Oblique Gravity Wave Propagation During Sudden Stratospheric Warmings

C. C. Stephan¹, H. Schmidt², C. Zülicke³, and V. Matthias³

¹Max Planck Institute for Meteorology, Hamburg, Germany, ²Leibniz Institute of Atmospheric Physics, Kühlungsborn, Germany, ³Potsdam Institute for Climate Impact Research, Potsdam, Germany

Abstract  Gravity waves (GWs) are important for coupling the mesosphere to the lower atmosphere during sudden stratospheric warmings (SSWs). Here, a minor SSW is internally generated in a simulation with the upper-atmosphere configuration of the ICOsahedral Nonhydrostatic model. At a horizontal resolution of 20 km the simulation uses no GW drag parameterizations but resolves large fractions of the GW spectrum explicitly, including orographic and nonorographic sources. Consistent with previous studies, the simulated zonal-mean stratospheric warming is accompanied by zonal-mean mesospheric cooling. During the course of the SSW the mesospheric GW momentum flux (GWMF) turns from mainly westward to mainly eastward. Waves of large phase speed (40–80 m s⁻¹) dominate the eastward GWMF during the peak phase of the warming. The GWMF is strongest along the polar night jet axis. Parameterizations of GWs usually assume straight upward propagation, but this assumption is often not satisfied. In the case studied here, a substantial amount of the GWMF is significantly displaced horizontally between the source region and the dissipation region, implying that the local impact of GWs on the mesosphere does not need to be above their local transmission through the stratosphere. The simulation produces significant vertically misaligned anomalies between the stratosphere and mesosphere. Observations by the Microwave Limb Sounder confirm the poleward tilt with height of the polar night jet and horizontal displacements between mesospheric cooling and stratospheric warming patterns. Thus, lateral GW propagation may be required to explain the middle-atmosphere temperature evolution in SSW events with significant zonally asymmetric anomalies.

1. Introduction

Sudden stratospheric warmings (SSWs) are marked by a sharp temperature increase of several tens of kelvins in the middle to upper stratosphere (30–50 km) over the course of several days (Butler et al., 2015). Warmings are commonly classified as “minor” when the zonal-mean 10-hPa meridional temperature gradient between 60°N and 90°N reverses, and as “major” when in addition the zonal-mean 10-hPa zonal wind at 60°N reverses (Schoeberl, 1978). The interaction of planetary waves with the polar vortex is the well-established mechanism for the initiation of SSWs (Matsuno, 1971).

Gravity waves (GWs) are the major contributor to the mesospheric cooling that has been observed to accompany stratospheric warming (Quiroz, 1969; Labitzke, 1972). Polar cap temperatures from the Aura Microwave Limb Sounder (MLS) averaged north of 60°N show a joint occurrence of a warm stratosphere (10 hPa) and a cold mesosphere (0.01 hPa) in 71% of major warmings in 2004–2015 (Zülicke et al., 2018). The causal chain for this stratosphere-mesosphere coupling is well understood: under undisturbed winter conditions GWs propagate through stratospheric westerlies, which filter the GW pseudo-momentum flux (GWMF) to be predominantly westward in the mesosphere. The westward drag induces a northward and downward mesospheric residual circulation, which dynamically controls the winter stratopause (Holton, 1983). In contrast, when stratospheric winds weaken or change to easterly directions, the mesospheric GW forcing results in less warming or even a cooling. During the recovery phase of SSWs, significant downward transport of nitric oxide from the mesosphere to the stratosphere has been observed (Randall et al., 2006, 2009; Hauchecorne et al., 2007), with consequences for the reduction of stratospheric ozone (Bailey et al., 2014; Pérot et al., 2014).

The focus of this study is on the role of GWs in shaping mesospheric temperature variations during SSWs. We analyze a minor SSW in a simulation with a 20-km horizontal resolution and a model top at 150 km. At this...
resolution a substantial fraction of the GW spectrum is resolved. Therefore, we turn off any additional GW parameterizations to isolate the effects of resolved GWs, accepting that this choice precludes any tuning of the circulation by adjusting the parameterized GW drag.

