, and by implication, 2004 and 2009. The extended warming will be seen to differ noticeably from the other more standard cases in the inferred role that nonorographic GWD plays in driving key features of the observed mesospheric response.

2. Description of NOGAPS-ALPHA

Here we summarize aspects of the system relevant to this work. Additional details are provided by Hoppel et al. [2008], Eckermann et al. [2009] and McCormack et al. [2009].

2.1. Forecast Model Component

Our primary tool is the spectral forecast model component of the NOGAPS-ALPHA forecast assimilation system [Eckermann et al., 2009]. The model is initialized by analysis fields summarized in section 2.2. As done by Siskind et al. [2007], we performed all the forecasts here at L74 (74 levels) which place the model top at \(9 \times 10^{-5}\) hPa (a pressure altitude of 115 km). Since the analysis which provides the initial conditions only extends to 0.005 hPa, we initialize the higher levels, by progressively relaxing extrapolated analysis values to climatology, using the algorithm described by Eckermann et al. [2004]. The forecasts can be run at different resolutions; here we use either T79 or T239, that is, truncation at total wave numbers of 79 or 239; these correspond to latitude-longitude spatial resolutions on the Gaussian grid of 1.5° × 1.5° and 0.50° × 0.50°, respectively. These changes in horizontal resolution are accompanied by corresponding changes in, for example, the model’s mean spectral orography and the subgridscale orographic statistics used by the orographic GWD parameterization. At the bottom boundary, we use analyzed ground and sea surface temperatures, ice concentrations and snow depths.

A complete description of the current suite of physical and chemical parameterizations implemented in the NOGAPS-ALPHA forecast model is given in section 2.1 of Eckermann et al. [2009], and only those details relevant to the present study are mentioned here. The forecast model currently lacks relevant thermospheric physics (e.g., downward conduction of heat, far ultraviolet heating, molecular diffusion), and so we ignore forecast model results at heights above the 0.001 hPa level. A number of different parameterizations of subgridscale orographic and nonorographic GWD are available for use in the model. Here we use a Palmer et al. [1986] scheme to parameterize unresolved orographic GWD, and the multiwave Lindzen-based scheme of Garcia et al. [2007] to parameterize unresolved nonorographic GWD, the latter using the tuned parameter settings of Eckermann et al. [2009]. In this study, we use a collection of forecasts with different model resolutions and parameterized gravity wave settings to assess the impact of gravity waves on the mesosphere during different conditions.

2.2. Data Assimilation

In concert with the forecast model, we make extensive use of high-altitude global data assimilation products, or “analyses,” generated by NOGAPS-ALPHA production runs that coupled the forecast model to the Naval Research Laboratory (NRL) Atmospheric Three-Dimensional Variational (3DVAR) Data Assimilation System (NAVDAS) [Daley and Barker, 2001]. NAVDAS provides an optimal estimate of the initial conditions for the forecast model that is based, in this case, upon stratospheric and mesospheric observations, a short-range forecast, and the statistical assumptions about the errors in the background and in the observations. In addition to these initial conditions, the analysis fields also provide a specification of the evolution of the middle atmosphere during the two SSW events that we use to validate various forecasts.

NOGAPS-ALPHA data assimilation products debuted by Hoppel et al. [2008], where temperatures from the Microwave Limb Sounder (MLS) on the Aura satellite and from the SABER instrument on the TIMED satellite were assimilated to 0.01 hPa during January–February 2006. Figure 4 of Hoppel et al. [2008] compare the analysis fields with temperatures from SABER. Eckermann et al. [2009] extended the capability by assimilating newer versions of the MLS and SABER temperature retrievals up to 0.002 hPa, as well as SABER ozone and water vapor mixing ratios, for the period May–July 2007. The system described in sections 2 and 3 of Eckermann et al. [2009] was subsequently frozen as a production configuration, and used to generate 6-hourly analyses beyond July 2007 out to April 2008, and to generate a reanalysis of the period November 2005 to April 2006. Coy et al. [2009] and McCormack et al. [2009] presented science studies based on these 2006 reanalysis fields, while Nielsen et al. [2010] show results from products extending beyond July 2007. We use these products to study and model the SSWs of 2006 and 2008.

The production configuration of Eckermann et al. [2009] uses a spectral forecast model resolution of T79L68: that is, triangular truncation at total wave number 79, and 68 model layers extending to 0.0005 hPa. These forecast assimilation runs are computationally expensive, and thus impractical to regenerate at a range of different horizontal and vertical resolutions. To permit forecasts at different model resolutions, analysis fields on the T79L68 model grid are
also interpolated onto a uniform latitude-longitude grid and output at a range of reference pressure levels. These regularly gridded isobaric versions of the analysis output can then be read in as initial conditions for the forecast model running at any horizontal and vertical resolution, by interpolating them onto the model’s quadratic Gaussian grid and hybrid \( \sigma - p \) vertical layers using terrain-following hydrostatic balance constraints. We use this latter capability here to conduct forecasts at different model resolutions.

2.3. Transport Diagnostics

Since our objective is to understand how changing mesospheric conditions may ultimately affect trace constituent transport, we employ a number of wave forcing diagnostics of the zonally averaged momentum budget based upon the Transformed-Eulerian-Mean (TEM) formulation. Starting with the residual circulation defined by equations (3.5.1) and (3.5.2) of Andrews et al. [1987], the zonal momentum equation takes the form

\[ \frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = \mathbf{F} + \nabla \cdot \mathbf{F}_g + B + X, \]

where \( \mathbf{v} \) and \( \nabla \mathbf{v} \) represent the meridional and vertical components of the residual circulation, respectively, and \( \mathbf{F} \) is the mean zonal wind. Other terms are the density, \( \rho_0 \), the radius of the earth, \( a \), and the latitude \( \phi \). We will focus primarily on the terms on the right hand side of (1) that represent forcing from waves. The vector \( \mathbf{F} \) is the Eliassen-Palm (EP) flux, representing the forcing by resolved planetary and gravity waves. Following the discussion by McCormack et al. [2009], we have found it convenient to express \( \mathbf{F} \) as having two components. The component which is due mainly to planetary waves is given by

\[ \mathbf{F}_p[\phi, z] = \rho_0 \cos \phi \left[ \nabla \cdot \mathbf{v} \right] \nabla \cdot \mathbf{v} = \rho_0 \cos \phi \left[ \frac{\partial}{\partial \theta} \left( u \frac{\partial v}{\partial \theta} \right) \right]. \]

where \( z \) is the log-pressure vertical coordinate, \( f \) is the Coriolis parameter and \( \theta \) is potential temperature. Primed variables denote terms due to planetary waves and overbars denote an average around a latitude circle. The other component of the EP flux which results from smaller-scale, higher-frequency gravity waves is given by

\[ \mathbf{F}_g[z] = \rho_0 \cos \phi \left[ -u'w' \right]. \]

