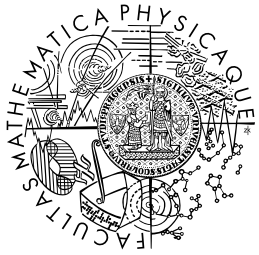

FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

DOCTORAL THESIS

Petr Šácha

New Perspective on the Role of Gravity Waves in the Stratospheric Dynamics and Variability

Department of Atmospheric Physics

Supervisor of the doctoral thesis:	doc. RNDr. Petr Pišoft, Ph.D.
Study programme:	Physics
Study branch:	Meteorology and Climatology

Prague 2016

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In ..... date ..... signature of the author

Title: New Perspective on the Role of Gravity Waves in the Stratospheric Dynamics and Variability

Author: Petr Šácha

Department: Department of Atmospheric Physics

Supervisor: doc. RNDr. Petr Pišoft, Ph.D., Department of Atmospheric Physics

Abstract: This thesis is concerned with the role of internal gravity waves (IGWs) in the stratospheric dynamics and variability demonstrating the effect of spatiotemporal distribution of their activity on the stratospheric dynamics and transport. The first part introduces a theoretical overview of the most recent as well as classical approaches used for description of the wave-mean interaction in the middle atmosphere. Methodology for an IGW analysis from the GPS radio occultation density data is described in the next chapter and the advantages of utilization of density data are listed. The third chapter presents results describing the peculiar dynamics and anomalous IGW activity in the Eastern Asia/Northern Pacific region. An important part is dedicated to a discussion of accuracy limits and usability of different IGW activity proxies. The possible impact of the localized IGW activity is investigated using a mechanistic middle and upper atmosphere model in the last chapter. Sensitivity simulations are used to demonstrate an important role of the spatial distribution of IGW activity for a formation of planetary waves and for the longitudinal variability of the Brewer-Dobson circulation. Implications for the middle atmospheric and climate change research are discussed along with consequences for parameterizations of IGWs in global climate models.

Keywords: stratosphere, internal gravity waves, GPS radio occultation, wave-mean flow interaction

At first, I would like to thank my supervisor Petr Pišoft for all of his support and guidance and for giving me the opportunity to study such an interesting subject. Huge thanks go also to my supervisors at short-term research stays Ulrich Fölsche and Christoph Jacobi for making my stays there productive and enjoyable, and for introducing me to the international research level. I would also like to thank Jan Bednář and Jiří Mikšovský and my fellow postgraduate students Aleš Kuchař and Jan Karlický for many productive discussions and their help in variety of issues I ran across. Especial thanks are to Friederike Lilienthal, who spent a lot of her time translating my thoughts into the world of numerical simulations and who was always ready to correct my texts and to comment my ideas.

But, biggest and most heartily thanks go to my wife Nikola and daughters Mariana and Anežka. Should it not be you, this thesis might be much more detailed, but my life would be inexpressibly poorer!

This dissertation was supported by GA CR under grant 16-01562J, by the Charles University Grant Agency under grant n. 108313 and the author was also partly supported by a grant from the DAAD, and AKTION scholarship programs.

Contents

Introduction

This thesis is concerned with the manifestation of internal gravity waves (IGWs) in the stratosphere and with consequent effects on its large-scale circulation and climate variability. We demonstrate the influence of spatial distribution of the IGW activity on the stratospheric dynamics and transport.

The atmosphere is conventionally divided into regions according to the vertical temperature gradient. Andrews et al. [1987] have defined the term middle atmosphere as a region from the tropopause to the homopause (10–110 km) containing the stratosphere, mesosphere and lower thermosphere, where the same physical processes govern the dynamics and similar approximations are valid. In this part of the atmosphere we can neglect effects of latent heating as well as molecular diffusion and electromagnetic forces. In the middle atmosphere, the evolution of a system is governed by a three-way interaction between dynamics, physics (microphysics, phase changes, radiative transfer) and chemical processes. By the term dynamics we understand physical processes described by the momentum and thermodynamic equations.

Further in this text the emphasis is placed on processes connected with large-scale flows that fall within the scope of geophysical fluid dynamics (GFD). Gravity waves in GFD are usually described as a group of wave motions in a fluid where the restoring force is the gravity (or a so-called reduced gravity). Internal gravity waves (IGWs), as a part of them existing in continuously stratified fluids, are of special importance in the atmosphere for their effects on atmospheric composition, circulation and dynamics in general.

In the atmosphere, IGWs are a naturally occurring and ubiquitous, though intermittent (in the sense of larger amplitude wave packets [e.g. Hertzog et al., 2012; Wright and Gille, 2013] phenomenon. They can be pictured as being composed of oscillating interacting particles with mutually connected phase. The waves can propagate both horizontally and vertically under a continuous interplay between gravity and inertia and an exchange between potential and kinetic energy [Cushman-Roisin, 1994]. A characteristic feature of their propagation is that the group and phase velocity are always perpendicular. IGWs are very important for atmospheric dynamics influencing atmospheric thermal and dynamical structure such as its angular momentum distribution. For example, one of their crucial roles is that they transport angular momentum from the ground upwards. That helps to synchronize the rotation of the upper parts of the atmosphere and also in general, they contribute to restore the equilibrium and to gain energetically more favourable conditions.

The importance of gravity waves for the dynamics of the middle atmosphere has been recognized for quite a while [Hines, 1960] and the research in this area is still very active and connected with some of the most challenging issues of Earth climate and atmospheric science. In recent years, the significance of IGWs has been particularly recognised. For example, Ern et al. [2014] pointed out their role in the formation of the quasi-biennial oscillation (QBO) that, together with the findings of Marshall and Scaife [2009], can link them to the European winter surface climate. Ern et al. [2011] suggested that IGWs strongly interact with the polar night jets not only in the mesosphere, but also in the stratosphere,

they influence the stratosphere-troposphere exchange (STE) through affecting the tropopause inversion layer [Kunkel et al., 2014]. IGWs have a direct influence on the middle atmospheric climate and the acceleration of the Brewer–Dobson circulation (BDC) [e.g. Garcia and Randel, 2008] and recently, Demirhan Bari et al. [2013] provided evidence of the effects of IGW activity on the three-dimensional (3D) transport structure in the middle atmosphere.

Consideration of IGW related processes is necessary for a proper description and modelling of the middle (up to the beginning of 21st century reviewed comprehensively by Fritts and Alexander [2003] and upper atmospheric dynamics (see, e.g., the review by Smith [2012]). But, as the mesoscale waves, IGWs have horizontal scales from tens to thousands kilometers, cannot be fully resolved by general circulation models and have to be parameterized. In one of the first numerical attempts to demonstrate the importance of IGWs for the middle atmospheric dynamics, Holton [1983] concluded with formulating some requirements on IGW related information (e.g. spatial and seasonal distributions of phase speeds, amplitudes etc.) to improve global climate model parameterization schemes. However, only recently satellite and other global datasets with improved resolution and novel analysis methods together with high-resolution global models have been tightening the constraints for the parameterizations that can improve the treatment of these waves in climate models [Alexander et al., 2010; Geller et al., 2013].

Complex understanding and unbiased modeling of the 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 [Calvo et al., 2015]. Long-term and decadal changes of the middle atmosphere have been found to be largely owing to greenhouse gas changes and thus reflect climate change [e.g. Lübken et al., 2013; Jacobi et al., 2015] yet it has been difficult to match observations with the robust model stratospheric change projections determined to a large extent by a speeding-up of the Brewer-Dobson circulation (BDC) [Kidston et al., 2015]. Although weather systems are not able to penetrate into the stratosphere there is 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]. Better representation of the stratosphere could improve the long-range [Hardiman and Haynes, 2008] and also short range forecast skills [Gerber et al., 2012].

There are many ways how the middle atmosphere can affect tropospheric conditions. In the last decades a lot of effort has been put forward to study the vertical coupling of dynamics between the stratosphere and the troposphere and consequently possible dynamical downward links that can be responsible for, e.g., observed signals of solar variability in the troposphere [Haynes, 2005]. There are two important theories: a) the piecewise potential vorticity inversion [Bishop and Thorpe, 1994; Hartley et al., 1998] making use of non-locality of the inversion of elliptical operator in analogy to electrostatics and b) the downward control principle [Haynes et al., 1991], which gives a connection between the zonal force exerted above some level and the vertical mass flow across that level. However, those are the steady state theories giving us no information on the causality of the problem. In a modelling study, Hardiman and Haynes [2008] have proven that,

though the density exponential decay with height, the downward propagation of information over larger vertical distances is possible due to the nonlinear processes involved and due to a dynamical sensitivity of the stratosphere.

Before that, a traditional view was that the middle atmosphere responds passively to the forcing from the troposphere, with upward propagating planetary waves (PWs) dominating the driving in the lower stratosphere (LS) and IGWs playing an important role higher in the upper stratosphere and mesosphere. In a frame of this thesis we will examine a theory endeavoring to explain the origin of forces in the LS. We demonstrate that IGWs can significantly affect the dynamics (e.g. PW creation) and transport already in the LS. IGWs are the most natural, immediate and fastest way for vertical communication of information in the atmosphere (excluding acoustic and acoustic-gravity waves affecting higher layers predominantly) and so, the IGW activity reflects almost immediately conditions in the troposphere and even at the surface. Giving observational evidence that IGWs are breaking (locally) already in the LS and showing numerically their possible dynamical effect, we will literally enclose the circle of mutual interactions and causality between the stratosphere and troposphere.

The thesis is structured as follows: In the first chapter we give a brief introduction to the flow description in the middle atmosphere and to the theory of wave-mean flow interactions and elucidate the physics standing behind diagnostic methods used in this thesis. Following chapters are based on papers and results published in impacted journals.

Second chapter is concerned with the methodology of IGW observations. After a brief review, the basics of Global Positioning System (GPS) exploitation by radio occultation (RO) are explained. A novel methodology for IGW activity analysis from GPS RO dry density profiles [Šácha et al., 2014] is described and its advantages are shown. Third chapter presents a global study of IGW activity in the LS with special focus on the Eastern Asia/Northern Pacific (EA/NP) region as a newly discovered IGW activity hotspot [Šácha et al., 2015].

Fourth chapter presents results from sensitivity mechanistic model simulations to illustrate a possible effect of the EA/NP region on the large-scale circulation and transport in the stratosphere. The results demonstrate an important role of the spatial distribution of IGW activity for the formation of PWs, polar vortex stability and induced residual circulation anomalies [Šácha et al., 2016]. Finally, in Conclusions we give a summary of our results and discuss the consequences of our results for atmospheric science and climate change studies.

1. A brief theoretical insight into the wave – mean flow interaction theory

Although the Navier-Stokes equations are an approximation to some order, they are far more complicated than necessary for a description of the middle atmospheric phenomena considered in this thesis. As stated for example by Andrews et al. [1987], the most general set of equations used in the middle atmosphere dynamics are the so called primitive equations, which we get after a serie of simplifications of the equations of motion for a gas in a rotating frame. They consist of two momentum balance equations, one in the zonal and one in the meridional direction, the hydrostatic balance and continuity equation for mass and the thermodynamical relation between diabatic heating and material rate of change of the potential temperature. Derivation, summary of approximations made and discussion of consistency of the hydrostatic primitive equations can be found in White et al. [2005].

Despite the simplifications, primitive equations are still a complicated set and further approximations are needed to obtain analytical solutions. There are many ways how to simplify the primitive equations. For example, by making suitable geometrical (beta-plane etc.) and dynamical (e.g. small Ekman, Rossby or temporal Rossby number) approximations we may obtain, following for example Holton [1992] or Cushman-Roisin [1994], the quasi-geostrophic (QG) potential vorticity equation for nonlinear motions in a continuously stratified fluid on a beta-plane, which should fully capture the large-scale, slower motions at least in the extratropical region [Andrews et al., 1987].

Additional method, which can help us to guide the governing equations to an analytically resolvable form, is the linear wave theory. This theory is used to linearize the equations (e.g. by neglecting the advective processes). It starts by decomposing the dependent variables into a mean and an eddy (perturbation) part. Thus, some kind of averaging, which helps us to determine the mean part, is needed. Although the circulation in the middle atmosphere shows a great degree of variability with altitude, latitude and longitude, the longitudinal variations can be at the leading order neglected. Therefore, zonal averaging is commonly utilized as an example of the Eulerian mean (average taken over one coordinate while the value of other three spatio-temporal coordinates is held fixed) in the middle atmosphere and in the stratosphere, especially.

Another averaging approach is a particle following Lagrangian mean taken over a trajectory or over a specified ensemble of fluid particles. Particularly, in the middle atmosphere, a combination of the Eulerian and Lagrangian approach called generalized Lagrangian mean [GLM; for details see e.g. McIntyre, 1980; Craik, 1988; Bühler, 2014] became very popular. In GLM, independent variables remain the same as in the Eulerian approach while the Lagrangian information is stored in an additional displacement field. Displacements are dependent on the Eulerian position and time as well. Under the assumption of well behaving disturbances, GLM theory extends easily to finite-amplitude disturbances and

therefore it seems to be a natural extension of small-amplitude theory [Bühler, 2014], which will be dealt with in further text.

To explain the evolution of a longitudinally averaged atmosphere we have to take into account systematic effects of eddies (waves). As a phenomenological note, I would like to point out the difference between eddies and waves. In spite of the fact that in analytical description all of the eddies (deviations from a mean state) are created by the waves and these two terms merge, not all of the eddy terms have to have a wave origin in reality and one has to be careful to use the terms correctly. Waves are characterized by their dispersion and polarization relations. In dependence to the assumptions made, a lot of different wave types can be derived (Rossby waves, Kelvin waves, internal gravity waves, etc., see e.g. Andrews et al. [1987] or Cushman-Roisin [1994] for derivation).

The theory of wave mean-flow interaction is a highly developed theory and despite its limitations (e.g. small amplitude of disturbances, zonally symmetric basic flows) it still provides a useful quantitative framework for understanding of the circulation of the middle atmosphere [Haynes, 2005]. One of the most important results of the wave mean-flow interaction theory is a so-called Charney-Drazin nonacceleration (or noninteraction) theorem. The noninteraction theorem itself has many versions involving different wave types and different assumptions; see e.g. Boyd [1976]. One of its most illustrative interpretations is that the zonal mean flow is not accelerated in the presence of a steady and non-dissipative wave field [Bühler, 2014].

The theorem has a very simple form in the so-called transformed Eulerian mean (TEM) set of equations [see e.g. Andrews et al., 1987]:

$$\nabla \cdot \vec{F} = 0, \tag{1.1}$$

where $\vec{F}$ is the Eliassen-Palm (EP) flux. For more realistic situations (presence of sources, time varying wave amplitudes etc.) the theorem was derived by Andrews and McIntyre [1978] under the name generalized Eliassen-Palm theorem in a form:

$$\frac{\partial A}{\partial t} + \nabla \cdot \vec{F} = D + O(a^3), \tag{1.2}$$

where the time derivative of A, the so-called wave-activity density, represents a measure of wave transience and D contains the frictional and diabatic effects. To elucidate equations (1.1) and (1.2) we have to come back to the TEM equations and the ideas behind their derivation.

The TEM formulation enables us to rewrite the zonally-averaged thermodynamic and momentum equations so that only diabatic eddy heat fluxes appear in the thermodynamic equation. Under the assumption of conservative and steady waves we aim to subtract the "untrue" eddy heat fluxes stemming from the improper averaging. This is achieved by modification of the mean meridional and vertical advection $(\overline{v}, \overline{w})$:

$$\overline{v}^\star = \overline{v} - \rho_0^{-1} \cdot \left(\rho_0 \frac{\overline{v\theta}}{\overline{\theta}_Z} \right)_Z, \tag{1.3}$$

$$\overline{w}^\star = \overline{w} + (a \cos\phi)^{-1} \left(\cos\phi \frac{\overline{v\theta}}{\overline{\theta}_Z} \right)_\phi. \tag{1.4}$$

Here $\overline{\theta}$ is the zonally averaged potential temperature, ρ_0 the background density, z is the log-pressure height, a the mean radius of the Earth, ϕ the latitude and $\overline{V\theta}$ the meridional eddy heat flux. The velocities $(\overline{v}^\star, \overline{w}^\star)$ constitute the residual mean meridional circulation and under the above-mentioned assumptions (conservative, steady waves) the second term in their definitions (1.3, 1.4) is equivalent to a Stokes drift at a leading order [Bühler, 2014]. Dunkerton [1978] showed that for steady and conservative waves the residual-mean circulation is equivalent to a Lagrangian-mean circulation to the second order [Bühler, 2014] of a wave amplitude.

After substituting the residual mean velocities into zonally averaged momentum equations we can regroup the eddy terms to form a single forcing term $\nabla \cdot \vec{F}$ in the transformed mean zonal momentum equation. Consequently we define the E-P flux vector as (quasi-geostrophic beta plane approximation):

$$\vec{F} = \left(0, -\rho_0 \overline{v'u'}, \rho_0 f_0 \overline{v'\theta'}/\theta_{0,Z}\right) \tag{1.5}$$

Using geometric formalism, Maddison and Marshall [2013] have shown that the fundamental object of the eddy-mean flow interaction is a rank-two momentum flux tensor (E-P flux tensor). In the appropriate limit only, the E-P flux tensor is restricted to the classical E-P flux vector. For the QG approximation, Maddison and Marshall [2013] also show the underlying mathematical principle behind E-P flux and residual mean theories – exploitation of the gauge freedom. The form of QG equations allows us to add rotational terms to the eddy flux tensor and in particular, the QG potential vorticity (PV) equation is absolutely unaffected by addition of anti-symmetric components to the eddy flux tensor. Naturally, such a gauge freedom allows for different choices of additional terms resulting in various types of residual-mean equations and E-P flux vectors. Some interesting choices of anti-symmetric tensors (leading e.g. to a form of the Plumb [1986] E-P flux) are shown in Maddison and Marshall [2013].

Now, the idea behind the E-P theory will be illustrated from a rather kinematical point of view using a trajectory analysis. Starting with a graphical illustration of the difference between Eulerian and Lagrangian mean velocity (Stokes drift) for a steady, vertically propagating PW1 (PW with zonal wavenumber equal to one) in a uniform basic flow on a beta-plane bounded by vertical walls at 90^o and 30^o, the physics behind the non-acceleration theorem is illustrated and the artificiality of Eulerian mean considerations is demonstrated following the situation solved analytically by Matsuno [1980].

In Fig. 1.1 a horizontal velocity field associated with a stationary PW1 is illustrated. Note that in the frame connected with the uniform basic westerly zonal flow U, the wave propagates westward with the phase speed $c = -U$. Substituting geopotential in the following form (eq. 1.6) into the QGPV equation, the 3D structure of the horizontal velocity wave field can be derived as geostrophic winds connected with the PW1 with amplitude A (1.7, 1.8):

$$\phi = A e^{\frac{Z}{2H}} \sin y \sin\left(kx + mz\right), \tag{1.6}$$

$$u = -\frac{A e^{\frac{Z}{2H}}}{f} \cos y \sin\left(kx + mz\right), \tag{1.7}$$

$$u = \frac{Ae^{\frac{Z}{2H}}}{f} \sin y \cos\left(kx + mz\right), \tag{1.8}$$

$$y = 3\left(\varphi - \frac{\pi}{6}\right). \tag{1.9}$$

Here ϕ is the perturbation geopotential, u and v are perturbation velocity components, f is constant Coriolis parameter, H is the constant scale height and φ latitude in radians. The meridional coordinate y is defined so as it increases monotonically between the vertical walls and is equal to zero at 30^o and to π at 90^oN. Note that the geopotential perturbation attains the maximum in the middle of the channel while wanishing at the boundaries. Such a distribution of PW1 associated perturbations resembles approximately what is actually observed in the winter troposphere and lower stratosphere [Miyakoda, 1963].

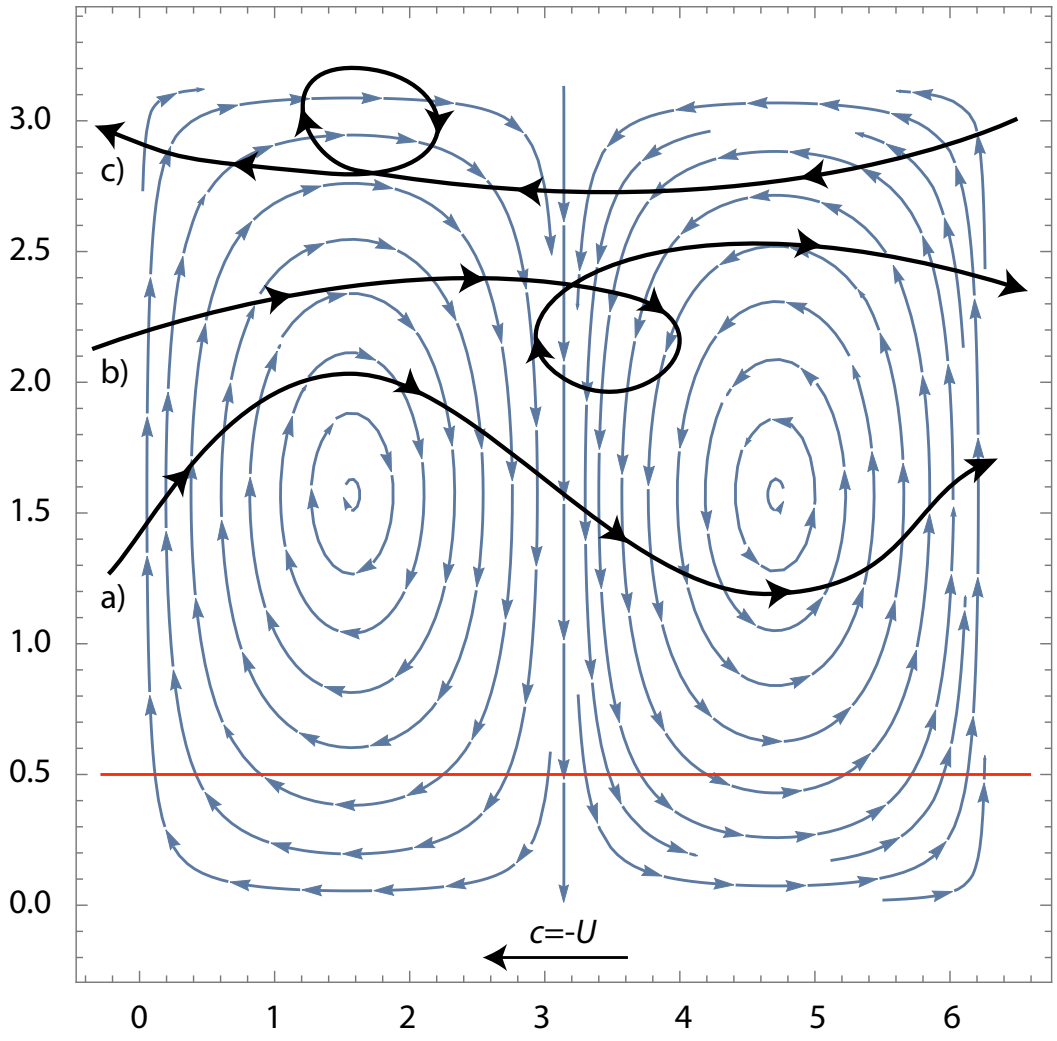

Figure 1.1: Illustration of a horizontal wind perturbation field (indicated by blue arrows) associated with the upward propagating PW1 in a channel bounded by 30^o and 90^oN. The vertical axis is for y (1.9) and the horizontal axis represents longitude in radians. Red curve represents an illustration of the Eulerian mean and black curves illustrate particle trajectories.

