ROLE OF GRAVITY WAVES IN VERTICAL COUPLING DURING SUDDEN STRATOSPHERIC WARMINGS

ERDAL YİĞİT¹, ALEXANDER S. MEDVEDEV²,³

ABSTRACT

Gravity waves are primarily generated in the lower atmosphere, and can reach thermospheric heights in the course of their propagation. This paper reviews the recent progress in understanding the role of gravity waves in vertical coupling during sudden stratospheric warmings. Modeling of gravity wave effects is briefly reviewed, and the recent developments in the field are presented. Then, the impact of these waves on the general circulation of the upper atmosphere is outlined. Finally, the role of gravity waves in vertical coupling between the lower and the upper atmosphere is discussed in the context of sudden stratospheric warmings.

Subject headings: gravity wave, thermosphere, ionosphere, sudden stratospheric warming, general circulation model, parameterizations, keyword

1. INTRODUCTION

The lower atmosphere, where meteorological processes take place, is the primary source of internal atmospheric waves: gravity waves (GWs), planetary (Rossby) waves, and solar tides. These waves can propagate upward and influence the dynamics and thermal state of the middle and upper atmosphere (see e.g., the reviews of Fritts & Alexander 2003; Laštovička 2006; Yığit & Medvedev 2015). Waves transfer their energy and momentum to the mean flow via breaking and dissipative processes, such as radiative damping, eddy viscosity, nonlinear diffusion, molecular diffusion and thermal conduction, and ion drag (Yığit et al. 2008). Sudden stratospheric warmings (SSWs) are spectacular events that disturb the circulation in the winter hemisphere. They affect not only the stratosphere, but their influence extends to the mesosphere and thermosphere. In the upper atmosphere, plasma processes, such as Joule and auroral heating, ion friction are important processes that shape the morphology and dynamics. Thus, interactions between the lower and upper atmosphere should be considered within the framework of the atmosphere–ionosphere system.

Such coupled upper atmosphere–ionosphere system is subject to the following internal and external influences:

- Meteorological effects that encompass internal wave impacts and transient processes of lower atmospheric origin,
- Internal processes due to nonlinearity,
- Space weather effects that are associated with the solar and magnetospheric phenomena.

Among the meteorological effects, we distinguish a direct influence of internal GWs on the upper regions of the atmosphere. Although transient events such as SSWs are technically categorized as stratospheric processes, and, thus take place above the region of weather-dominated phenomena, they are often referred to as meteorological effects in the context of the upper atmosphere research.

The thermosphere–ionosphere system is highly nonlinear. In the real atmosphere, ion and neutral parameters vary simultaneously, and the resulting changes in the heating ought to contain higher order terms, which is indicative of the nonlinear nature of the system (Yığit & Ridley 2011a). The atmosphere-ionosphere system is subject to the influence of space weather, which can enhance these nonlinear processes and impact the upper atmosphere (Prölls 2011, and references therein).

In this paper, we report on the recent advances in understanding the meteorological effects in the upper atmosphere, focusing primarily on the links between SSWs, small-scale GWs, and thermosphere–ionosphere dynamics.

2. INTERNAL GRAVITY WAVES

Internal gravity waves are characteristic features of all stably stratified planetary atmospheres. GWs in the upper atmosphere have been studied for more than 50 years since the early work of Hines (1960). Their importance for the general circulation of the middle atmosphere has been greatly appreciated (e.g., Garcia & Solomon 1985; Becker 2011). However, despite the previous theoretical approaches to GW propagation into the thermosphere (Klostermeyer 1972; Hickey & Cole 1988), only since recently, the role of GWs in coupling the lower and upper atmosphere is being increasingly acknowledged (Yığit et al. 2009; Vadas & Liu 2009; Fritts & Lund 2011; Yığit et al. 2012a; Miyoshi et al. 2014; Hickey et al. 2011, 2010; Heale et al. 2014).

GWs are always present in the lower and upper atmosphere, however their amplitudes and dynamical importance differ with height. Wave energy is proportional to air density, and, therefore, a conservatively propagating harmonic has a larger amplitude in regions with lower density. In the troposphere, GW amplitudes are relatively small, however, their dynamical importance increases with height, and can no longer be neglected in the middle and upper atmosphere.

We next discuss basic principles of how GW processes are represented in atmospheric models, reviewing the underlying assumptions and limitations.