It is useful to consider the divergence of (2) and (3) as representing the forcing term from planetary and gravity waves, respectively. The advantages of separating the total EPFD into these two components was discussed by Miyahara et al. [1986]. They did note that there can be some ambiguity in this separation because the heat flux term \( \nabla \cdot \mathbf{F}_g \) in (2) can also be due to gravity waves of larger scales, specifically inertial gravity waves. However, they also showed that for waves with wave number greater than 5, the neglect of this term is “not crucial” and that (3) is a useful approximation for the gravity wave momentum forcing. For cases that we discuss below where we have inferred very large values of \( \mathbf{F}_g \), we have performed a spectral analysis and confirmed that the contributions to (3) are almost entirely from waves with wave number greater than 5. Thus for the rest of the paper, we will call these two forcing terms the planetary wave EP flux divergence (EPFD) and the gravity wave EPFD, although we will typically present them in units of acceleration by dividing them by \( \rho_0 \cos \phi \) according to (1).

The second term on the right-hand side of (1), \( X \), is the zonal forcing term from parameterized gravity waves which are not resolved by the model and thus not captured in equation (3). In our model and in the discussion which follows, \( X \) is also broken out into two separate components which are calculated separately. The first is the orographic component, \( X_{\text{og}} \), from the parameterization of Palmer et al. [1986] [see Siskind et al., 2007]. The second component, \( X_M \), is the non orographic GWD parameterization (hereinafter also called MGWD where M stands for “multiwave”), taken from Garcia et al. [2007] and tuned by Eckermann et al. [2009] in the context of NOGAPS-ALPHA. Eckermann et al. [2009] present the results of a series of tuning experiments which were used to reduce the forecast model bias. These tuning experiments were performed for June 2007, with specific emphasis on the summer mesosphere. While we do not retune the scheme here for our Northern winter simulations, we later investigate to what extent the tuned parameter settings of Eckermann et al. are valid for these winter conditions. Note that \( \nabla \cdot \mathbf{F}_g \), \( X_{\text{og}} \), and \( X_M \) are just the zonal components of these GW forcing terms.

In this study, we ignore the potential contributions of meridional gravity wave flux divergences to mean momentum balance and transport, based on the usual theoretical arguments of their lesser impacts due to larger mean terms in the meridional momentum equation and less critical-level filtering and flux anisotropy [e.g., Schoeberl et al., 1983]. This implicitly assumes gradient wind balance, which generally holds to a good approximation in the extratropical winter stratosphere and mesosphere [McLandress et al., 2006]. Nonetheless, since zonal asymmetries during SSWs might invalidate such assumptions, we plotted the profiles meridional momentum flux divergences of resolved gravity waves in our high-resolution T239 simulations. These plots (not shown) suggested meridional values were typically only 25% or less of the zonal values, thus justifying to first order our neglect of these terms in what follows.

We thus isolate four separate forcing terms of the zonal mean flow: planetary wave EPFD (\( \nabla \cdot \mathbf{F}_p \)), resolved gravity wave EPFD (\( \nabla \cdot \mathbf{F}_g \)), parameterized OGWD (\( X_{\text{og}} \)), and parameterized MGWD (\( X_M \)). The magnitudes and in some cases, even the sign, of these four terms can vary significantly for differing dynamical states in the winter mesosphere. In these winter simulations \( X_{\text{og}} \) is always westward which for the present context means that mountain waves act to decelerate a westerly wind. Planetary wave EPFD is generally westward because the wintertime planetary waves are dominated by stationary Rossby waves, although transient waves can produce eastward momentum. As a consequence of the downward control principle (DCP) [Haynes et al., 1991; Garcia and Boville, 1994], a westward force will drive a poleward and downward circulation and ultimately increase the winter
temperatures above their radiative equilibrium value. On the other hand, gravity wave EPFD and parameterized MGWD can either force deceleration (drag) or acceleration of \( \tau \) depending upon both the biases that are introduced into the gravity wave spectrum due to critical-level filtering by winds at lower altitudes as well as the speed and direction of the mean flow within which the waves break.

[23] The discussion above only applies to the T79 simulations. Our diagnostics for the T239 simulations are slightly different. Since finer-scale terrain is explicitly resolved at T239, the subgridscale terrain variance and resulting \( X_O \) are smaller relative to the T79 simulations. Likewise the T239 model should explicitly resolve a greater portion of the multiwave gravity wave momentum flux than at T79, although it is not clear to what specific degree. Rather than retune the T239 gravity wave drag parameterization as described by Eckermann et al. [2009] for T79, we decided to deactivate the MGWD parameterization entirely for the T239 simulations (i.e., \( X_M = 0 \)). Thus the T239 simulations only have three forcing terms: \( \nabla \cdot \mathbf{F}_P \), \( \nabla \cdot \mathbf{F}_g \), and \( X_O \), all of which will differ to some extent from their T79 counterparts.

## 3. Model-Data Comparisons

### 3.1. Outline of Different Forecast Periods

[24] We ran the forecast model for five specific 6 day periods. These five periods, summarized in Table 1, were selected to span a range of different dynamical states of the middle atmosphere. Over the course of these periods, the model gradually diverges from the analysis, reflecting the degradation of forecast skill over the 6 days in question. The degree of degradation, and the resulting discrepancy with the analysis, is interpreted here as a diagnostic of how well the physics in the model matches that in the atmosphere.