Note that the zonal mean zonal velocity (Eulerian average; red curve in

Fig. 1.1) is trivially equal to U everywhere in the domain, because the velocity perturbations cancel each other out in the zonal average. But, as depicted by black curves in Fig. 1.1, Lagrangian mean zonal velocity (particle following time mean over a period of the wave in this case) is different from U. Particles have the Lagrangian mean zonal velocity greater than U in the middle of the channel and lower near the boundaries. Analytical derivation of trajectories in the horizontal plane can be found in Matsuno [1980].

There are two competing physical mechanisms acounting for the Stokes drift (difference between Lagrangian and Eulerian mean velocity). The first cause of this difference, which is dominant in the middle of the channel, is connected with the meridionaly variable wave amplitude (eq. 1.7, 1.8). The situation can be easily illustrated by considering a particle starting at a point b) in Fig. 1.1 with velocity U, initially. During its passage through the wave field it is displaced northward at first, where it experiences stronger magnitude of eastward flow connected with the consequent wave phase. Then the particle is being advected southward where it experiences weaker magnitude of westward flow than before, resulting in a net eastward drift.

For illustrative purposes, the trajectory b) in Fig. 1.1 is outlined for a longer time interval than the corresponding wave period. On the contrary, trajectory a) is depicted exactly for one wave period. A particle starting initially at a) has a strongest westerly Stokes drift, because, due to the phase shift between u and v, it never experiences westward velocity fluctuations during the passage through the wave field. The Eulerian mean is absolutely unable to reflect the different velocity values at the actual particle positions during the passage (this is the idea behind GLM theory and the so-called lifting map [Bühler, 2014].

Although this mechanism is active in the whole domain, due to the decrease of the magnitude of a gradient of wave amplitude (decreasing from the middle towards boundaries), the second Stokes drift mechanism begins to gain importance near the boundaries. This mechanism is connected with a different time sampling of velocity fluctuations during the passage of the wave field. Namely, due to the westward phase speed of the PW1 relative to the eastward background flow U, the relative phase speed is higher and the phase changes more quickly for a particle having eastward velocity perturbation while more slowly for the same particle in a later time, when having westward velocity perturbation. Therefore, the westward velocity perturbation will be sampled more often during the wave period and the resulting Lagrangian mean velocity of particles starting around c) is lower than U.

Both of the above discussed competing kinematical mechanisms are identified analytically by Matsuno [1980], equation 2.20. Matsuno [1980] was the first to analytically derive trajectories for a steady vertically propagating PW using the Eulerian information only. He proved that, when projected into the meridional plane, the trajectories (up to the $O(A^2)$) are cyclic and have a form of a tilted ellipse. This is an important result - since the parcels are not displaced permanently in the vertical, the mean temperature field cannot be affected, despite the observed Eulerian mean vertical motions! In Fig. 1.2 we present a slightly modified Fig. 4 of Matsuno [1980], where we can most easily illustrate the physics acting behind the non-acceleration theorem and the artificiality of Eulerian mean considerations.

In Fig. 1.2, parcel trajectories are drawn as circles instead of inclined ellipses (vertical distances covered by the parcels are exaggerated to be comparable with meridional displacements), for simplicity, but with a realistic sense of rotation. Vertical bounds are depicted with a thick line and doted beyond. Potential temperature, as illustrated, is monotonically increasing with height, but the wave amplitude is assumed constant with height, for simplicity.

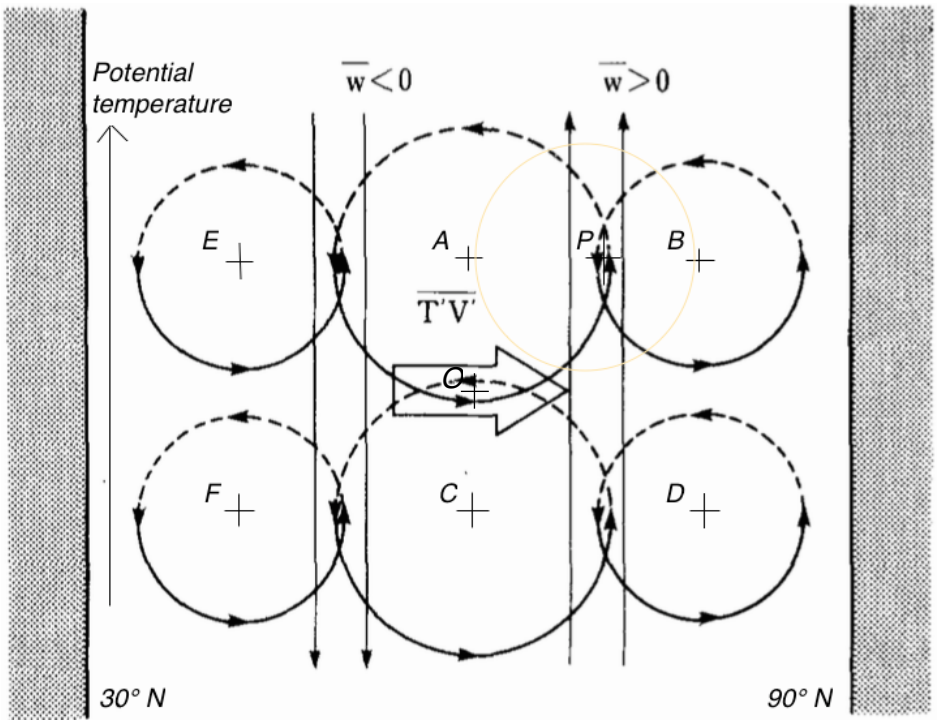

Figure 1.2: Illustrative scheme showing how the observed zonal mean vertical motions and eddy heat flux are generated by cyclic motions of air parcels (slightly modified version of Fig. 4 in Matsuno [1980]). Dashed lines indicate where the air parcel is cooler than its surrounding. Capital letters A–F indicate the mean basic positions of air parcels, O, P mark observer's position and the yellow circle shows a group of basic positions of air parcels that pass by the point P during a wave period.

Now, if we consider the situation observed by the observer at O, we see that during a wave period he observes northward moving particles with higher potential temperatures ($T > 0$) while lower potential temperature values are connected with southward moving particles ($v < 0$). So, the observer at O gets a net northward eddy heat flux ($\overline{Tv} > 0$). This feature originates from the vertical gradient of potential temperature, particularly. Due to the meridional variations of wave amplitude, the mean eddy heat flux is largest at the center of the domain leading to mean eddy heat flux convergence at higher and divergence at lower latitudes.

Observer P measures a complete wind rose of wind directions during a wave period (from basic points depicted by yellow circle in Fig. 1.2). Due to the meridionally variable wave amplitude, only the vertical wind speed fluctuations do not cancel out. Because the air parcel from the basic position A has bigger amplitude of velocity fluctuations than from the position B, observer P measures

a mean (Eulerian mean) upward vertical motion. This situation holds in the whole northern half of the channel and is opposite ($\overline{w} < 0$) in the southern half. So, we have adiabatic heating from the mean downwelling together with eddy heat flux divergence in the southern half and adiabatic cooling together with the eddy heat flux convergence in the northern half of the domain. This cancellation of the eddy heat transport and mean vertical motion effects led many authors to explanations of Charney-Drazin theorem using causal theories between the Eulerian mean eddy fluxes and Eulerian mean circulation. (e.g. Andrews et al. [1987]: "..under Charney-Drazin theorem eddy fluxes do not generally vanish, but induce a nonzero Eulerian mean circulation ($\overline{v}, \overline{w}$) that precisely cancels their effect."). But, having the knowledge of the underlying parcel trajectories, the Eulerian mean budget considerations seem somewhat artificial [Matsuno, 1980].

Clearly, the "artificiality" comes from the process of averaging. It would be more useful if one can consider some averaging type respecting better the physical problem under consideration. For example, as the advection operator was modified to reduce the eddy buoyancy fluxes in the derivation of TEM equations, we can redefine the averaging operator for the same purpose. This is the idea leading to a thickness-weighted averaging (for details see Maddison and Marshall [2013]. We can demonstrate its principle by assuming a dynamical equation for the buoyancy b:

$$\left(u^a b\right)_{,a} = \Theta, \tag{1.10}$$

where Θ is a source term, index a spans independent coordinates (x, y, z, t), comma indicates a derivative with respect to the indexed coordinate and Einstein summation convention is assumed. After application of the thickness-weighted average (along material lines with $b = const.$, $b^{'} = 0$ by definition), the average buyoancy equation becomes

$$\left(\overline{u^a} b\right)_{,a} = \overline{\Theta}. \tag{1.11}$$

No eddy terms appear in (1.11) and hence, no additional residual circulation needs to be defined. Moreover, after derivation of the thickness-weight-averaged momentum equation, we get the EP flux tensor directly related to the flux of potential vorticity [Maddison and Marshall, 2013]. The relative simplicity of the resulting equations comes at the cost of a careful and sophisticated definition of a volume-form-weighted average and further modifications have to be applied in the case of non-monotonical variations of buoyancy with height (as it is often the case in atmosphere).

In this thesis (Chapter 4) we use a 3D residual circulation approach derived by Kinoshita and Sato [2013], where the time averaging is being used. However, the time averaging brings additional complications. There is not any completely right choice of the averaging period as there is a whole spectrum of waves present in the atmosphere. Also, stationary waves cannot be treated [Sato et al., 2013]. The latter can be partly overcomed by using an extended method of the Hilbert transform to estimate the envelope function of momentum or energy fluxes of the wave field as a substitute for the temporal or spatial averaging [Sato et al., 2013]. Kinoshita and Sato [2013] and Sato et al. [2013] gave evidence that their residual circulation is approximately equal to Lagrangian mean velocities. However, there is a controversy with Noda [2014], who formulated a set of so-called Generalized TEM (GTEM) equations applicable also to unsteady waves. Noda [2014] shows that the residual circulation and EP fluxes are very different between TEM and

GTEM for a strongly unstable wave. The GTEM approach was not applied in our analyses (Chapter 4), since additional information is needed compared to the TEM diagnostic (frictional forcing and diabatic heating) and it was not available in our model outputs.

Another promising approach are the recently derived Local Transformed Eulerian Equations by Gaßmann [2014]. In her approach, by utilizing the basic conservation laws for entrophy (potential temperature), energy (Bernoulli function) and angular momentum (PV), Gaßmann [2014] is able to decompose the Eulerian wind field into dynamically active and inactive part – in each time step, without any averaging being used. Notably, the time evolution of inactive winds in the stratosphere can be traced to the closed elliptical trajectories (Fig. 2b in Gaßmann [2014]) and the active winds give us immediate 3D information about the non-conservative part of a mass flux that would lead to mean state changes afterwards, should any space or time average be computed.

2. Observations of gravity waves using GPS-RO density profiles

A variety of observation techniques has already been applied in the research of wave disturbances in the atmosphere. Those include radiosonde and rocketsonde measurements, balloon soundings, radar and lidar observations and other remote sensing measurements. In the last two decades, the remote sensing with occultation methods has undergone a remarkable development. GPS signals are exploited by the RO methods and are often utilized for studies of IGWs. In the future, the potential of these sounding techniques will most likely grow due to increasing numbers of transmitters and receiver platforms [Wickert et al., 2009], especially with the upcoming launch of the COSMIC 2 mission.

Wu et al. [2006] categorized the GPS RO as a Line of Sight (LOS) sensor characterized by excellent vertical and coarse horizontal (due to LOS-smearing) resolution. Therefore, the GPS RO measurements are mostly sensitive to IGWs with a small ratio of vertical to horizontal wavelength. The observations are limited at high altitudes by ionospheric residual errors and at low altitudes by strong water vapour effects. Hence we focus on the height domain between 8 km and 40 km. Nevertheless, filtering the RO data for ionospheric correction remains a factor influencing the spectral density of the signal prior to the standard density or temperature retrieval. The highest accuracy is found in the lower stratosphere where it is usually better than 1 K [Steiner and Kirchengast, 2000]. The GPS RO technique provides atmospheric profiles with global coverage under all weather and geographical conditions together with self-calibration ability and long-term stability. That makes the GPS RO an almost perfect tool for atmospheric monitoring [Foelsche et al., 2008].

Using GPS RO, the IGW description can be retrieved in a series of steps. At first, a height profile of atmospheric refractivity index is derived from bending angles. In this step local spherical symmetry is assumed and therefore a limited horizontal resolution of about 300 km is common to limb scanning methods [Preusse et al., 2008]. Then, the refractivity index can be directly related to the density of dry air. Temperature profiles, which are usually used, are computed only subsequently from the density profiles using the hydrostatic balance and the state equation for dry air. This approach, termed "dry air retrieval", neglects the contribution of water vapour to radio refractivity, but in the height range of interest for our study this effect can be neglected [Foelsche et al., 2008].

Research on atmospheric waves using GPS RO data expanded since the papers of Tsuda et al. [2000] and Steiner and Kirchengast [2000]. According to the linear theory of IGWs, a separation between a small wave-induced fluctuation and background field has to be performed, if vertical profiles of any state quantity are used for the detection of the gravity wave parameters. The choice of the background state significantly affects the results and it is quite a complicated issue because an artificial model could never perfectly reflect the real state of atmosphere. Lots of authors who have studied IGWs retrieved from the GPS RO dry temperature profiles have utilized different methods for determining the background temperature. For example, Steiner and Kirchengast [2000] and

many others have applied various sorts of band pass filters with cut-offs at some specified vertical wavelengths. E.g., Gubenko et al. [2012] or Vincent et al. [1997] have used approximations by low order polynomials for the stratospheric levels. The analysis of the tropopause region is always connected with problems because of the artificial enhancement of the wave activity (even when using the advanced band-pass filter associated with different vertical wavelengths, see Alexander et al. [2011]. Schmidt et al. [2012] suggest two possible approaches to solve this problem. Those are a separation of the profile into tropospheric and stratospheric parts and application of a filter for each region or, more appropriately, the double filtering method. The question of the background separation is an important part of our study too and it is discussed in details in the following sections.

In this chapter, we present a new method for the analysis of IGWs using density profiles. Atmospheric density is the first quantity of state gained during the retrieval of GPS RO data and it is not burdened by any additional assumptions (e.g., hydrostatic balance). Moreover, we would like to stress out another advantage, which has not been discussed yet – unlike temperature profiles, the density evolution with the height is theoretically inferable by means of statistical physics. Separation of the density background has not only a physical basis but it can also be computed partly analytically using our method presented in next section.

To our knowledge this is the first study where atmospheric density profiles from GPS RO are explicitly used for the study of IGWs. Liou et al. [2003, 2005] and Pavelyev et al. [2006, 2009], however, followed a methodologically similar approach by studying vertical refractivity gradients based on RO data. That is unexpected since the density data should be very accurate. For example, refractivity profile intercomparisons of nearly collocated profiles with radiosondes show differences of less than 0.7 % below 30 km altitude [Schreiner et al., 2007]. Another instrument that gives direct information about density is the lidar. There are several observational studies of IGW induced density fluctuations [e.g. Wilson et al., 1991; Sica and Russell, 1999] and theoretical papers about the determination of density fluctuation from the returned power [Gardner et al., 1989], and estimations of the response of neutral density layers to gravity waves perturbations [Gardner and Shelton, 1985; Chiu and Ching, 1978].

The next subchapter introduces the methodology and the data sets used to retrieve the density perturbations profiles and to analyse IGWs. The results are described in the section 2.2 and they are discussed with concluding remarks in section 2.3.

2.1 Methodology for IGW activity analysis from density profiles

2.1.1 Separation of the background density profiles

In deriving the separation method we will make use of two basic facts – the fact that the general form of the density decrease with height is known from theory and that the variations of the background Brunt-Vaissala frequency squared (N_0^2) are substantially small [Steiner and Kirchengast, 2000; Tsuda et al., 2000]. The latter leads in various studies [e.g. Steiner and Kirchengast, 2000; Tsuda et al.,

2011; Gubenko et al., 2012] to a simplification of equations where N_0^2 is replaced within the area of interest with one constant value. Care must be taken when the area includes more than one layer because there is a jump of N_0^2 at the boundaries (in our case at the tropopause).

Let us have a background density profile $\rho_0(z)$ and assume that all departures from the background density are due to the response to the IGW induced wind perturbations and are governed by the continuity equation. By assuming horizontal homogeneity and neglecting the diffusion and chemical effects, [Gardner and Shelton, 1985] have shown that the density response can be written as

$$\rho\left(\vec{r},t\right) = e^{-\chi} \cdot \rho_0\left(z-\xi\right). \tag{2.1}$$

Here $\rho\left(\vec{r},t\right)$ is the perturbed density at position $\vec{r}$ and time t and χ, ξ are solutions of partial differential equations (for details, see Gardner and Shelton [1985]) that are related to the wind divergence and to the vertical displacement, respectively.

Unlike continuous lidar measurements, the GPS RO data provide a snapshot of the perturbed vertical density profile $\rho_0(z)$. Additionally, we will make an assumption that the background density profile could be expressed analytically in the form

$$\rho_0\left(z\right) = \hat{\rho}_0\left(z\right) exp\left[-a\left(z\right) \cdot z\right], \tag{2.2}$$

where $\hat{\rho}_0$ is the background density generally at the lower boundary of our vertical profile, z is the vertical distance from the lower boundary and $a(z)$ is a coefficient of the background density exponential decay. In the real atmosphere, the coefficient $a(z)$ would be primarily influenced by the ambient temperature.

In general, the buoyancy frequency at any vertical level is

$$N^2 = -\frac{g}{\bar{\rho}}\frac{d\bar{\rho}}{dz}. \tag{2.3}$$

For our case, we will define our own slightly altered expression of the background Brunt-Vaissala frequency squared

$$N_0^2\left(z\right) = -\frac{g}{\rho_0\left(z\right)}\frac{d\left(\rho_0\left(z\right)\right)}{dz}. \tag{2.4}$$

By substituting the background density in the form defined by (2.2) and by omitting the notation of functional dependencies and introducing the notation $\frac{da(z)}{dz} = a'$ we gain

$$N_0^2 = -g\left(-a - za'\right). \tag{2.5}$$

Now it comes to the previously mentioned fact that N_0^2 generally evolves very weakly inside layers and with moderate jumps on boundaries. Hence, the frequency should be suitable for approximation by selected analytic functions. First, we must create a vertical profile of the perturbed local buoyancy frequency squared according to

$$N^2\left(z\right) = -\frac{g}{\rho\left(z\right)}\frac{d\left(\rho\left(z\right)\right)}{dz}, \tag{2.6}$$

which is illustrated in Fig. 2.1 Then, by fitting the perturbed frequency, we have the formula for N_0^2 and we can substitute it back to (2.5).

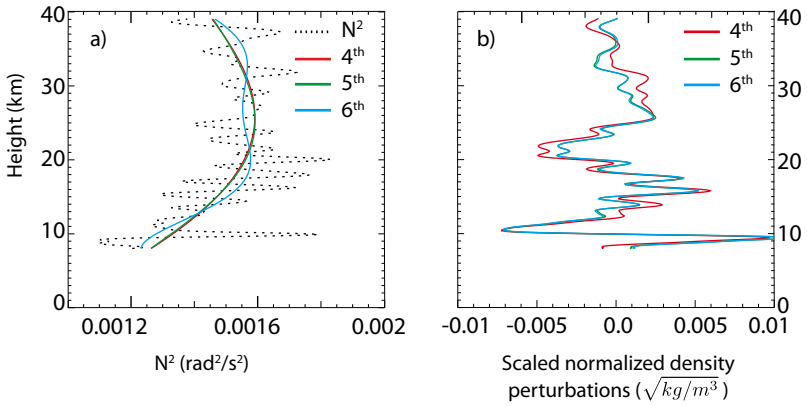

Figure 2.1: a) Various (polynomial) fits of the local buoyancy frequency squared ($N^2(z)$), b) corresponding scaled normalized density perturbations.

Further steps will be illustrated on the simple example of fitting the perturbed local buoyancy frequency squared by a fourth order polynomial (see Fig. 2.1) in the form

$$A_4 z^4 + A_3 z^3 + A_2 z^2 + A_2 z + A_0. \tag{2.7}$$

After that we can substitute into (2.5) to get the first order differential equation, which can be (after making a substitution $y = a \cdot z$) directly solved to gain an expression for $a(z)$. Then going back to (2.2) we have:

$$\rho_0(z) = \hat{\rho}_0 exp\left(-\frac{\frac{A_4 z^5}{5} + \frac{A_3 z^4}{4} + \frac{A_2 z^3}{3} + \frac{A_1 z^2}{2} + A_0 z}{g} - C\right), \tag{2.8}$$

where A_i and C are constants.

Consequently, we fit the perturbed profile $\rho(z)$ with the previously derived function of the background density supposed dependence on the height, using the specified curve fitting procedure MPFIT introduced by Markwardt [2009]. For a better and faster convergence of iterations, the numerical values of A_4, A_3, A_2, A_1, A_0 can be used as the first guess from the previously done fit of the perturbed local buoyancy frequency squared. For ρ_0, the density value at the lower boundary can be chosen and the integration constant C as a correction to ρ_0 can be assumed to be zero as a first guess.

2.1.2 Data description and methodology

The method of the density profile separation and identification of IGWs is illustrated using data from the Constellation Observing System for Meteorology, Ionosphere and Climate [FORMOSAT3/COSMIC Anthes et al., 2008]. We have analysed a sample of 60 observational events spatio-temporally nearest (prior as well as subsequent) to the 2011 Tohoku earthquake. From this sample, one representative RO profile was chosen from 2011-03-11 04:08:29 UTC. This occultation event was used for illustration of the results where only a single profile is used. The dry temperature profile of this event is depicted in Fig. 2.2, note the inversion layer around 10 km altitude and the significantly perturbed tropopause. The

whole sample of 60 events was averaged for the analysis and results described in subsection 2.2.2.

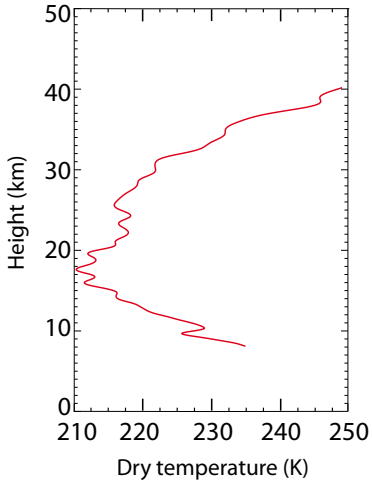

Figure 2.2: The dry temperature profile corresponding to the occultation event 2011-03-11 04:08:29 UTC.