2.1. Principles of parameterization of gravity wave processes in global atmosphere models

Spatial scales of GWs are considerably smaller than the planetary radius. Their sources are highly intermittent, and
propagation is strongly dispersive. Therefore, the GW field in the thermosphere is highly irregular and transient. Unlike with distinct larger-scale planetary waves, it appears as an ever changing “sea of waves” with occasional well-defined and detectable packets. In many applications, such chaotic wave field and its influence on the larger-scale flow can be conveniently described in terms of statistical quantities devoid of the phase information. Examples of the most widely used statistical characteristics for the GW field are the variance $\phi'^2$, vertical flux of horizontal momentum $\vec{u}'w'$, sensible heat flux $w'T'$, etc, where $w'$, $T'$, and $\phi'$ are the deviations of vertical velocity, temperature and of any field variable from the corresponding mean values, respectively.

General circulation models (GCMs) have spatial resolutions usually much coarser than the scales of GWs. Only few GCMs have endeavored to perform simulations with grids small enough in an attempt to resolve at least a part of the GW spectrum (e.g., Miyoshi & Fujiwara 2008; Miyoshi et al. 2014). In most simulation studies, the effects of subgrid-scale GWs have to be parameterized. This means that (1) the average effects must be presented in terms of statistical quantities similar to the described above, and the quantities have to be functions of the background flow. In other words, the parameterization has to self-consistently capture responses of the wave field to the evolution of the resolved larger-scale flow.

(2) Parameterizations should preferably be based on first principles, that is, they should rely on rigorous laws of physics rather than on a set of empirically introduced (tuning) parameters. Obviously, no parameterization can be devoid of such parameters as they are a substitute for an unknown. But the lesser the number of tunable parameters, the more sophisticated the parameterization is.

(3) Parameterizations must be verifiable. This condition means that they have to provide quantities, which can be compared with observations. For instance, GW-induced heating/cooling rates are hard to measure, but temperature variances $T'^2$ can be.

2.2. Assumptions and limitations in gravity wave parameterizations

In modeling, it is assumed that the majority of GWs are generated in the lower atmosphere. Amplitudes of those excited in the upper layers and propagating downward decrease exponentially with height together with their influence on the mean flow. Therefore, (1) only harmonics propagating upward are considered in parameterizations. This assumption allows one to omit a detailed consideration of the wave reflection, and to (2) apply the Wentzel-Kramers-Brillouin (WKB) approximation. Under the WKB method, (3) only those harmonics are considered whose vertical wavelengths are much shorter than vertical variations of the background fields. Mathematically, the latter can be expressed as $k_z H \gg 1$, where $k_z$ is the vertical wavenumber and $H$ is the density scale height. This limitation becomes very restrictive in the thermosphere, because fast (and long vertical-wavelength) harmonics have more chances to penetrate from tropospheric heights. In the real world, GWs propagate obliquely with respect to the surface. However, because $k_z \gg k_h$ for most harmonics, $k_h$ being the horizontal wavenumber, parameterizations (4) usually assume vertical-only propagation. Limitations of this approximation in the middle atmosphere have been recently discussed in the work by Kalisch et al. (2014), and higher-order effects have been found with a scheme employing ray tracing (Song et al. 2007). A special care should be taken with parameterizations extending to the thermosphere, where longer vertical wavelength harmonics (lower $k_z$) tend to propagate to from below. In other words, all gravity waves accounted for by a parameterization must remain within their grid columns. Finally, (5) all column-based parameterizations employ a steady state approximation. That is, transient processes of wave propagation assume an instantaneous response to changes in the forcing below. This approximation is suitable for modeling the general circulation, however, implications of time delay due to the finite group speed of wave packets should be carefully weighted for simulations of more rapid processes.