[25] Figure 1 summarizes the key atmospheric state variables characterizing these five periods as defined by the NOGAPS-ALPHA analysis. Period 1 (15–20 January 2008) was relatively undisturbed: the midstratosphere was cold (\( T = 195–200 \) K near 10 hPa), the westerly zonal winds were strong and planetary wave activity was weak. Indeed, using the meteorological analyses available at NASA/GSFC web site (http://acdb-ext.gsfc.nasa.gov/Data_services/met/ann_data.html) we see that the stratosphere at 10 hPa in mid-January 2008 was actually colder than in 2005, a year commonly used to represent quiescent conditions [e.g., Siskind et al., 2007]. Further, a comparison of the analysis in Figure 1 for 2008 with the NASA/GSFC analysis at 10 hPa shows excellent agreement. Thus Period 1 represents a good

### Table 1. Nomenclature for 6 Day Forecast Runs for 2006 and 2008

<table>
<thead>
<tr>
<th>Period</th>
<th>Dates</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15–20 Jan 2008</td>
<td>dynamically quiet</td>
</tr>
<tr>
<td>2</td>
<td>20–25 Jan 2008</td>
<td>minor warming</td>
</tr>
<tr>
<td>3</td>
<td>20–25 Jan 2006</td>
<td>major warming, initial phase</td>
</tr>
<tr>
<td>4</td>
<td>25–30 Jan 2006</td>
<td>major warming, extended phase</td>
</tr>
<tr>
<td>5</td>
<td>2–7 Feb 2006</td>
<td>major warming, extended phase</td>
</tr>
</tbody>
</table>

### Figure 1.
Pressure-time sections of the atmosphere from the NOGAPS-ALPHA analysis for the two stratospheric warming events discussed in the paper. (a, c, e) The period of the minor warming which commenced on 20 January 2008 and (b, d, f) the period of the major warming which commenced on 20 January 2006 showing zonal mean temperature at 80°N (Figures 1a and 1b), zonal wind at 60°N (m s\(^{-1}\)) (Figures 1c and 1d; the zero wind line is bold), and the amplitude of perturbation wave 1 geopotential height (m) (Figures 1e and 1f). The vertical dash-dotted black lines mark the beginning of each of the five 6 day periods that are the focus of this paper; they are also indicated at the bottom. The rationale behind the selection of these periods is discussed in the text.
baseline case. A comparison of the forecast model simulations with Figures 1a, 1c and 1e for Period 1 also allows us to assess how well the tuned MGWD settings of Eckermann et al. [2009] perform in dynamically quiet winter conditions. During Period 2 (20–25 January 2008), a minor SSW occurred. This is clearly seen in Figure 1a with the descent of the warm stratopause and the development of a noticeable cool region in the overlying mesosphere. Further, as seen in Figures 1c and 1e, the zonal winds at 60N decreased sharply, coincident with the development of a large wave 1 that peaked at the 1 hPa level. Our analysis of Period 2 can most readily be compared with the results of Coy et al. [2005] who also modeled a minor SSW; the present study includes a more complete representation of gravity wave drag.

[26] Periods 3–5 cover different phases of the unusual 2006 major SSW [Manney et al., 2008; Hoppel et al., 2008; Coy et al., 2009]. Like the 2008 event, the 2006 SSW commenced on 20 January and Period 3 (20–25 January 2006) represents the initial phase of the event where the stratopause temperature maximum descends from near 1 hPa to near 10 hPa (see Figure 1b). Despite the similarities in the onset date, it is clear from Figures 1a and 1b that there are many differences between the 2006 and the 2008 events. Thus at 10 hPa, the temperatures in 2008 were higher than in 2006 whereas at 100 hPa, 2006 became much warmer than in 2008. A mesospheric cooling is also evident in Period 3. While the cooling penetrated to lower altitudes in 2006 than in 2008, the actual mesospheric temperatures in 2006 did not become as low as they did in 2008. Indeed, despite being only a minor SSW, the magnitude of the 2008 mesospheric cooling (over 40K) was larger not only relative to 2006, but also to other events which have been studied. For example, the Ren et al. [2008] results for the major SSW warming of 2002 only yielded coolings near 20K. Similarly the SABER data modeled by Coy et al. [2005] showed a 10–15K cooling that was limited to a narrow vertical layer near 60 km. This highlights the ways in which mesospheric coolings can differ from event to event, and points out the lack of a simple correlation between the magnitudes of stratosphere warmings and the associated mesospheric coolings.

[27] We also see in Figures 1c and 1d a dramatic wind reversal in Period 3 which penetrates to well below the 10 hPa level, unlike the 2008 wind reversal which was confined to the upper stratosphere and mesosphere. Also unlike the 2008 event, the 2006 SSW was preceded by minor events earlier in January that weakened the zonal winds significantly compared with the undisturbed state seen in Period 1. Finally, the 2006 event evolved in dramatically different ways than the 2008 event. The 2008 event was relatively short lived; after several days, the stratopause returned to 1 hPa, the westerly polar jet reformed and the large wave 1 decayed. By contrast in 2006, the stratospheric easterlies persisted for almost a month [Manney et al., 2008]. Most unusual was the reformation of the stratopause temperature maximum at 80 km. This maximum then descended to near 64 km by the 2nd week of February (see Figure 1b). Finally, Figure 1f shows that there was almost no planetary wave activity in the stratosphere after the 2006 SSW although, as noted previously [Siskind et al., 2007; Manney et al., 2008] there was a relatively weak wave 1 near 80 km. This unusual extended phase is sampled by Periods 4 (25–30 January) and 5 (2–7 February).

[28] For each of these 5 periods, we performed three different forecast simulations. One with only orographic GWD (labeled “O” in all subsequent discussion). This connects back to the approach of Siskind et al. [2007], who specifically investigated the role of OGWD for the 2006 event. The second set of simulations (labeled “MO”) includes both the parameterized OGWD and multiwave GWD (MGWD), the latter as discussed above. Both the O and MO simulations are at T79 resolution; the third set is at T239 which only includes parameterized OGWD (T239O). As noted above, the T239O runs offer an additional approach toward quantifying the relative roles of planetary waves, resolved gravity waves and unresolved waves. Table 1 summarizes the 5 periods.