Results are computed using the dry density and temperature profiles only. The vertical extent of our analysis is 8–40 km and it includes the areas of two basic inversion algorithms for GPS RO – geometrical optics (GO) and radio holography (RH). At the COSMIC Data Analysis and Archive Center RH (Full spectrum inversion method in this case) is applied from ground to the upper troposphere and GO from there to the top of the profile. The vertical resolution is therefore variable across the height profile, but not less than 1.5 km [Melbourne, 2005]. For details about the algorithms and the discussion of their usage possibilities see e.g., Tsuda et al. [2011].

The used methodology is limited because the tropopause is included in the investigated vertical range too. In the tropopause region assumptions of the Wentzel, Kramers and Brillouin (WKB or ray tracing) theory, that background velocity and Brunt-Vaissala frequency vary slowly over a wave cycle [Fritts and Alexander, 2003], are violated. Using this theory we are treating the wave packets as particles moving along rays [Sutherland, 2010] and using the WKB approximations we can relate the wave frequency to a wave spatial characteristics and background atmosphere properties through a dispersion relation as shown by Fritts and Alexander [2003]. Polarization relations used for the determination of the IGW parameters from temperature or density profile as proposed by Gubenko et al. [2011] are also a product of WKB approximations.

In the region where the ray theory is not valid, we cannot consider the amplitude envelope as slowly evolving and the relative phase, amplitude and spatial extent of the wave packet may change significantly as it evolves over time. Therefore the results are separated into the parts including the whole profile (8–40 km) and only the stratosphere (tropopause – 40 km). In the whole vertical range, the typical analysis method of vertical wave number spectral density [e.g. Steiner and Kirchengast, 2000; Tsuda and Hocke, 2002; Tsuda et al., 2011] wouldn't give much sense as a consequence of the facts discussed above.

Relying only on the linear theory, we will present the results of our method

for background separation in the extensive region using the continuous wavelet transform and its skeleton as used by Chane-Ming et al. [2000]. The wavelet transform was computed using the Morlet wavelet and algorithms proposed by Torrence and Compo [1998].

After the subtraction of the background density, the resulting density perturbations are further normalized and scaled by the square root of the background density to conserve the kinetic energy as suggested by Hines [1960] and discussed by Sica and Russell [1999]. Nevertheless, for the results in the stratospheric region alone, we use only the normalized density perturbations. Thus, we are able to compare directly the vertical wavenumber spectra of the density and temperature perturbation and evaluate them with theoretical model spectra. This distinction, where needed, is emphasized in following text.

2.2 Advantage of using density profiles and a remark on the connection between density and temperature backgrounds

The goal of this subchapter is to present the results of application of our method for the background separation from GPS RO density profiles. However, the issue of background determination is not satisfactory solvable and it is also not possible to identify with certainty the ideal background profile. Therefore, where possible, the results are shown for more types of the Brunt-Vaissala frequency background profile fits and then their qualities are discussed.

2.2.1 The background issue

In this subsection, the fits of the Brunt-Vaissala frequency squared are shown and discussed. Consequently, after application of our method, the resulting normalized and scaled density perturbation profiles are given. Additionally, the corresponding forms of the temperature background are computed using the equation of state for dry air and the hydrostatic balance. Reversely, background density profiles are calculated for various forms of background temperature functional dependence on altitude.

In the stratospheric region (in our case from tropopause up to 40 km), the results are derived using polynomial fits of $N^2(z)$, starting with the second order polynomial and ending with the fourth order. The second order is the lowest appropriate one to capture the supposed non-linear decay of N_0^2. On the other hand, higher order than the fourth order polynomial can oscillate and even the fourth order polynomial may gain a wave-like form in this narrow region. Taking into consideration the slow changes of N_0^2 in the stratosphere and the limited vertical extend of our region, we must be careful not to exaggerate the polynomial order, which could result in filtering of waves with vertical wavelengths in the interval of the IGW likely appearance.

Chane-Ming et al. [2000] introduced a review of studies focused on the vertical wavelengths of dominant modes. In the area of our vertical range, the dominant vertical wavelengths were smaller than 10 km. Considering the limited height range, Wang and Alexander [2010] suggested to limit the analysis to vertical

scales up to 15 km. Steiner and Kirchengast [2000] and most of other relevant studies applied the analysis threshold of IGWs around 10 km of the vertical wavelength. The lower boundary of analysed IGW vertical wavelengths should be determined by the Nyquist frequency arising from the vertical resolution, but since the discussion about the analysis range for studying IGWs is still open [Luna et al., 2013] we will examine a broad range of vertical scales (we will come back to the question of suitable IGW vertical wavelengths from GPS RO analysis in Appendix to this chapter showing an additional robust analysis of relation between the temperature and density profiles.

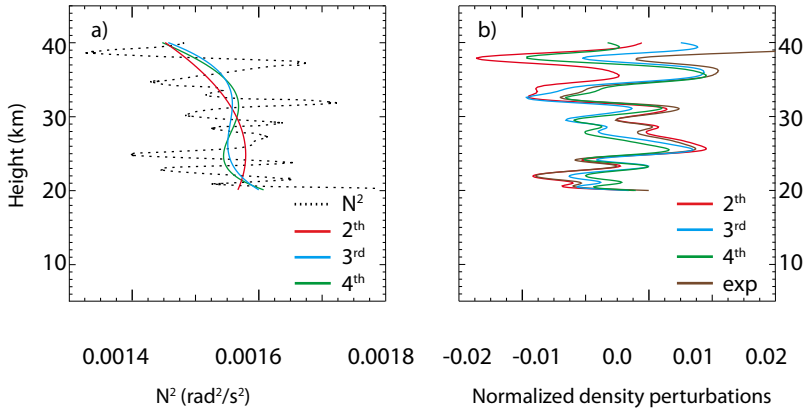

Figure 2.3: a) Various (polynomial) fits of the local buoyancy frequency squared ($N^2(z)$) in the stratosphere, b) corresponding normalized density perturbations.

Having the N_0^2 profile in the form of a polynomial, the analytical form of the density background can be easily inferred from (2.8). An illustration of fits to the perturbed local buoyancy frequency squared and corresponding density perturbations are shown in Fig. 2.1 and 2.3. In Fig. 2.3 we added also the density perturbations resulting from the subtraction of the background in the form

$$\rho(z) = \rho_0 exp(-az), \tag{2.9}$$

corresponding to an isothermal atmosphere. The other forms of the density background which we use are not so easily transferable in the temperature space.

For example, even when we consider the density background derived from linear N_0^2 profile

$$\rho(z) = \rho_0 exp\left(-\frac{\frac{A_1 z^2}{2} + A_0 z}{g} - C\right). \tag{2.10}$$

The background temperature profile can be derived as follows. Assuming that the background state is the same for all quantities of state and that the background state is in hydrostatic balance, we use the equation of state for dry air and the hydrostatic balance equation. After solving the differential equations, the resulting background temperature profile has a quite complicated form:

$$T_0(z) = exp\left[\frac{(A_1 z + A_0)^2}{2A_1 g}\right] \cdot K - \frac{exp\left[\frac{(A_1 z + A_0)^2}{2A_1 g}\right] g^{\frac{3}{2}} \sqrt{\frac{\pi}{2}} Erf\left(\frac{A_1 z + A_0}{\sqrt{2A_1 g}}\right)}{\sqrt{A_1} R}, \tag{2.11}$$

where K is an integration constant, g is the gravitational acceleration and R is the gas constant.

If we analyse an inverse problem and try to find out which background density profile corresponds to frequently used temperature backgrounds, we discover that the polynomial fits of temperature are leading to an unphysical background in density space, generally looking like:

$$\rho_0(z) \approx \frac{exp[-f(z)]}{pol}, \tag{2.12}$$

where f is non specific function of z and *pol* denotes the original polynomial. Otherwise, if the temperature background is written using goniometric functions, the resulting density profile behaves like an exponential, but $f(z)$ has a very complicated form. Supporting information on the correspondence between density and temperature profiles is given in the appendix to this chapter.

In the region from 8 km to 40 km, the task of the background separation is more complicated because we have to deal with the jump of N_0^2 across the tropopause. For demonstration of our method, we have chosen to fit the N^2 profile with polynomials of fourth, fifth and sixth order. The corresponding density background forms can be easily inferred from (2.8). The local buoyancy frequency squared fits and corresponding density perturbations are then illustrated in Fig. 2.1.

As we can see in the left of Fig. 2.1, the sixth order polynomial gains a wave-like pattern with vertical wavelength of approximately 20 km, which is out of the interval of vertical scales normally considered to be gravity wave induced in the GPS RO literature. Taking into account the temperature profile shown in Fig. 2.11, we can see on the right of Fig. 2.1 that the tropopause (around 18 km) does not cause enhancement of the gravity wave activity unlike the inversion layer around 10 km altitude. The method isn't able to assign correctly such meteorological phenomena with sharp changes of characteristics (as inversion layers) to the background making them an artificial source of IGWs. In general, the perturbations are not biased and are centred around zero and due to the scaling, the decrease of the wave activity in the stratosphere is clearly visible.

2.2.2 Comparison with temperature – in the stratosphere

In this section, we present a comparison between the normalised density perturbations resulting from the polynomial fits of $N^2(z)$ and the fit of the temperature profile with a cubic polynomial as used, e.g., by Gubenko et al. [2008]. We have computed the mean vertical wavenumber power spectrum for the sample of sixty profiles. The vertical wavenumber spectra of the normalized density and temperature perturbations are further compared with the theoretical shape of saturated spectra for density perturbations induced by IGWs as derived by Senft and Gardner [1991].

The mean power spectral densities (PSD) computed for various backgrounds and dry temperature for sixty profiles are depicted in Fig. 2.4. The presented range of vertical wavelengths is from 20 km to 400 m. The lower frequency limit was chosen due to the vertical extent of the examined area and the higher frequency limit was chosen as 2 times the Nyquist frequency of the interpolation

of data to the height profile. According to Luna et al. [2013], the adequate choice of the wavelength cut-off for studying gravity waves through RO measurements is still open. For example, Wang and Alexander [2010] limited their analysis to vertical scales between 4 and 15 km, Tsuda and Hocke [2002] applied no minimum wavelength limit but Markwardt and Healy [2005] argued that in the altitude region below 30 km only temperature fluctuations with vertical wavelengths more than 2 km can be safely interpreted as originating from small-scale atmospheric waves.

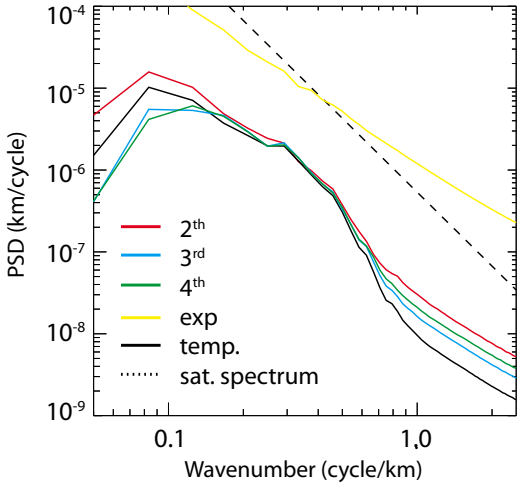

Figure 2.4: Vertical wavenumber power spectral density for the normalized density perturbations from different backgrounds and temperature normalized perturbations compared with the theoretical saturated spectra.

In Fig. 2.4 in the lower wavenumber area, the influence of the background separation is dominant with higher order fits giving lower powers. The fourth order fit has the maximum shifted from roughly 12.5 to 7.5 km unlike the other density and temperature fits. That is a clear consequence of the higher possibility for this order fit to gain a wavelike form in this area (see Fig. 2.3, part a) and therefore the decrease of power in the longwave part. For determination if such a longwave mode is caused by IGW, more information (temperature, velocity and so on) would be needed and then the polarization relations would have to be examined.

Between the vertical wavelength of about 10 km and 2.5 km, the spectra of density from all three fits and the temperature spectrum are similar and their slopes are in good agreement with the theoretically predicted slopes. An important feature emerges at approximately 2.5 km, where the temperature fluctuation spectrum begins to decrease more rapidly than the density spectra regardless of the fit order. This should clearly be the consequence of the usage of hydrostatic balance in the temperature data retrieval, which excludes non-hydrostatic waves and according to Steiner and Kirchengast [2000] acts to suppress wave amplitudes. The exponential fit (isothermal atmosphere) is failing to give correct orders of PSD values as well as to capture the feature of saturated theoretical spectrum.

2.2.3 Comparison with temperature – over the whole profile

Since we cannot rely on the theory (derived using WKB approximations) in the full vertical extend, we have chosen the continuous wavelet transform (CWT) [Torrence and Compo, 1998] method for the analysis of IGWs. Using CWT, we can study the behaviour of our method for the background separation in the tropopause region and track the gravity wave activity behaviour depending on altitude. Further, for the determination of dominant modes and for studying their development with height, we have applied a method of reducing CWT to its skeleton. We have used the same setting for drawing the spectral lines as Chane-Ming et al. [2000].

In Figs. 2.5, sections a, b, c, the CWTs are shown for the scaled normalized density perturbations, resulting from different background fits. For comparison, CWTs of the unscaled normalized density perturbations resulting from 6th and 5th order polynomial fit are depicted in Fig 2.5, sec. d, e. Finally, the normalized temperature perturbations resulting from the separation of the background in the form of a sixth order polynomial are shown in Fig. 2.5, sec. f.

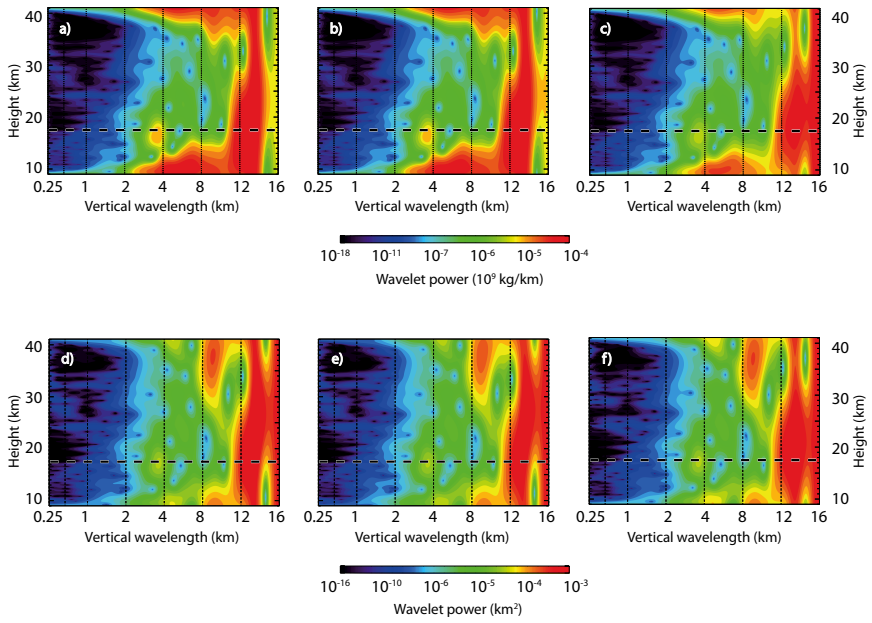

Figure 2.5: Wavelet power spectra of various perturbations resulting from different background states: Scaled normalized density perturbations resulting from the 6th (a), 5th (b) and 4th (c) order polynomial fits; Normalized density perturbations resulting from the 6th (d) and 5th (e) order polynomial fits; Normalized dry temperature perturbations resulting from the 6th (f) order polynomial fit.

The differences between CWTs of different fits as well as between the scaled and unscaled perturbations point to characteristic features. The differences between different fits are, as expected, most pronounced in the region of the longer vertical wavelengths. This is even more evident comparing the skeletons of CWT in Fig. 2.6. In general, the difference between the scaled and unscaled perturbations is mainly connected with the theoretically predicted decrease of wave activity in the stratosphere and the shift of wave activity to longer wavelengths with height due to gradual filtering of small wavelength IGWs above the tropo-

sphere [Fritts and Alexander, 2003]. That is captured better by the CWT of the scaled perturbations. Nevertheless, there are also unexpected differences between the scaled and unscaled CWTs in the region of wavelengths around 16 km, where the amplitude is around one order higher and constant with height for the scaled case. The temperature CWT behaves similarly to that of the unscaled density perturbations.

Spectral lines of the CWTs maxima are plotted in Fig. 2.6 in the same order as in Fig. 2.5. From comparison of Fig. 2.6, sec. a – c, we can infer that the dominant mode with the wavelength around 16 km is probably caused by low quality separation of background. We can argue this, because in the cases of the temperature perturbations (Fig. 2.5 and 2.6, sec. f) and fourth order $N^2(z)$ fit (Fig. 2.5 and 2.6, sec. c), where this mode is the strongest, it has height-constant wavelength regardless of the variable local stability.

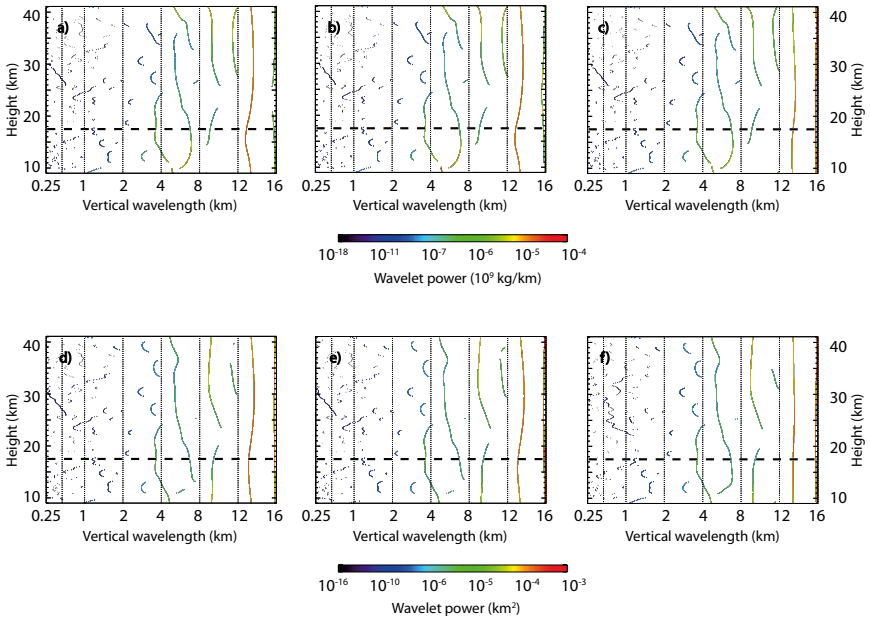

Figure 2.6: Wavelet power spectra skeletons of various perturbations resulting from different background states: Scaled normalized density perturbations resulting from the 6th (a), 5th (b) and 4th (c) order polynomial fits; Normalized density perturbations resulting from the 6th (d) and 5th (e) order polynomial fits; Normalized dry temperature perturbations resulting from the 6th (f) order polynomial fit.

Other dominant modes occurring across the different fits and across the whole profile have scales around 13 km and approximately from 5 to 7 km. The skeletons should be compared with the N_0^2 profiles in Fig. 2.1, sec. a, because according to, e.g. Chane-Ming et al. [2000], the change of the vertical wavelength of individual modes with height should be inversely proportional to the local stability.

In our case, in the upper stratosphere, the decreasing wavelength of modes with height suggests the role of non-linear wave processes acting to limit the amplitude, which is reflected by the cessation of the wavelength grow. On the other hand, in the upper troposphere/lower stratosphere region where, in our case, the stability is growing, we can find increasing wavelength with height, which is to see for example for the mode with scale of 5–7 km in Fig. 2.6, sec. a.

This suggests the possible enhancement of this mode's amplitude by means of wave-wave interactions or possible amplification resulting from some instability in the tropopause region.

2.3 Summary and conclusions

The aim of this chapter was to present and evaluate a method for the analysis of IGWs from GPS RO density profiles. After a description of the method for the density background separation, the methodology suitable for evaluation and comparison of resulting IGW analysis from density profiles with those from temperature profiles was discussed with regard to the theory.

In the process of deriving our method in section 2.1, we described the quantities defined by equations (2.2), (2.3) and (2.4) as the Brunt-Vaissala or buoyancy frequency. Although it helps to introduce and to understand the separation method, it might be also seen as phenomenologically inconsistent. Such a description would be correct for a fluid where the density is independent of pressure and the density of a fluid particle could be considered unchanged when displaced vertically [Sutherland, 2010]. In our case this is not true. Therefore, a different naming convention should be used for these quantities in future works. For example, using the density scale height (H_ρ) definition [e.g. Sutherland, 2010], (2.3) can be written like

$$\frac{g}{H_\rho}. \tag{2.13}$$

Nevertheless, the naming convention has not any influence on the results.

In section 2.1, the problems of the background choice are discussed, correspondence between analytical forms of density and temperature background profiles is examined and the unphysicality of polynomial temperature fit is shown by solving related differential equation.

A very important result is shown in subsection 2.2.2. In the high wavenumber region from approximately 0.4 cycle/km (2.5 km vertical wavelength), the power spectrum of the normalized fluctuations derived from dry temperature GPS RO data has lower values than for normalized density fluctuations regardless of the background type. Steiner and Kirchengast [2000] argued that the likely causes of a tendency of increasing underestimation of spectral power toward higher wavenumbers are the GO technique, the local spherical symmetry assumption, and hydrostatic equilibrium assumption. The latter is confirmed by our results.

Steiner and Kirchengast [2000] stated that above mentioned causes of underestimation are acting to suppress the wave amplitudes and therefore to lower the PSDs. But this is not the only implication. Filtering the data with the hydrostatic balance in the standard temperature retrieval leads to the fact that in the equations of motions of the filtered field also the dynamic pressure and density fluctuations are in hydrostatic balance. That results in vanishing of vertical accelerations. The consequence is that the whole group of non-hydrostatic IGWs is filtered out. According to Sutherland [2010], those are waves with frequency close to the buoyancy frequency and waves with phase line slopes significantly different from zero. It should be noted that these waves with smaller ratio of

horizontal to vertical wavelength may be already filtered by the local spherical symmetry assumption.

In the subsection 2.2.3, the results of CWT and its skeleton do not reveal noticeable differences between the dry temperature and density based IGW analyses. Rather the differences between the scaled and unscaled cases are more important and the role of subjective background choice has visible influence on the longer modes. We do not discuss the propagation or sourcing of waves, which could be tempting especially in the case of modes emerging or ending suddenly in the profile. The GPS RO measurement is a "snapshot" of the real atmosphere and no information from the cotangent space is included. Hence, using just a simple profile, we cannot say if or in which direction the mode propagates. Special care must be taken also when interpreting the results of the CWT skeleton because the vertical behaviour of modes could be also a consequence of fluctuation retrieval rather than of physical processes. That can be seen in Fig. 2.6 from comparison of the dominant modes between 4 and 8 km vertical wavelength in the upper stratosphere.