Parameterizations compute vertical profiles of a specified statistical quantity characterizing the GW field, such as horizontal velocity variance $u'^2$ (e.g., Medvedev & Klaassen 1995), or vertical flux of horizontal momentum $u'^2w'$ (e.g., Yiğit et al. 2008). The former is convenient for comparison with observations of GW spectra. The latter is physically more lucid, because $\rho u'w'$ is an invariant in a non-dissipative atmosphere. In GCMs, sources are specified by (1) prescribing the corresponding quantity at a certain level $z_0$ in the lower atmosphere, or (2) calculating it interactively using large-scale fields resolved by the model as an input. The latter is sometimes called “parameterization of gravity wave sources”. Because mechanisms of wave excitation in the lower atmosphere are numerous, each requires a separate approach. To date, physically based schemes suitable for GCMs have been developed for GWs excited by convection (Chun & Baik 2002; Beres et al. 2004), flow over topography (McFarlane 1987), and fronts (Charron & Manzini 2002). In most other modeling studies, spectra at a source level are prescribed based on observational constraints, or simply tuned to obtain desired simulated fields. A comprehensive comparison of GW fluxes in observations and modeling has recently been performed by Geller et al. (2013). Although many GCMs use time-independent source spectra, GW excitation can undergo large changes during transient events, such as SSWs. Therefore, the importance of such variations should be explored and their possible impacts on the general circulation have to be taken into account in whole atmosphere GCMs.

In the middle atmosphere, the main mechanism of GW obliteration is nonlinear breaking and/or saturation that occurs when amplitudes become large. Therefore, most GW parameterizations developed for middle atmosphere GCMs (starting from that of Lindzen (1981)) have in common that they terminate harmonics, whose amplitudes reach a certain instability threshold. Exceptions are the approaches of Hines (1997) (“Doppler spread”) and Medvedev & Klaassen (1995) (“nonlinear diffusion”), which sought to describe the underlying physics. The former is based on the assumption that harmonics are Doppler shifted by varying wave-induced wind directly to very short scales where they are removed by molecular diffusion. When averaged over wave phases, this parameterization, however, yields the very same termination of harmonics employing ad hoc chosen criteria. The approach of Medvedev & Klaassen (1995) is based on the concept of “enhanced diffusion” (Weinstock 1976; Weinstock et al. 2007). It takes into account Doppler shift by larger-scale harmonics in the spectrum, and erosion by shorter-scale ones. For parameterization purposes, Doppler shift can be neglected, the coefficient of eddy-induced diffusion is self-consistently cal-
GW parameterizations suitable for thermosphere GCMs must account also for damping by molecular diffusion, thermal conduction, and ion friction. This is usually done by incorporating the respective dissipation terms into the complex dispersion relation in the form of imaginary parts of frequencies. The first parameterization of this kind has been proposed by Matsuno (1982), and the most recent derivation for molecular diffusion and thermal conduction has been performed by Vadas & Fritts (2005). This approach is based on the assumption that dissipation is relatively weak, where the degree of “weakness” depends on the characteristics of the harmonic and the background flow. This constitutes another limitation on GW parameterizations. Molecular viscosity grows exponentially with height in the thermosphere, and, eventually, the dissipation terms can significantly exceed all other balancing terms in the equations for waves. This means that GWs degenerate into other types (“viscous waves”), and can no longer be considered within the parameterization framework.

We illustrate the principles outlined above, and discuss some general details of implementation into a GCM using the extended GW parameterization (Yiğit et al. 2008).

2.3. The extended nonlinear spectral gravity wave parameterization

The word “extended” denotes that the parameterization has been extended to account for wave propagation in the thermosphere in accordance with the requirements outlined above (Yiğit & Medvedev 2013). It solves the equation for the vertical structure of the horizontal momentum flux (per unit mass) $\overline{ww'}$ associated with the harmonic $j$ from a given spectrum of waves:

$$\frac{d\overline{ww'}_j}{dz} = -\left(\frac{\rho_j}{\rho} + \beta^j_{tot}\right)\overline{ww'}_j,$$

(1)

where $\beta^j_{tot}$ is the total vertical damping rate acting on the harmonic. If propagation is conservative ($\beta^j_{tot} = 0$), then the flux $\rho u w' w'_j$ is constant with height. The total damping rate for a given harmonic is the sum of the rates due to various dissipation processes affecting the propagation and acting simultaneously

$$\beta^j_{tot} = \beta^j_{non} + \beta^j_{mol} + \beta^j_{ion} + \beta^j_{rad} + \beta^j_{eddy} + \ldots$$

(2)