3.2. January 2008: Periods 1 and 2

[29] In this section we compare our 3 forecast model simulations with the analyzed temperatures, zonal winds and planetary wave 1 for Periods 1 and 2. Figure 2 shows height-latitude plots of the zonal average analysis temperatures for 19 and 24 January 2008. The 19 January temperature field is representative of Period 1, the dynamically quiet period; the 24 January temperature is representative of Period 2, near the peak of the minor warming.

[30] Figure 3 plots differences in zonal and diurnal mean temperature on 19 January between three model forecasts (F) and the analysis (A). The three forecasts (T79O, T79MO and T239O) were initialized on 15 January at 0000 UT and sampled at 0000, 0600, 1200 and 1800 UT on 19 January (forecast hours 96, 102, 108, 114) which correspond to the 6 hour update cycle of the NOGAPS–ALPHA analysis. This comparison serves to establish a baseline for forecast model temperature differences under dynamically quiet conditions. The forecast-minus-analysis (F–A) maps in Figure 3 show similar patterns for all three model forecasts, i.e., a warm bias in the upper stratosphere, a cool bias in the midmesosphere and a warm bias in the upper mesosphere. This pattern is also seen in other simulations that were initialized on other dates during the period between 10 and 15 January (not shown). The inclusion of parameterized MGWD has only a modest effect for Period 1. The T79MO forecast is warmer in the 0.1–0.01 hPa region relative to the T79O; this improves the agreement near 0.1 hPa but worsens it near 0.01 hPa.

[31] The more noticeable differences from the analysis, the upper stratospheric warm bias and the slight lower mesospheric cold bias, are independent of whether MGWD is included so we conclude this pattern is not sensitive to parameterized MGWD. It may, however, be connected to our
Figure 3. Difference between forecasted and analyzed temperatures for the 5th day (forecast hours 96–120) of Period 1 (corresponding to 19 January 2008). Shown are the temperature differences between the 5th day of the forecast model integration and the daily averaged analysis for (a) the T79O simulation, (b) the T79MO simulation, and (c) the T239O simulation. For both forecast and analysis, the daily average is taken over output for 0000, 0600, 1200 and 1800 UT.

Palmer et al. [1986] parameterization of OGWD which may be slightly overestimating this drag. One clue to this is that the T239O forecast is slightly better than the two T79 forecasts. Recall from section 2.3 that we expect the parameterized OGWD in the T239 model to play a smaller role relative to T79 because more of the subgridscale orographic variance is explicitly resolved and as we will show, this is indeed the case. If our parameterized OGWD is biased high, we might therefore expect reduced bias in the T239O simulation. Further, the sense of the discrepancy is also consistent with the fact that the Palmer et al. [1986] parameterization will tend to produce larger OGWD than other schemes such as, for example, the Webster et al. [2003] approach which partitions the low-level drag into gravity wave generating and flow-blocking components. The winds in the T239O forecast are about 5 m s$^{-1}$ stronger than in the two T79 simulations, and agree somewhat better with the analysis in the upper stratosphere and lower mesosphere (not shown). This is consistent with less OGWD and thus less dynamical heating through OGWD–driven descent.

Figure 4 shows the zonal mean temperature change from 24 January (a daily average) relative to 20 January for the analysis and three model forecasts as a function of altitude and latitude. In general the three model forecasts in Figure 4 are in good qualitative agreement with the analysis. One difference is that none of the models completely capture the >40 K cooling between 0.3 and 0.1 hPa seen in the analysis; model coolings are smaller and centered above the 0.1 hPa level. Differences between the T79MO and T79O calculations are evident at the higher altitudes, with the T79MO model simulating more cooling than the T79O calculation. The T79MO simulation also seems to be in better agreement with the analysis at the higher altitudes. For example, the −20 K contour in Figure 4 that extends above 0.01 hPa level in the analysis only does so in the T79MO simulation. Our results are thus consistent with those of Ren et al. [2008] in that we also find nonorographic GWD (here called MGWD; see section 2.3) to be important in governing the vertical extent of the mesospheric cooling. However, compared with Ren et al. [2008] MGWD effects are more localized in our results. For example, their model showed effects at all altitudes; MGWD effects in our model simulations are generally limited to the midmesosphere and above. Finally, the relative shallowness of the mesospheric coolings documented by Coy et al. [2005] and Siskind et al. [2005] compared with this one may suggest that MGWD effects differ from event to event.

[33] Figure 5 shows the evolution of the zonal winds at 60°N for the three model forecasts that can be compared to the analysis shown in Figure 1. The results are generally consistent with the picture presented in Figure 4. The model forecasts tend to underestimate the penetration of the easterly winds into the upper stratosphere relative to the analysis. The effects of parameterized MGWD in the T79MO simulation are seen most clearly in the generation of westerly winds above 0.01 hPa level, reaching 60 m s$^{-1}$ on 26 January at 0.001 hPa, somewhat greater than corresponding analysis winds which reach 50 m s$^{-1}$ (see Figure 1c). The winds in the
could have occurred by chance, were very low ($<10^{-4}$ for $p = 0.0018$ hPa). We thus conclude that over the 4 day period in question, the increase in Wave 1 in the T79MO model relative to T79O model is statistically significant. We also performed the $t$ test for each of the three models paired with the analysis. At 0.0018 hPa, the probabilities were slightly larger, but less than 0.024 and thus still small enough that meaningful comparisons between the models and the analysis can be made.

