Influence of the spatial distribution of gravity wave activity on the middle atmospheric dynamics

Petr Šácha\textsuperscript{1}, Friederike Lilienthal\textsuperscript{2}, Christoph Jacobi\textsuperscript{2}, and Petr Pišoft\textsuperscript{1}

\textsuperscript{1}Department of Atmospheric Physics, Faculty of Mathematics and Physics, Charles University, V Holesovickach 2, 180 00 Prague 8, Czech Republic
\textsuperscript{2}Institute for Meteorology, University of Leipzig, Stephanstr. 3, 04103 Leipzig, Germany

Abstract. Analysing GPS radio occultation density profiles, we have recently pointed out a localised area of enhanced gravity wave (GW) activity and breaking in the lower stratosphere of the east Asian–northwestern Pacific (EA/NP) region. With a mechanistic model of the middle and upper atmosphere, experiments are performed to study the possible effect of such a localised GW breaking region on large-scale circulation and transport and, more generally, a possible influence of the spatial distribution of gravity wave activity on middle atmospheric dynamics.

The results indicate the important role of the spatial distribution of GW activity for polar vortex stability, formation of planetary waves and for the strength and structure of zonal-mean residual circulation. Furthermore, a possible effect of a zonally asymmetric GW breaking in the longitudinal variability of the Brewer–Dobson circulation is analysed. Finally, consequences of our results for a variety of research topics (e.g. sudden stratospheric warming, atmospheric blocking, teleconnection patterns and a compensation mechanism between resolved and unresolved drag) are discussed.

1 Introduction

Consideration of gravity wave (GW)-related processes is necessary for a proper description and modelling of the middle (as reviewed comprehensively by Fritts and Alexander (2003) and upper atmospheric dynamics (see, e.g., the review by Smith, 2012). However, only recently satellite and other observational data sets with improved resolution and novel analysis methods together with high-resolution global models have been tightening the constraints for the parameterisations that can improve the treatment of these waves in climate models (Alexander and Shepherd, 2010; Geller et al., 2013). Complex understanding and unbiased modelling of middle atmospheric conditions is vital for climate research and there is strong evidence that coupling between chemistry and dynamics in the stratosphere is essential for surface climate variability and climate change in both hemispheres (Manzini, 2014; Calvo et al., 2015). There is also a wide recognition of dynamical links between the stratosphere and troposphere with a potential to significantly affect conditions at the surface (Haynes, 2005; Kidston et al., 2015). Hence, better representation of the stratosphere could improve the long-range as well as short-range forecast skills (Hardiman and Haynes, 2008; Gerber et al., 2012).

The Brewer–Dobson circulation (BDC) was discovered based on a distribution of trace gases by Brewer (1949) and Dobson (1956). Using transformed Eulerian-mean equations, Dunkerton (1978) derived the first dynamically consistent two-dimensional (2-D) picture of the mean-transport streamlines for the middle atmosphere, which is often used as a basic BDC concept. However, Demirhan Bari et al. (2013) found the 3-D structure of circulation in the middle atmosphere to be in good correspondence with tracer fields, especially in relation to the zonal wave-1 pattern observed in the stratosphere and mesosphere, although their study did not give a comprehensive dynamical explanation of the discovered circulation structures (enhanced downward branch of BDC over northeastern Asia, wave-1 pattern).

Planetary waves (PWs) are usually thought to be created in the troposphere and then vertically propagated into the...
middle atmosphere. The theoretical possibility of PW creation by zonally asymmetric GW breaking was first numerically analysed by Holton (1984) and later on, e.g., by Smith (2003) and Oberheide et al. (2006), and experimentally verified by Lieberman et al. (2013). There is building agreement in the literature on the role of wave activity in preconditioning sudden stratospheric warming (SSW; e.g. Ayarzagüena et al., 2011) events.

SSW belongs to the most pronounced atmospheric phenomena, as they cause abrupt changes of the middle atmospheric circulation and tracer distribution, and they also affect tropospheric weather patterns (e.g. Manney et al., 2009; Kuroda, 2008; Lehtonen and Karpechko, 2016). SSW dynamics and their impact differ whether a split or a displacement of the stratospheric polar vortex takes place (Sevour et al., 2016). It has been observed that displacements are connected with dominating wave-1 activity, whereas vortex splits correlate with stronger wave-2 activity (e.g. Kuttippurath and Nikulin, 2012). Generally, most attention is paid to the role of upward propagating PWs in preconditioning SSW (Hoffmann et al., 2007; Nishii et al., 2009; Alexander and Shepherd, 2010).

The two open questions regarding the dynamics of SSW are what types of wave phenomena are responsible for the SSW triggering and what are the necessary basic state conditions. There are two main triggering theories discussed – anomalous tropospheric upward wave fluxes or non-linear resonance in connection to the vortex geometry (Albers and Birner, 2014). Furthermore, there is growing observational evidence that GW amplitudes are enhanced prior to SSW (Ratnam et al., 2004; Wang and Alexander, 2010; Yamashita et al., 2010), and GWs are recognised as playing an important role in a wide range of SSW-related processes (e.g. mesospheric cooling, stratopause separation and recovery; Dunkerton and Butchart, 1984; Richter et al., 2010; Limpasuvan et al., 2012; France and Harvey, 2013; Chandran et al., 2013; Siskind et al., 2010; Albers and Birner, 2014). Recently, Ern et al. (2016) showed that for both polar vortex event types, the GW activity distribution displays strong longitudinal structure and day-to-day variation. Ern et al. (2016) have also found an indication for a dynamical role of the poleward propagating mid-latitude GWs and they suggest the need for global models to reflect the non-vertical propagation and to include physical GW sources that are as realistic as possible.

However, the majority of studies are focussed on the modulation of GWs by PWs (e.g. Cullens et al., 2015) and on the GW impact on the upper stratosphere–mesosphere region. Šácha et al. (2015) indicated a possible GW breaking in the lower stratosphere. Indeed, model experiments with gravity-wave drag (GWD) parameterisation showed that orographic GWD in the lower stratosphere can significantly affect the development of SSW (Pawson, 1997; Lawrence, 1997) and the large-scale flow in the lower stratosphere and troposphere in general (McFarlane, 1987; Alexander and Shepherd, 2010; Sandu et al., 2016).

McLandress et al. (2012) found changes of PW drag resulting from artificial enhancements of the orographic GW sources in the parameterisation. This was called a compensation process and was further statistically confirmed by Cohen et al. (2014), who interpreted it as a response of the resolved waves to maintain a “sensible” stable circulation. Such a response is expected, since all processes in the atmosphere are driven by the tendency to reach an energetically more favourable, stable state. In addition to the stability constraint, Cohen et al. (2014) proposed two additional mechanisms using a potential vorticity (PV) concept, PV-mixing and refractive index interaction.