In summary, the analysis of IGWs with the GPS RO density profiles bears advantages over the analysis from dry temperature profile. Primarily, unlike the temperature, the density background has a familiar form and the consequent analysis is not restricted to hydrostatic waves only.

Appendix – Modification of Šácha et al. [2014] conclusions

Introduction and method

Dry temperature data from Global Positioning System radio occultation (GPS RO) are widely utilized for IGW analyses [e.g Steiner and Kirchengast, 2000; Gubenko et al., 2008; Faber et al., 2013; Hindley et al., 2015] and the potential of this sounding technique is most likely to grow [Wickert et al., 2009] in the future, especially with the upcoming launch of the COSMIC-2 constellation [Cook et al., 2016]. However, Šácha et al. [2014] argued for utilizing of the GPS RO dry density profiles, because density is the first quantity of state obtained in the GPS RO retrieval without the hydrostatic balance utilization. Also, from a basic theory, the density background has a familiar form. In this appendix some conclusions given in the second chapter are modified and a strong and clear argument for utilization of the dry density instead of dry temperature GPS RO data is provided. Thanks to the related analysis, we have pointed out some flaws in the OPS5.4 version of reprocessed GPS RO data produced by Wegener Center in Graz. Those were used by Šácha et al. [2014] and supported the incorrect conclusions on the spectral differences between dry temperature and density data and on the filtration of non-hydrostatic IGWs.

To compare IGW related information, it is necessary to subtract identical backgrounds from the dry temperature and density profiles. Following the method and notation from the second chapter we assume that the background atmosphere is in a state of hydrostatic balance. Therefore only two quantities of state (ρ_0, T_0, p_0) are independent (in the stratosphere) and we can express background pressure as a function of the background density:

$$p_0(z) = \hat{p_0} - g \int_z^{z_0} \rho_0(z)\, dz. \tag{2.14}$$

Note that we have chosen to integrate the hydrostatic balance equation from bottom up. Than by substituting into the equation of state for dry air and using (2.2) we have a background temperature in the form:

$$T_0(z) = \frac{p_0(z)}{R\rho_0(z)} = \frac{\hat{T}_0}{\exp(-a(z))} - \frac{g}{R} \cdot \frac{\int_z^{z_0} \rho_0(z)\, dz}{\rho_0(z)}, \tag{2.15}$$

where $\hat{T}_0$ is the value of the background temperature at the lower boundary. (2.15) is a general expression which can be further simplified depending on the form of $a(z)$ or better on the type of an analytical function chosen for fitting $N_0^2(z)$. For fitting the dry temperature profile with (2.15) we can use numerical values of constants A_i, C from previous fits (see the density background separation method introduced in 2.1.1 and iterate for $\hat{T}_0$ only.

This procedure ensures that we have the density and dry temperature perturbations as a product of subtraction of identical backgrounds. This is a necessary condition for future analysis and comparison of the resulting perturbations.

Preliminary results and general conclusion

Steiner and Kirchengast [2000] and Šácha et al. [2014] argued that the utilization of hydrostatic balance in the GPS RO dry temperature retrieval can act to suppress wave amplitudes and therefore to lower the PSD of dry temperature perturbations. However, we must note that this hypothesis is not correct as the hydrostatic equation does not have any filtering effect on the amplitude of oscillations. It just shifts the phase of harmonics (easy to inspect by substituting a monochromatic plane wave or more general wave types into the hydrostatic equation). Thus, there should not be any difference between PSD from density and coupled dry temperature perturbations.

In practice, during the dry temperature retrieval (or pressure at first step of the retrieval), the hydrostatic balance is initialized at the upper boundary going from upper to lower levels. There are two possible sources of discrepancies between dry temperature (or pressure) and density profiles. The first one stems from the estimation of pressure value at the upper boundary and the second stems from the vertical shift due to the discretization of the hydrostatic balance equation in the vertical. But, these two factors should not play any role in PSD of disturbances. So, the correspondent dry temperature and density PSDs (Fig. 2.4) should be similar. The only possible difference can stem from non-uniformity of the correspondent dry temperature and density backgrounds due to the numerical implementation of the method.

In the ongoing research done by a collective of authors as in Šácha et al. [2014] it has been confirmed that the background separation method is very accurate. Only in the lower vertical wavenumber region (from around 8 km vertical wavelength), bigger differences can be found in the spectral amplitude. This corresponds to the harmonics limited by the vertical extent of the analysis (20–35 km in this case). The spectral extent, where the density and temperature PSD are in agreement, can be used as an indicator for vertical wavelengths allowed for the analysis. On the other side, in the higher wavenumber region (approximately 1 km and smaller vertical wavelengths) oscillations emerge in the PSD differences that are connected with the data quality. Those differences cannot be connected

with the quality of the above-described method, because the method itself does not affect the small-scale structure of perturbations.

The preliminary results show that the differences between dry temperature and density PSDs in Fig. 2.4 are spatio-temporally variable and are connected with some yet undetermined flaws in the OPS5.4 version of reprocessed GPS RO data produced by Wegener Center in Graz. In the newer version, OPS5.6, the differences between the dry temperature and density PSDs are almost negligible (up to a few % relative to the respective modal amplitudes) confirming a good performance of the method of identical background separation.

General conclusion for the issue of utilization of the dry temperature versus density GPS RO data is that the GPS RO dry temperature does not represent the true temperature that would correspond to the observed atmospheric density (GPS RO dry density data). The GPS RO dry temperature profile is just a density derivative that would be equal to the true temperature only in the case of the atmosphere being in the state of a hydrostatic balance. More illustratively, if we assume that the background state is in the hydrostatic balance, then the GPS RO dry temperature perturbations would be equal to the real temperature perturbations only if the perturbations would be created by hydrostatic waves exclusively. Otherwise, in the presence of non-hydrostatic IGWs, the GPS RO dry density perturbations (observable) would correspond to different temperature perturbations, if we were able to measure density and temperature simultaneously.

This is a clear message, that GPS RO temperature data should not be used for IGW analyses. In future work, it would also be desirable to quantify the effect of deviation from the hydrostatic balance assumption on the total quality of the GPS RO temperature data, because the GPS RO data are frequently used in the climate research as well as assimilated into the numerical weather prediction models. Such an analysis could be done e.g. by comparing temperature profiles from fully non-hydrostatic IGW resolving models with temperature profiles derived from model density profiles using the same retrieval as for the GPS RO data.

3. Enhanced internal gravity wave activity and breaking over the Northeastern Pacific – Eastern Asian region

A number of observational studies have examined the climatological structure of the IGW characteristics using RO data from Challenging Minisatellite Payload [e.g. Tsuda et al., 2000; Schmidt et al., 2008], Formosat Satellite Mission 3 – Constellation Observing System for Meteorology, Ionosphere, and Climate (FORMOSAT–3–COSMIC) [e.g. Alexander et al., 2009; Horinouchi and Tsuda, 2009; Wang and Alexander, 2009] or using High Resolution Dynamics Limb Sounder (HIRDLS) [e.g. Ern and Preusse, 2012] or Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) – Thermosphere Ionosphere Mesosphere Energetics Dynamics (TIMED) satellite data [e.g. Zhang et al., 2012]. But, as noted by Wang and Alexander [2010], despite significant advances in our understanding of IGW and their effects in different regions of the atmosphere in the past few decades, observational constraints on their parameters are still sorely lacking, especially for momentum fluxes and IGW propagation direction.

This chapter is focused on an analysis of the IGW activity in the lower stratospheric region bounded roughly by a tilted ellipse with end points near 30°N 120°E and 60°N 180°E. As we will show in this paper, this area is often a part of a larger region of enhanced IGW activity or breaking, or anomalies of some fields (ozone, wind) that, however, also may reach slightly beyond this region, which is a consequence of large-scale dynamical processes. Therefore, in the following our region of interest is roughly defined as this region, but we would refrain from presenting rigid boundaries.

The chapter is structured as follows: the remainder of the Introduction provides a list of results of observational studies of the IGW activity or characteristics relevant to the region of interest. Description of data and methodology used for processing of the GPS RO data are provided in the next section. The following subchapter 3.1 starts with an introduction of climatology of the stratospheric low annual cycle area and then, primarily, results of the analysis of IGW activity and stability in the region of interest are presented. The Results section smoothly passes into the Discussion subchapter 3.4, where the results of the wind direction and its change analysis are presented together with discussion of possible wave sources. Possible reasons why this IGW hotspot area was not discovered before are discussed in the Discussion section together with the appropriateness of potential energy of disturbances (E_p) as a wave activity proxy and with possible implications for the middle atmospheric dynamics (longitudinal variability of BDC, creation of planetary waves with effects on the polar vortex stability and stratosphere–troposphere exchange). The summary and conclusions are presented in the last section of this chapter.

3.1 Satellite studies of wave activity

There have been a number of studies that dealt with IGW activity globally. In the following we will give a brief summary of the results bearing some information on the IGW activity, characteristics and peculiarities over our region of interest. Alexander et al. [2008] using 2006/2007 northern hemispheric winter data from FORMOSAT–3/COSMIC found that the E_p of IGWs (vertical wavelength < 7 km) is mostly related to the subtropical jet stream with some regional-scale contributions from orography. Among other areas they have found a 2006/2007 winter mean 17–23 km E_p maximum above Japan and suggested that it is due to orographic waves coinciding with strong sub-tropical jet wind speeds in this area.

McDonald et al. [2010] studied geographic variation of the RMS (root mean square) temperature difference between pairs of FORMOSAT–3/COSMIC profiles, and the maps of RMS at 15 and 30 km altitude show enhancements in the northern hemispheric subtropics in the vicinity of the Gulf of Mexico and the Kuroshio stream. These regions have previously been identified as regions of strong gravity wave activity associated with convection [Preusse et al., 2001; Jiang et al., 2004; Preusse and Ern, 2005]. Their results demonstrate that IGW activity dominates the variability observed in stratospheric temperature at time and spatial scales often used in validation studies, and they suggested that much of the seasonal variability observed at higher altitudes may be due to changes in the IGW propagation conditions in the lower stratosphere, while the wave field may be particularly affected by changes in the zonal wind field between 15 and 25 km; in particular, longer horizontal wavelength waves (with smaller phase speed c_p) may be preferentially removed by critical level filtering in this region.

Wang and Alexander [2010] presented global maps of seasonal mean IGW amplitude, vertical and horizontal wavelength, intrinsic frequency to Coriolis parameter ratio and horizontal wave propagation. Contoured maps of December 2006 to January 2007 vertical wavelengths in the altitude range of 17.5–22.5 km reveal that in our area of interest the leading mode has shorter vertical and horizontal wave-length than in other equivalent latitudes areas while having comparable amplitude. This might suggest higher buoyancy frequency (stronger stratification) values in the area of interest, or this could be an effect of different IGW sources or different background wind structure in this region.

Faber et al. [2013] calculated momentum fluxes connected with IGWs together with their vertical and horizontal wavelengths using sets of three collocated FORMOSAT–3/COSMIC RO profiles. Their results show that in summer 2006 as well as in winter 2006/2007 there are regions of longer dominant vertical and horizontal wavelength and corresponding increased momentum flux above our region of interest in the altitude range of 20–30 km.

Wright and Gille [2013] used the S–Transform method to allow for the detection of multiple overlapping waves, and they found that including these waves changes the observed distribution of IGW momentum flux. An overall 68% increase in measured momentum flux was observed for the 20–30 km altitude range, with significant regional variability. Among others, they found enhancement of the relative importance of the Himalayas on a global scale

Wright et al. [2011] have compared HIRDLS, COSMIC and SABER for the detection of stratospheric IGWs and concluded that, taking into account different

vertical resolution of these instruments, all of them reproduce each other's results for magnitude and vertical scale of waves in approximately 50 % of cases, although COSMIC has a positive frequency and temperature variance bias. This should be supposedly due to the higher vertical resolution.

Using data from HIRDLS, Ern and Preusse [2012] computed spectral distribution in terms of horizontal and vertical wavenumber for IGW momentum flux. They labelled the southern edge of our area of interest as a deep convection area, and at 25 km altitude they computed mean IGW momentum fluxes (from June to August) to range from -2.7 to -3.2 log 10 Pa from south to north in the area, and these values do not stand out in comparison to values in other regions. For November, Ern et al. [2011] found one of the largest IGW momentum flux values north of Japan at the altitude of 30 km. Their analysis of SABER/TIMED gravity wave momentum flux reveals increasingly interesting patterns with altitude up to 70 km above the region of interest.

Zhang et al. [2012] studied the activity of IGWs from SABER/TIMED temperature profiles between January 2002 and December 2009. For vertical wavelengths between 2 and 10 km integrated over a layer between 21 and 45 km they found in all seasons only small values of E_p, not exceeding 1.5 J·kg^{-1} above the area of interest. Further south they identified the region of Southeast Asia as a region with large E_p values corresponding to the tropical deep convection.

Ern et al. [2013] introduced projects having addressed IGWs in the priority program Climate and Weather of the Sun–Earth System. They showed global distributions of IGW (vertical wavelength range 4–10 km) momentum fluxes derived from SABER temperature data at 25 km altitude, averaged over the years 2002–2006 for the months of January, April, July and October. They did not find exceptionally large fluxes above the north-western Pacific region in any season. Again, south of the area, in April, July and October there is a clearly visible local maximum region likely due to IGWs generated by deep convection.

John and Kumar [2012] studied IGW activity from the stratosphere to the lower mesosphere in terms of their potential energy. Averaging SABER data from 2003 to 2006 and in the region 20–60 km they presented monthly mean global maps of IGW potential energies. Maximal values are ranging up to 120 J·kg^{-1} (above the Andes and above Scandinavia in winter) but in the area of interest the IGW potential energy does not exceed 20 J·kg^{-1} in any month and interestingly there are not expected areas of higher E_p connected with deep convection. Potential energy averaged in the 60–80 km height range exhibits almost the same distribution across the globe.

Baumgaertner and McDonald [2007] highlighted the role of background winds in the gravity wave distributions and there are some papers where information about wind patterns above the area of interest can be found. Oberheide [2002] computed zonal and meridional geostrophic wind fields in the altitude range of 20–90 km from the Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere and we can see on the 9 November 1994 by zonal winds at 1 hPa (above the upper boundary of RO analysis) that the jet avoids the region of interest. Verkhoglyadova et al. [2014] computed geostrophic wind maps from a simulated data set based on COSMIC RO for the upper troposphere–lower stratosphere region and for example in January 2007 one can see at the 200 hPa constant dry pressure surface the highest zonal wind speeds starting above Japan and contin-

uing further eastward above the Pacific. This feature also emerges in January 2009 as shown by Scherllin-Pirscher et al. [2014] (their Fig. 2). The 200 hPa level is below the usually defined lower boundary of GPS RO IGW analyses, but conditions of the upper troposphere influence the upward propagation of IGWs into the stratosphere.

3.2 Data and methodology

To describe the background climatology over our region of interest, we have analysed the MERRA [Modern Era Reanalysis for Research and Applications; Rienecker et al., 2011] and JRA-55 [Japanese 55-year Reanalysis; Ebita et al., 2011] series for the 1979–2013 time interval. The data were analysed on a monthly basis in the horizontal resolution of $1.25° \times 1.25°$. We have studied temperature, geopotential height, ozone mixing ratio and zonal wind. To examine the annual cycle amplitudes, we used a continuous wavelet transform [CWT; e.g. Torrence and Compo, 1998; Percival and Walden, 2006] that delivers the amplitude of a detected oscillation as a function of both frequency and time. To study spatial variation of the annual cycle amplitude, we applied an extension of the CWT – the pseudo-2D wavelet transform Pišoft et al. [2009]; Pisoft et al. [2011]. The analysis was computed using the Morlet mother wavelet, with the wavenumber Ω_0 equal to 6. To construct the 95 % confidence intervals, we have employed a significance test by Torrence and Compo [1998]. For the interpretation only statistically significant results outside the cone of influence were selected.

To investigate the IGW activity in the area of interest compared to the other regions, we have analysed L2 level FORMOSAT-3/COSMIC RO data [Anthes et al., 2008] on a $3° \times 3°$ grid from 2007 to 2010. GPS RO data proved to be a very useful tool for atmospheric monitoring and studies. They are frequently used for analyses of the IGWs in the upper troposphere–lower stratosphere region. GPS data are characterised by a good vertical resolution providing atmospheric profiles with global coverage under all weather and geographical conditions [Foelsche et al., 2008]. Atmospheric density, as was described in the previous chapter, is the first quantity of state gained in the GPS RO retrieval process and is not burdened by any additional assumptions, as it is the case, e.g., with temperature. According to the linear theory of the IGWs, a separation between a small wave–induced fluctuation and background field has to be performed. As shown by Markwardt and Healy [2005], small-scale fluctuations of dry temperature RO profiles can be interpreted with certainty as IGWs, when the vertical wavelength is equal or greater than 2 km. To separate the perturbations, we applied a method described in Šácha et al. [2014] for the density background separation. The method is based on fitting the buoyancy frequency height profile and on the consequent analytical derivation of the background density functional dependence on altitude. After the background separation and normalisation of the disturbances, we obtain one normalised density perturbation height profile for each occultation.

Šácha et al. [2014] argues that the usage of density profiles bears many advantages, for example the inclusion of non-hydrostatic waves. Those are the waves with frequencies close to the buoyancy frequency and with phase line slopes significantly different from zero Sutherland [2010]. These waves are often too small to be resolved by circulation models and according to CCMVal [2010] such un-

resolved (10–1000 km) IGWs play a significant role in stratospheric circulation, driving nearly a quarter of the stratospheric circulation in comprehensive models.

The lower boundary of analysed IGW vertical wavelengths should be normally determined by the Nyquist frequency arising from the vertical resolution of the occultation technique, but since we are more interested in the relative distribution of IGW activity than the absolute values we are not making any vertical wavelength cut-off at the lower boundary. We assume the noise to be almost independent of the geographical location, which is, however, generally not true Markwardt and Healy [2005] and so our calculated distributions of IGW activity can be partly affected by the spatio-temporal distribution of noise. The discussion on adequate choice of the wavelength cut-off for studying gravity waves through RO measurements is still open Luna et al. [2013]; thus, we decided not to make any cut-off even at the upper boundary of vertical wavelengths. Considering the geographically variable vertical extent of our analysis (from the tropopause up to 35 km), we must note that the IGW modes with vertical wavelengths comparable or greater than the vertical range (≈ 15 km and more) will be increasingly underestimated with increasing tropopause altitude. But because these modes have vertical wavelengths many times longer than the most energetic modes (2–5 km in the lower stratosphere; Fritts and Alexander [2003]), we do not expect this underestimation to significantly affect our results. The geographical distribution of IGW activity is almost unchanged in an analysis using a vertical wavelength threshold proven to be the right choice for GPS RO data, as it is shown in Appendix to chapter 2.

Because single GPS measurement provides just a snapshot of an actual atmospheric state without any direct information from the cotangent-phase space (e.g. velocity), the choice of derivable diagnostic quantities is quite restricted. Many authors used the energy density as a measure for wave activity [e.g. Tsuda et al., 2000; Ratnam et al., 2004; De la Torre et al., 2006; Hei et al., 2008]. We have computed the mean potential energy density of disturbances per unit mass ($\bar{E}_p$) using the formula provided by Wilson et al. [1991]:

$$\bar{E}_p = \frac{1}{2} N^2 \langle \xi^2 \rangle = \frac{1}{2} \left(\frac{g}{N} \right)^2 \left\langle \left(\frac{\rho'}{\rho_0} \right)^2 \right\rangle , \tag{3.1}$$

where ξ is a wave induced Lagrangian displacement, N the Brunt–Vaissala frequency, ρ' the density perturbation and ρ_0 the background density; $\langle ()^2 \rangle$ denotes a variance. For a single profile, the variance is computed in altitude and the symbol $\langle \rangle$ means averaging across the whole altitude range of the analysis (from the tropopause up to 35 km). To obtain the results also at particular height levels, the variance is then computed across the ensemble of occultations belonging to the grid at those levels.

Tsuda et al. [2000] referred to VanZandt [1985] for a theoretical evidence that the ratio between kinetic (mostly horizontal) and potential energy is approximately constant and equal to the spectral index p (roughly 5/3 to 2) and then concludes that under the linear theory it is possible to estimate total energy from temperature observations only. If this is true for temperature perturbations it should then naturally stand for density as well. Tsuda et al. [2000] further compared the climatological behaviour of potential (from GPS Meteorology experiment) and kinetic energy of disturbances (from middle and upper atmosphere, MU, radar) around Japan and found reasonable agreement.

Ratnam et al. [2004] validated potential energy of disturbances computed from the GPS RO dry temperature profiles versus radiosondes data. They found that at most of the heights, values estimated from ground-based instruments are showing higher values. One of the reasons for this is the lower vertical resolution of GPS RO compared with radiosondes or lidar. This situation should slightly improve when using the density profiles, mainly due to the higher spectral power at lower vertical wavelengths Šácha et al. [2014]. Although, our new results presented in the previous subchapter suggest that the analysis of density profiles gives, in generall, 10% higher $\bar{E}_p$ values than using dry temperature profiles. Another source of reduction is the viewing geometry with respect to the wave fronts [e.g. Lange and Jacobi, 2003].

Nevertheless, Tsuda et al. [2000], among other approximations made by Van-Zandt [1985], whose discussion is beyond the scope of this manuscript, did not explicitly mention a crucial assumption of theoretical saturated spectra in the interval used for the derivation of the constancy of wave kinetic to potential energy ratio. For a single IGW in the rotating frame of reference, however, the equipartitioning holds between the kinetic energy in the x–z plane and the sum of kinetic energy in the y direction and potential energy (consequence of the out of phase motion in y direction vs. in the x–z plane). So, for a single IGW, the partitioning ratio between wave kinetic and potential energy is dependent on its intrinsic frequency Bühler [2014]. Approximately, we can write

$$\bar{E}_p = (1 - \frac{f^2}{\hat{\omega}^2})\bar{E}, \tag{3.2}$$

where $\bar{E}$ is the mean wave energy, f the Coriolis parameter and $\hat{\omega}$ the intrinsic frequency. We will come back to the question of how good a proxy for wave activity potential energy is in the discussion (and in Appendix to this chapter), after which we will see in our results an indication of possible masking of regions with important IGW activity.

This is why we have employed two additional characteristics (novel in GPS RO studies) to access the stability of the wave field. A stratified fluid may become unstable if disturbances can overcome the stabilising effect of buoyancy by drawing kinetic energy from the mean flow. The necessary condition for such a dynamical instability is Sutherland [2010]

$$Ri_g = \frac{N_0^2}{s_0^2} < 1/4. \tag{3.3}$$

Here Ri_g is the gradient Richardson number, N_0 the background stratification frequency and s_0 the background wind shear. Senft and Gardner [1991] have estimated the gradient Richardson number by appropriately scaling the variance of the vertical gradient of relative density perturbations, assuming that the background wind shear is negligible, and using the polarisation relations for IGWs they had

$$Ri_g = \frac{g^2}{N^4} \left\langle \left[\frac{\partial}{\partial z} \left(\frac{\rho'}{\rho_0} \right) \right]^2 \right\rangle. \tag{3.4}$$

In our analysis we search for local values instead of computing the vertical variance.