[35] The appearance of a secondary high-altitude planetary wave, driven by zonally asymmetric multiwave GWD (MGWD), was also noted in model simulations of Meyer [1999] and Liu and Roble [2002]. One additional signature of this secondary wave growth is a phase change with altitude above and below the wave amplitude minimum, which is evident in Figure 6b with all three model forecasts in generally good agreement with the analysis. Since the degree of wave amplitude growth is the result of MGWD, it suggests that a comparison of calculated wave 1 amplitude with the analysis can be used as an additional constraint on the magnitude of parameterized MGWD in the forecast models.
Diagnostics of the four zonal momentum forcing terms for the T79MO model, averaged from 50°–70°N. Dashed T79 model lines are an average for Period 1 (15–20 January 2008), representing the dynamically quiet period before the 2008 SSW. Solid lines are for Period 2 (20–25 January 2008), representing the period of the minor SSW. Figure 7c shows the parameterized orographic and multiwave gravity wave drag (OGWD and MGWD) separately.

much as the overall magnitude of the mesospheric cooling, discussed above. Accordingly, the smaller high-altitude wave 1 amplitude in the T79O simulation can be attributed to insufficient explicitly resolved MGWD. Likewise the fact that the T79MO model wave 1 is larger than the analysis may suggest that the parameterized MGWD is somewhat overestimated at these altitudes. Both conclusions are consistent with the comparison of the model winds with the analyzed winds discussed in Figure 5. Additional forecasts we have done (not shown) with an efficiency (ε) of parameterized MGWD [see Eckermann et al., 2009] that is smaller than that used in the T79MO forecast yielded a value of high-altitude wave 1 which was intermediate between the T79O and the T79MO forecasts, further confirming the link between high-altitude planetary wave amplitude and parameterized MGWD.

We can understand the overall temperature, wind and planetary wave changes by analyzing the four momentum forcing terms discussed in section 2.3 for the models both before and during the SSW. These four terms, including the three GWD terms (ωc, ωm and resolved gravity wave EPFD) and the planetary wave EPFD, and their sum are shown in Figure 7 for the 2008 T79MO simulation, averaged over the latitude range 50°N–70°N. We show results for both the quiescent period before the warming (Period 1) and during the warming (Period 2). Figure 7a shows that in the stratosphere there is an increase in the westward forcing from the planetary wave from Period 1 to Period 2 while in the mesosphere the change in planetary wave forcing is generally of opposite sign. In the lower mesosphere the parameterized OGWD in Figure 7c is reduced due to the increased filtering of mountain waves by the weakened winds the stratosphere. In the upper mesosphere, the parameterized MGWD in Figure 7c switches sign from westward to eastward, consistent with the results of Ren et al. [2008]. MGWD as a function of longitude (not shown) shows large variations which have a wave 1–like signature. Thus although the overall average is near zero, for a wide band of longitudes from 80°E–220°E, values up to +20 m s⁻¹ d⁻¹ are seen with negative values at most other longitudes. This is consistent with the increased planetary wave 1 discussed in the context of Figure 6 according to the zonally asymmetric GWD forcing theory of Smith [1996, 1997]. The reduction in parameterized OGWD and the eastward shift of both the planetary wave EPFD and the parameterized MGWD all act in concert to reduce the poleward and downward residual circulation that normally heats the winter mesosphere and thus all contribute to the mesospheric cooling. The resolved gravity wave EPFD in Figure 7b is generally small at T79 resolution with a slight eastward component during the SSW in the upper mesosphere. The combined effect of all the changes in the upper mesosphere is to switch the total zonal forcing from westward before the SSW to weakly eastward during the SSW.

Figure 8 shows the same diagnostics as Figure 7, but for the T239O simulations. While the overall total forcing before and after the SSW is generally similar, some differences are immediately apparent. For example, during Period 1 (before the SSW), the net westward forcing near the stratopause and in the lower mesosphere is about 25–30% less at T239 resolution than at T79 resolution (see Figures 7d and 8d). The largest contribution to this difference is the 25–50% reduction in the parameterized OGWD in the T239 model relative to the T79 model, consistent with the smaller fraction of orographic variance whose OGWD must be parameterized in the T239 model. However, near the stratopause, we do not see a compensating increase in resolved gravity wave EPFD at T239 relative to T79 (see Figures 7b and 8b) so the net effect of increased model resolution is a smaller total westward forcing at these altitudes. This is consistent with slightly lower upper stratospheric temperatures in Period 1 as we discussed in the context of Figure 3.

In the upper mesosphere, the net T239O forcing during Period 2 is less in the T239O simulation than in the T79MO simulation due, in part, to the absence in the T239O run of

Figure 7. Diagnostics of the four zonal momentum forcing terms for the T79MO model, averaged from 50°–70°N. Dashed T79 model lines are an average for Period 1 (15–20 January 2008), representing the dynamically quiet period before the 2008 SSW. Solid lines are for Period 2 (20–25 January 2008), representing the period of the minor SSW. Figure 7c shows the parameterized orographic and multiwave gravity wave drag (OGWD and MGWD) separately.

Figure 8. Same as Figure 7 but for the T239O model, and only three forcing terms are shown since parameterized multiwave GWD is not included.
parameterized MGWD. The net total forcing is near zero above the 0.01 hPa level in Figure 8d whereas in Figure 7d, it is about 5–15 m s$^{-1} d^{-1}$ in the eastward direction. Thus the T79MO calculation produces a deeper mesospheric cooling than does the T239O calculation and the faster westerlies seen in Figure 5 in the T79MO simulation relative to the T239O simulation reflect that.


As discussed by Manney et al. [2008] and evident in Figure 1, the major SSW of 2006 commenced on 20 January with a reversal of the zonal mean zonal winds at 10 hPa at 60°N (several minor SSWs with wind reversals at 1 hPa occurred earlier in the month). This wave 1 event was one of the strongest and most prolonged SSWs on record to that time (a similar prolonged event in 2009 was even stronger; e.g., Manney et al., 2009b). The rationale for the 3 different initialization dates for the forecast model was discussed in section 3.1. We will also use the terms “initial phase” for Period 3 (20–25 January) and “first and second extended phases” for Periods 4 (25–30 January) and 5 (2–7 February), respectively. Note that since our interest is in mesosphere’s response to this SSW and how it might impact observed NO descent, we are not investigating the dynamics of the SSW [e.g., Coy et al., 2009] or the preconditioning which likely occurred before 20 January [Liu and Roble, 2005; Hoffman et al., 2007]. Here we will use the forecasts as well as our reanalysis of this period to study the role of GWD in creating the unusual conditions seen in the mesosphere during this period. As we will show, the forcing of the zonal momentum of the middle atmosphere during the extended phases is quite different from that encountered during both the initial phase of this SSW as well as anything seen throughout the minor SSW of 2008.