Early theoretical work showed that some aspects of stratospheric warming and mesospheric cooling during SSWs can be explained based on the Langrangian-mean vertical motion related to planetary waves alone (Matsuno & Nakamura, 1979). Simulations of the 2002 warming with the NOGAPS-ALPHA model produced mesospheric cooling without a GW drag parameterization, but the spatial structure of the cooling improved when mesospheric drag in the form of Raleigh friction was turned on (Coy et al., 2005). Previous modeling studies have emphasized the importance of GWs for mesospheric variability during SSWs. The relative importance of eastward and westward GWMF for coupling the mesosphere to the lower layers of the atmosphere depends on the dynamics of the SSWs and the time of the warming cycle (Siskind et al., 2010). In simulations with the Whole Atmosphere Community Model (WACCM), parameterized nonorographic GWs were key for determining the mesospheric circulation during the recovery phase of a SSW, which featured an elevated stratopause (Ren et al., 2011). Tests with the Canadian Middle Atmosphere Model (CMAM; McLandress et al., 2013) and the Hamburg Model of the Neutral and Ionized Atmosphere (HAMMONIA; Meraner et al., 2016) also emphasized the importance of parameterized nonorographic GWs at times when stratospheric westerlies were weak and hindered the vertical propagation of orographic GWs into the mesosphere. TIME-GCM/CCM3 model simulations of self-generated SSWs with mesospheric cooling were analyzed by Liu and Roble (2002). They found GW-induced circulation changes in the mesosphere and lower thermosphere to have substantial impacts on the abundance of atomic oxygen. They pointed out, however, that the parameterization of GWs introduces some uncertainties, for instance, in the wave breaking altitudes, as GW specifications are often tuned so that the model results agree with reference climatologies.

A numerical study with resolved GWs was performed by Zülicke and Becker (2013). They used an advanced parameterization of turbulence to permit an explicit description of GWs in the Kühlungsborn Mechanistic Circulation Model (KMCM) at a T120 spectral truncation. At this resolution the shortest resolved horizontal wavelength is about 333 km. Thus, their simulation mainly included waves from the stormtrack, as convection was not parameterized. Zülicke and Becker (2013) found the mesosphere to respond without much delay to changes in the stratosphere, which they explained with a fast upward propagation of GWs from the stratosphere. Furthermore, they hypothesized that a quick mesospheric response may help explain why general circulation models with parameterized as well as resolved GWs can successfully simulate the stratosphere-mesosphere coupling during SSWs. This would suggest that a simplified treatment of GWs including the assumption of straight upward propagation, as is common in parameterizations, is sufficient for correctly simulating the stratosphere-mesosphere coupling during SSWs, at least in the zonal-mean sense.

The notion of stratosphere-mesosphere coupling through GWs usually relies on a zonal-mean view. However, it is well known that substantial zonal asymmetries exist in the propagation and selective transmission of GWs during SSWs (Dunkerton & Butchart, 1984). For instance, the magnitude and occurrence of GWs correlate with the location and strength of the polar vortex in reanalysis data, lidar and satellite observations (Yamashita et al., 2010). Moreover, Yamashita et al. (2013) argued that poleward propagation of GWs from low latitudes is important for GW variations during SSWs, even though, unlike our study, they focused on the recovery phases of elevated stratopause events. Their ray tracing results are based on observed background conditions but homogeneous distributions of sources, as there was no information available on the true GW sources. Oblique propagation over such long distances implies that the negligence of the former is not only problematic for parameterizations in high-resolution climate models but also in coarse models.