In this study, we focus on the physical mechanism and structure of the atmospheric response to zonally asymmetric forcing represented by an artificially injected GWD in the stratosphere. We are following Šácha et al. (2015), who described a localised area of enhanced GW activity and breaking in the lower stratosphere over the east Asian–northwestern Pacific (EA/NP) region and discussed possible implications of this GW hotspot for large-scale dynamics and transport. By artificially enhancing the GWD in a 3-D mechanistic circulation model of the middle atmosphere, we examine the hypothesis that such a robust breaking region plays a role in forcing the longitudinal variability of the BDC and can generate PWs. Further, we investigate possible implications of the GWD distribution and particular forcing components (direction of the force) for polar vortex stability.

The structure of the paper is as follows: in Sect. 2 we describe the model and sensitivity simulation set-up together with the observational motivation and justification for an artificial GWD enhancement.

The section closes with a brief description of tracer data used in the Supplement to this study. Section 3 starts with an illustration of the geopotential response to different GWD injections with particular focus on effects in the polar region. We also present the dynamical impact, structure and modes of PWs generated by the artificial GWD. Finally, we show the differences of the BDC due to the geometry of the GWD modulation and analyse the 3-D residual circulation spatial patterns in relation to the GWD distribution. In Sect. 4 we give a summary of our results, discuss potential implications of our findings and outline future directions of our work.

2 Data and methodology

2.1 Model description and configuration

We use the middle and upper atmosphere model (MUAM), which is a non-linear 3-D mechanistic global circulation model. It has a horizontal resolution of $5^\circ \times 5.625^\circ$ and 56 layers in the vertical extending to an altitude of about 160 km in log-pressure height (Pogoreltsev et al., 2007). At
1000 hPa, at the lower boundary of the model, we prescribe stationary PWs of wave numbers 1, 2 and 3 obtained from decadal monthly mean ERA Interim (ERAi) temperature and geopotential reanalysis data (ECMWF, 2016). Up to an altitude of 30 km, the model zonal-mean temperature is nudged to ERAi zonal-mean temperature. This is necessary because MUAM does not feature a detailed accounting of the troposphere, including orography and radiation processes (3-D water vapour, surface albedo, etc.). However, the troposphere is necessary for stationary PW forcing and the generation and propagation of travelling PWs and tides, and therefore it cannot be neglected. The assimilation of stationary PWs and zonal-mean temperatures is not only active during the spin-up of 330 model days but also during the 30-day analysis period. The effect of nudging during the analysis period is dependent on the strength of the artificial forcing. In the reference simulation the nudging effect is lower than 1 K day$^{-1}$ everywhere, and in the simulation with the strongest forcing, its zonal mean locally reaches magnitudes of around 2 K day$^{-1}$ (as shown in Fig. 1c). Because in MUAM simulations only the zonal-mean temperatures are nudged to the zonal mean, nudging has no direct effect on the wave structure of the response to the forcing, but it is likely to reduce the magnitude of the zonal-mean response.

The time step of the model is 225 s following a Matsuno (1966) integration scheme. For simulations, the model starts with a globally uniform temperature profile and no wind. During a spin-up period, the mean circulation is built, and PWs and tides are generated. After that, a time interval of 30 model days with a temporal resolution of 2 h is analysed. Since the lower boundary conditions are taken as a decadal-mean January mean (mean value of all January data in 1 decade), this interval refers to an average January climatological state. Monthly zonal means of wind, temperature and GWD are given in Fig. 1. Owing to the constant forcing with time in the lower atmosphere, the standard deviation of temperature within these 30 days is smaller than 3 K near the stratopause and mesopause, and smaller than 1 K elsewhere. The standard deviation of the zonal wind is the largest within the jets reaching 4 m s$^{-1}$ in the summer easterlies. These values do not have a meteorological meaning and are provided here to demonstrate that MUAM has rather small variability within the analysis interval.

GWs are parameterised after a linear Lindzen-type scheme Lindzen (1981) updated as described in Fröhlich et al. (2003) and Jacobi et al. (2006). GWs are initialised at an altitude of 10 km with six different phase speeds ranging from 5 to 30 m s$^{-1}$, each propagating in eight different azimuth angles, and with GW vertical velocity amplitudes with an average value of 0.01 m s$^{-1}$. As input for the GW parameterisation scheme, we modified the GW source function to reflect a distribution based on the mean January field of the potential energy of disturbances computed from FORMOSAT3/COSMIC radio occultation density profiles between the tropopause and 35 km altitude taken from Šácha et al. (2015). The GW weights are calculated from these data by dividing the potential energy at each grid point by its global mean. This set-up has a positive impact on some climatological features in MUAM. Nevertheless, the effect on the horizontal distribution of the GWD in the stratosphere is negligible. We will refer to this set-up as the reference simulation. Zonal (gcu) and meridional (gcv) flow acceleration as well as the heating due to breaking or dissipation of GWs (gt) is calculated by the parameterisation scheme.

To examine and to demonstrate the effect of spatial distribution of the GW activity, we performed a set of sensitivity simulations (Table 1) with artificially changed GWD imposed on the model by modulating the GW parameterisation output. Note that this change of GWD is only added after the spin-up so that only the 30 model days incorporate GWD changes. Thus, the simulation period also includes the temporally delayed response for the adaption from reference conditions to enhanced GWD (gcu/gcv/gt) values. The naming convention (Table 1) is given by “Gcu + distribution + gcv”, where the basic value of gcu of 0.5 m s$^{-1}$ day$^{-1}$ is not stated.

The enhancement is performed for a certain 3-D box in the lower stratosphere (about 18–30 km) above the EA/NP region (37.5–62.5°N, 112.5–168.8°E), according to the area of enhanced GW activity described by Šácha et al. (2015). This refers to the “box” distribution in Table 1 (an example is shown in Fig. 2, left panel). There are no exceptional GWD values in the reference simulation in this region. In a second version we additionally averaged the respective GWD parameters in the same latitude range as the box but across all longitudes. This way we obtain a zonally uniform distribution, i.e. a ring of enhanced GWD parameters instead of a box but with a smaller local magnitude. We refer to this configuration as ring or “Zon” simulations (see Table 1). For all simulations, the GWD parameters outside the box or the ring, respectively, remain unchanged and are not influenced by the enhancement. We are not smoothing the boundaries of the artificial enhancement area and the step between artificial and background GWD values is dependent on the horizontal location, the time step and, most importantly, the altitude level. To illustrate the sudden and localised effect of GW breaking, we have chosen to enhance the GWD in our simulations stepwise and rather abruptly. As suggested by Cohen et al. (2013), such a sharp change (as at the boundaries of our enhancement) leading to dynamic instabilities is likely to induce compensation processes.

Although it is impossible to directly compute the GW drag force from current satellite measurements alone (Alexander and Sato, 2015), Ern et al. (2011) gave a methodology to estimate absolute values of a “potential acceleration” caused by GWs (maximum zonal-mean values of 3 m s$^{-1}$ day$^{-1}$ below 40 km). Using ray-tracing simulations, Kalisch et al. (2014) gave an estimate of around 20 m s$^{-1}$ day$^{-1}$ for a zonally averaged GWD in the lower stratosphere. In our model simulations we are injecting three values of the artificial
Figure 1. Mean January zonal means of temperature (a), zonal wind (b), GW-induced heating (d) GW-induced zonal wind acceleration (e) and meridional wind acceleration (f) for the reference simulation. Additionally, mean January zonal-mean nudging strength for the strongest GWD injection (SSWbox simulation in Table 1) is shown (c).