The main processes accounted for by the scheme include, correspondingly, nonlinear breaking/saturation ($\beta^j_{non}$), molecular diffusion and thermal conduction ($\beta^j_{mol}$), ion friction ($\beta^j_{ion}$), radiative damping ($\beta^j_{rad}$), and eddy diffusion ($\beta^j_{eddy}$) as suggested in the work by Yiğit et al. (2008). The term $\beta^j_{ion}$ is parameterized after the work by Medvedev & Klaassen (2000), and comprises the effects of other harmonics on a given harmonic. Thus, the total wave field is not a simple collection of independent waves, but of interacting ones. The word “nonlinear” in the name of the parameterization signifies this. Dissipation of a harmonic is strongly affected by changes in the background wind as the vertical damping is inversely proportional to the intrinsic phase speed of the harmonic, i.e., $\beta^j \propto (c_j - u)^{-n}$, where the exponent $n$ differs for various dissipation mechanisms (see e.g., Yiğit et al. 2008, 2009, 2012a; Yiğit & Medvedev 2013). If the flux $\rho u w' w'_j$ changes with height, the wave momentum is transferred to the mean flow by means of an acceleration or deceleration, which is often called “wave drag”

$$a_j = -\frac{1}{\rho} \frac{d\rho u w' w'_j}{dz}.$$
Miyoshi et al. (2014)’s recent simulations with a whole atmosphere GW-resolving GCM have confirmed Yiğit et al. (2009)’s predictions of the appreciable dynamical effects of lower atmospheric GWs on the general circulation of the thermosphere above the turbopause. Figure 2 presents the divergence of momentum fluxes (a in equation 3) due to the resolved portion of GW spectra (with horizontal scales longer than 380 km) calculated for solstice conditions (Miyoshi et al. 2014, Figure 3) as in the GCM modeling by Yiğit et al. (2009). Considering the various approximations and limitations of the extended parameterization, and, especially, uncertainties with specifying GW sources, the two distributions in Figures 1 and 2 appear to be in a good qualitative and quantitative agreement. There are also some differences between the two simulations. In particular, in the Southern Hemisphere MLT, the high-resolution simulations show a region of eastward GW, which is only present at the Southern Hemisphere high- and low-latitudes in the parameterized simulation. Two possible sources of the discrepancies are the source spectrum and effects of the background winds on the propagation and the resulting dissipation. Overall, both simulation studies demonstrated that, due to propagation conditions in the middle atmosphere, most of the thermospheric GW activity concentrate at high-latitudes, where solar tides modulate local time variations of GW drag. This, and further analyses of the simulations with the high-resolution model provided evidences that thermospheric effects of GWs can be successfully parameterized in lower-resolution GCMs.

Thermal effects of GWs are two-fold: (a) heating due to conversion of the mechanical energy of dissipating harmonics into heat, and (b) heating and cooling associated with the downward sensible heat flux $w^T$ induced by these waves (Medvedev & Klaassen 2003; Becker 2004). Magnitudes of the former in the thermosphere are comparable with those due to the Joule heating, while the latter is comparable with the cooling rates due to molecular thermal conduction (Yiğit & Medvedev 2009), which suggests that the thermal effects of GWs cannot be neglected in the upper atmosphere. Yiğit & Medvedev (2010)’s GCM simulations with the extended scheme have demonstrated that the variations of thermospheric GW effects are appreciable. GWs propagate to higher altitudes during high solar activity, but produce weaker drag than during periods of low solar activity. Their observations have later been qualitatively verified by the satellite observations of Park et al. (2014).

4. SUDDEN STRATOSPHERIC WARMINGS

4.1. Characteristics

Sudden stratospheric warmings first discovered observationally by Scherhag (1952) are transient events during which the eastward zonal mean zonal winds weaken, or even reverse the direction at 60°N (geographic) at ~ 30 km (10 hPa), followed by the significant warming of the winter North Pole (90°N) (Labitzke 1981; Andrews et al. 1987). Since the 1950s, as the interest in studying SSWs has grown, the classification of SSWs has evolved (see Butler et al. 2015, for a comprehensive discussion). Essentially, there are two commonly accepted types of warmings: a minor and a major warming. The warming is major if the equator-to-pole temperature gradient reverses poleward of 60° latitude in addition to the reversal of the zonal mean zonal winds at 60° N at 10 hPa (Labitzke 1981). If the westerly mean zonal wind weakens but does not reverse the direction, i.e., the stratospheric
Gravity Waves during Sudden Warnings