Another characteristic is used to access the overturning instabilities. According to Sutherland [2010] we define σ, the maximum growth rate of disturbances arising from Rayleigh–Taylor convective instability, as

$$\sigma^2 = \frac{g}{\rho_0}\left(\frac{d\rho_0}{dz} + \frac{\partial\rho'}{\partial z}\right). \tag{3.5}$$

The value of σ is real when waves drive the fluid to be overturning. Nevertheless, for convective instability to occur (for the waves to overturn and break), the convection growth should be faster than the wave frequency Sutherland [2010]. For hydrostatic waves, only the overturning condition suffices to access the stability.

In reality there is interplay between different types of instabilities and interactions and therefore we do not use these characteristics to find areas or seasons where values exceed the exact threshold and turbulence and mixing occurs. In fact, it is natural to expect the values of those characteristics to be below their threshold because of the effect of the observational filter Lange and Jacobi [2003], and also the probability that the occultation takes place at the exact time IGW breaking is low. Instead, we use these characteristics as a comparative hint to study the geographical distribution of possible IGW effects on the stratosphere, mainly comparing values in the region of interest with other areas. The lower the Ri_g value and the higher the value of σ^2, the higher the probability that breaking of IGWs is underway and that waves are interacting with the mean state in a region.

For all results, to illustrate large intervals of irregularly distributed values, we have followed the approach applied by Pisoft et al. [2013]. For the frequency and IGW characteristics, we use a colour scale derived from the relative frequency of the detected values. The results were sorted according to their values and subsequently split into 100 equally large groups. The groups define intervals for particular colours and every colour then represents the same number of identified values. Using this approach, even subtle features of the illustrated fields can be displayed. On the other hand, this technique leads to a highly non-uniform scaling that has to be reflected in the subsequent interpretation of the results.

3.3 IGW activity and the EA/NP hotspot

Here, we present selected results consisting of a) presentation of the anomaly over the region of interest using the wavelet analysis and climatology of the annual cycle, b) IGW analysis describing the spatial distribution of potential energy of the disturbances, c) analysis of the Richardson number and sigma indicating wave breaking and d) study of possible wave sources with the use of cumulative wind rotation analysis. Supplementary and more detailed results of analyses a)–d) are presented in the Supplement.

3.3.1 Anomaly over the Northern Pacific / Eastern Asia region

Fig. 3.1 illustrates the distribution of the wavelet power linked to the annual cycle amplitude in the temperature and the zonal wind field at 30 hPa using the MERRA reanalysis. A region of anomalous small amplitudes is seen across the

northern Pacific and eastern Asia. This anomaly is found for levels from about 50 up to 10 hPa for temperature and up to 1 hPa for zonal wind (for details see Figs. S.1–S.3 in the Supplement).

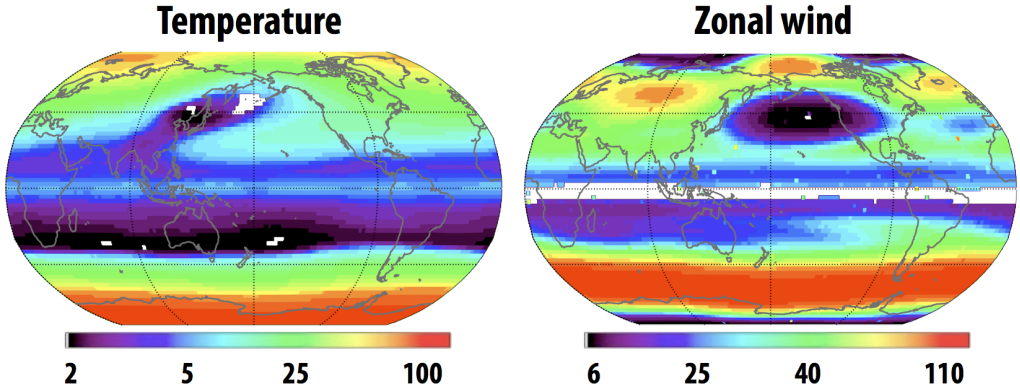

Figure 3.1: Annual cycle amplitudes in the temperature (left) and zonal wind (right) series at 30 hPa. The non-linear color-scale used represents square root of the wavelet power in K for temperature and in m/s for zonal wind.

The region where the anomaly is detected can also be linked to other distinct characteristics found in climatological fields. Considering the mean annual cycle of the temperature, zonal wind and ozone, we see very specific patterns in the stratosphere over the northern Pacific and East Asia region coinciding with the area of interest in this study.

Figure 3.2 presents the 1979–2013 mean seasonal averages at 30 hPa from MERRA (for other levels see Figs. S.4–S.6 in the Supplement). In northern hemispheric winter and (weaker expressed) in spring and autumn, there is an eastward wind minimum over the northern Pacific. The westerly jet is shifted northward, which results in a wave-1 pattern. A similar pattern can be found in the southern hemispheric mid-latitudes where the jet stream is shifted slightly poleward at similar longitudes.

Seasonal averages of the temperature series at the 30 hPa (Fig. 3.2) reveal a region of the high autumn and winter temperatures over the area of interest. A similar structure is found from about 100 hPa up to 10 hPa (for details see Fig. S.5 in the Supplement). A similar anomaly is found also in the Southern Hemisphere south of Australia, but most strongly in southern hemispheric spring. For levels above about the 10 hPa there is no longer such a well-marked pattern in the temperature field as in the lower levels. This can be connected to the higher temperatures in lower levels, which lift the 10 hPa level with respect to other locations. In the zonal wind field, the jet stream northward shift is still clearly visible even at the 10 hPa.

Figure 3.2 also illustrates the mean annual cycle in the ozone series at 30 hPa. There is a region of enhanced ozone concentrations above the area of interest during the whole year except for northern hemispheric summer. Maxima in the ozone concentration fields are shifted northward above the area of interest in northern hemispheric winter, again producing the wave-1 pattern. During autumn and spring, the maxima of concentration are found over the area of interest. In general the ozone concentrations above the area of interest are among the highest in

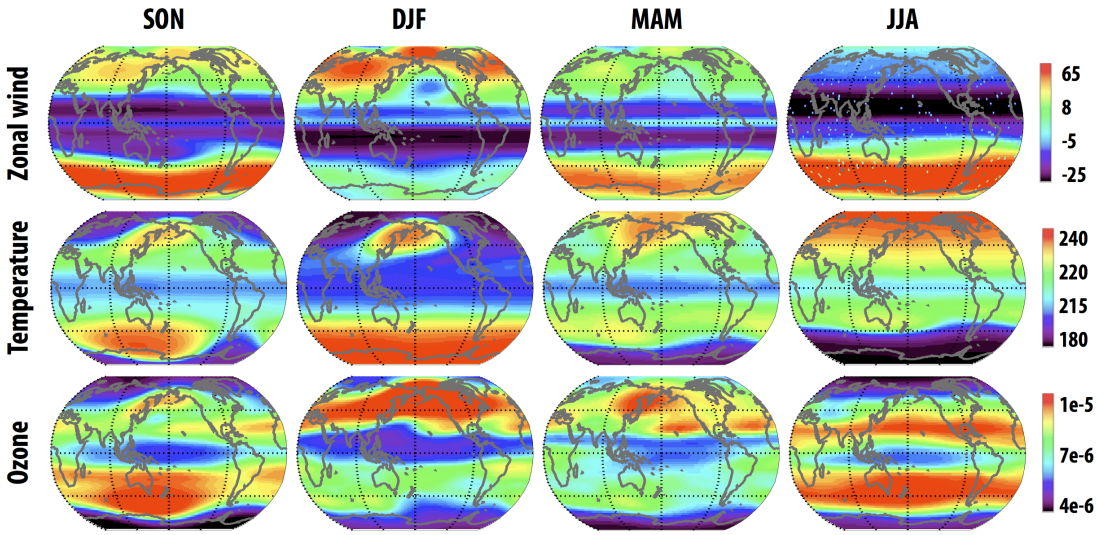

Figure 3.2: Seasonal averages in zonal wind in m/s, temperature in K and ozone mass mixing ratio in mg/kg for 1979–2013 time period using MERRA series (non-linear color scale used).

northern hemispheric mid to polar latitudes.

Specific patterns found in the mean climatology in Fig. 3.2 can be considered as the signature of an enhanced downwelling of the equatorial air reaching more northward over the northern hemispheric Pacific. In the next section we are looking for a stronger IGW activity above the analysed region, because such localised stronger wave activity region could hypothetically lead to a longitudinally distinct intensified branch of the Brewer–Dobson circulation. Also, the wave-1 pattern in the ozone, temperature and zonal wind fields at 10–30 hPa suggests that the area of interest could play an important role in the dynamics of the polar vortex making it zonally asymmetric and forcing the wave-1 structure.

3.3.2 IGW Analysis

It should be noted that the time interval of the analysis of the MERRA data set is much longer than the interval of the following IGW analysis. For consistency, we computed annual cycle amplitudes and mean seasonal averages also for the period 2007–2010 (not shown here). The results show that the above-described features are similar for this short period comparing to the original one. Kozubek et al. [2015] studied the longitudinal distribution and long-term trend of northern hemispheric stratospheric winds and identified a dominant effect of the Aleutian high (our region of interest is located at its western border in the stratosphere). They found that the trends of meridional winds connected with the Aleutian high are independent of sudden stratospheric warmings (SSW) or QBO. They observed intensification of the winds in the period of ozone depletion deepening (1970–1995) and weakening of the winds in the period of ozone recovery (1996–2012). However, there is an indirect dependence of the winds on QBO, as the solar cycle influence is pronounced mainly for the west phase of QBO.

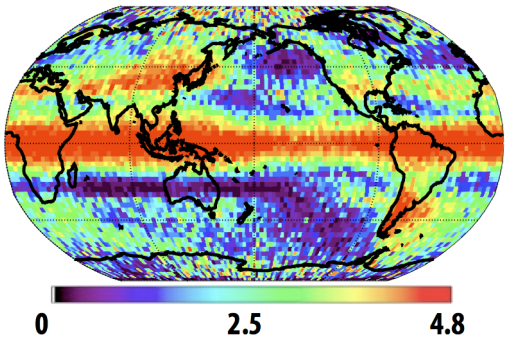

Figure 3.3: Annual mean of the potential energy in J/kg averaged across the whole vertical profile for the studied time period 2007–2010 (non-linear color scale used).

Figure 3.3 shows the mean potential energy averaged across the whole vertical profile and averaged over 4 years (2007–2010). The $\bar{E}_p$ values over eastern Asia are as large as in the equatorial area (including a possible Kelvin wave contribution there) or in regions with significant topography (e.g. the Andes). The area of enhance $\bar{E}_p$ spreads out from the Himalayas to the east and above the region of interest.

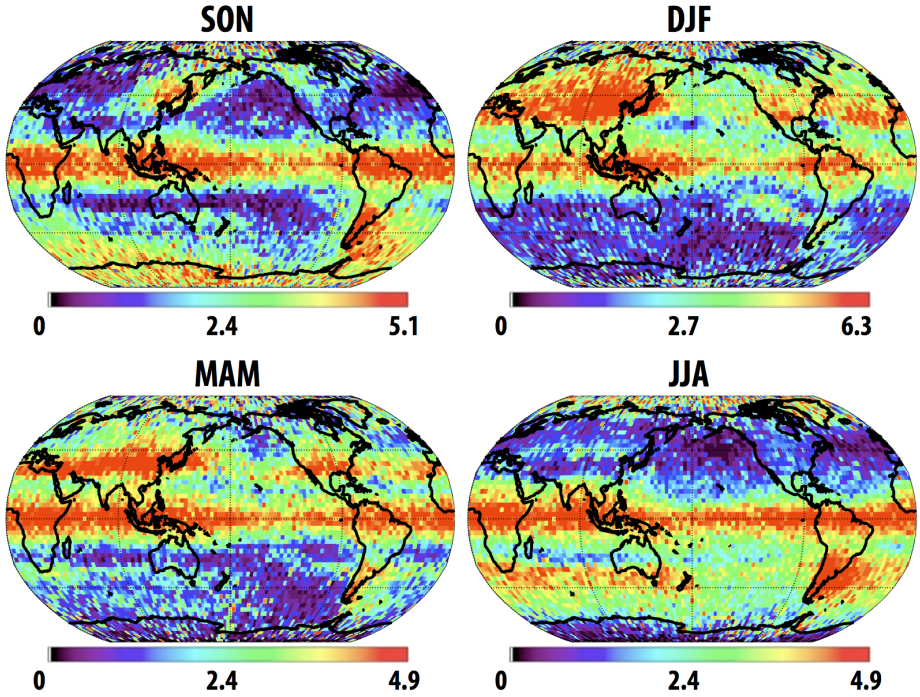

Figure 3.4: Seasonal means of the potential energy in J/kg averaged across the whole vertical profile for the studied time period 2007–2010 (non-linear color scale used).

Other important features are visible in the seasonal means of the $\bar{E}_p$ averaged across the whole vertical extent (Fig. 3.4). While in winter and spring the area of interest lies in the region of high $\bar{E}_p$ values extending over the whole latitude circle (except for the northern Pacific) and strengthening eastward of the Himalayas, in

summer there is a region of low values over the broad region of the Pacific Ocean. Finally, in autumn, there is a unique and localised area of the highest $\bar{E}_p$ values (ranging up to 6 J·kg^{-1}) in the Northern Hemisphere.

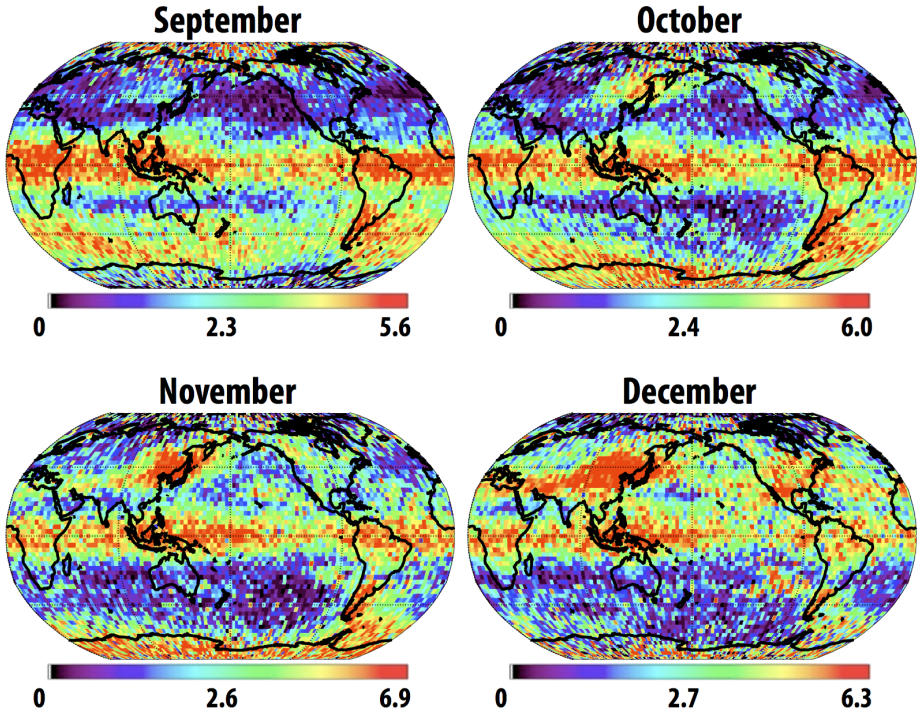

Figure 3.5: Selected monthly means of the potential energy in J/kg averaged across the whole vertical profile for the studied time period 2007–2010 (non-linear color scale used).

Analysing this pattern on a monthly basis (Fig. 3.5), we may see that in September there is an area of low $\bar{E}_p$ above the northern Pacific while over north-eastern Asia there is a weakly defined area of higher $\bar{E}_p$ values around 2.5 J·kg^{-1}. This region expands eastward above Japan in October and the wave activity strengthens. In November, there is a well-defined area of northern hemispheric maximal $\bar{E}_p$ above the region of interest. In December, we can see a typical winter pattern of high $\bar{E}_p$ values ranging approximately from Himalayas while strengthening towards the analysed region.

In fields for individual pressure levels (Fig. S.7 in supplement), the described patterns over the area of interest are detected from 70 hPa up to the 6 hPa level. All these results support the hypothesis of special IGW activity in the analysed area.

The specific annual cycle of the stratospheric conditions over the area of interest described in the previous chapter is in accordance with this wave activity cycle. Wave activity generally precedes changes of the residual circulation and its dynamical effects [e.g. Kuchar et al., 2015] and in the distribution of temperature and zonal wind the region is pronounced most significantly in winter. However, to quantify the role of such localised wave activity maxima for the specific dynamics leading to the anomalous annual cycle amplitudes in this region, it would be necessary to have information on time evolution of the wave activity and distribution of dissipative processes in the region to quantify a wave-mean flow interaction in

the sense of generalised Eliassen–Palm theorem [for zonal averages, see e.g. Andrews et al., 1987, equation 3.6.2]. But this approach cannot be applied having the information from observations only.

Therefore, further analysis is focussed on the stability of the wave fields in order to find regions of possible wave breaking and interaction with the mean state of the atmosphere. This can help to quantify the relative importance of this region and may give a better view on its dynamical effects on the stratospheric circulation. The analysis is followed by a discussion of the types and sources of IGWs in this region especially in October and November, based on study of prevailing surface wind directions and the rotation of the wind (change of wind direction) with height.

3.3.3 Wave breaking indication

Figure 3.6 presents the distribution of the gradient Richardson number at three levels between 30 and 10 hPa in April and May. The fields are shown for those months when the $\bar{E}_p$ values in the area of interest do not stand out against other regions. One can see from Fig. 3.6 that at each altitude considered the gradient Richardson number (Ri_g) is small over the area of interest. Further analysis showed that, except for March, June, July, August, and September, the lowest Ri_g values are over the area of interest at all levels from 70 up to 6 hPa (for details see Figs. S.8–S.11 in the Supplement). This suggests that the wave/wind fields in this region are closer to dynamical instability than in other regions. This interesting pattern emerges in October and November (Fig. S.8 in Supplement) and at some pressure levels also in the winter season (Fig. S.11 in the Supplement).

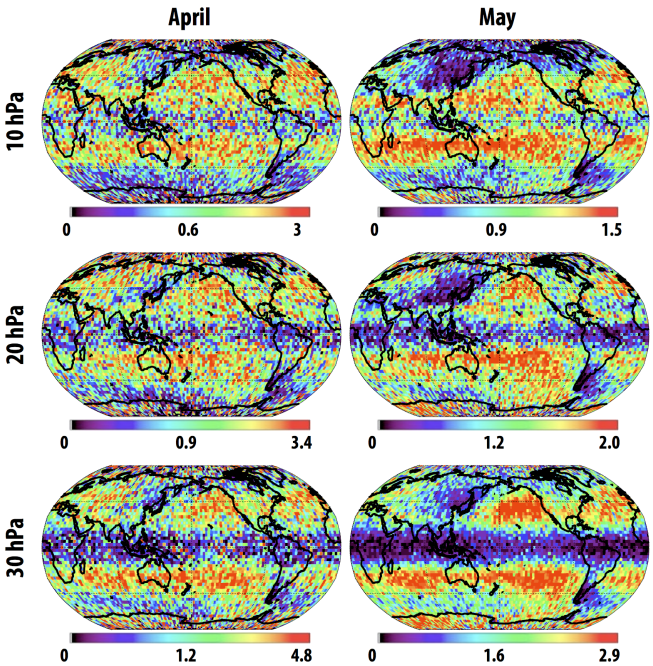

Figure 3.6: Selected monthly means of the gradient Richardson number at 10, 20 and 30 hPa for the studied time period 2007–2010 (non-linear color scale used).

Furthermore, low Ri_g values are located not only above the analysed region but also extending along the western Pacific coast and then southward along the

eastern Pacific coast to Northern America. This is in agreement with the wave-1 pattern of the jet stream and the jet location shown in Fig. 3.2: Low Ri_g values are detected even in the areas where the wave activity is not especially strong. Suitable background conditions, i.e. strong wind shear, can bring the waves close to instability and allow them to affect the mean flow.

Results of the analysis of the sigma, i.e. the Rayleigh–Taylor convective instability disturbances growth rate are illustrated in Fig. 3.7. The distribution of the maximum detected in the sigma squared profiles (uppermost row) is presented together with the distribution of selected secondary maxima in each grid box. The results are calculated in such a way that within all of the occultation profiles in a selected grid box and time interval we look for the highest sigma squared value, consider it (noting its altitude as well) as the first maximum and then we look for the second highest value, etc. The secondary values are often found in different profiles or are from the same profile but located at lower altitudes. Positive values of sigma squared indicate convective instability, so that we observe negative values only; however, small magnitudes of σ indicate regions of weak stability. Figure 3.7 further accentuates the importance of the area of interest. The maximum σ^2 values detected over the region of interest (often together with its coastline extension) are a dominant feature of the maps. These findings indicate a vertically robust and persistent breaking of IGWs. Also the altitudes of the sigma squared maxima over the region of interest are among the lowest detected ones (around 25 km altitude, for details see Figs. S.12 and S.13 in the Supplement), suggesting that breaking begins at lower levels in this region. Another interesting result is the detection of high sigma squared maximum values in the stratosphere over the summer polar latitudes. These are found for the first maximum but are better visible for secondary maxima at lower altitudes. This finding is in line with Baumgaertner and McDonald [2007], who attributed the small amount of summertime potential energy to lower level critical filtering.

3.3.4 Possible wave sources

The IGW spectrum is shaped not only by different sources but also reflects tropospheric background conditions contributing to filtering of various gravity waves Fritts and Alexander [2003]. This can be easily applicable to the orographic gravity waves that are critically filtered when the wind speed is zero. This condition is fulfilled in the case of directional shear exceeding 180°. Above regions where this is fulfilled, one can rule out the possibility of orographic gravity wave modes contributing to the observed IGW activity. Vice versa, regions of small wind rotation (change of wind directions between levels) from the lower levels are favourable for vertical propagation of orographic waves. Alexander et al. [2009] argued that higher stratospheric orographic wave activity is expected to be observed when there is relatively little rotation between the surface and 400 K isentropic surface.