Figure 2. Two examples of the GWD enhancement horizontal distribution imposed between approx. 20 and 30 km of log-pressure height. Left panel: box distribution (Box0.1 simulation). Right panel: ring distribution (Zon0.1 simulation). Colours indicate GW-induced zonal acceleration [m s$^{-1}$ day$^{-1}$]. Depending on the GW type and on the direction of background winds, the GWD also has a meridional component, which is usually poorly constrained by observations. We performed simulations with three different values of meridional GW-induced acceleration ($-0.5$, $-0.1$, $0.1$ m s$^{-1}$ day$^{-1}$). The direction of the zonal and meridional GW-induced accel-
erations were chosen according to the prevailing directions of horizontal winds in the EA/NP region in January (see Šácha et al., 2015) assuming that the majority of GWs in the EA/NP region are of orographic origin (in January). On this basis we argue that the 5 : 1 ratio between the zonal and meridional GW-induced acceleration is the most realistic and therefore we choose the Box0.1 (and Zon0.1) simulation as a representative conservative enhancement for most of the analyses in this paper. A comprehensive discussion of our sensitivity simulation set-ups is given in the Discussion section.

2.2 Residual circulation

To highlight the importance of the stratospheric research in the EA/NP region, we present in the Supplement the 1978 to 2008 average total ozone January mean distribution from the ozone Multi-Sensor Reanalysis version 1 (MSR1; van der A. et al., 2015) data (Temis, 2016). Additionally, in the Supplement, we illustrate the similarity in the vertical structure and longitudinal variability of the model residual circulation and of zonal cross sections of the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) methane volume-mixing ratio profiles (KIT, 2016; see von Clarmann et al., 2009; Plieninger et al., 2015). However, the comparison of the distributions must be done with care, since the tracer distributions result from several different processes in the atmosphere, namely advective transport, mixing and chemical reactions (Garny et al., 2014). Also, the residual velocities are closely related to Lagrangian-mean velocities up to O(α²) only for small amplitude α steady waves (Bühler, 2014).

In the Sect. 3.3, we study consequences of the GW hotspot for the longitudinal variability of the residual circulation (and BDC consequently) by means of the time mean 3-D residual circulation according to Kinoshita and Sato (2013). The time averaging inserts additional uncertainty in the 3-D residual circulation concept. Unlike Demirhan Bari et al. (2013), who based their analysis on monthly means and daily eddies, we are employing a 5-day-running average on the 6 h MUAM output fields. This configuration gives the strongest zonally averaged Stokes drift from several choices of the running mean, but it is still smaller (not shown) than the value of the Stokes drift resulting from transformed Eulerian-mean equations, which is computed in this study according to Hardiman et al. (2010) for log-pressure height vertical coordinate models.

3 Results

To establish the timescales of the response, in Fig. 3 we show Hovmöller diagrams of the zonal-mean zonal wind and its variance. The time evolution is presented at the 6.25 hPa level (around 35 km log-pressure height, 13th model level). This level was chosen for our analysis because it is above the location of the artificially modified area and above the nudging extent; therefore, it contains the atmospheric response only. In Fig. 3a, a Hovmöller diagram is given for the zonal-
mean zonal wind at 6.25 hPa level in the reference simulation documenting that the model is essentially steady. Figure 3b shows the time evolution of variance of the zonal-mean wind anomaly (Box0.1—reference) and Fig. 3c shows the time evolution of zonal-mean zonal wind for the 10box simulation. We can see the response building up in the first 7 days after the GWD injection and after that the structure of the response remains quasi-steady, with small variations of the magnitude only.

In contrast to this, the zonal-mean zonal wind time evolution from the SSW simulations (Fig. 4a and 4b) do not reach a steady state in the course of the 30-day simulation and therefore the results based on those simulations are presented at particular time steps or in animations in the Supplement. Results of other simulations (Table 1) are averaged across the quasi-steady state (7th–30th day of the simulation) and are supplemented with the estimate of statistical certainty or standard deviation of the mean.

Except for the SSW simulations, our study is focussed mainly on the mean response to a monthly mean GWD distribution, because from observational analyses we usually have information on the GW activity distribution on a monthly or seasonal basis (Šácha et al., 2015). The short-term response, which would arguably more relevant to the real atmosphere taking into account the intermittency of large-amplitude GWs (e.g. Hertzog et al., 2012; Wright et al., 2013), is not well captured by the mechanism of constant GWD injection, which is discussed in the final section. Still, there are some interesting results concerning the short-term response mentioned in the course of this study, e.g. note the agreement with the timescale of the transient response build-up in Fig. 11d of Cohen et al. (2014), where this timescale has been related to the life cycle of the PW breaking.

### 3.1 Atmospheric response to variations in GWD and SSW

Figure 5a1 shows the mean (7th–30th day) horizontal wind and geopotential field at the 6.25 hPa level (13th model level) for the reference simulation and the remaining plots in the first row show anomalies (i.e. differences from the results of this run) caused by different components of GWD with artificial values corresponding to the Box0.1 simulation. The second row (Fig. 5a2–d2) shows horizontal wind and geopotential anomalies for the 10box (Fig. 5a2) and Box0.1 (Fig. 5b2) simulations and differences between simulations with conservative GWD enhancements (Fig. 5c2 and d2). The third row (Fig. 5a3 through d3) shows the same as the second row, but for the artificial ring GWD configuration. Note the different scaling of the colour bars, which is chosen according to the maximal and minimal value of geopotential (anomaly), so that the labels of the colour bar provide direct information on the magnitude of the differences in geopotential response.

The anomalies and differences are analysed with special focus on the polar vortex response, since it will be shown below that the dynamical response to GWD changes is strongest in the polar region. This comparison demonstrates not only the importance of the role of the longitudinal distribution of the zonal-mean drag force but also highlights an important and different effect of each of the individual GWD components.

From comparison of Fig. 5b1, c1, and d1, we see that among the GWD components modified in the Box0.1 simulation the response to the gcu component is the strongest. It induces a dipole-structured anomaly with negative geopotential anomaly downwind from the region of GWD enhancement and positive anomaly north of this region (Fig. 5b1). The gt component alone induces a positive anomaly of smaller magnitude northward and upstream of the area (Fig. 5c1). In contrast, meridional drag induces a negative geopotential anomaly northward and downwind of the area, which has the smallest magnitude of all three components, but is still significant (Fig. 5d1).