4.2. Mechanism of the sudden warming

In the winter (solstice) period, the Northern Hemisphere stratosphere is dominated by westerly jets whose strength increases with altitude. Quasi-stationary planetary waves can propagate vertically upward, provided that the mean zonal flow satisfies the conditions for vertically propagating wave modes. For these waves, the zonal wind has to fulfill the following condition (Holton & Hakim 2012, Equation (12.16)):

$$0 < \bar{u} < u_c,$$

where the Rossby critical velocity $u_c$ is defined in terms of the characteristics of the background atmosphere and wave by

$$u_c \equiv \beta \left( k_h^2 + \frac{f_0^2}{4 N^2 H^2} \right),$$

where $k_h^2 = k^2 + l^2$ is the horizontal wavenumber that depends on the zonal $(k = 2\pi/\lambda_x)$ and the meridional $(l = 2\pi/\lambda_y)$ wavenumbers; $f = f_0 - \beta y$ is the beta-plane approximation for the Coriolis parameter, and $\beta = \frac{\partial f}{\partial y}$ is the meridional gradient of the Coriolis parameter. The condition (4) suggests that planetary waves can propagate vertically only in the presence of westerly winds that are weaker than a certain critical value $u_c$, which depends on the horizontal scale of the wave. Dynamical conditions are, therefore, favorable for the vertical propagation of planetary waves in the winter Northern Hemisphere with prevalent mean westerly winds. This condition is important for understanding the propagation of GWs, which are also affected by the mean wind distributions. Namely, before the warming, the stratospheric zonal mean winds are eastward. They filter out a significant portion of the eastward directed GWs, favoring the upward propagation of harmonics with phase velocities directed westward. During the warmings, the decelerating westerlies increase the chances of GWs with eastward horizontal phase speeds to propagate to higher altitudes (Yigit & Medvedev 2012).

In the winter stratosphere, waves are rapidly attenuated, thus decelerating the mean zonal flow. For the occurrence of SSWs, a large-scale wave transience, in particular, rapid temporal changes of planetary wave activity are also important. They maintain the convergence of the westward momentum flux, leading to strong polar night jet deceleration and poleward meridional flow enhancement (Andrews et al. 1987). Additionally, radiative forcing sustains a cold winter North Pole with negative equator-to-pole mean temperature gradient, that is, $\frac{\partial T}{\partial y} < 0$. The rapid deceleration of the stratospheric mean flow implies a decreasing (positive) vertical gradient of the zonal flow between the troposphere and stratosphere. From the thermal wind relation $\frac{\partial u}{\partial z} \sim -\frac{\partial T}{\partial y}$, this decrease implies a rise of temperature at the winter pole, meaning that the equator-to-pole mean temperature gradient becomes less negative. During a major warming, this gradient even reverses due to the reversal of the vertical gradient of zonal mean wind. The strong polar night jet deceleration leads to a departure from the thermal wind balance, and the poleward meridional flow, which is caused by the Coriolis force associated with the westward forcing, is induced to recover this balance. This enhancement of the Brewer-Dobson circulation ultimately results in an adiabatic warming at Northern Hemisphere high-latitudes.
5. OBSERVED CHANGES IN THE UPPER ATMOSPHERE DURING SUDDEN STRATOSPHERIC WARMINGS

Given the rapid and strong local changes in the circulation and thermal structure of the stratosphere during SSWs, the natural questions that bear in mind are (1) how high the effects of the warming propagate in altitude, and (2) to what extent the changes in the upper atmosphere can be associated with the sudden warmings. Planetary waves cannot propagate directly to much higher altitudes, but the stratosphere and mesosphere are closely connected via circulation and by GWs and tides. As sudden warmings and the associated dynamical changes in the stratosphere occur over relatively long time scales (e.g., ~ 10 days) compared to the periods of internal waves, lower atmospheric wave disturbances have sufficient time to propagate to higher altitudes, provided that propagation conditions are favorable. Therefore, one ought to expect a certain degree of coupling between the stratosphere and higher altitudes, probably beyond the middle atmosphere.

How can one associate observed upper atmospheric changes with SSWs? Essentially, a ground-to-upper atmosphere observation with a single instrument is beyond the capabilities of the current technology. For the purposes of observational analysis, SSW events/periods ought to be identified. For this, an appropriate description of stratospheric dynamics is needed in the first place. This representation could be, for example, obtained from numerical forecast models that assimilate in-situ and remote-sensing data, such as the European Centre for Medium-Range Weather Forecast (ECMWF) analyses, and produce the required global fields of atmospheric parameters. Then, observational data can be investigated together with the numerical model output (e.g., Pancheva et al. 2008).