Figure 9 presents height-latitude plots of zonal mean and daily mean temperatures for the analysis and (second through fourth rows) nine forecasts. Each date corresponds to the 6th day of Periods 3, 4, and 5 (which sample the 2006 event). For each period, there is a T79O model (second row), a T79MO model (third row) and a T239O model (fourth row). Each model output represents the average of 0000, 0600, 1200, 1800 and 2400 UT for the indicated date. The contours are at 10K intervals.
overestimate of MGWD for the 2008 event, here, for the 2006 event, it appears we have an underestimate of MGWD. This disparity from one event to the other is not surprising since our MGWD parameterization is not tuned for specific meteorological events, but rather, is designed to produce the best overall climatology [Eckermann et al., 2009].

In the 30 January column of Figure 9 (Period 4), the transition to the extended phase of the SSW is seen. The SSW has descended to 100 hPa although the overall temperature in the stratosphere is now reduced relative to Period 3. There is also now the beginning of the high-altitude warm layer at about 0.004 hPa. This is best captured in the T79MO forecast, partially captured in the T239O forecast but only weakly with the T79O model.

By 7 February (Figure 9 right, Period 5), the high-altitude warm layer is well established with temperatures near 250K. The rest of the stratosphere above the 50 hPa level has cooled below 220K. All three forecast simulations show the high-altitude warm layer, although the temperature maximum in the T79O simulation falls well below the analysis. Both the T79MO and T239O forecasts come closer to the analysis, but are still about 20K too cold. Since, as we will show, gravity waves are an important part of the momentum budget in this situation, this model-data discrepancy may reflect shortcomings in the model’s treatment of GWD.

Figure 10 shows the zonal mean zonal wind profiles from the forecasts and the analysis, averaged from 55°N–65°N for the last day of Periods 3 (Figure 10a), 4 (Figure 10b) and 5 (Figure 10c). In Figures 10a–10c, the T79O simulations are clearly the furthest from the analysis in the upper mesosphere region at 0.01–0.001 hPa, although interestingly, the sign of this F-A bias changes between Period 3 and Periods 4–5. Thus the lack of MGWD (either parameterized or explicitly resolved) means the T79O winds are too weak in Period 3, consistent too little eastward forcing and an underestimated mesospheric cooling and consistent with the zonal mean temperatures in Figure 9. For Periods 4 and 5, the T79O winds are now too strong. As we will see, the sign of the parameterized MGWD and $F_g$ wave forcing switches to negative (westward) for the extended phases; since the T79O model does not include parameterized MGWD, it undoubtedly underestimates the drag on the westerly wind. The T79MO and T239O forecasts do better, but their mean winds are also too weak for Period 3 and too strong for Period 5, suggesting that the gravity wave forcing is even greater at those times than simulated by any of the model forecasts.

[44] A secondary wave 1 amplification at the altitude of the reformed stratopause has been noted for this event [Siskind et al., 2007; Manney et al., 2008]. Figure 11 presents plots of analyzed geopotential height anomalies at 0.01 hPa in Hovmöller form for the entire January–February 2006 period. A clear change in behavior ensues as the SSW proceeds. Prior to 20 January, a westward propagating wave 1 disturbance with an approximate 8–10 day period is seen, somewhat similar to the disturbance documented by Palo et al. [2005] in SABER temperature data. Figure 11 reveals a dramatic change in character after 20 January with a large-amplitude quasi-stationary wave 1 disturbance on 25–30 January that transitions abruptly to a weaker antiphased quasi-stationary wave 1 feature in February.

Figure 10. Mean zonal winds for the last day of (a) Period 3 (b), Period 4, and (c) Period 5. The associated dates are indicated in the upper left. Figure 10a also provides the key to the four curves.

Figure 11. Time-longitude plot of the perturbation geopotential height for 0.01 hPa for the period 1 January to 28 February 2006 (y axis is day number of 2006).
Manney et al. [2008] suggest that enhanced wave 1 amplitudes on 25–30 January may be forced in situ by zonally asymmetric GWD much as we discussed in the context of the January 2008 event (Figure 6). This appears not to be the case here. Figure 12 compares wave 1 geopotential height amplitudes, averaged over the period 27–30 January for the T79O, T79MO and T239O simulations and the analysis. Both the T79O model and the T79MO model give essentially the same profile which suggests it is not driven by zonally asymmetric nonorographic GWD. Rather the wave enhancement may be a combination of forcing from below combined with some in situ instability. Evidence for this is given in Figure 13 which presents EP flux vectors, calculated from the analysis, for wave 1, as well as their divergence (colored contours), superimposed upon a height–latitude plot of the analyzed zonal mean zonal wind for 27 January. By this date, the westerly zonal wind is beginning to reform above the stratospheric easterlies. The arrows show significant propagation of EP flux into this region from lower latitudes and altitudes (e.g., near 40° at 5 hPa). In addition, there is a suggestion of some divergence (wave source) at the 0.5 hPa level. Note that this level corresponds closely with the local minimum in wave 1 amplitude shown in Figure 12. The grey shading in Figure 13 indicates regions where $\partial^2 \phi / \partial z^2 < 0$, thus indicating regions of potential instability near the wave source region. A detailed analysis of the planetary wave structure associated with this event is beyond the scope of this paper. Nonetheless, Figure 13 suggests that the enhanced wave 1 may have its origin at lower altitudes, either channeled up from lower latitudes in the stratosphere by the perturbed zonal wind profile, or generated by instabilities near the 50 km region.