The respective geopotential responses in the Box0.1gcu and the Box0.1gt simulations have almost exactly opposite features, as the positive gt enhances geopotential in the upwind and northward direction from the GWD region, while artificial northward deceleration has an opposite effect. Although we used a non-linear model, the additivity of effects of different GWD components (Fig. 5b1, c1, and d1) seems to hold reasonably well as can be seen from the Box0.1 anomaly (Fig. 5b2), where the forcing is constituted by identical artificially enhanced GWD components. Also, the differences between simulations with different meridional drag (compare Fig. 5c2 and d2) show the same pattern as induced by the meridional drag only (Fig. 5d1). The distribution of the response to the meridional component suggests that a box gcv enhancement in this geographical position can influence the geopotential response in the area of the location of the Aleutian high.

Another two important results are visible from the comparison of the plots in the second and third row of Fig. 5. First, there are much bigger anomalies for the box enhancements (second row) than for the corresponding ring enhancements (third row). This is true locally as well as in the zonal mean (compare the sum of geopotential responses given in the legend for Fig. 5a2, b2, a3 and b3). In the box simulations (Fig. 5a2 and b2), the response is typically dominated by a rather meridionally oriented dipole pattern with a localised positive geopotential anomaly at the centre of the polar vortex and a negative geopotential anomaly at the location of Aleutian high. In the corresponding ring simulations (Fig. 5a3 and b3) the geopotential response is more zonally uniform.

Second, there are large and significant differences (50 or 25 % of the magnitude of the anomaly) between box simulations with slightly different set-ups of the meridional drag (Fig. 5b2 vs. Fig. 5c2 and d2, respectively), whereas this is not true for ring GWD enhancements (few percent; see Fig. 5b3 vs. Fig. 5c3 and 5d3). Unlike the box enhancements,
ring enhancements are almost insensitive to the different versions of GWD in the meridional direction. The difference between Zon0.1, Zon0.5 and Zon0.1pos simulations is very small and not significant.

As noted above, the magnitude of the geopotential response is larger for the box enhancements than for the ring enhancements. For the Box0.1 simulation, the geopotential anomaly at the 6.25 hPa level reaches about 20 gpm in a monthly mean. The horizontal wind anomaly for the Box0.1 simulation (Fig. 5b2) reaches maximal values slightly below 1 m s⁻¹. Anomalies for the 10box simulation (Fig. 5a2; 20 times bigger eastward deceleration than for Box0.1) are almost exactly 20 times stronger and show a very similar dipole pattern. Although locally the difference between these two simulations may seem to be linear, this comparison is misleading, since both simulations (10box and Box0.1) have different ratios between the strength of GWD components. This means, for example, that the drag force has different orientation between these two simulations.

Unexpectedly, the box simulations lead to anomalies that would contribute to weakening rather than amplification of the Aleutian high. Based on the results and discussion of Šácha et al. (2015), who argued that the EA/NP hotspot (high GW activity already in October/November) may play a role in the onset of the winter circulation in the stratosphere in this region, we expected a positive contribution of the GWD response to the background climatology (e.g. contribution to the unusually hot temperatures in the stratosphere in the EA/NP region by induced subsidence).

In Fig. 4a and b, we presented a time evolution of the zonal-mean zonal wind at 6.25 hPa for the SSWbox and SSWzon simulations with signs of a wind reversal at polar latitudes at particular time steps suggesting an occurrence of a minor SSW. We show additional results from the SSWbox and SSWzon simulation in two animations of the geopotential and horizontal wind field response at 6.25 hPa (animation 1a and 1b in the Supplement). Also, in Fig. 6 we show two snapshots of the geopotential and horizontal wind field response at 280 h after the GWD injection to illustrate the situation at a developed stage of the SSW. In response to a strong GWD enhancement in a box we observed a vortex displacement (Fig. 6a and animation 1a in the Supplement), and in response to a strong GWD enhancement in a ring we obtain a vortex-split-like event (Fig. 6b and animation 1b in the Supplement).

In the SSWbox simulation (animation 1a), immediately after the spin-up period when the GWD starts to be artificially modified (injection of GWD), a geopotential ridge begins to form above the northern Pacific (northward from the GWD area). This anomaly strengthens and shifts a little westward above Siberia, where, within approx. 5 days of the GWD injection, we observe an evolution of a pressure high. All the time the vortex is shifting towards the northern boundary of North America, where it stays till the end of the simulation.
Figure 5. Mean geopotential and horizontal wind vectors at the 13th model level (6.25 hPa) for the reference simulation and differences for the sensitivity simulations with different GWD set-up. From top left (index 1a) to bottom right (index 3-D): (a1) reference simulation overlaid with an illustration of the box area, (b1) reference–Box0.1gcu, (c1) reference–Box0.1gt, (d1) reference–Box0.1gcv, (a2) reference–10box, (b2) reference–Box0.1, (c2) Box0.1–Box0.5, (d2) Box0.1–Box0.1pos, (a3) reference–10zon, (b3) reference–Zon0.1, (c3) Zon0.1–Zon0.5 and (d3) Zon0.1–Zon0.1pos. Colours indicate geopotential height (gpm). Note the different scaling of the respective plots. Arrows refer to horizontal wind [m s\(^{-1}\)] with unity arrows given below the individual plots. The statistical significance of the mean geopotential differences was computed by a \( t \) test and regions with \( p \) values < 0.05 are stippled. The sum of geopotential difference across the plotted area is given in the legend to each plot.

In the SSWzon simulation, we observe a slow (compared to the SSWbox simulation) creation of a pressure high above the northern Pacific together with a high-pressure ridge above the northern Atlantic. This pressure high is almost stationary (in contrast to the SSWbox) leading to the vortex split approx. 10 days after the injection. This is a potentially very interesting result suggesting that a symmetric forcing favours vortex split and localised forcing favours displacement events, but the robustness of this claim needs to be tested in future work for various initial vortex states.

For illustration, in Fig. 6 we show the geopotential field and horizontal wind speed 280 h after the GWD injection, when the vortex split develops (Fig. 6b) and the vortex displacement is in its mature state (Fig. 6a). The vortex displacement event develops more quickly, as seen from comparison of Fig. 4a and b or in animations 1a and 1b in the Supplement. However, both events have limited vertical extent, and
do not disturb the entire vortex (only up to 60 km of log-pressure height; not shown).

3.2 Creation of planetary waves and dynamical impact

In this section we compare PW activity and amplitude structure of the leading PW modes between reference, box and ring simulations. We show results of the Eliassen-Palm (E–P) flux diagnostics and Fourier transform (FT) analysis of geopotential anomalies.

Figure 7 shows the mean (7th–30th day) E–P flux and its divergence for the reference simulation (Fig. 7a), Box0.1 simulation anomalies (Box0.1–reference; Fig. 7b), Zon0.1 simulation anomalies (Fig. 7c), the difference between the Box0.1 and Zon0.1 simulations (Fig. 7d) and mean E–P flux and its divergence for the 10box simulation (Fig. 7e) and respective anomalies (Fig. 7f). Note that we show the E–P flux divergence as a force per unit area (units [kg m$^{-1}$ s$^{-2}$]), not as an induced acceleration (units [m s$^{-2}$]), as in Hardiman et al. (2010); otherwise, upper stratospheric and mesospheric effects would dominate the plots due to the density decrease with height. The statistical significance of the mean E–P flux divergence differences has been computed by a $t$ test and regions with $p$ values < 0.05 are stippled.