The deceleration of the stratospheric eastward zonal flow during sudden warmings leads, ultimately, to an upward circulation in the mesosphere that results in mesospheric cooling (Liu & Roble 2002). Such direct link between these two regions has motivated a number of scientists to investigate the details of stratosphere-mesosphere changes during warmings. Based on temperature and geopotential height data obtained from the Sounding the Atmosphere using Broadband Emission Radiometry (SABER) instrument of the Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) satellite and the VHF radar horizontal winds, Pancheva et al. (2008) have investigated planetary wave-induced coupling in the stratosphere-mesosphere during the major warming of 2003/2004 winter Northern Hemisphere.

Yuan et al. (2012) studied the response of the middle-latitude mesopause region to the 2009 major SSW, using a sodium Doppler wind-temperature lidar. They have discovered anomalous behavior of the mean temperature and zonal winds around the mesopause during the warming, and have concluded that it was due to a direct impact of the major warming on the middle-latitude mesopause. The 2009 SSW has been one of the strongest warming events that has been recorded. The features around the mesopause during SSWs can be largely characterized in terms of an “elevated stratosopause”, which forms around 75–80 km after the SSW occurrence and then descends (Manney et al. 2009). The role of GWs and planetary-scale waves in the time evolution of the elevated stratosopause have been investigated by a number of authors (e.g., Siskind et al. 2010; Chandran et al. 2011; Limpasuvan et al. 2012).

Vertical coupling between the stratosphere and the lower thermosphere has been studied in the low- and middle-latitude Northern Hemisphere winter of 2003/2004 based on the temperature data from SABER/TIMED (Pancheva et al. 2009). According to Goncharenko & Zhang (2008)’s analysis of the Millstone Hill incoherent scatter radar (ISR) ion temperatures data, warming in the lower thermosphere and cooling above 150 km were revealed during a minor SSW. Using data from the Jicamarca ISR, Chau et al. (2009) have detected significant semidiurnal tidal variations in the vertical E × B ion drifts in the equatorial ionosphere during the winter 2007-2008 minor warming. Using temperature measurements from the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) on board European Space Agency’s (ESA) Envisat satellite measurements, Funke et al. (2010) have demonstrated observational evidence for the dynamical coupling between the lower and upper atmosphere during the 2009 major SSW. Based on TEC (total electron content) data retrieved from a worldwide network of GPS observations, Goncharenko et al. (2010) have found a significant local time modulation of the equatorial ionization anomaly (EIA) induced by SSWs. Using the European Incoherent Scatter (EISCAT) UHF radar, Karihar et al. (2010) have detected short-term variations in the upper atmosphere during the 2009 major warming. In their analysis of Fabry-Perot and incoherent scatter radar data, Conde & Nicolls (2010) have identified that the period of reduced neutral temperatures at 240 km, which corresponded closely to the main phase of the warming.

More recently, analyzing the Constellation Observing System for Meteorology Ionosphere and Climate (COSMIC) data, Pancheva & Mukhtarov (2011) have found a systematic negative response of ionospheric plasma parameters (f0F2, h0F2, and n0) during the warmings of 2007–2008 and 2008–2009. An illustration of their results for the mean zonal mean electron density (in MHz) at 300 km are seen in Figure 4, where the 2007–2008 and 2008–2009 warming events are shown on the left and right panels, respectively. The response to the warming is negative and mainly occurs in the low- and middle-latitude region. Liu et al. (2011) used neutral mass density observations from the CHAMP (Challenging Minisatellite Payload) and GRACE (Gravity Recovery and Climate Experiment) satellites to study the thermospheric variations during the 2009 major warming. They have found a substantial decrease of the mass density, and concluded that it was potentially associated with a strong thermospheric cooling of about 50 K. Goncharenko et al. (2013) have investigated the day-to-day variability in the middle-latitude ionosphere during the 2010 major warming using the Millstone Hill ISR. They have discovered that semidiurnal and terdiurnal tidal variations were enhanced during the SSW. Jonah et al. (2014) have used a suit of observational data from GPS, magnetosphere, and meteor radar in order to investigate the response of the magnetosphere and ionosphere to the 2013 major SSW. Analyzing long-term data of the global average thermospheric total mass density derived from satellite orbital drag, Yamazaki et al. (2015) showed density reduction of 37 % at 250575 km during SSW period that can be associated with lower atmospheric forcing. Recently, using data from GPS and ionosonde stations, Fagundes et al. (2015) investigated the response of the low- and middle-latitude ionosphere in the Southern Hemisphere to the 2009 major SSW and found that during the warming, TEC was depressed following the SSW temperature peak.