We studied the wind direction and its change between 975 hPa and higher levels. The results are shown in Fig. 3.8 for November 2008. All results are presented in Figs. S.14–S.15 in the Supplement. The analysis shows that in the region of interest the wind direction changes only little. Only in June, July and August are the orographic gravity waves likely to be completely filtered out of the

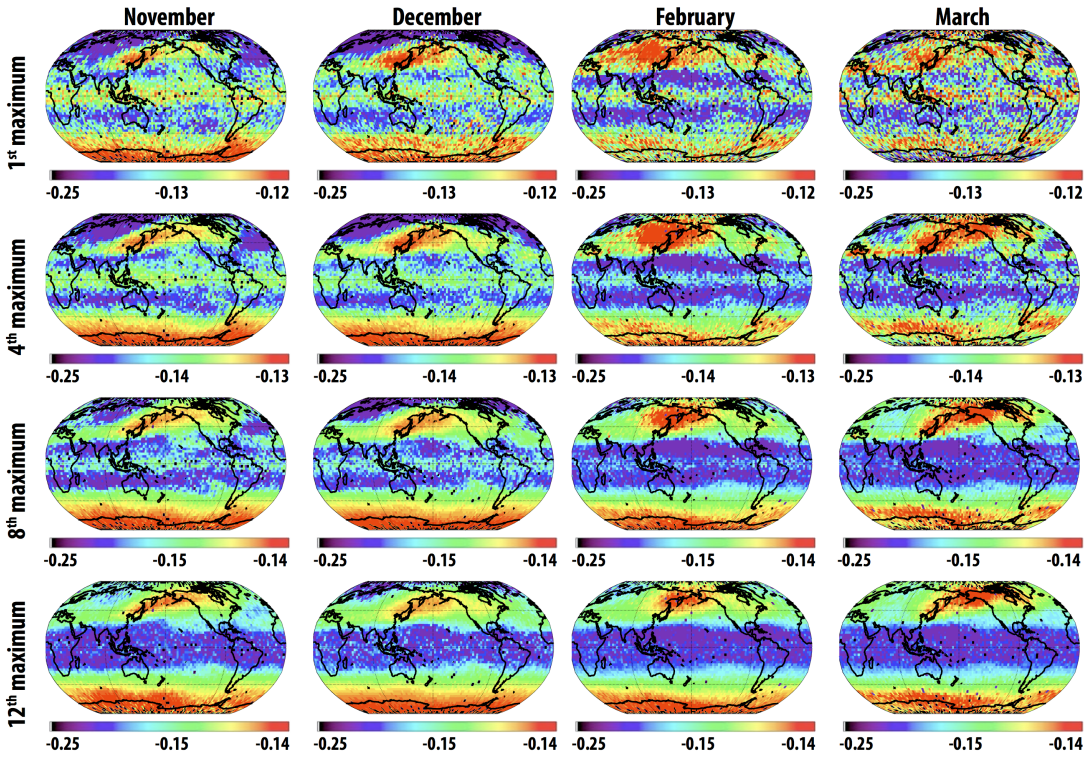

Figure 3.7: Selected monthly means of primary and selected secondary (i.e., higher order) sigma squared maxima in s^{-2} for the studied time period 2007–2010 (non-linear color scale used).

spectra before reaching the lower stratosphere 100 hPa level above the analysed region.

The direction of the surface winds suggests orographic formation of the IGWs due to the topography of Japan, Sachalin, Korean Peninsula or eastern Asia coastline. The significance of this topography is enhanced by the contrast with the ocean surface. Considering the optimal conditions for propagation into the stratosphere and the topography suitable for emitting large amplitude orographic waves, we conclude that a significant part of the measured IGW spectra in the area of interest may consist of orographic IGWs. This is in line with findings of Alexander et al. [2008] regarding this region.

Another source of IGWs in this region, which is of special interest because of their autumn appearance, is convective activity connected with the Kuroshio current. Its role in forcing IGWs has been pointed out, e.g., by McDonald et al. [2010] and Jia et al. [2014]. Ern and Preusse [2012]; Zhang et al. [2012]; Ern et al. [2013] labelled an area to the south of the region of interest as a deep convective region with large gravity wave activity. Taking into account that the meridional propagation is predominantly northward [Horinouchi and Tsuda, 2009], we may conclude that these waves could play a role in the region of interest mainly in spring and summer.

Alexander et al. [2009] argued that the FORMOSAT-3/COSMIC observed IGW variance in the stratosphere likely depends upon a combination of orographic waves, Doppler shifting of tropospheric source waves and possibly some in situ stratospheric wave generation. Considering the wind field in the region of interest

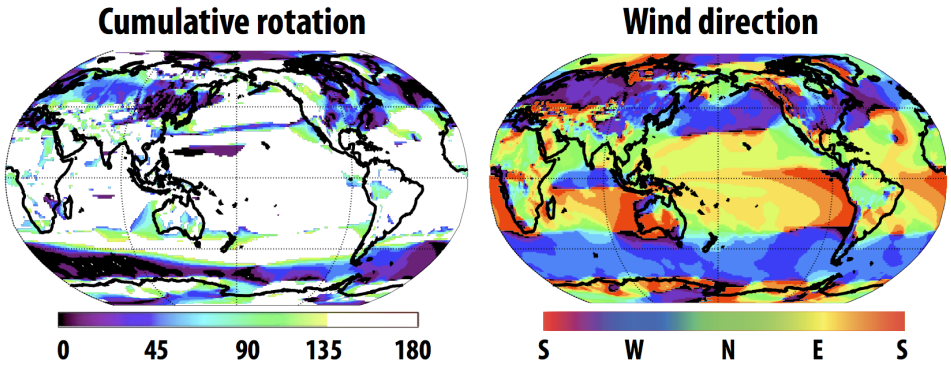

Figure 3.8: Cumulative rotation of wind from 975 hPa to 30 hPa (left) and prevailing wind direction in the level of 975 hPa (right). Computed from JRA-55 for November 2008.

(seasonally dependent location of the subtropical westerly jet and the polar front jet in the upper troposphere–lower stratosphere), the Doppler shifting must play a role in amplifying wave amplitudes while propagating upwards, and it may also be responsible for one of the possible reasons for the enhanced wave activity in this region.

Finally, in connection with the jet location above the region of interest, we can expect a strong contribution to the IGW spectrum from spontaneous emission processes Plougonven and Zhang [2014]. Evidence for this claim can be found, e.g., in Hirota and Niki [1985] and Fukao et al. [1994], who analysed middle and upper atmospheric radar data (located at Shigaraki, Japan, falling into our region of interest) and who found inertia IGWs propagating upward and downward from the jet stream. Orographic waves were also identified.

As a curiosity, according to Mohri [1953], during the colder season the subtropical jet stream reaches its maximum intensity south of Japan, while the polar front jet is located north of the Tibetan Plateau. Moreover, these jets sometimes merge (mainly in winter) and create extremely thick frontal layers. Such episodes can become a very interesting and unique source of IGWs in this area.

3.4 Discussion

The specific dynamics, as well as the IGW hotspot area, has, to our knowledge, not been identified in previous observational or theoretical studies. In the following, we shall discuss possible reasons for this, together with a discussion of the usage of $\bar{E}_p$ as a wave activity proxy, and possible implications of such a unique wave activity region for the large-scale dynamics and transport in the middle stratosphere.

3.4.1 Unmasking the IGW hotspot area

There are a lot of satellite studies of the global distribution of IGW activity. But autumn, the season of most unique wave activity in the region of interest, is often left out because the main focus is usually laid on the winter season. Generally, to our knowledge, there are only a few monthly analyses of the mean annual cycle of

gravity wave activity. This can account for masking the significance of the region of interest.

One of the unique aspects of our analysis is also that we were purposefully looking for anomalous wave activity in the region of interest, which was hypothesised due to anomalies found in the zonal wind, temperature and ozone fields.

The methodology of our analysis is novel through the use of GPS RO density instead of temperature profiles. Although the differences between IGW characteristics derived from the dry temperature and density profiles have not been studied in details yet, Šácha et al. [2014] demonstrated that in comparison with temperature the density perturbation spectra have higher power spectral density in the shorter vertical wavelength range below roughly 2.5 km. Comparison in the larger wavelength region is generally inconclusive because the power spectral density there is dominated by the background separation method. How this could influence the results of our analysis will partly be discussed in the next paragraphs.

3.4.2 Wave activity proxy

Šácha et al. [2014] summarised that, unlike using GPS RO dry temperature profiles, density perturbations include contributions from non-hydrostatic waves and the obtained wave amplitudes are not additionally suppressed by the hydrostatic assumption in the retrieval. Nevertheless, the spatial distribution of non-hydrostatic waves, to our knowledge, has not been fully quantified yet. Thus, it is possible that the magnitude of the difference between temperature and density spectra can change with location. Šácha et al. [2014] found differences between temperature and density PSD (power spectral density) in the same region as we have analysed. However, in other locations and basically in all vertical wavenumber regions where the PSD slope of perturbations disagree with the slope of theoretical saturated spectra, the magnitude of differences between power spectral densities of temperature and density perturbations may vary; moreover, the validity of $\bar{E}_p$ as a wave activity proxy is crucially connected with the agreement or disagreement of vertical wavenumber spectra and theoretical saturated spectra.

There is a long tradition of using mean wave energy as a measure for wave activity in the IGW studies using GPS RO data. However, strictly speaking, the wave energy is not a conserved quantity during the wave propagation Bühler [2014]. To estimate it using only instantaneous temperature or density measurements, a constant ratio between mean wave kinetic and potential energy has to be assumed VanZandt [1985]. Then the key question is whether the vertical wavenumber spectrum obtained from the GPS RO occultation can be considered as saturated to fulfil the underlying essential assumption of universal saturated spectra of IGWs.

Theory of IGW spectra generally assumes saturation at each wavenumber Smith et al. [1987], and although the observed spectra in the middle atmosphere are influenced by the distribution of tropospheric sources and filtering due to propagation in different conditions, they appear to be close to theoretical ones when observed over long times Fritts [1984].

Smith et al. [1987]; Allen and Vincent [1995]; Fritts and Alexander [2003] noted that mean energy and characteristic vertical wavenumber of gravity waves

may experience significant variations with time and geographical location. This is due to the variations in the energy of motions at vertical wavenumbers lower than the characteristic wavenumber [around 3 rad·km^{-1} in the lower stratosphere Fritts and Alexander, 2003; Tsuda et al., 1994] caused by variable IGW sources, mean wind, and static stability. Even VanZandt [1985] noted that lee waves are likely to be important for small values of the characteristic vertical wavenumber.

For temperature profiles in the lower stratosphere, Steiner and Kirchengast [2000] documented that the average GPS/MET vertical wavenumber spectra have amplitudes smaller than a saturated one. There is a tendency of increasing discrepancy with the theoretically saturated spectra towards higher wavenumbers. Steiner and Kirchengast [2000] observed dominant fluctuations occurring at wavelengths near 3 and 5 km. But assuming a saturated spectrum, dominant fluctuations should have the largest wavelengths allowed by the methodology. It can be easily shown (the proof is given in Appendix at the end of this chapter) that for two profiles with the same value of $\bar{E}_p$ but with different spectral contributions in the regions of disagreement with the theoretical slope, the resulting wave activity (mean wave energy; another theoretical inaccuracy discussed in the Appendix) is different if the ratio between kinetic and potential energy changes with vertical wavenumber (direct proportion in the mid-frequency approximation).

So, to use $\bar{E}_p$ theoretically correctly as a wave activity proxy, the interval of IGW wavelengths used in the computation should be bounded by the longest and shortest vertical wavelength where the vertical wavenumber power spectra density of disturbances has a slope similar to the theoretically predicted one. This would, however, substantially limit the spectral extent of GPS RO IGW analysis. In addition, Luna et al. [2013] showed that the uncertainty in potential energy derived from GPS RO is influenced more by the choice of integration limits (lower and upper boundary of the vertical region) and maximal and minimal vertical wavelength of disturbances allowed by the filter.

Another approach for quantifying the wave activity, as proposed by Ern et al. [2004], is to compute momentum flux (vertical flux of horizontal pseudo-momentum according to the naming convention by Bühler [2014]), which is a conserved quantity. This approach is theoretically consistent, but it is limited by introducing additional approximations such as a sinusoidal dominant wavelength in the mid-frequency range. The advantage of this method could be seen in the results of [e.g. Wright and Gille, 2013], where the enhancement of significance of wave activity around the Himalayas could be related to the better representation of lee wave activity with smaller slope of the phase lines (higher ratio between kinetic and potential energy), whose activity would be underestimated using $\bar{E}_p$. On the other hand, another large mountain range, the Andes, is oriented perpendicular to the mean winds. Thus, there is a shift to non-hydrostatic waves in comparison to the lee waves from the Himalayas Durran [2003]. This makes the Andes better visible in the $\bar{E}_p$ analyses due overestimation of IGW activity, although concurrently their amplitudes could have been smoothed using the ordinary GPS temperature data Steiner and Kirchengast [2000], which is not the case in our analysis. In this discussion, to show our point, for simplicity we avoid discussing the effect of observational geometry with respect to the wave orientation, which could otherwise be a leading source of differences between observed IGW activity especially when contrasting Himalayas and Andes.

In our analysis we compute monthly averages of occultations on a 3° × 3° grid. The number of profiles does not exceed a few tens of profiles per month in one grid box. This is not sufficient to remove individual qualities of each profile power spectrum by averaging. It is also highly probable, as discussed above, that these individual qualities (e.g. a peak at larger vertical wavelengths) would remain dominant in the locations with, e.g., unchanged meteorological conditions in the analysed month, even if the ensemble were large enough. This all leads us to the conclusion that the wave activity could be underestimated (or overestimated) and regions, e.g., with prevailing inertia gravity waves could be masked when analysing the $\bar{E}_p$ distributions. This could be the case near the northern and north-eastern Pacific coast, where additional analysis using the characteristics based on the rate of change of density perturbations with height (gradient Richardson number and sigma) reveals wave fields close to instability as seen, e.g., in Fig. 3.7.

3.4.3 Implications for the large-scale dynamics

Smith [2003] argued that zonal asymmetries in gravity wave activity contribute to zonal asymmetries in mesospheric geopotential height and might also have an effect on the 3–D transport in the mesosphere. Demirhan Bari et al. [2013], using the approach of the 3–D residual circulation proposed by Kinoshita et al. [2010], recently investigated the longitudinal variations in the time-mean transport by the Brewer–Dobson circulation. In Fig. 3 in their paper they showed that the mean January residual circulation at 60°N (northward boundary of our region of interest) penetrates to the lowest altitude (below 20 km) around 140°E (the meridian crossing our region of interest). This is in line with our findings of robust and persistent breaking of IGWs starting at anomalous low levels in the region of interest. Also the distributions of atmospheric quantities indicate robust downwelling of tropical air masses (enhanced branch of Brewer–Dobson circulation) reaching deeper into the stratosphere in the region of interest than elsewhere.

To investigate this in detail, we are currently preparing a study analysing model runs of a middle atmospheric mechanistic model with means of 3–D EP (Eliassen–Palm) flux and residual circulation diagnostic formulated by Noda [2014]. The study will focus mainly on the structure of dynamics and transport in the region of interest. Another issue that we would like to address is the causality of processes behind the specific dynamics in this region. In particular, we are interested in whether the specific IGW activity in autumn alone could be the reason for anomalously high temperatures and enhanced downwelling in winter. On the other hand, the complexity of the atmospheric system favours feedback mechanisms between waves and mean state, as the anomalous wave activity is probably a result of favourable meteorological conditions for IGW sources and propagation.

Another hypothetical importance of this region is the possibility that such a confined IGW breaking region creates planetary waves. This theoretical possibility was mentioned, e.g., by Smith [2003], who suggested that momentum forcing associated with breaking gravity waves that have been filtered by planetary-scale wind variations below acts to generate planetary waves in the middle and upper

mesosphere. This is supported for example by the jet wave-1 pattern in Fig. 3.2 (again a question of causality). The sourcing of planetary waves by the region of interest could be exceptional through its low altitude (breaking starting already below 20 km altitude).

Thus, in a further paper, we shall investigate possible formation and propagation directions of planetary waves caused by such a localised IGW forcing in model simulations. Although preliminary results (not shown here) suggest that the poleward propagating mode is smaller than the equatorward propagating one, it can have influence on the polar vortex stability. It is up to further studies to analyse the connection between the interannual variation in strength of wave activity in the region of interest in the autumn or winter season and the SSW occurrence.

The equatorward propagating planetary wave modes created in November, as described by Ortland [1997], can play an important role in the stratosphere-troposphere exchange in the tropical region, where they break in easterly winds. Such a process is supported by results from Riese et al. [2014] indicating that the longitude of maximal passage of tropical air into the stratosphere in autumn corresponds to our region of interest. Furthermore, Škerlak et al. [2015] identified an area south/southeast of the region of interest to have maximum tropopause fold frequency (in Dec-Jan-Feb 1979–2012). Finally, Berthet et al. [2007] used back trajectories driven by large-scale analysed wind fields to investigate troposphere to stratosphere transport and found, particularly in autumn, a region of significant transport south of the region of interest.

Finally, our findings can also have implications for weather forecast and climate change. Horinouchi [2014] studied the synoptic variability of precipitation and moisture transport roughly in the same area as ours. This study concluded that from the dynamical point of view, it is arguably more meaningful to view the synoptic variability as initiated in the upper troposphere and to place the formation of surface quasistationary fronts as an aspect of this variability. For climate change and the debate about acceleration of Brewer–Dobson circulation it seems more and more clear that it is necessary to consider the Brewer–Dobson as longitudinally variable as shown by Demirhan Bari et al. [2013]. This is supported by the presented results in the sense of the wave pumping hotspot and anomalous dynamics area suggestive of enhanced downwelling.

3.4.4 Summary and conclusions

Using the GPS RO density profiles, we analysed the IGW activity in an area of low annual cycle amplitude in the north-eastern Pacific–eastern Asia coastal region. Enhanced IGW potential energy values were found uniquely in the lower stratosphere region in this area in October and November. Convective and dynamical instability indicators suggest robust wave breaking in this region starting at anomalously low levels, and this was detected also in spring.

Possible IGW sources were examined, analysis of prevailing surface winds and wind direction change revealed ideal conditions for sources and vertical propagation of orographic waves. Other sources contributing to the enhanced wave activity in this region are likely a convective activity connected with the Kuroshio current, Doppler shifting of vertically propagating waves and in situ wave gener-

ation in the upper troposphere–lower stratosphere (e.g. geostrophic adjustment). The latter mechanism can become a very interesting and unique source of IGWs during the episodes of merging jets.

The reasons why this particular IGW activity hotspot was not discovered before as well as why the specific dynamics of this region have not been pointed out are discussed. There are possible subjective reasons like the best visibility and uniqueness of this hotspot in autumn or motivated research based on previous assumption of the wave activity in this region on the one hand and possible objective physical mechanisms on the other hand. Those mechanisms can be separated into two groups that can influence each other: first, there are differences between our study and other studies that use GPS RO dry temperature data and there are differences with studies using different wave activity proxies.

The novel usage of GPS RO dry density data in this study belongs to the first group. The dry density data are not filtered by the hydrostatic balance and contain more information than the dry temperature data in the shorter vertical wavelength range. For the larger vertical wavelength region, the comparison is generally inconclusive because of the dominance of the background separation method effect Šácha et al. [2014]. By relating the density and temperature backgrounds with the hydrostatic assumption, we are currently preparing a further study to reveal the differences and their spatio-temporal variations in the whole analysed spectrum. The study should help to quantify the contributions of the noise and non-hydrostatic IGWs to the differences between usage of the dry density and temperature data.

In the second group of mechanisms, which can influence the comparability of our results with those of other studies, the inaccuracy of using mean potential energy of disturbances as a wave activity proxy is an important point. Its discussion is related not only to the region of interest, but also the possibility of underestimating/overestimating the wave activity depending on the dominant IGW modes using $\bar{E}_p$ is important in a general sense. The possibility of underestimating/overestimating the wave activity must be kept in mind when comparing $\bar{E}_p$ fields with the distributions of pseudomomentum flux or of the wave breaking indicators employed additionally in our analysis. For example, the regions with prevailing inertia IGWs can be masked using $\bar{E}_p$.

We have also discussed possible consequences of the region of interest on the middle atmospheric dynamics and transport (e.g. Brewer–Dobson circulation), linkage to the conditions in the troposphere and the necessity to consider the real geographical and seasonal distribution of IGWs together with 3–D residual circulation diagnostics to learn about its change in a changing climate. In a next chapter, these hypotheses will be examined in detail.

Appendix - Proof of inaccuracy of using E_p as a wave activity proxy

In this appendix, the proof is given that using $\bar{E}_p$ inappropriately, the mean wave energy may be over or underestimated. But, at first, we must note that the mean energy of disturbances is not a good wave activity proxy either, because it is not conserved in the presence of mean flows Bühler [2014]. The conserved quantity derivable from the mean energy is a so-called pseudoenergy (for

definition see Bühler [2014]), but from observations we generally do not have enough information available for its computation. Also, since the pseudoenergy is frame-dependent, it is not desirable to use it for analyses of wave-mean interactions Bühler [2014]. For IGWs (in the presence of rotation) the mean energy does not satisfy energy equipartition and on top of that the partitioning ratio between wave kinetic and potential energy is dependent on its intrinsic frequency.

Now we consider an imaginary situation illustrated in the Figure 3.9, where PSDs of two vertical profiles have the same value in the vertical wavelength range, where the slope of their PSD corresponds to the theoretical saturated one. But outside this range the slopes differ with each other and with the theoretical slope as well. The region of disagreement can be divided into two small spectral regions $d\lambda_a$ and $d\lambda_b$ with equal sizes. In the interval a the second profile (blue) has higher PSD than the first one (red) and in the interval b it is otherwise. Both profiles have the same mean potential energy of disturbances. Mathematically, the situation can be summarized as follows:

$$\bar{E}^a_{p1} - \bar{E}^a_{p2} = \bar{E}^b_{p2} - \bar{E}^b_{p1}, \tag{3.6}$$

where subscripts a and b determine the interval of averaging. If the averaging intervals a and b span a wavenumber interval of the same length each, then we can compute the mean energy simply as an arithmetic average of the mean value in the two intervals.

$$\bar{E}^{a+b} = \frac{\bar{E}^a + \bar{E}^b}{2} = \frac{\bar{E}^a_k + \bar{E}^b_k + \bar{E}^a_p + \bar{E}^b_p}{2} \tag{3.7}$$

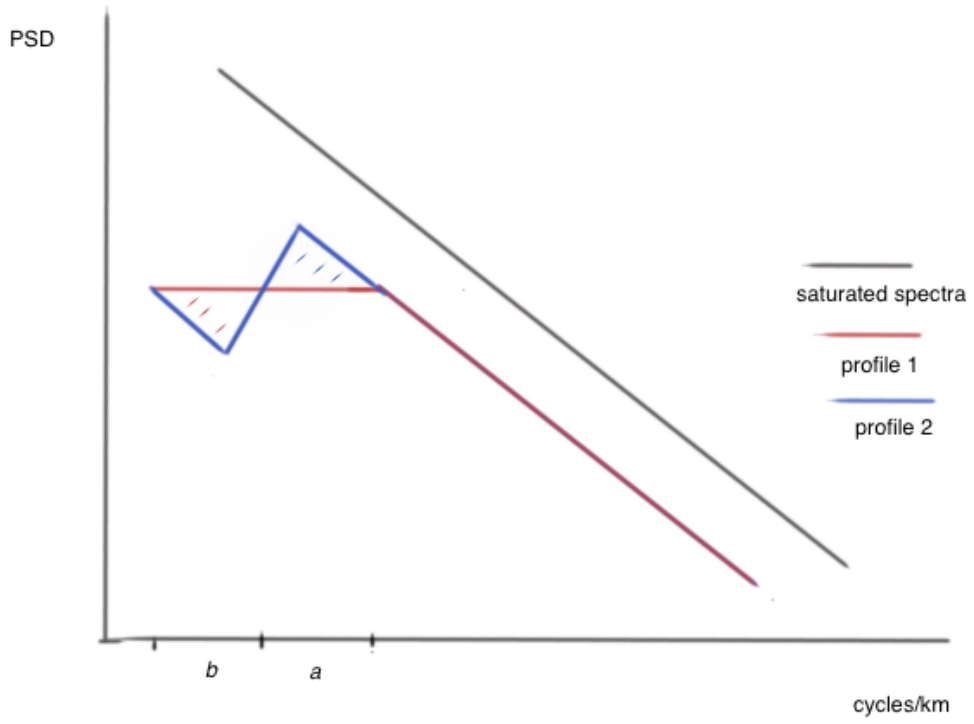

Figure 3.9: Illustration of PSDs of two vertical profiles having the same value of $\bar{E}_p$ with different modal contribution to it.