In Fig. 7b, for the Box0.1 and reference simulation differences, we find an anomalously weak E–P flux convergence (positive difference to the reference simulation) centred at the equatorward flank of the GWD enhancement area and an anomalous convergence in a broad area around 60° N. This pattern is similar for the Zon0.1 simulation anomalies (Fig. 7c), but much weaker and with the anomalous convergence starting more poleward. It is also similar in the 10box simulation anomalies (Fig. 7f), but much stronger in magnitude (approx. 20 times). In all of those simulations, this anomalous pattern is limited in altitude and only slightly exceeds the vertical boundaries of the GWD area (especially in the polar region).

Taking into account the reference E–P flux field (Fig. 7a), the anomalies can be caused by two different mechanisms. The first one is an indirect mechanism, when the artificial GWD drag modifies the winds causing changes (with respect to the reference simulation) in propagation conditions for PWs propagating from below (for more details on the refractive index interaction see Cohen et al., 2014). According to this mechanism, the E–P flux and its divergence anomalies and differences (Fig. 7b, c, d, f) would be associated with a stronger poleward and weaker/stronger upward propagation of PWs in the 10box/Box0.1 simulation along the northern edge the polar night jet in the northern part and northern boundary of the GWD area. The E–P flux divergence anomaly at the southern flank of the GWD would be associated with a suppression of upward and equatorward PW propagation elsewhere (between approx. 40 and 60° N) in the latitudinal extent of the GWD region. But this mechanism fails to explain some features in Fig. 7, e.g. the E–P flux divergence emerging in the E–P flux field in the 10box simulation (Fig. 7e). Therefore, although the changes in the refractive index will definitely be present in the artificial GWD simulations, we have an indication that another mechanism is dominant.

This second mechanism is a stratospheric analogue of the evidence given by Holton (1984) that a zonally asymmetric GW breaking possibly generates PWs in the mesosphere. In the stratosphere the artificial GWD enhancement in a box would cause displacements of fluid particles (in the initially balanced predominantly zonal flow) and thus generate a broad spectrum of waves depending on background
Figure 7. Mean E–P flux vectors (kg s$^{-2}$; arrows are scaled according the relative distances of the plot) and its divergence (colours in kg m$^{-1}$ s$^{-2}$) for Box0.1 (a), its anomalies (b), Zon0.1 anomalies (c), difference between the Box0.1 and Zon0.1 simulation (d), mean E–P flux and its divergence for 10box (e) and its anomaly (10box–reference; f). Note that scales are adjusted for each subfigure, except the plots of anomalies (b, e) sharing the same scaling. In panels (a)–(f) contours of zonal-mean zonal wind from the respective simulations are overlaid with an increment of 10 m s$^{-1}$. All panels are overlaid with selected contour of gravity-wave-induced zonal acceleration to illustrate the location of artificial GWD. The statistical significance of the mean E–P flux divergence differences was computed by a t test and regions with p values < 0.05 are stippled.

Conditions and the geometry of the drag region. We can find support for this mechanism from the E–P flux difference between Box0.1 and Zon0.1 simulations (Fig. 7d).

In the previous section, we have shown that the box enhancement induces a stronger zonal-mean geopotential response than the corresponding ring enhancement. Therefore, we can assume that the first mechanism has bigger effect in the box simulations, which is true for the E–P flux divergence difference (Fig. 7d). However, considering E–P flux vectors, Fig. 7d reveals that there are not only differences in magnitude between Box0.1 and Zon0.1 E–P flux anomalies, but also that the Zon0.1 simulation lacks the horizontal component of the anomalous E–P flux, with biggest differences in the latitudinal band encompassing the artificial GWD area. This latitudinal band is not significant in Fig. 7b, c and d, because the plotted t test results are based on the difference of the E–P flux divergence (not on the magnitude of the E–P flux vector difference). From Fig. 7b, d, f, we see that the anomalous PWs are generated at the southern flank of the GWD area and propagate predominantly northward (with a small downward component), where they cause anomalous convergence between 60 and 80$^\circ$ N.
For the conservative Box0.1 simulation, the anomalies in the E–P flux divergence are about 5 % of the reference values. Zon0.1 E–P divergence anomalies (Fig. 7c) reach only 1–2 % of the reference values, locally. Anomalies of the 10box simulation (Fig. 7f) exhibit the same pattern as Box0.1 anomalies, but the magnitude is much stronger – more than 50 % of the reference E–P flux divergence values. Therefore, we observe an influence of the 10box GWD enhancement also in the mean field in Fig. 7e, where the artificial GWD box demonstrates itself as an E–P flux divergence area on the southern flank of the GWD enhancement region. This is another supporting argument that the box enhancement generates PWs, with further evidence given below.

In Fig. 8, E–P flux diagnostics are presented at particular moments (1 and 5 days) after the GWD injection for the SS-Wbox and SSWzon simulations. The anomalous E–P fluxes in those highly non-linear simulations absolutely overcome the reference fields, so that we can directly observe the generation and propagation of PWs generated by the artificial GWD. However, for these simulations the structure of the E–P flux divergence area changes with time and also the propagation directions of PWs created in this region are time dependent. Therefore, we have chosen to present snapshots from the 1st and 5th day to demonstrate particular features of the box GWD enhancement. For interested readers, the full time evolution is given in animation 2 in the Supplement.

In Fig. 8a, b, one can clearly see the generation of PWs by the box enhancement; 5 days after the GWD enhancement (Fig. 8b), the E–P flux divergence region extends almost over the whole GWD area. Anomalous PWs propagate equatorward, poleward and upward with two major E–P flux convergence regions around 30 °N and between 60 and 80 °N. A total of 1 day after the GWD injection (Fig. 8a), the E–P flux divergence area is located at the southern flank of the GWD and generates horizontally, southward propagating PWs only. In Fig. 8a, in the majority of the GWD region, we can also see the first mechanism (refractive index interaction) being active, as the GWD region influences propagation of PWs propagating from below. This is the most dominant effect of the ring enhancement (Fig. 8c, d), where in the SSWzon simulation we can hardly observe any anomalous PW generation and the dominant effect of this ring enhancement is altering the propagation conditions for the upward propagating PWs from the troposphere. There is a weaker propagation through the GWD region, with deflection of PWs northward and southward at the southern GWD flank.

Further indication of the creation of PWs by the GWD region is provided by the FT analysis of geopotential anomalies at the 6.25 hPa level. FT provides information about the representation of different harmonics in the anomalous wave activity revealed by the E–P flux diagnostic, and about the spatiotemporal distribution of their amplitudes. The mean (7th–30th day) latitudinal structure of reference amplitudes of leading PW modes is given in Fig. 9a, b. Anomalous amplitudes (Box0.1–reference simulation) are presented in Fig. 9c, d and differences from the Zon0.1 simulation are shown in Fig. 9e, f. To quantify the dispersion of the monthly mean differences, the dotted lines show the standard deviations.