Overall, these studies suggest that (1) a variety of instru-
Gravity Waves during Sudden Warnings

Figure 4. SSW event and ionospheric variations. The upper panels show the 2007/2008 (left) and 2008/2009 (right) SSWs according to the UK Met Office (UKMO) model. In the lower panels, the associated variations of the mean zonal mean electron density (in MHz) retrieved from COSMIC at 300 km during the two warmings are seen. Adopted from Pancheva & Mukhtarov (2011, Figures 1, 2).

6. UPPER ATMOSPHERIC CHANGES DURING SUDDEN STRATOSPHERIC WARMINGS

Observing dynamical changes, e.g., with satellites and radars, cannot provide sufficiently detailed information on characteristics and physical mechanisms of vertical coupling. Observations may, and in fact do raise new questions, which can be investigated by models. A powerful tool is general circulation models (GCMs) that solve the coupled governing equations of atmospheric and ionospheric physics in time and three-dimensional space. GCMs generate a full set of field parameters that can be diagnosed in detail in order to investigate the physical mechanisms that shape the state and evolution of the atmosphere. Therefore, global models can provide an unprecedented insight in vertical coupling processes between the different atmospheric regions. One should nevertheless be aware of the limitations of GCMs, such as, resolution, boundary conditions, and parameterizations.

As discussed in section 3, GWs have a profound effect on the dynamical (Yiğit et al. 2009, 2012a; Miyoshi et al. 2014; Vadas et al. 2014), thermal (Yiğit & Medvedev 2009; Hickey et al. 2011) and compositional (Walterscheid & Hickey 2012) structure of the upper atmosphere. The state of the background middle atmosphere plays a crucial role in modulating the propagation of GWs into the thermosphere. Given that SSWs modify the middle atmospheric circulation significantly, how can they influence the upper atmosphere in the context of GW propagation and dissipation in the thermosphere? Resolving this science question requires the use of comprehensive GCMs with appropriate representation of small-scale GWs. The general circulation modeling of Yiğit & Medvedev (2012) using the extended GW parameterization of Yiğit et al. (2008) has demonstrated GW propagation and appreciable dynamical effects in the upper thermosphere during a minor warming. The universal time (UT) variations of the GW-induced zonal mean root-mean-square (RMS) wind fluctuations (in m s⁻¹) and zonal mean GW drag (in m s⁻¹ day⁻¹) are shown in Figures 5a,b. GW-induced RMS wind fluctuations are given by \( \sigma = \left( M^{-1} \sum_j M \mathbf{u}_j^2 \right)^{1/2} \), where \( M \) is the number of harmonics in the spectrum and variance \( \mathbf{u}_j^2 \) is related to the wave amplitude as \( \sqrt{\mathbf{u}_j^2} = |\mathbf{u}_j| \). The GW RMS wind fluctuations are a proxy for the subgrid-scale GW activity as the fluctuations induced by all waves in a GW spectrum are taken into account and do not represent the resolved wind fluctuations. In the course of the warming, GW activity increases by a factor of three exceeding 6 m s⁻¹ in response to weakening of the polar night jet. In addition to persistently large values in the lower thermosphere, modulation of the GW activity is seen higher in the thermosphere. Following the increase of GW activity, (eastward) GW drag increases in the thermosphere during the warming as well.

The effects of GWs in the upper atmosphere during SSWs are not confined to only those in a zonal mean sense. Recently, Yiğit et al. (2014) have investigated the details of GW temporal variations in the thermosphere during a minor warming simulated with a GCM. They modeled that GW drag and
its variability is dramatically enhanced in the thermosphere during the warming and thus lead to a $\sim \pm 50\%$ modulation of small-scale and short-term variability in the resolved thermospheric winds, where the small-scale variability has been evaluated by subtracting the contributions of the large-scale tides. Recently, Miyoshi et al. (2015) have demonstrated with a GW-resolving GCM that the SSW has major dynamical and thermal impact on the upper atmosphere, substantially influencing the global circulation. Changes in the mean zonal wind produce a feedback on GWs by modifying filtering, dissipation, and propagation conditions.