As noted above, the partitioning ratio between wave kinetic and potential energy is dependent on wave's intrinsic frequency and therefore under midrange frequency approximation [Fritts and Alexander, 2003] also directly on its vertical wavelength. When averaging across the spectral intervals a and b, the partitioning ratio is the same for both profiles but differs with spectral interval (ratio X).

$$\frac{\bar{E}^a_{p1,2}}{\bar{E}^a_{k1,2}} = C_a, \frac{\bar{E}^b_{p1,2}}{\bar{E}^b_{k1,2}} = C_b \tag{3.8}$$

and

$$\frac{\bar{E}^a_{p1,2}}{\bar{E}^a_{k1,2}} = X \cdot \frac{\bar{E}^b_{p1,2}}{\bar{E}^b_{k1,2}}, X = \frac{C_a}{C_b}. \tag{3.9}$$

From the equality of the partitioning ratios between the two profiles 3.8 we have:

$$\bar{E}^a_{p1} \cdot \bar{E}^a_{k2} - \bar{E}^a_{p2} \cdot \bar{E}^a_{k1} = 0. \tag{3.10}$$

Finally we are ready to prove that the mean energy differs between the two profiles:

$$\bar{E}^{a+b}_1 - \bar{E}^{a+b}_2 = \frac{1}{2}\left(\bar{E}^a_{k1} + \bar{E}^b_{k1} - \bar{E}^a_{k2} - \bar{E}^b_{k2}\right). \tag{3.11}$$

Now we use 3.9 to introduce back potential energies into 3.11 by substituting for kinetic energies averaged in the interval b. After some computation we have:

$$\begin{aligned}\bar{E}^{a+b}_1 - \bar{E}^{a+b}_2 &= \frac{1}{2}\left(\bar{E}^a_{k1} - \bar{E}^a_{k2} + X\frac{\bar{E}^a_{k1}\bar{E}^b_{p1}}{\bar{E}^a_{p1}} - X\frac{\bar{E}^a_{k2}\bar{E}^b_{p2}}{\bar{E}^a_{p2}}\right) = \\ &= \frac{1}{2}\cdot\frac{X\left(\bar{E}^a_{k1}\bar{E}^b_{p1}\bar{E}^a_{p2} - \bar{E}^a_{k2}\bar{E}^b_{p2}\bar{E}^a_{p1}\right) + \bar{E}^a_{k1}\bar{E}^a_{p1}\bar{E}^a_{p2} - \bar{E}^a_{k2}\bar{E}^a_{p1}\bar{E}^a_{p2}}{\bar{E}^a_{p1}\bar{E}^a_{p2}}.\end{aligned} \tag{3.12}$$

Now we factor out $\bar{E}^a_{k1}\bar{E}^a_{p2}$ from the first and third term in 3.12 and $\bar{E}^a_{k2}\bar{E}^a_{p1}$ from the second and fourth term to get:

$$\bar{E}^{a+b}_1 - \bar{E}^{a+b}_2 = \frac{1}{2}\cdot\frac{\bar{E}^a_{k1}\bar{E}^a_{p2}\left(X\bar{E}^b_{p1} + \bar{E}^a_{p1}\right) - \bar{E}^a_{k2}\bar{E}^a_{p1}\left(X\bar{E}^b_{p2} + \bar{E}^a_{p2}\right)}{\bar{E}^a_{p1}\bar{E}^a_{p2}}. \tag{3.13}$$

Using 3.10 we see that the numerator can be non-trivially equal to zero only if:

$$X\bar{E}^b_{p1} + \bar{E}^a_{p1} = X\bar{E}^b_{p2} + \bar{E}^a_{p2}. \tag{3.14}$$

But from the construction of our thought experiment and from 3.6 this could be true only if $X = 1$ meaning that the partitioning ratio must have been constant for the whole vertical wavelength spectra, which is not true.

4. Influence of the spatial distribution of gravity wave activity on the middle atmospheric circulation and transport

This is a prime chapter of the thesis, because, building upon the research from previous chapters, with a mechanistic model for the middle and upper atmosphere, we performed sensitivity simulations to study a possible effect of a localized IGW breaking hotspot in the Eastern Asia/Northern Pacific region (hypothetical influence discussed in the subsection 3.4.3.) and also more generally, possible influence of spatial distribution of gravity wave activity on the middle atmospheric circulation and transport.

Consideration of IGW related processes is necessary for a proper description and modeling 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 datasets with improved resolution and novel analysis methods together with high-resolution global models have been tightening the constraints for the parametrizations that can improve the treatment of these waves in climate models [Alexander and Shepherd, 2010; Geller et al., 2013].

Complex understanding and unbiased modeling of the 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 and also short range forecast skills [Hardiman and Haynes, 2008; Gerber et al., 2012].

BDC was first discovered based on the distribution of trace gases by Brewer [1949] and Dobson [1956]. Using the transformed Eulerian mean equations, Dunkerton [1978] derived a first dynamically consistent two-dimensional picture of the mean-transport streamlines for the middle atmosphere that is often used as a basic BDC concept. However, Demirhan Bari et al. [2013] found a 3D structure of the circulation in the middle atmosphere to be in good correspondence with tracer fields, especially in relation to the zonal wave-one 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 north-eastern Asia, wave-one pattern).

PWs are usually thought to be created in the troposphere and then vertically propagating into the middle atmosphere. The theoretical possibility of PW creation by zonally asymmetric IGW breaking was first numerically analyzed 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 a building agreement

in the literature on the role of wave activity in preconditioning sudden stratospheric warming [SSW; e.g. Ayarzaguena et al., 2011] events.

SSWs belong to the most pronounced atmospheric phenomena, as they cause abrupt changes of 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 impacts differ whether a split or a displacement of the stratospheric polar vortex takes place [Seviour et al., 2016] and it has been observed that displacements are connected with a dominating wave-one activity while vortex splits correlate with stronger wave-two activity [e.g. Kuttippurath and Nikulin, 2012]. Generally, most attention is paid to the role of upward propagating PWs preconditioning SSWs [Hoffmann et al., 2007; Nishii et al., 2009; Alexander and Shepherd, 2010].

The two open questions regarding the dynamics of SSWs 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 nonlinear resonance in connection to the vortex geometry [Albers and Birner, 2014]. Also, there is growing observational evidence that GW amplitudes are enhanced prior to SSWs [Ratnam et al., 2004; Wang and Alexander, 2010; Yamashita et al., 2010], and IGWs are acknowledged to play a 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].

However, when the IGW-PW interaction is considered, the majority of studies are concerned by the modulation of IGWs by PWs [e.g. Cullens et al., 2015] and about the IGW impact on the upper stratosphere/mesosphere region. Šácha et al. [2015] indicated a possible IGW breaking in the lower stratosphere. Indeed, model experiments with gravity wave drag (GWD) parameterization showed that orographic GWD in the lower stratosphere can significantly affect the development of SSWs [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 IGW sources in the parametrization. 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 the zonally asymmetric forcing represented by an artificially injected GWD in the stratosphere. We are following Šácha et al. [2015], who described a localized area of enhanced IGW activity and breaking in the lower stratosphere over the Eastern Asia/North-western Pacific (EA/NP) region and discussed possible implications of this IGW hotspot for large scale dynamics and transport. By artificially enhancing the GWD in a 3D mechanistic circulation

model of the middle atmosphere we examine the hypothesis that such a robust breaking region plays a role in forcing longitudinal variability of the BDC and can generate PWs. Further, we investigate possible implications for the polar vortex stability and the role of the GWD distribution and of the artificial forcing components (direction of the force).

The structure of the chapter is as follows: In section 4.1 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 this study. Section 4.2 starts with an illustration of a 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 the 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 3D residual circulation spatial patterns in relation to the GWD distribution. In section 4.3, we give a summary of our results, discuss potential implications of our results for a variety of research topics (Sudden Stratospheric Warmings, atmospheric blocking, teleconnections and a compensation mechanism between resolved and unresolved drag) and outline future directions of our work. In the appendix to this chapter, we present supporting information on the BDC longitudinal structure in November.

4.1 Data and methodology

4.1.1 Model description and configuration

We use the Middle and Upper Atmosphere Model (MUAM), which is a nonlinear 3D mechanistic global circulation model. It has a horizontal resolution of $5^o \times 5.625^o$ and extends in 56 vertical layers up to an altitude of about 160 km in log-pressure height [Pogoreltsev and Jacobi, 2007]. At 1000 hPa, 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 account for, e.g., orography or some radiative processes including 3D water vapor or surface albedo. However, the troposphere is necessary for stationary PW forcing and the generation and propagation of traveling 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 everywhere and for the simulation with strongest forcing it reaches locally to magnitudes around 2 K/day in a zonal mean (as shown in Fig. 4.1, part c). 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 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 analyzed. Since the lower boundary conditions are taken as a decadal mean January mean, this interval refers to an average January climatological state. Monthly zonal means of wind, temperature and GWD are given in Fig. 4.1. 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 largest within the jets reaching 4 m/s in the summer easterlies. These values do not have a meteorological meaning and are provided here to demonstrate that MUAM has a rather small variability within the analysis interval.

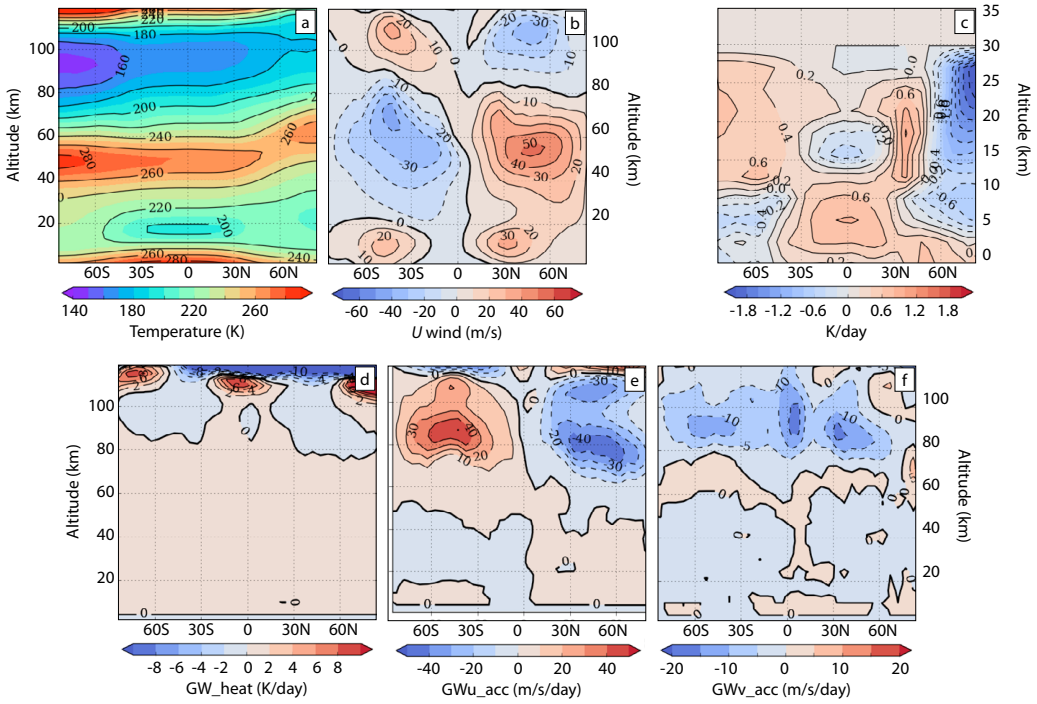

Figure 4.1: Mean January zonal means of temperature (a), zonal wind (b), GW induced heating (d) GW induced zonal wind (e) and meridional wind (f) acceleration for the reference simulation. Additionally, January mean zonal mean nudging strength for the strongest GWD injection (SSWbox simulation in Table 4.1) is shown (c).

IGWs are parameterized after a linear Lindzen [1981] type scheme updated as described in Fröhlich et al. [2003] and Jacobi et al. [2006], and they are initialized at an altitude of 10 km with six different phase speeds ranging from 5 to 30 m/s, each propagating in eight different azimuth angles, and with GW vertical velocity amplitudes with an average value of 0.01 m/s. As input for the IGW parametrization scheme, we modified the IGW 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

Name	Distribution of the artificial GWD	Artificial gcu per grid-point of the artificial area (m/s/d)	Zonal mean gcu in the altitude of artificial GWD (m/s/d)	Artificial gcv per grid-point of the artificial area (m/s/d)	Zonal mean gcv (m/s/d)	Artificial gt per grid-point of the artificial area (K/d)
Ref	$\sim$	$\sim$	0.011	$\sim$	-0.001	$\sim$
Box0.5	box	-0.5	-0.073	-0.5	-0.085	0.05
Zon0.5	ring	-0.073	-0.073	-0.085	-0.085	0.05
Box0.5pos	box	-0.5	-0.073	0.1	0.018	0.05
Zon0.1pos	ring	-0.073	-0.073	0.018	0.018	0.05
Box0.1	box	-0.5	-0.073	-0.1	-0.016	0.05
Zon0.1	ring	-0.073	-0.073	-0.016	-0.016	0.05
Box0.1gcu	box	-0.5	-0.073	$\sim$	-0.001	$\sim$
Box0.1gcv	box	$\sim$	0.011	-0.1	-0.016	$\sim$
Box0.1gt	box	$\sim$	0.011	$\sim$	-0.001	0.05
10box	box	-10	-1.706	-0.1	-0.016	0.05
10zon	ring	-1.706	-1.706	-0.016	-0.016	0.05
SSWbox	box	-70	-12.018	-0.1	-0.016	0.05
SSWzon	ring	-12.018	-12.018	-0.016	-0.016	0.05

Table 4.1: Sensitivity simulation names and GWD settings for zonal wind drag (gcu), meridional wind drag (gcv) and heating due to IGWs (gt) within the box. Note the gcu enhancements are negative because the drag is westward directed. The distribution describes whether the artificially enhanced GWD is implemented only for certain longitudes (Box) or zonally uniform (Zon). The tilde "$\sim$" indicates that values are unchanged w.r.t. the reference simulation.

IGW weights are calculated from these data by dividing the potential energy at each grid point by its global mean. This setup has a positive impact on some climatological features in MUAM. Nevertheless, the effect on the horizontal distribution of GWD in the stratosphere is negligible. We will refer to this setup as the reference simulation. Zonal (gcu) and meridional (gcv) flow acceleration as well as the heating due to breaking or dissipation of IGWs (gt) is calculated by the parametrization scheme.

To examine and to demonstrate the effect of spatial distribution of the IGW activity we performed a set of sensitivity simulations (Table 4.1) with artificially changed GWD imposed on the model by modulating the IGW parametrization 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 4.1) is given by "Gcu+distribution+gcv", where the basic value of gcu of 0.5 m/s/day is not stated.

The enhancement is performed for a certain 3-dimensional 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 IGW activity described by Šácha et al. [2015]. This refers to the "box" distribution in Table 4.1 (an example is shown in Fig. 4.2, 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 zonally within the same latitude range like the box. 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 4.1). 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 localized effect of IGW 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.

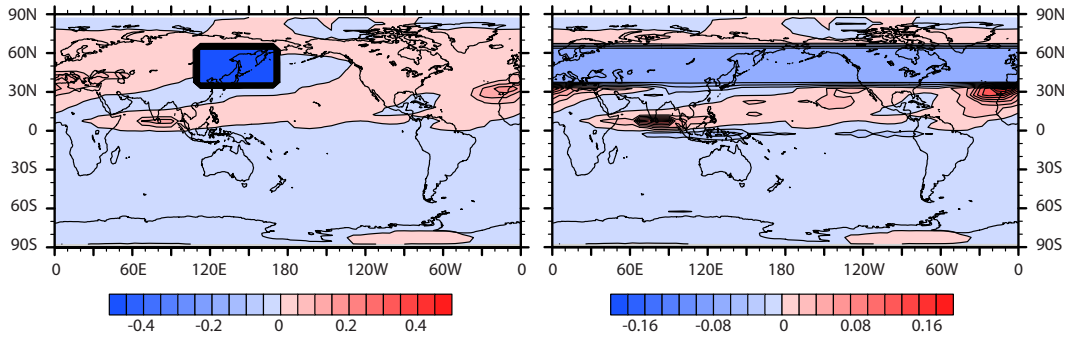

Figure 4.2: Two examples of the GWD enhancement horizontal distribution imposed between approximately 20 and 30 km of log pressure height. Left panel: box distribution (Box0.1 simulation). Right panel: ring distribution (Zon0.1 simulation). Colors indicate IGW induced zonal acceleration [m/s/day]

Although it is impossible to directly compute the IGW 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 IGWs (maximum zonal mean values of 3 m/s/day below 40 km). Using ray tracing simulations, Kalisch et al. [2014] estimated a zonal averaged GWD to be around 20 m/s/day in the lower stratosphere. In our model simulations we are injecting three values of additional artificial zonal component of GWD, -0.5 m/s/day as a conservative enhancement and -10 m/s/day to demonstrate a big impact of the injection. In addition, an extreme case with -70 m/s/day is added to force substantial circulation changes.

Depending on the IGW type and on the direction of background winds the GWD has also a meridional component, which is usually poorly constrained by observations. We performed simulations with three different values of meridional IGW induced acceleration (-0.5 m/s/day, -0.1 m/s/day, 0.1 m/s/day). Directions of the zonal and meridional IGW induced acceleration were chosen based

on the assumption that the majority of IGWs in the EA/NP region in January will be of orographic origin and taking into account the prevailing directions of horizontal winds in the EA/NP region (see Šácha et al. [2015]). On this basis we argue that the 5:1 ratio between the zonal and meridional IGW 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.

4.1.2 Residual circulation

To highlight the importance of the stratospheric research in the EA/NP region and to show the robustness of our claim of an enhanced branch of the BDC in this 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]. Additionally, in the Supplement and in the Appendix, the comparison is shown between the vertical structure and longitudinal variability of the residual circulation and zonal cross sections of selected tracers from Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) volume mixing ratio profiles [KIT, 2016; see von Clarmann et al., 2009; Plieninger et al., 2015]. However, the interpretation of the differences 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 $O(\alpha^2)$ only for small amplitude α steady wave [Bühler, 2014].

In the subsection 4.2.3, we study consequences of the IGW hotspot for the longitudinal variability of the residual circulation (and BDC consequently) by means of the time mean 3D residual circulation according to Kinoshita and Sato [2013]. The time averaging inserts additional uncertainty in the 3D 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 hourly MUAM output fields. This configuration gives the strongest zonally averaged Stokes drift from several choices of the running mean. But, it is still smaller 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.

4.2 Results

To establish the timescales of the response, in Fig. 4.3 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 35.5 km log-pressure height, 13th model level). This level was chosen for our analyses because it is above the location of the artificially modified area and above the nudging extent, so it contains the atmospheric response only. In Fig. 4.3 section a, a Hovmöller diagram is given for the zonal mean zonal wind at 6.25 hPa level in the Ref simulation documenting that the model is essentially steady. Fig. 4.3 sec. b shows the time evolution of

variance of the zonal mean wind anomaly (Box0.1–Ref) and Fig. 4.3 sec. c shows the time evolution of zonal mean zonal wind for the 10box simulation. We can see the buildup of the response until approx. day 7 after the GWD injection and after that the structure of the response remains quasi-steady, with small variations of the magnitude only.

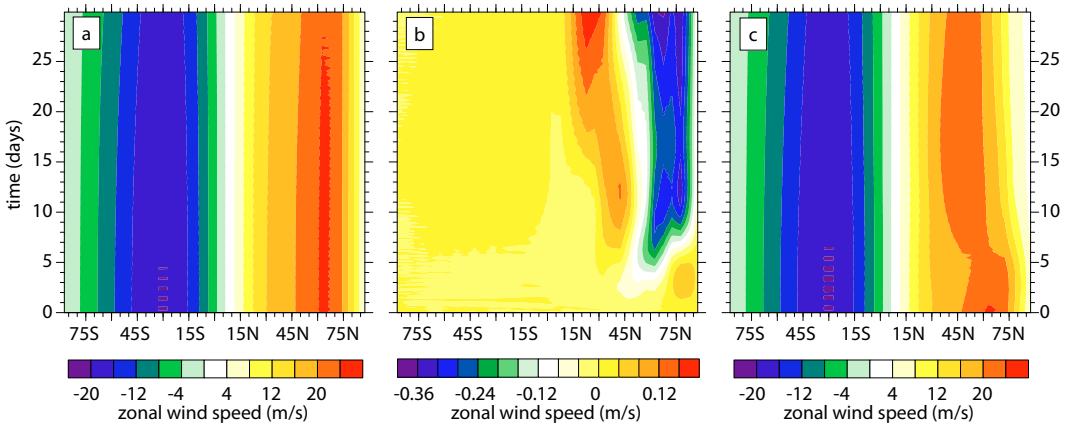

Figure 4.3: Hovmöller diagram of the zonal mean zonal wind for the Ref simulation (a), the zonal mean zonal wind difference with Box0.1 (b) and the zonal mean zonal wind for the 10box simulation at the 6.25 hPa level.

In contrast to this, the zonal mean zonal wind time evolution from the so-called SSW simulations (Fig. 4.4, sec. a and b) do not reach a steady state in the course of the 30 days simulation and therefore the results based on those simulations are presented at particular time steps or as animations in the Supplement. Results of other simulations (Table 4.1) 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.