The wave-1 geopotential amplitude is anomalously enhanced for a box GWD (Box0.1–reference; Fig. 9c). The amplitude anomaly is positive starting at the northern flank of the artificial GWD (37.5–62.5 °N) and further poleward. The maximum is gained between 70 and 75 °N. Another smaller, but still significant, region of positive wave-1 amplitude anomaly is located around 30 °N south of the GWD. Another smaller negative wave-1 amplitude anomaly lies inside the GWD area. In the Box0.1 simulation, wave-2 (Fig. 9c) has a pronounced negative amplitude anomaly inside the latitudinal belt encompassing the enhancement region. For wave-3 (Fig. 9d), we find positive anomalous amplitudes starting from central latitudes of the GWD region and ending around 80 °N, although inside the GWD region the positive anomaly is locally not significant. There is a negative wave-3 amplitude anomaly starting at the southern flank of the GWD region with the end around 10 °N. The effect on wave-4 amplitudes is almost negligible (Fig. 9d). The ring enhancement in the Zon0.1 simulation has a negligible effect on amplitudes of harmonics, as is visible from the similarity of the Box0.1 anomalies (Fig. 9c, d) and differences with Zon0.1 simulations (Fig. 9e, f). These results suggest that the box GWD enhancement generates preferentially wave-1 and -3 modes in comparison to the reference and also the ring GWD configuration.

Another indication that the PWs are indeed generated by the GWD box enhancement is given in Fig. 10, where the time evolution of the anomalous wave-1 and wave-3 amplitudes is presented. Especially in the first approx. 7 days from the GWD injection, we can observe a slow propagation of anomalous wave-1 (Fig. 10a) and wave-3 (Fig. 10b) amplitudes from the GWD region to the north. For wave-3 this propagation is visible later than for wave-1 (from approx. day 3). The oscillating patterns in Fig. 10 most likely originate from a non-linear interaction between anomalously generated inertia GWS and solar tides (see e.g. Walterscheid, 1981). Those inertia GWS are responsible for propagation of the anomalous wave activity through the Rossby wave critical layer in the tropics, across the Equator, and into the Southern Hemisphere (Fig. 11).

3.3 Residual circulation response

The first row of Fig. 11 shows the mean (over the whole 30 days) residual circulation mass fluxes for the reference simulation and the snapshot at 5 days from the GWD injection for SSWbox simulation on the right. Mean (7th–30th day) anomalies and differences with the respective ring configuration are given in the second and third row for the Box0.1 simulation on the left and 10box simulation on the right. There are some remarkable results visible. First, even
for a conservative drag enhancement (Box0.1 simulation) there are significant (dashed) differences in the magnitude of the residual mass flux between box and ring GWD distribution of up to 3 % in the lower stratosphere (Fig. 11e). For the 10box simulation, the differences reach about 40 % and create a similar pattern as that for the conservative enhancement (Fig. 11f). The largest differences between the two artificial GWD configurations are found poleward from the GWD enhancement region in the altitude range between 20 and 30 km, corresponding approximately to the vertical extent of the area, and are associated with a stronger subsidence north of the enhancement region in the box simulations.

There is a smaller region of significant differences at the southern flank of the enhancement region associated with lesser downwelling in the box simulations. These two regions of significant differences together constitute a butterfly-like pattern in the box-ring differences centred at approx. 45° N (the centre of the enhancement region) and influencing a shallow BDC branch. Taking into account the reference field (see Fig. 11a), we can explain this pattern as being a faster northward advection starting at approx. 45° N and stronger subsidence northward of 60° N. On the other hand, there is less upwelling in the equatorial region (not significant for the Box0.1 simulation) and slower advection from the tropics. The continuity is satisfied through smaller downwelling south of 60° N.

We observe a similar but stronger pattern in the anomalies (Fig. 11c, d), with the mean residual circulation mass flux anomaly reaching up to 5 % for the Box0.1 simulation and
The fact that the mean response of the upper BDC branch is rather weak and for the most part not significant can be explained by the effect of the artificial GWD region acting like an obstacle for northward flowing wind. The GWD enhancement region (Fig. 11b, snapshot for a SSWbox run) is constantly flown around inducing a significant mean anomaly (Fig. 11c, d) with anomalous upwelling in its southern part and downwelling on the northern flank. But, the GWD region (obstacle) also creates a lee-wave-like pattern with oscillating anomalies in the upper stratosphere and in the SH. Considering a time mean, these anomalies are small and not significant, but, at particular time steps, the magnitude of the anomalies is comparable regardless of the BDC branch. Supporting information is given in animation 3 in the Supplement, which presents the time evolution of the zonal-mean residual-circulation-associated mass flux for the 10box simulation (on the left) together with its anomaly (on the right). One can see here the global nature of the response and gain insight into how quickly the residual circulation is affected by the anomalous forcing in the Northern Hemisphere. After few time steps, the response is constituted by a constant anomaly corresponding roughly to an accelerated shallow BDC branch sloping down from approx. 30 km at the North
Figure 10. Time evolution of a wave-1 (a) and wave-3 (b) amplitude difference with respect to the reference run, as given by the FT of geopotential height at approx. 35 km log-pressure height for Box0.1 simulation. Units are given in (gpm).

Pole to the lowest analysed levels at the Equator. Except for this region, the entire domain is dominated by anomalies seemingly descending downward from the mesosphere associated with the obstacle analogy.

The zonal structure of the induced flow, and possible consequences of the GW hotspot for the longitudinal variability of the BDC were studied by means of 3-D residual circulation analysis according to Kinoshita and Sato (2013). In total, a 5-day-running averaging was performed. Šácha et al. (2015) pointed out unusually high temperatures in the EA/NP region at 30 hPa in winter and concluded that there could be an enhanced downwelling above the EA/NP region, which penetrates to levels lower than elsewhere. This is in agreement with Fig. S in Demirhan Bari et al. (2013). Supporting results highlighting the importance of future research in this region are given in the Supplement. In Fig. S1 in the Supplement, we present a 30-year average January MSR total ozone column field with a total ozone column maximum located in the EA/NP region. In Fig. S2, longitudinal cross sections of MIPAS methane volume-mixing ratios indicate a peak of subsidence around 15 km in the EA/NP region (at 140° E) and the interesting massive upwelling branch east of it.

To evaluate the possible role of the GW activity in the longitudinal variability of the BDC, we present longitudinal cross sections of the reference 3-D vertical residual velocity and Box0.1 anomalies going from the northern to southern part of the artificial GWD (Fig. 12). From longitudinal cross sections of the reference vertical residual velocity (left side of Fig. 12), we see that MUAM vertical residual velocity field is dominated by a wave-2 pattern, with the maximum subsidence branch penetrating to the lower stratosphere in the EA/NP region and with an abrupt switch to upwelling on the east. Ridges and troughs of the wave show a characteristic westward tilt with height.