This study uses a state-of-the-art global high-top model with a sufficiently fine resolution to explicitly resolve a large fraction of the GW spectrum from orographic and nonorographic sources. By analyzing a simulated minor SSW we provide evidence that oblique propagation, which is typically neglected in GW parameterizations, is responsible for a substantial fraction of the mesospheric GWMF. Section 2 introduces the simulations, observational data and analysis methods. Section 3 presents the results, and section 4 contains a discussion and conclusions.
2. Data and Methods
2.1. Numerical Simulations

We performed a set of five free-running simulations initialized on 1 October in 1979, 1980, 1981, 1982, and 1983, respectively, and integrated until 30 April of the following year. As the simulated atmospheric evolution deviates relatively soon from the observed evolution, simulated SSWs are independent of any observed SSWs of these winters. None of the simulations generated a major SSW but all produced one or two minor warmings. For this study we selected the simulation initialized in 1983 because it had a particularly sharp and temporally well-isolated midwinter minor SSW.

All simulations are performed with the upper-atmosphere extension of the ICOsahedral Nonhydrostatic (ICON) general circulation model. ICON is based on a nonhydrostatic dynamical core and an icosahedral-triangular grid (Zängl et al., 2015; Giorgetta et al., 2018). The upper-atmosphere configuration extends the dynamical core from shallow- to deep-atmosphere dynamics and contains an additional upper-atmosphere physics package. The deep-atmosphere dynamics account for the spherical shape of the atmosphere and include the contribution of the horizontal component of the Earth's angular velocity to the Coriolis acceleration. A detailed description and model evaluation for the R2B4 grid (horizontal mesh size of about 160 km) can be found in Borchert et al. (2019). Our simulations use the R2B7 grid, which has a horizontal mesh size of about 20 km, with 180 vertical levels extending from the surface to the model top at 150 km. A Rayleigh damping layer (Klemp et al., 2008) starts at 120 km, well above the levels analyzed here. All simulations are initialized from analysis files of the Integrated Forecasting System developed at the European Centre for Medium-Range Weather Forecasts with an extrapolation to climatology above 75 km. Climatological boundary conditions are those described in Giorgetta et al. (2018). The integration time step is 30 s, and all physics schemes are called at this frequency, except for the radiation schemes, which are called every 30 min.

Averaged December zonal-mean temperatures of the 1983 simulation, as well as their averages over all simulations, compare reasonably well with the COSPAR International Reference Atmosphere (Fleming et al., 1990) available from the ftp server at /sparc/ref_clim/fleming (Figure 1). Simulated zonal winds are compared to the URAP climatology (Swinbank & Ortland, 2003) available online (https://www.sparc-climate.org/data-centre/). The summer mesopause in the simulations is on average too cold by 17 K compared to the 135.6 K in the COSPAR International Reference Atmosphere climatology and the polar night jet is located too far north. In 1983, however, the latter bias is particularly pronounced due to the meteorological situation leading up to the SSW, which began in December and involved a strengthening and poleward propagation of the polar night jet prior to the breakdown of the vortex at the beginning of January. Some of these deficiencies might be improved if we used a GW drag parameterization, but, as mentioned above, we deliberately refrain from this to isolate the effects of resolved GWs. Certainly, the circulation simulated for December 1983 to January 1984 is sufficiently close to the natural range of atmospheric variability to permit a study of the processes associated with the SSW.

2.2. Wave Analysis of Model Data
2.2.1. Computation of GW Pseudomomentum Flux

We compute the resolved GWMF using two methods, both of which are applied to instantaneous output every 6 hr. The computationally more expensive small-volume few wave decomposition technique S3D (Lehmann et al., 2012) is applied at two altitude levels. It fits three-dimensional sinusoidal waves to perturbation temperatures that are subset into small cubes centered at 50 and 80 km, respectively, and yields the net GWMF as well as local information on the full wave vector. The other method is based on wind and temperature quadratics (WTQ) and only provides the absolute GWMF without information on the spectral properties of the GWs. The latter technique is applied at all levels between 20 and 80 km.