The difference between the 2006 event and the minor SSW of 2008 is made clear in Figures 14 and 15 which present the momentum forcing terms for the T79MO simulation (Figure 14) and the T239 simulation (Figure 15). These can be compared with Figures 7 and 8 for the 2008 run, although here, each panel presents three curves for the three periods of the 2006 event. Starting with the planetary wave EPFD, in the stratosphere we see westward forcing in

---

**Figure 12.** Calculated wave 1 amplitude at 60°N for the period 27–30 January 2006 (initialized on 25 January, start of Period 4) for the three models (T79O, T79M, and T239) compared with the analysis.

**Figure 13.** EP Flux vectors from the analysis for 27 January 2006 for wave 1. The contours are the zonal wind with solid (dashed) curves representing westerly (easterly) values. The blue contours are regions of EP flux divergence, and the red contours are regions of convergence. Shading indicates regions where $\partial^2 \phi / \partial z^2 < 0$.

**Figure 14.** Diagnostics of zonal momentum forcing for the three 2006 T79MO simulations, averaged over 50°N–70°N. The format is the same as Figures 7 and 8. The solid line is for Period 3, the dashed line is Period 4, and the dash-dotted line is Period 5. Figure 14c also gives the associated dates and shows both the parameterized OGWD (curves without symbols) and MGWD (curves with symbols) separately.
Same as Figure 14 but for the T239O simulations
Calculated TEM $w^*$, area weighted and averaged using both methods, and the values generally agreed to from the latter (1994, equation (4)). This approach has the advantage of allowing the effects of each drag term on the residual vertical velocity to be explicitly separated, but requires the assumption of steady state which may not be valid in a transient situation such as an SSW. We computed $w^*$ using both methods, and the values generally agreed to ~10–20%.

Figure 15. Same as Figure 14 but for the T239O simulations for Periods 3, 4, and 5 (2006 event). Only three forcing terms are shown since parameterized MGWD is not included. Figure 15c identifies the dates and curves.

critical-level interactions. This is consistent with our earlier model simulations of this period [Siskind et al., 2007].

[47] The parameterized MGWD produces eastward forcing at the beginning of the SSW (Period 3) and then transitions to westward as the SSW proceeds into the extended phases. This is also seen in the resolved gravity wave EPFD (Figures 14b and 15b) which, after being weakly westward for Period 3, changes to large and westward by Periods 4 and 5. The large westward gravity wave EPFD is in sharp contrast to what was seen during the 2008 minor SSW and is a distinguishing feature of the extended phase of the 2006 major SSW. We interpret this as being due to the increasing westerly winds above 40 km which filter out gravity waves with eastward phase speeds. Our results are therefore consistent with the suggestion of Hauchecorne et al. [2007] that the fast westerly winds in the reformed strong vortex in the lower mesosphere would favor mesospheric penetration of gravity waves with westward phase speeds that, upon breaking, would ultimately drive a poleward and downward circulation. They are also consistent with results from Hoffman et al. [2007], who found enhanced turbulent energy dissipation rates from their radar data during this period of intensified westerly winds. Note that the T239O simulation shows much more westward gravity wave EPFD than the T79MO simulation, most notably in Period 5. This is a natural consequence of more fully resolved gravity waves in the T239O simulation compared with the T79MO calculation. The planetary wave EPFD shows the westward stratospheric forcing from the wave 1 in Period 3, similar to what was seen in 2008 (see Figures 7a and 8a).

[48] The total zonal forcing (Figures 14d and 15d) reflects the change from weakly westward forcing in the upper mesosphere and westward forcing in the stratosphere during Period 3 to much stronger westward forcing in the upper mesosphere during Periods 4 and 5 with near-zero forcing in the stratosphere during the extended phase. The lack of external momentum sources below 70 km (about 0.04 hPa) during the extended phase is consistent with the spin-up of the high-altitude vortex; that is, the atmosphere relaxes toward a radiative equilibrium state. In the T239O simulation, this westward forcing comes entirely from resolved waves; in the T79 model, the gravity wave EPFD is smaller, but the parameterized MGWD makes up the difference. Without this parameterized MGWD, forecasts at T79 have insufficient westward forcing of the upper mesosphere, yielding a cold bias characterized by excessively strong westerlies and consistent with the T79O results presented in Figure 9.

4. Vertical Motions

[49] Since a key motivation for this work was to understand the enhanced downward flux of NO$_2$ reported by Randall et al. [2009, 2006] during SSW events, it is useful to compute and study high-latitude TEM vertical velocities, $w^*$, for the several cases we have considered. Our approach is to directly evaluate the expression [Andrews et al., 1987, equation 3.5.1a]

$$w^* = v - \rho_0 c^{-1} (\rho_0 w_\|/\beta_z)$$

(4)

and then solve for $w^*$ from the continuity equation [Andrews et al., 1987, equation 3.5.2c]. A second approach is to integrate the so-called downward control expression of Garcia and Boville [1994, equation (4)]. This approach has the advantage of allowing the effects of each drag term on the residual vertical velocity to be explicitly separated, but requires the assumption of steady state which may not be valid in a transient situation such as an SSW. We computed $w^*$ using both methods, and the values generally agreed to ~10–20%.

Figure 16 shows the area-weighted $w^*$ from the T79MO simulations, averaged over the latitude range 65°N–85°N for the five periods we have considered: before and during the minor SSW of 2008 and for the three periods in the major SSW of 2006. Figure 16 shows uniform downwelling during the quiet period (Period 1). During Periods 2 and 3, this downwelling weakens, consistent with reduced GWD in the mesosphere and a relaxation toward radiative equilibrium. As we noted before, despite the more greatly disturbed

Figure 16. Calculated TEM $w^*$, area weighted and averaged over the latitude range 65°N–85°N and over 6 days for the T79MO models. The inset illustrates the specific period corresponding to each curve symbol.
stratospheric conditions in 2006 relative to 2008, it is too simplistic to then argue that the mesospheric cooling response is correspondingly greater. The mesospheric cooling penetrates to lower latitudes during 2006, but the change in $w^*$ is greater during Period 2 (minor SSW) than during Period 3 (major SSW). However, as we’ve noted, our model simulations underestimate the depth of the mesospheric cooling in 2006. This suggests that the perturbation in the $w^*$ in Period 3 may be underpredicted and may have been large enough to change $w^*$ from negative to positive. Support for this can be seen in the apparent upward buckling above 70 km of the contours of NO from the Atmospheric Chemistry Experiment (ACE) presented by Randall et al. [2006] in their Figure 1.