Šácha et al. (2015) hypothesised that the collocation of the GW hotspot and the peak of the downward BDC branch identified in the EA/NP region by Demirhan Bari et al. (2013) can be partly a consequence of the circulation induced by the GW breaking. But the results are rather contradictory. In agreement with the zonal-mean residual circulation analysis, we can see that in the southern part of the area (Fig. 12f), the GWD induces predominantly anomalous upward flow. Anomalous subsidence strengthens when going further northward (Fig. 12b, d). In line with the obstacle analogy, we observe subsidence in the eastern part of the GWD region only, while anomalous upward flow dominates the western part of the GWD region, and then again eastward and slightly above the anomalous subsidence area. Similar structure of an Eulerian-mean vertical velocity field has been found by Shaw and Boos (2012) as a response to an artificial torque placed in the troposphere around 30° N. These results show that GWs can contribute to longitudinal variations in the BDC and not only the downwelling but also upwelling patterns may be related with GWs.

The magnitude of the vertical residual velocity anomalies maximises around 2% of the reference value for the Box0.1 simulation (Fig. 12b, d, f). For the 10box simulation (Fig. S3), the distribution of upwelling and subsidence is identical and the magnitude reaches 30% locally. Physically, such an anomalous pattern can be explained by considering the dominant background horizontal northeastward wind together with the previously mentioned small obstacle analogy, with induced upward flow upwind and downward flow downwind from the GWD box. However, for the SSWbox simula-
Figure 11. Mean January zonal-mean residual circulation (streamlines for illustration of direction only) and its mass flux (colours, in kg m$^{-2}$ s$^{-2}$) on the left (from top to bottom): reference simulation (a), relative Box0.1–reference simulation anomaly (c), relative Box0.1–Zon0.1 simulation difference (e); and on the right (from top to bottom): snapshot of the SSWbox simulation at 5 days after the GWD injection (b), 10box–reference simulation relative anomaly (d), relative 10box–10zon simulation difference (f). Relative anomalies and differences are given in % of the reference or corresponding box simulation, respectively. The statistical significance of the mean residual circulation mass flux differences was computed by a $t$ test and regions with $p$ values < 0.05 are dashed.

In this paper, we presented results of a set of sensitivity simulations to find out the possible role of a localised GW to the MA circulation and transport. We can observe a completely different distribution variable with time, with subsidence dominating directly above the GWD area in the later stages of the simulation (animation 4 in the Supplement). When the artificial GWD is strong enough to induce significant dynamical changes (SSW simulations), the anomalies cannot be directly explained as being GW-induced because also the dynamical state of the atmosphere changes (e.g. the anticyclonic evolution in animation 1a). Therefore, the explanation of residual vertical wind cross section patterns for both SSW simulations is much more complicated and requires future research allowing for at least the GWD enhancement to reflect the changing background conditions.

4 Discussion and conclusions

In this paper, we presented results of a set of sensitivity simulations to find out the possible role of a localised GW...
hotspots and also, generally, to demonstrate the influence of the spatial distribution of GWD on the middle atmospheric dynamics. The focus was on a mean response to a steady GWD perturbation injected into climatological January condition. Except for the strongest GWD enhancement (SSW simulations; Fig. 2), all simulations (Table 1) have reached a quasi-steady state approx. 7 days after the GWD enhancement (Fig. 3). The average across this state was considered to be the mean response. Section 3.1 focussed on the mean geopotential response at the 6.25 hPa level (Fig. 5). Mean anomalies (differences with reference) were found to be largest in the polar region and larger for the box GWD enhancements (both globally and locally) than for the corresponding ring enhancements. The important role of the often omitted meridional GWD component, especially for the polar vortex response, was highlighted. Most importantly, for simulations with the strongest GWD enhancement (SSWbox and SSWzon; Table 1), we observed different types of polar vortex events, namely a vortex split in response to the ring GWD enhancement and a vortex displacement for a localised forcing (Fig. 6).

In Sect. 3.2 we studied the influence of the artificial GWD and of its distribution on PW activity. We found (Fig. 7) the mean (7th–30th day) E–P flux convergence anomaly centred at the equatorward flank of the GWD enhancement area and an anomalous convergence in a broad area around 60° N in response to the artificial GWD. The anomalies are bigger for the box enhancements. In the box simulations we have identified anomalous, predominantly horizontal PW propagation indicative of in situ PW generation. This is further supported by the results of FT analysis of the geopotential anomalies (Fig. 9), where, for the box simulations, we have found especially the wave-1 and also wave-3 mean amplitude to be anomalously enhanced. Also, the short-term response (Fig. 10) showed the origin of the enhanced amplitudes to lie in the GWD area.

Section 3.3 focussed on the residual circulation response. It was shown that there are significant differences in the zonal-mean residual circulation between different distributions of the same zonal-mean GWD (Fig. 11). A butterfly-like pattern in the box-ring differences was identified centred at approx. 45° N (the centre of the GWD region), with a stronger/weaker subsidence north/south of the enhancement region in the box simulations between 20 and 30 km log-pressure height. Evidence was given that the artificial GWD in our model acts like a small obstacle for the flow, which was further supported by the 3-D residual circulation analysis (Fig. 12). We have found downwelling to the northeast (downwind) and upwelling to the southwest (upwind) of the GWD box showing that GWs can contribute to longitudinal variations in the BDC.

The biggest limit of our analysis is naturally the artificiality of our GWD enhancement. The GWD enhancement introduces an additional artificial constant momentum sink in the model. The concept of the artificial GWD enhancement leaves us also no chance to reflect any feedback between GWs and background conditions (changes in background winds, evolving PW field, etc.). Therefore, for example, our simulation of a vortex displacement differs from reality by not reflecting the background changes, as the GWs are known to be significantly filtered during SSW (e.g. Holton, 1983; Limpasuvan et al., 2012). Considering the intermittent nature of GWs (e.g. Hertzog et al., 2012; Wright et al., 2013), another inaccuracy of our sensitivity simulation set-ups arises from the constancy of the artificial GWD. In particular in the EA/NP region, where we expect mountain wave forcing to be prominent in January, variations of more than an order of magnitude from day to day are to be expected (Schroeder et al., 2009). A multiple (during a month) pulse-like injection of the artificial GWD would be arguably more realistic, but at the expense of the absence of any steady re-
response during the whole simulation. It is also a question of what is a more realistic illustration of the GW effect on the atmosphere: a sudden GWD injection or a smooth increase and decrease with, e.g., a 10-day e-folding time to minimise the initial adjustment noise as proposed by Holton (1983). Furthermore, the spatial distribution of our artificial GWD is highly idealised (in both the horizontal and the vertical). We must note that we compare two “extreme” GWD longitudinal distributions only. It is also very likely that the sharp boundaries of the GWD enhancement in the 10box/zon and SS-Wbox/zon simulations are influencing some minor patterns of the response (e.g. the lee wave pattern in Fig. 10b).