S3D has previously been applied to ICON data in Stephan et al. (2019a, 2019b), where its underlying assumptions, benefits and drawbacks are discussed. Our choices for the free parameters of the method closely follow Stephan et al. (2019b). Thus, as the method is already extensively described in the literature, we only give a summary of the numerical steps:

- **Regridding.** Six-hourly three-dimensional winds and temperatures are horizontally interpolated from the triangular ICON grid to a Gaussian n512 longitude-latitude grid, corresponding to a resolution of ~20 km at the equator. The data are then vertically interpolated to geometric height levels with a 500-m spacing.
Figure 1. Average December conditions: zonal-mean December-mean (a, c, and e) temperatures and (b, d, and f) zonal wind from (a) the CIRA climatology, (b) the URAP climatology, (c, d) the ICON simulation analyzed in this study (initialized on 1 October 1983), and (e, f) an average over five ICON simulations that were initialized on 1 October of 1979, 1980, 1981, 1982, and 1983, respectively.

- **Background removal.** Temperature perturbations are derived by removing zonal waves of wavenumbers up to 12.

- **Splitting data into cubes.** The S3D algorithm performs three-dimensional sinusoidal fits to temperature perturbations inside small cubes. The cubes are 15 km deep and centered at 50 and 80 km, respectively. They span 13 grid points in the zonal direction and 7 grid points in the meridional direction, such that at 60°N a cube is approximately quadratic in its horizontal dimensions. The cubes are placed to create an overlap between neighboring cubes of five grid points in the zonal direction and three grid points in the meridional direction.

- **Sinusoidal fit.** Inside each cube we carry out two consecutive sinusoidal fits to find the leading two waves, which are defined as the two waves that explain the largest fraction of the variance. The most important aspect to note is that the vertical and horizontal wavelengths derived from these fits are not bound by the cube dimensions. Instead, S3D can retrieve wavelengths up to 11 times the horizontal cube size, even under the unfavorable conditions of spatially homogeneous, full spectra (Lehmann et al., 2012). Each sinusoidal fit yields the waves amplitude $\tilde{T}$ and wave vector $(k, l, m)$. The GWMF is then computed according to linear wave theory (cf. Ern et al., 2004) as
Here, $g$ denotes the gravitational acceleration, $N$ the buoyancy frequency, and $T_0$ the background temperature. For all waves we assume upward propagation.

Typical horizontal wavelengths identified by S3D range from about 250–1,800 km for the simulation considered here. These bounds correspond to the 10th and 90th percentiles, respectively. Typical vertical wavelengths are 4–20 km.

The WTQ GWMF is computed as (cf. Geller et al., 2013)

$$ F_x^2 + F_y^2 = \rho_0^2 \left( 1 - \frac{f^2}{\hat{\omega}^2} \right)^2 w'^2 \left( u'^2 + v'^2 \right) \left[ 1 + \frac{f^2}{\hat{\omega}^2} \right]^{-1}, $$

where $(u', v', w')$ are defined as deviations from an average over the surrounding 50 $\times$ 50 grid points in the zonal and meridional direction. The intrinsic frequency $\hat{\omega}$ is calculated according to

$$ \frac{f^2}{\hat{\omega}^2} = \left( \frac{fgw'N^2}{T_0'} \right)^2 \left( \frac{T'}{T_0} \right)^2. $$

In a comparison of global GWMFs derived from WTQ and S3D, Stephan et al. (2019b) found good agreement between the two methods. For the simulation analyzed here we compared WTQ and S3D maps at 50 and 80 km (not shown) to confirm the good match between the methods.

2.2.2. Ray Tracing

We perform backward ray tracing of mesospheric GWs to determine their possible pathways from the stratosphere (40 km) into the mesosphere (80 km), using the three-dimensional ray tracing scheme for internal GWs formulated by Hasha et al. (2008). The scheme accounts for wave refraction and horizontal propagation in spherical geometry. The wave vector at the end of a ray is taken from the S3D analysis at 80 km. The ground-based frequency and phase speed, respectively, are derived from the wave vector using the gravity wave dispersion relation. We trace the waves in spatially and temporally varying background conditions by interpolating from the 6-hourly model output to the position of the ray. Horizontal variability at scales shorter than 500 km is first removed from wind and temperature fields to filter out short-lived small features from the instantaneous model output. Waves dissipate when the background wind speed is equal to the wave’s phase speed. Unless a ray encounters such a critical level, we trace it back in time until it reaches 40 km. The actual origin of a ray could be anywhere along the calculated path, but we assume that most rays originate below 40 km.