[51] In any event, the extended phases of the 2006 event (Periods 4 and 5) are clearly distinguished by enhanced (rather than reduced) downwelling. Our results suggest latitudinally averaged descent rates exceeding 2.5 cm s$^{-1}$ in the 80–90 km region (0.01–0.001 hPa), more than twice that of the baseline Period 1 case. These large negative $w^*$ magnitudes are consistent with the large westward gravity wave forcing terms, both parameterized MGWD and resolved EPFD, illustrated in Figures 14b, 14c and 15b for pressures less than approximately 0.02 hPa (i.e., above the core of the fast westerly jet). Also note that as the extended phase evolves (2 February 2007), the downwelling remains small in the lower mesosphere. This is consistent with the very cold stable vortex that was observed and the near absence of any momentum forcing from parameterized OGWD. Based upon Figure 1 of Randall et al. [2006], it appears that much of the NO descent through the lower mesosphere occurred much later in February, as the warm layer began to descend to lower altitudes. Manney et al. [2009a] also discussed the timing of the descent in February with regards to CO transport.

5. Conclusion

[52] Using the NOGAPS-ALPHA forecast assimilation system, we have studied the dynamics that control the mesospheric response to four different stratospheric dynamical states: so-called quiet conditions, a minor SSW, the initial phase of a major SSW, and the extended phase of a major SSW. These states are listed above in order of increasing stratospheric disturbance, i.e., the extent to which the conventional westerly wintertime stratospheric zonal winds are reduced or even reversed. The first two states occurred in January 2008 and the latter two in January 2006 (the extended phase was sampled in two different periods, thus five periods of analysis for four dynamical states). We have characterized these four states by the different contributions from gravity wave and planetary wave zonal momentum forcing. Specifically, we have evaluated the relative roles of parameterized (unresolved) orographic and nonorographic (multiwave) gravity wave drag (OGWD and MGWD) and resolved gravity wave and planetary wave forcing (EPFD). By running the forecast model at two different horizontal resolutions (T79 and T239), we were also able to explore different combinations of explicitly resolved and parameterized gravity wave forcing.

[53] One consistent pattern that emerges from this study is the reduction and then disappearance of parameterized OGWD in the upper stratosphere and mesosphere as the stratosphere becomes progressively more disturbed. As the stratospheric westerly zonal flow progressively weakens or reverses, orographic gravity waves increasingly encounter stratospheric critical lines where they are completely absorbed. This removal of OGWD forcing from the stratosphere and mesosphere serves to remove a source of wave-driven mean downwelling leading to cooling at 50 km.

[54] For both the minor SSW and the initial phase of the major SSW, both of the other two gravity wave components, parameterized MGWD and resolved gravity wave EPFD, became more eastward relative to their quiet condition values. This transition also acted to reduce the downwelling and led to cooling. For the initial phase of the 2006 event, we tended to underestimate the overall mesospheric cooling, and we speculate that it may be due an underestimate of the parameterized MGWD for those conditions. For both the 2008 and 2006 SSWs, the T239O simulations consistently yielded forecasts which better reproduced the analysis than did the T79O (OGWD only) forecasts due to the greater ability of the higher-resolution T239 model to resolve the gravity waves which control the mesospheric momentum budget. However, for the 2006 event, the T239O forecasts still did not fully reproduce the magnitude of the initial mesospheric cooling or the subsequent high-altitude warming. This suggests that the T239 resolution may not be sufficient to resolve all the gravity waves (either orographic or nonorographic) needed to completely specify the zonal momentum budget.

[55] We have identified the extended phase of a major SSW as being dynamically distinct from the initial phase. The distinguishing feature of the extended phase was that the sign of both the parameterized MGWD and the explicitly resolved gravity wave EPFD switched from eastward to westward as the underlying zonal winds at lower altitudes switched from easterly to westerly. This spin-up of the lower mesospheric polar vortex is linked to the continued cooling at 50 km, which in turn, is linked to the suppression of orographic gravity wave transmission by the disturbed lower stratosphere over a 2–3 week period. Thus, although we do not have the data to prove it, we can speculate that several of the persistent warmings described by Manney et al. [2005] might have been associated with the development of a stratosphere at 80 km (see also Siskind et al. [2007] for additional historical speculation along these lines). In any event, as the extended phase develops, we deduce very large westward forcing at altitudes near the conventional upper mesosphere, approaching 60 m s$^{-1}$ d$^{-1}$, was shown to correlate with enhanced downwelling rates, over 2.5 cm s$^{-1}$. This is consistent with previous observations of enhanced descent of NO and CO for these times and locations.

[56] Finally, we evaluated the role of high-altitude planetary waves, both in terms of their overall momentum forcing and their interaction with gravity waves. There were significant differences between the two events. In 2008, for the minor SSW, we showed how the regrowth of planetary wave 1 in the model’s upper mesosphere could be used as a diagnostic of MGWD, much as the overall magnitude of the mesospheric cooling is used. However, in 2006, during the major SSW, reintensification of the high-altitude wave 1 seemed to have its origin at lower altitudes and was unconnected to MGWD. The momentum forcing from the planetary waves also differed between the two years. These differences may be part of the reason that, despite some similarities in the gravity wave forcing for the minor event in 2008 and the
Acknowledgments. Siskind et al.

References


N. L. Baker, Marine Meteorology Division, Naval Research Laboratory, Monterey, CA 93943, USA.
L. Coy, S. D. Eckermann, J. P. McCormack, and D. E. Siksd, Space Science Division, Naval Research Laboratory, 4555 Overlook Ave., SW, Washington, DC 20375, USA. (david.siksd@nrl.navy.mil)
K. W. Hoppel, Remote Sensing Division, Naval Research Laboratory, 4555 Overlook Ave., SW, Washington, DC 20375, USA.