In future work it is therefore necessary to take into account more realistic GWD distributions to address, for example, the efficiency of PW creation. For example, it is possible that a configuration of GWD, taking into account the EA/NP and, e.g., the Greenland GW activity hotspot, would favour enhanced wave-2 instead of wave-1 activity, and for comparison a chessboard-like or random distribution of GWD might be more appropriate. Generally, the fact that the PW activity depends on the longitudinal GWD distribution (Fig. 7) suggests that the rate of compensation between resolved and unresolved drag (Cohen et al., 2013, 2014) can be variable in dependence on the GWD distribution influencing the efficiency of PW creation.

Another motivation for future research is to concentrate on the position of the GW hotspots relative to the climatological stationary wave location in the stratosphere and to analyse the interaction between the GWD effects and the climatological waves. For example, the EA/NP hotspot lies in the region of the phase transition between a trough and a ridge of the climatological wave-1, and our results show (Fig. 8) an anomalous amplification of wave-1 amplitude for a box GWD enhancement in this region. The importance of standing waves for polar vortex strength is well recognised (Watt-Meyer and Kushner, 2015; Yamashita et al., 2015).

In the atmosphere, the most natural, immediate and fastest way for communication of information in the vertical are the GWs (apart from acoustic and acoustic-gravity waves with effects much higher in the atmosphere). We can argue that any change in the troposphere resulting in changes of source, propagation or breaking conditions for GWs will almost immediately influence the distribution of GWD in the stratosphere, with possible effects demonstrated in our paper (in situ generation of PWs in the lower stratosphere, anomalous vertical movements, etc.). For example, on the interannual scale, the occurrence and strength of the EA/NP GW hotspot can be dependent on the Pacific Decadal Oscillation (PDO) phase and can play a role in the relationship between PDO and SSW occurrence frequency (Kren et al., 2015; Woo et al., 2015; Kidston et al., 2015).

There are more conclusions relevant for the SSW research in our results. It is a common methodology (see (Albers and Birner, 2014) for a review of SSW preconditioning concepts) to estimate, e.g., the relative impact of GWs and PWs on polar vortex preconditioning from zonal-mean values of zonal forces only. But our results show that the dynamical effect of forcing depends also on its distribution. The impact connected with a localised area connected with a higher value of drag can be much stronger than one would expect from the zonal-mean value only. Importantly, we have found that for a sufficiently strong artificial zonal-mean zonal force there is a vortex split response to the ring artificial GWD configuration and vortex displacement for a localised forcing. We aim to investigate this in more detail and also for more realistic forcing distributions, but it seems to be clear at this stage that the SSW type may be determined also by the geometry of the forcing, not only by the vortex geometry. On the other hand, vortex geometry can to a large extent influence the distribution of the forcing, e.g. spontaneous emission processes connected with the jet (Plougonven and Zhang, 2014).

Blocking connection with SSW is a well-known correlation (e.g. Andrews et al., 1987; Martius et al., 2009; Nakamura et al., 2016; Albers and Birner, 2014) but the mechanisms standing behind it are still rather elusive. The geographical location and evolution of the stationary positive geopotential anomaly with anomalous anticyclonic horizontal winds upstream of the GWD area is a remarkable feature of the atmospheric response to a localised GWD (Fig. 5), suggesting that GWs can be one of the missing mechanisms behind this relationship. This is connected with the important role of the meridional GWD component, especially for the polar vortex response. Interestingly, this feature becomes apparent for the localised enhancement only and has an almost negligible effect in simulations with ring enhancements. To our knowledge, the effect of the meridional component of GWD on the middle atmospheric circulation has not yet been studied. Also, horizontal GW propagation is neglected in most climate model parameterisations (Kalisch et al., 2014). Thus, it is not surprising that there are only few modelling constraints regarding the horizontal propagation directions, although some information is available from ray-tracing simulations (Preusse et al., 2009). In most studies based on satellite data, GW propagation directions have not been analysed, because the information needed for such computation (e.g. hodograph analysis) is not available for most of the global observational instruments and their combinations (Wang and Alexander, 2010).

Finally, regarding polar vortex effects, the anomalous PW generation and breaking may be the physical justification for disturbing the vortex in its central levels, which was a mechanism hypothesised by Scott and Dritsche (2005). Traditionally, PWs are thought to be generated in the troposphere and propagate up on the polar vortex edge. But, as Scott and Dritsche (2005) pointed out, when wave amplitudes become large and non-linear effects become important, the notion of upward propagation ceases to be appropriate. Therefore, they considered an option of some in situ disturbance at a given level, with a possible explanation being what we propose – localised GW breaking inducing anomalous PW activity.
Regarding residual circulation, a general conclusion of this paper is that for the same magnitude of an artificial zonal-mean zonal force (zonal-mean meridional force as well) there are significant differences (depending on the magnitude of the GWD enhancement) in the zonal-mean residual circulation between different distributions of this force (localised vs. zonally uniform). Also our results indicate that the distribution of GWD may play a role in zonal asymmetries of the BDC. This is a clear signal that in the research of future BDC changes from climate models we need to be concerned not only merely with the magnitude or latitude–height profile of the zonal-mean GWD but also by its zonal distribution. In particular, the models should be able to mimic the main GW activity hotspots and account for the meridional propagation of GWs (Yamashita et al., 2013; Ern et al., 2016). This suggests the need for improvement especially in the non-orographic GW parameterisation (though non-orographic GWs are usually assumed to have significant effect at higher altitudes than in the vertical range analysed in this paper), since many global climate models use a globally uniform gravity wave source function (Geller et al., 2013).

5 Code availability

MUAM model code is available from the authors upon request.

6 Data availability

MIPAS methane volume-mixing ratio profiles (KIT, 2016; von Clarmann et al., 2009; Pleninger et al., 2015) are produced by Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research – Atmospheric Trace Gases and Remote Sensing and are available after registration through https://www.imk-asf.kit.edu/english/308.php. MSR total ozone (Temis, 2016; van der A. et al., 2016) are publicly available through ESA, Tropospheric Emission Monitoring Internet Service (TEMIS) on http://www.temis.nl/protocols/o3field/o3mean_msr2.php. ERA-Interim temperatures and geopotential heights data (ECMWF, 2016) have been provided by ECMWF through http://www.ecmwf.int/en/research/climate-reanalysis/era-interim.

The Supplement related to this article is available online at doi:10.5194/acp-16-15755-2016-supplement.

Acknowledgements. This study was supported by the Charles University Grant Agency project nos. 108313 and 1474314, by the grant no. SVV/2017/2708, by the program of Charles University prevouk no. 45 “Physics”, by GA CR under grant 16-01692J and by Deutsche Forschungsgemeinschaft under grants JA 836/30-1 and 836/32-1. The authors would like to thank Edwin P. Gerber, two anonymous referees and the co-editor Martin Dameris for their unbelievable effort and attitude towards improving the manuscript.

Edited by: M. Dameris
Reviewed by: E. P. Gerber and two anonymous referees

References

Dunkerton, T. J. and Butchart, N.: Propagation and selective transmission of internal gravity waves in a sudden


ECMWF: ERA-Interim, http://www.ecmwf.int/en/research/climate-reanalysis/era-interim, last access: 10 June 2016.