2.3. Observational Data

As we will demonstrate below, the zonal-mean mesospheric cooling is strongly related to the zonal-mean stratospheric warming in the simulation analyzed here. The horizontal signatures of these respective temperature anomalies, however, exhibit substantial longitudinal offsets. To underpin the credibility of these “twists” we analyze data from the MLS onboard the Aura satellite, which reveal that similar twists can occur in nature as well.

Zonal-mean MLS data have been used by Manney et al. (2009) for a comprehensive documentation of the major SSW of 2009. Lee et al. (2009) studied the vertical structure of SSWs from 2005 to 2009 based on MLS data by identifying the Northern Annular Mode at heights up to 0.001 hPa. Zülicke et al. (2018) relied on 60–90° polar cap temperatures from MLS at 10 and 0.01 hPa to examine the coupling between the stratosphere and the mesosphere.

Here, we attempt to resolve zonal and meridional structures in MLS Level 2 Version 4 temperatures (Livsey et al., 2015; Waters et al., 2006), available online (https://mls.jpl.nasa.gov/). The vertical resolution changes from 4 km at 10 hPa to 14 km at 0.01 hPa, while the respective temperature precision changes from ±0.6 to ±3.3 K. We analyze 2004–2018 pressure level maps of daily temperature with a resolution of 5° latitude and 10° longitude. This spatial sampling reduces gaps due to insufficient coverage. One part of the observational analysis is based on maps at 10 hPa (the stratospheric level) and 0.01 hPa (the mesospheric level). To derive anomalies, we first compute for each longitude separately the 60–90° area-weighted polar cap average, then eliminate the seasonal cycle based on multiyear seasonal averages, as detailed in
3. Results

The minor SSW in ICON is marked by a rapid warming of the stratosphere accompanied by a simultaneous cooling of the mesosphere, as can be seen from zonal-mean temperatures at 80°N (Figure 2a). It thus presents a case of strong coupling between the stratosphere and the mesosphere. The high stratopause before the SSW is likely related to a model bias: the height of the stratopause increases with latitude to the north of 60°N (Figures 1c and 1e). Beginning around 20 December, the warm stratopause and the cold mesopause descend at similar rates until the maximum warming is reached just below 50 km during the first couple of days in January. At this time the coldest layer in the mesosphere is situated at about 80 km. After the first week of January temperatures start to recover, although the stratopause does not regain its previous sharp definition. The SSW cannot be classified as major, as the zonal-mean zonal wind at 60°N and 10 hPa does not reverse (not shown). One may expect this from Figure 2a, which shows only a weak warming below 10 hPa. Nevertheless, the 10-hPa meridional temperature gradient reduces rapidly by about 20 K in 10 days and even reverses its sign in the first week of January (Figure 2b).
Figure 3. Horizontal structure of the SSW event: 5-day means of (a, b) temperatures (colors) and geometric height at (a) 1 hPa and (b) 0.01 hPa; (c, d) S3D zonal GWMF (colors) and zonal wind (dashed: negative; solid: positive) at (c) 1 and (d) 0.01 hPa. East is to the right. The region over the pole where S3D GWMF was not computed is shaded gray in (c) and (d). The orange circles in the middle panel of the top row mark the latitudes 80°, 70°, 60°, and 50°. Longitudes are shown in red.
The following analysis focuses mainly on 50 km (corresponding to about 1 hPa) and 80 km (corresponding to about 0.01 hPa), as these levels present the strongest temperature anomalies and are thus useful for studying the coupling between the stratosphere and the mesosphere. We will mainly discuss the three episodes marked in Figure 2b: the pre-warming phase (20–24 December), the peak-warming phase (30 December–3 January), and the recovery phase (9–13 January). The polar vortex at 1 hPa shifts off the pole into the 180–90° W sector during the peak-warming phase 30 December to 3 January (Figure 3a). This displacement of the vortex is associated with a 1-hPa zonal wind reversal near the pole (Figure 3c). The 1-hPa zonal GWMF is predominantly westward with notable zonal asymmetries in magnitude; peaks are located near westerly wind maxima. GWMF magnitudes diminish with a weakening of the 1-hPa westerlies. During 30 December to 3 January weak eastward GWMF is present where 1-hPa winds have turned easterly near the pole.