The upper atmosphere above the turbopause has a great amount of variability owing to the simultaneous influences of meteorological and space weather processes (Matsuo et al. 2003; Anderson et al. 2011; Yiğit & Ridley 2011b; Yiğit et al. 2012b). Often, separating the components and sources of variability in observations is a challenging task. Thus, following their observations of an SSW, Kurihara et al. (2010) have concluded that understanding the link between SSWs and thermal and dynamical changes in the upper atmosphere-ionosphere requires investigations of GW-mean flow interactions processes. GCM studies can greatly supplement these efforts.

Predictions with GCMs indicate that small-scale GWs can substantially contribute to the variability of the upper atmosphere. Also, recent modeling studies with a whole-atmosphere GCM have shown an enhancement of the semidiurnal variation in the ionospheric $E \times B$ drifts during the 2009 major warming (Jin et al. 2012). This increase has been interpreted as a consequence of the semidiurnal tidal amplification in the lower atmosphere. Further investigations that incorporate a fully two-way coupled thermosphere-ionosphere under the influence of lower atmospheric waves are required in order to assess the significance of the lower atmosphere in the context of upper atmosphere variability. In characterizing the upper atmosphere processes, the variability is always defined with respect to some appropriate mean. Therefore, the quantity of variability is not uniquely defined, and care should be taken when comparing one study to another. In a broader context, the presence of any kind of variability restricts scientists' ability to predict the future state of the atmosphere, and it is crucial to determine the sources of variability and quantify the magnitude thereof.

7. CONCLUSIONS

This paper has briefly reviewed the current state of knowledge and most recent developments with understanding the role of GWs in vertical coupling during SSWs. The observed upper atmosphere changes during SSWs have been presented. An emphasis was placed on the processes above the mesopause, and on how they can be studied with general circulation models.

The geosciences community increasingly recognizes that the effects of lower atmospheric gravity waves extend beyond the middle atmosphere into the atmosphere-ionosphere system and are of global nature. Similarly, sudden stratospheric warmings were used to be looked upon as stratospheric phenomena, but now compelling observational evidences of their signatures in the thermosphere-ionosphere are being routinely provided.

With the rapid progress in the field of atmospheric coupling, further key science questions on the role of GWs in coupling atmospheric layers arise:

- What are the spectra of gravity waves in the lower and upper atmosphere? How do they change during the different phases of SSWs?
- How well do GW parameterizations describe wave spectra and reproduce their effects during SSWs?
- What is the relative role of GW- and electrodynamical coupling between atmospheric layers in the variability of the atmosphere-ionosphere system during SSWs?
- What are the effects of GWs on the circulation and thermal budget of the upper atmosphere during major sudden stratospheric warmings?
- Do GWs in the upper atmosphere affect the development of sudden stratospheric warmings, or they are a mere reflection of processes in the lower atmosphere?
- Do GWs have a role in latitudinal coupling in the thermosphere during SSW events?

This is certainly an incomplete list of scientific questions, answering which requires concerted observational, theoretical,
Gravity Waves during Sudden Warnings

and modeling efforts from scientists of both lower and upper atmosphere communities.

COMPELLING INTERESTS

The authors declare that they have no competing interests.

AUTHOR’S CONTRIBUTIONS

Both authors have participated in writing all sections, read, and approved the final manuscript.

The work was partially supported by German Science Foundation (DFG) grant ME2752/3-1. EY was partially supported by NASA grant NNX13A036G and NSF grant AGS 1452137.

REFERENCES

Anderson, C., Davies, T., Conde, M., Dyson, P., & Kosch, M. J. 2011, J. Geophys. Res., 116
Andrews, D. G., Holton, J. R., & Leovy, C. B. 1987, International geophysics series, Vol. 40, Middle Atmosphere Dynamics (Academic press)
Becker, E. 2004, J. Atmos. Sol.-Terr. Phys., 66, 683

... (continued)
Yiğit, E., Ridley, A. J., & Moldwin, M. B. 2012b, J. Geophys. Res., 117, doi:10.1029/2012JA017596

Yuan, T., Thurairajah, B., She, C.-Y., et al. 2012, J. Geophys. Res., 117, doi:10.1029/2011JD017142