In the mesosphere warm temperatures are present over the pole in the pre-warming phase 20–24 December (Figure 3b). Close to the pole zonal winds are westerly, and zonal GWMF is westward. Mesospheric cooling occurs at midlatitudes in the east, away from the displaced polar vortex in the stratosphere. In the same area mesospheric winds have also reversed and zonal GWMF is eastward during 20–24 December. During 25–29 December stratospheric zonal winds have reversed at midlatitudes and the mesospheric cooling in the east is stronger than during 20–24 December (not shown). These anomalies are qualitatively consistent with the simplified picture of vertical GW propagation that is adopted in many parameterizations: the upward propagation of eastward propagating GWs into the mesosphere is no longer inhibited by westerly stratospheric winds. Owing to the Coriolis force, eastward GW drag in the mesosphere results into upward motion, which cools the mesosphere. An immediate consequence is the well-established stratosphere-mesosphere coupling during SSWs where a warming in the stratosphere is accompanied by simultaneous cooling in the mesosphere.

The rays in Figure 4 illustrate the great spatial extent that GWs may travel before reaching the mesosphere. We traced backward the GWs present during 30 December to 3 January at 80 km in 45–0° W, 70–80° N, which corresponds to the peak in eastward GWMF (Figure 3d). Figure 4 shows 200 rays traced backward to z ≥ 40 km. Note that the tracing is performed in time-varying background conditions, not using the average fields shown in Figures 3c and 3d. The rays shown in Figure 4 must be interpreted with caution, as uncertainties exist in the initial wave vector, taken from S3D, as well as in the tracing itself (Hasha et al., 2008). Nevertheless, results suggest that GWs are able to propagate substantial meridional and zonal distances.

Figures 5c and 5d show maps of stratospheric and mesospheric temperature anomalies observed by MLS. The SSMC days have been selected from the respective time series of polar cap temperature anomalies (Figures 5a and 5b; see section 2.3 for definitions). On 30 January 2010 the area of 0.01-hPa mesospheric cooling is shifted about 90° westward relative to the warming in the stratosphere at 10 hPa, whereas a shift in the opposite direction occurred on 6 January 2015. Thus, patterns of mesospheric cooling do not always...
Temperature anomalies observed by the MLS during SSW events in (a, c) 2010 and (b, d) 2015. (a, b) Zonal means of temperature anomalies at 10 hPa (solid red lines) and 0.01 hPa (dashed blue line); (c, d) maps. The thin vertical lines in (a) and (b) indicate days of strong stratosphere-mesosphere coupling (SSMC days). Please refer to the text for further definitions.

Considerable changes in the direction of mesospheric GWMF at 90°E to 90°W during the SSW do not only manifest in the spatial distributions (Figure 3d); they are even more pronounced in phase-speed propagation-direction spectra of the mesospheric GWMF (Figure 6). Prior to the SSW, GWMF is westward (20–24 December) north of 50°N. Southeastward GWMF becomes dominant at 40–50°N during 20–24 December. With weakening stratospheric winds across all latitudes, eastward GWMF also emerges in the spectra at 50–60°N and 60–70°N during 30 December to 3 January. During this peak-warming phase mesospheric GWMF is predominantly eastward at 60–80°N, as expected from Figure 3d.

The spectra of Figure 6 reveal that GWs with large phase speeds (~45–85 m s⁻¹) account for most of the eastward GWMF during the peak-warming phase 30 December to 3 January. At this time the GWMF associated...
Figure 6. Temporal evolution of mesospheric GWMF spectra in the sector 90°E to 90°W: S3D-derived GWMF as a function of propagation direction (east is to the right, north is at the top) and ground-based phase speed (bins of 5 m s⁻¹; spider web circles mark 30, 60, and 90 m s⁻¹, respectively). Columns are averaged over the 5-day intervals given at the top and rows are averaged over the latitude bands given on the left.

with phase speeds <30 m s⁻¹ is considerably weaker compared to the less disturbed states of the polar vortex, for example, 20–24 December or 9–13 January, when the largest GWMF magnitudes are associated with slow waves. This important result implies that during the peak of the SSW nonorographic waves dominate over orographic waves in the mesosphere, in line with previous results (Liu & Roble, 2002; Ren et al., 2011; Yamashita et al., 2013).
Based on the ray tracing results of Figure 3b and the spectra of Figure 6, we hypothesize that mesospheric GWs at 70–80°N during the peak phase of the SSW originate from the midlatitudes. This idea is supported by Figure 7, which shows zonal winds in three 45° zonal sectors between 0 and 135°W for the different phases of the warming. From 5 December (not shown) onward, the axis of the polar night jet assumes an increasingly stronger tilt toward the pole with height, while the jet core moves northward from about 57°N (5–9 December) to about 70°N in the sector 315–360°E 0–45°W (20–24 December). This evolution is qualitatively very similar in the three zonal sectors. This entire time, the strongest 20-km GWMF is centered at about 40–60°N. These latitudes coincide with the preferred location of the meandering 20-km zonal wind maximum (Figure 8). Moreover, 20-km GWMF closely follows these wind maxima, suggesting that the GWMF in Figure 7 stems from sources along the Northern Hemisphere storm track. Thus, these GWs may be generated by convection as well as by imbalanced flow, which is consistent with high ground-based phase speeds. To support this hypothesis, we trace the rays shown in Figure 3b back to an altitude of 20 km. Twenty-nine percent of the rays encounter a critical level above 20 km. Thirteen percent of the rays that reach 20 km originate north of 80°N, 42% between 60°N and 80°N, 31% between 40°N and 60°N, and 14% south of 40°N, which is consistent with the above-mentioned origins in Figure 8.
During the peak-warming phase 30 December to 3 January and above 20-km altitude, GWMF maxima are meridionally confined to a narrow band along the jet axis. The local GWMF maxima in the vertical direction imply meridional and zonal wave propagation. During 30 December to 3 January the westward jet has diminished and easterlies appear in the mesosphere and upper stratosphere near the pole at 45–135°W. Winds below 30 km, however, remain westerly, which would inhibit the upward propagation of orographic waves. Yet, waves with sufficiently high eastward phase speeds (>40 m s$^{-1}$) would not encounter any critical levels if they propagated, for instance, from 40 km at 60°N to 80 km at 80°N, which is consistent with the GWMF distribution in Figure 7 and the ray paths in Figure 3 and would serve as an explanation for the high eastward phase speeds in Figure 6.

To quantify the importance of oblique GW propagation we next consider GWMF spectra computed for 90° longitude sections at 40–80°N during 30 December to 3 January (Figure 9). The spectra labeled “simulated” at 80 km show the familiar eastward GWMF, particularly over the Atlantic sector at 270–360°E, where the strongest cooling occurs (Figure 9b). The stratospheric GWMF spectrum at 270–360°E contains a substantial
Figure 9. Importance of lateral GW propagation: (a) the row “simulated” shows S3D-derived GWMF spectra averaged over the period 30 December to 3 January and 60–80° N as a function of propagation direction (east is to the right, north is at the top) and phase speed (bins of 5 m s⁻¹; spider web circles mark 30, 60, and 90 m s⁻¹, respectively). The columns show different longitudinal sections; “vertical” shows the portion of the simulated spectra that is able to propagate vertically from z = 20 km; that is, this portion would not encounter critical levels; “difference” shows the differences between simulated and vertical and the missing percentage of the integrated GWMF. (b) As in (a) but for 80 km.
fraction of eastward GWMF, but westward GWMF still dominates (Figure 9a). This is consistent with weak westerly winds at 50 km (Figure 3c and 7). Interestingly, Figure 9a shows that GWMF at 180–270°E is predominantly northward, which is consistent with the rays in Figure 4.

Under the assumption of instantaneous upward propagation from an altitude of 20 km, substantial fractions of the GWMF spectra would be erased by critical level filtering. To compute the spectra for vertical propagation in Figure 9, labeled “vertical,” we remove those parts of the simulated spectra that would encounter a critical level between 20 and 50 or 20 and 80 km, respectively, under the assumption of instantaneous upward propagation and using background wind profiles averaged over 10° × 90° latitude-longitude. In doing so, a separate calculation was performed every 6 hr, based on the spectra and wind field at each time. Thus, between 20 and 50 km as well as between 20 and 80 km, critical level filtering would mainly affect eastward propagating waves. As a result up to 40% of the integrated GWMF, calculated as the difference between simulated and vertical spectra at 50 km, is estimated to be due to oblique propagation. In the mesosphere filtering would have the largest impact at 270–360°E. Thus, most of the eastward GWMF likely stems from sources that are not vertically below the peaks of the mesospheric GWMF. On average over 50% of the GWMF at 80 km are likely due to oblique propagation.

4. Discussion and Conclusion

In this study we analyzed a minor SSW in a upper-atmosphere ICON simulation with a 20-km horizontal resolution and no parameterized GW drag. During the course of the SSW the mesospheric GWMF turns from mainly westward to mainly eastward. In addition, the contribution of slow phase speed waves to the total GWMF decreases dramatically and waves of phase speeds >40 m s⁻¹ dominate the eastward GWMF during the peak phase of the warming. The strong colocation of GWMF maxima with the local maxima in the zonal wind (Figure 7) suggests that GWs from the storm track converge into the polar night jet, as was previously pointed out by Yamashita et al. (2013).

We calculated that a large fraction of the GWs must have traveled along slanted paths to reach the middle atmosphere if their sources are in the lower stratosphere or the troposphere. In the mesosphere at 80 km this oblique propagation accounts locally for up to 60% of the total GW pseudo-momentum flux (GWMF; Figure 9). Ray tracing supports that the GWs present at 80 km in the locations of the strongest mesospheric cooling may have propagated substantial zonal and meridional distances (Figure 3).

Our simulation suggests that lateral propagation of nonorographic waves from the midlatitudes is relevant for creating local mesospheric temperature extremes. This is not contradicting the more usual zonal-mean perspective on SSWs. Yet it is evident, at least in the case of our simulation and the observed situations in Figures 5c and 5d, that these local anomalies contribute strongly to the zonal-mean cooling. The role of localized GW sources in triggering SSWs by preconditioning the polar vortex was previously discussed by Ern et al. (2016), who also argued that the zonal-mean view of SSWs may be insufficient to understand the dynamics of SSW events, as GW variability interacts strongly with the zonal and meridional background wind.

On the one hand, GW propagation is affected by the planetary wave background. On the other hand, GW drag may in turn modulate the background. With regard to the zonally shifted anomalies in the stratosphere and mesosphere it is not obvious which direction of this mutual interaction is more important. Additional sensitivity studies would be necessary to address the problem of cause and effect. A study that examines the statistics of observed cases of zonally shifted temperature anomalies in MLS data is underway. A better understanding of zonal asymmetries may help with the interpretation of results from radar soundings that indicated a reversal of the mesospheric local zonal wind well before zonal-mean anomalies appeared in the stratosphere (Hoffmann et al., 2007). High-resolution simulations with high model tops and explicitly resolved GWs, as used in this study, provide a useful tool for studying the dynamics of explicitly resolved GWs under such conditions.

References