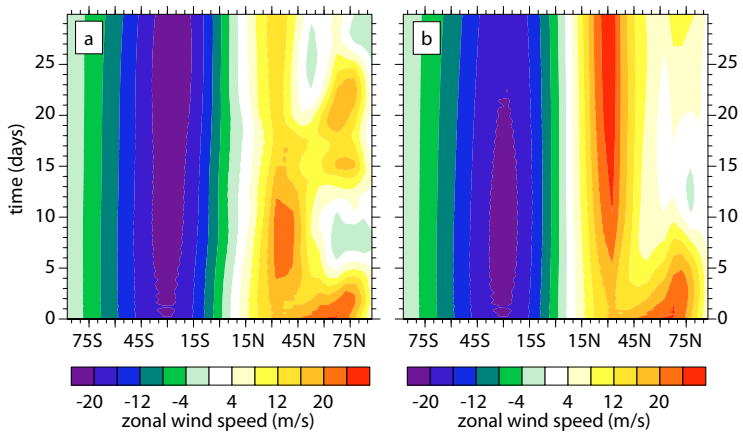

Figure 4.4: Hovmöller diagram of the zonal mean zonal wind for the SSWbox simulation (a) and the SSWzon simulation (b) at 6.25 hPa.

Except for the SSW simulations, our study is focused mainly on the mean response to a monthly mean GWD distribution, because from observational analyses we usually have information on the IGW activity distribution on a monthly or seasonal basis [Šácha et al., 2015]. The short-term response, which would be

arguably more relevant to the real atmosphere taking into account the intermittency of large amplitude GWs [e.g. Hertzog et al., 2012; Wright and Gille, 2013], is not well captured by the mechanism of constant GWD injection, which will be discussed in the final section. Still, there are some interesting results mentioned in the course of the study, e.g. note the agreement with the time scale of the transient response build up in Fig. 11d of Cohen et al. [2014], where it is related to the life cycle of the PW breaking.

4.2.1 Atmospheric response to variations in GWD and SSW

Fig. 4.5, sec. 1A shows the mean (7th–30th day) horizontal wind and geopotential field at the 6.25 hPa level (13th model level) for the Ref 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. 4.5, sec. 2A–2D) shows horizontal wind and geopotential anomalies for the 10box (Fig. 4.5, sec. 2A) and Box0.1 (sec. 2B) simulations and differences between simulations with conservative GWD enhancements (Fig. 4.5, sec. 2C and 2D). The third row (Fig. 4.5, sec. 3A–3D) shows the same as the 2nd row, but for the artificial ring GWD configuration. Note the different scaling of the color bars, which is chosen according to the maximal and minimal value of geopotential (anomaly), so that the labels of the color bar give direct information on the magnitude of the differences in geopotential response.

The anomalies and differences are analyzed with special focus on the polar vortex response, since it will be shown below that the dynamical response on 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 it also highlights an important and different effect of each of the individual GWD components.

From comparison of Figs. 4.5, sec. 1B, 1C and 1D, 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. 4.5, sec. 1B). The *gt* component alone induces a positive anomaly of smaller magnitude northward and upstream of the area (Fig. 4.5, sec. 1C). In contrast to that, 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. 4.5, sec. 1D).

The respective geopotential responses in the Box0.1gcv 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 nonlinear model, the additivity of effects of different GWD components (Fig. 4.5, sec. 1B, 1C and 1D) seems to hold reasonably well as can be seen from the Box0.1 anomaly (Fig. 4.5, sec. 2B), where the forcing constitutes of exactly these components. Also, the differences between simulations with different meridional drag (compare Fig. 4.5, sec. 2C and 2D) show the same pattern as induced by the meridional drag only (Fig. 4.5, sec. 1D). The distribution of the response to the

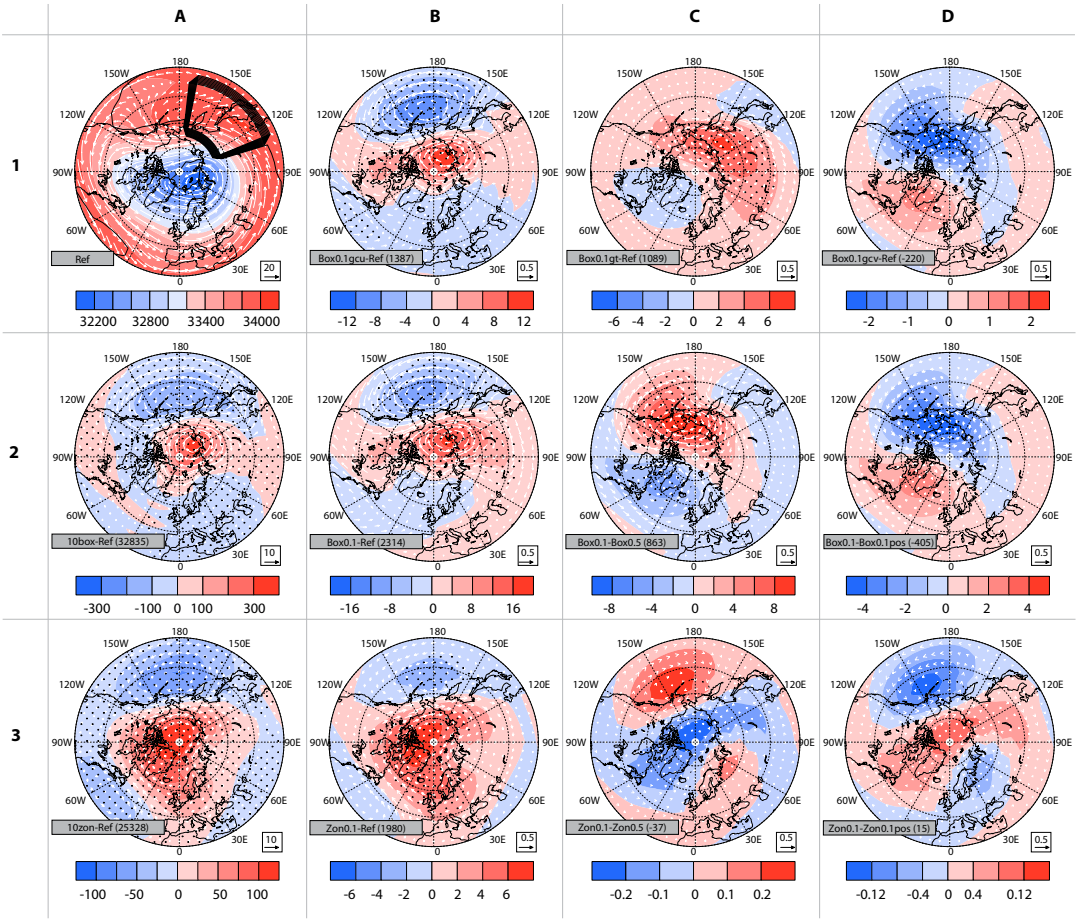

Figure 4.5: Mean geopotential and horizontal wind vectors at the 13th model level (6.25 hPa) for the reference simulation and differences for the sensitivity simulations with different GWD set-up. From top left (index 1A) to bottom right (index 3D): 1A) reference simulation overlaid with an illustration of the box area, 1B) reference-Box0.1gcu, 1C) reference-Box0.1gt, 1D) reference-Box0.1gcv, 2A) reference-10box, 2B) reference-Box0.1, 2C) Box0.1-Box0.5, 2D) Box0.1-Box0.1pos, 3A) reference-10zon, 3B) reference-Zon0.1, 3C) Zon0.1-Zon0.5 and 3D) Zon0.1-Zon0.1pos. Colors indicate geopotential height (gpm). Note the different scaling of the respective plots. Arrows refer to horizontal wind (m/s) 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.

meridional component suggests that a box *gcv* enhancement in this geographical position can influence the geopotential response in the area of location of the Aleutian High.

Another two important results are visible from the comparison of the plots in the second and third row of Figure 4.5. Firstly, 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 response given in the legend for Fig. 4.5, sec. 2A, 2B, 3A and 3B). In the box simulations (Fig. 4.5, sec. 2A and 2B) the response is typically dominated by a rather meridionaly oriented dipole pattern with a localized positive geopotential anomaly at the center of the polar vortex and negative geopotential anomaly at the location of Aleutian High. In the correspondent ring simulations (Fig. 4.5, sec. 3A and 3B) the geopotential response is more zonally uniform.

Secondly, there are large and significant differences (50 % or 25 % of the magnitude of the anomaly) between box simulations with slightly different setups of the meridional drag (Fig. 4.5, sec. 2B vs. 2C and 2D, respectively), while this is not true for ring GWD enhancements (few percent; see Fig. 4.5, sec. 3B vs. 3C and 3D). 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. 4.5, sec. 2B) reaches maximal values slightly below 1 m/s. Anomalies for the 10box simulation (Fig. 4.5, sec. 2A), 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 ratio 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 IGW 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. 4.4, sec. a 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. Now we show additional results from the SSWbox and SSWzon simulation in the form of animations of geopotential and horizontal wind field response at 6.25 hPa (Animation 1a and 1b in the Supplement) and a snapshot at 280 hours after the injection to show the response at a developed stage of the SSW.

In response to a strong GWD increase in a box we observed a vortex displacement (Fig. 4.6, sec. a and Animation 1a in the Supplement) and in response to a strong GWD increase in a ring we obtain a vortex split like event (Fig. 4.6, sec. b 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, from approximately 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 Northern America where it stays till the end of the simulation.

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 approximately ten days after the injection. This is a potentially very interesting result suggesting that a symmetric forcing favors vortex split and localized forcing favors displacement events, but the robustness of this claim needs to be tested in future work for various initial vortex states.

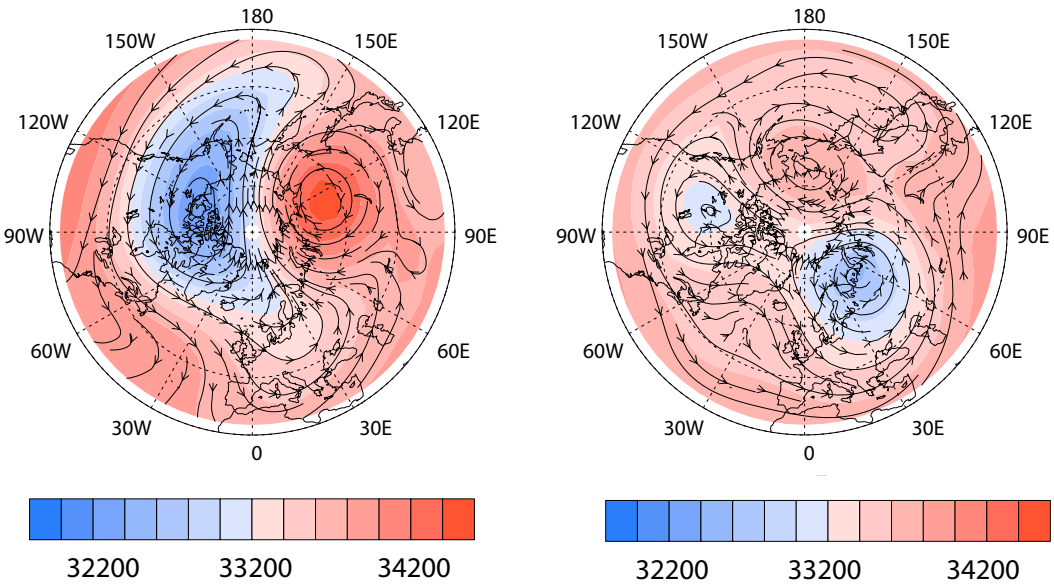

Figure 4.6: Geopotential (colors, given in gpm) and horizontal winds (stream lines, given in m/s) for the SSWbox (left) and SSWzon (right) simulation at the 13th model level (6.25 hPa) at 280 hours after the injection.

For illustration, in Fig. 4.6 we show the geopotential field and horizontal wind speed 280 hours after the GWD injection, when the vortex split develops (Fig. 4.6, sec. b) and the vortex displacement is in its mature state (Fig. 4.6, sec. a). The vortex displacement event develops more quickly, as seen from comparison of Fig. 4.4, sec. a and b or from the 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).

4.2.2 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 E-P flux diagnostics and Fourier transform (FT) analysis of geopotential anomalies.

Fig. 4.7 shows the mean (7th–30th day) E-P flux and its divergence for the Ref simulation (Fig. 4.7, sec. a), Box0.1 simulation anomalies (Box0.1-Ref; Fig. 4.7, sec. b), Zon0.1 simulation anomalies (Fig. 4.7, sec. c), the difference between the Box0.1 and Zon0.1 simulations (Fig. 4.7, sec. d) and mean mean E-P flux and its divergence for the 10box simulation (Fig. 4.7, sec. e) and respective anomalies (Fig. 4.7, sec. f). Note that we show the E-P flux divergence as a force per unit area (units $kg{\cdot}m^{-1}s^{-2}$), not as an induced acceleration (units $m{\cdot}s^{-2}$), as in Hardiman et al. [2010], because 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. 4.7, sec. b, for the Box0.1 and Ref simulation differences, we find an anomalously weak E-P flux convergence (positive difference to the Ref simulation) centered at the equatorward flank of the GWD enhancement area and an anomalous convergence in a broad area around 60^oN. This pattern is similar for the Zon0.1 simulation anomalies (Fig. 4.7, sec. c), but much weaker and with the anomalous convergence starting more poleward. It is also similar in the 10box simulation anomalies (Fig. 4.7, sec. f), 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. 4.7, sec. a), 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 Ref 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. 4.7, sec. b, c, d and 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 in the GWD region. But this mechanism fails to explain some features in Fig. 4.7, e.g. the E-P flux divergence emerging in the E-P flux field in the 10box simulation (Fig. 4.7, sec. e). 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 associated with the evidence given by Holton [1984] that a zonally asymmetric IGW breaking possibly generates PWs in the mesosphere. In connection with that 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 of different types depending on background conditions and on the geometry of the drag region. We

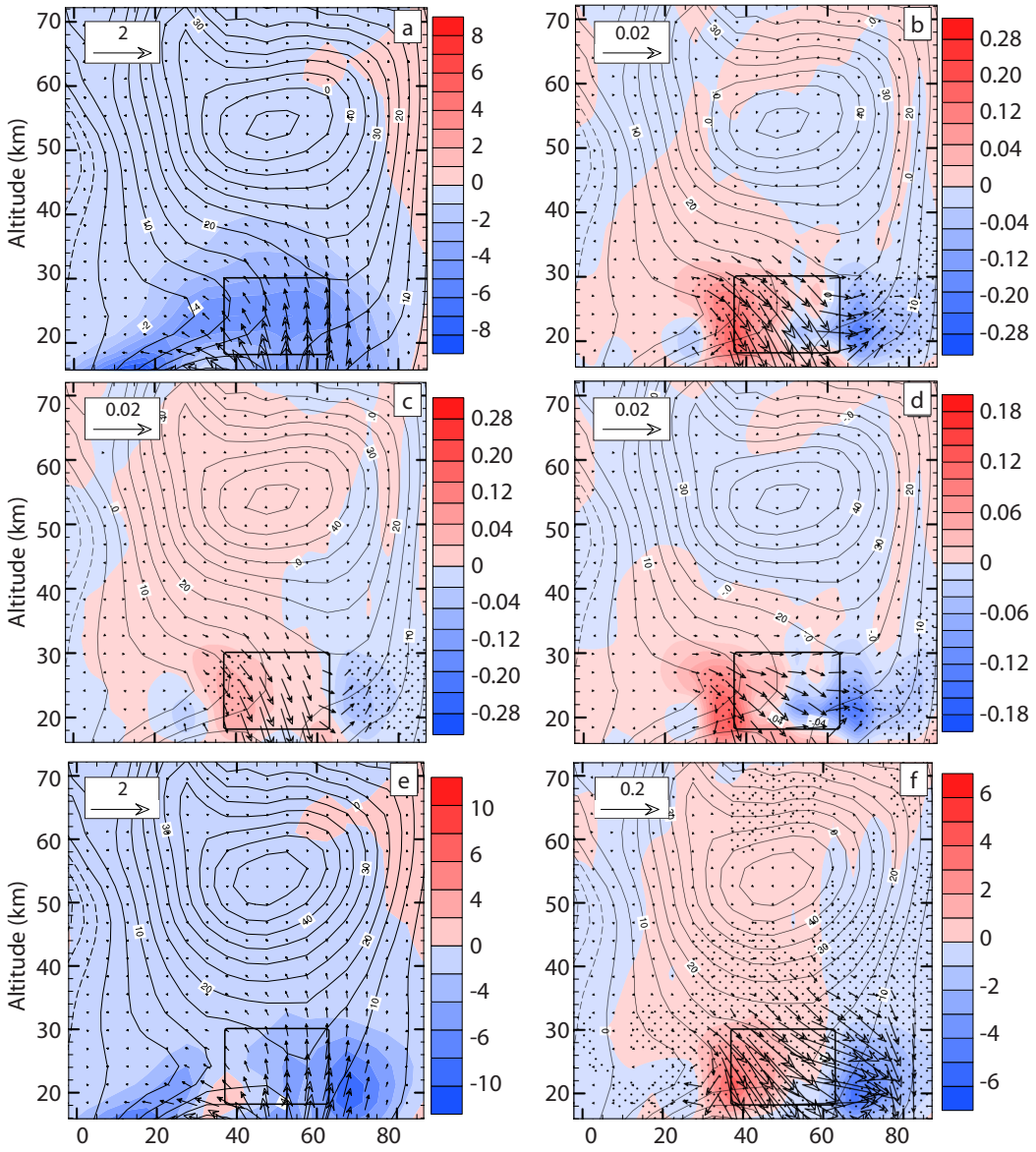

Figure 4.7: Mean E-P flux vectors ($kg \cdot s^{-2}$, arrows are scaled according the relative distances of the plot) and its divergence (colors in $kg \cdot m^{-1} s^{-2}$) for Box0.1 (a), its anomalies (b), Zon0.1anomalies (c), difference between the Box0.1 and Zon0.1 simulation, mean E-P flux and its divergence for 10box (e) and its anomaly (10box-Ref) (f). Note that scales are adjusted for each sub-figure, except the plots of anomalies (b,c) 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. 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.

can find support for this mechanism from the E-P flux difference between Box0.1 and Zon0.1 simulation (Fig. 4.7, sec. d).

In the previous section, we have shown that the box enhancement induces a stronger zonal mean geopotential response than the corresponding ring enhancement. So 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. 4.7, sec. d). However, regarding E-P flux vectors, Fig. 4.7, sec. d 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. 4.7, sec. b, 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. 4.7, sec. b, d and 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 60–80^{o}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. 4.7, sec c) reach only 1–2 % of the reference values, locally. Anomalies of the 10box simulation (Fig. 4.7, sec. f) 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. 4.7, sec. e, 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. 4.8, E-P flux diagnostics is presented at particular moments (1 and 5 days) after the GWD injection for the SSWbox and SSWzon simulations. The anomalous E-P fluxes in those highly nonlinear 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. So, 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 as Animation 2 in the Supplement.

In Fig. 4.8, sec. a and b, one can clearly see the generation of PWs by the box enhancement. Five days after the GWD enhancement (Fig. 4.8, sec. b), 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^{o}N and between 60 and 80^{o}N. One day after the GWD injection (Fig. 4.8, sec. a), the E-P flux divergence area is located at the southern flank of the GWD and generates horizontally, southward propagating PWs only. In Fig. 4.8, sec. a, in the majority of the GWD region, we can see also 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. 4.8, sec. c and d), where in

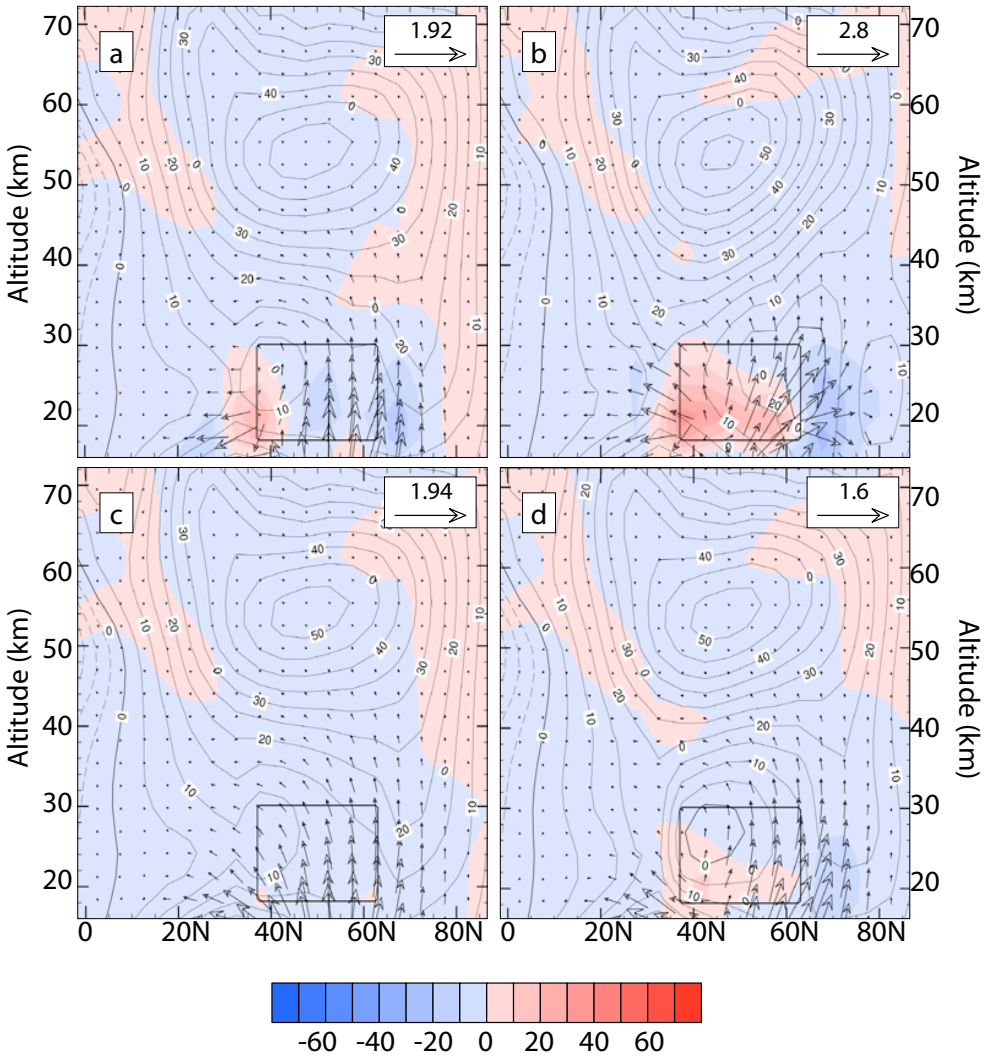

Figure 4.8: E-P flux vectors ($kg \cdot s^{-2}$, arrows are scaled according the relative distances of the plot) and its divergence (colors in $kg \cdot m^{-1}s^{-2}$) for the SSWbox simulation at 1 day (a) and 5 days (b) after the GWD injection, for SSWzon 1 day (c) and 5 days (d) after the GWD injection. In all panels contours of zonal mean zonal wind from the respective simulation and time step are overlaid with an increment of 10 m/s. All panels are overlaid with selected contour of gravity wave induced zonal acceleration to illustrate the location of the artificial GWD.

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.

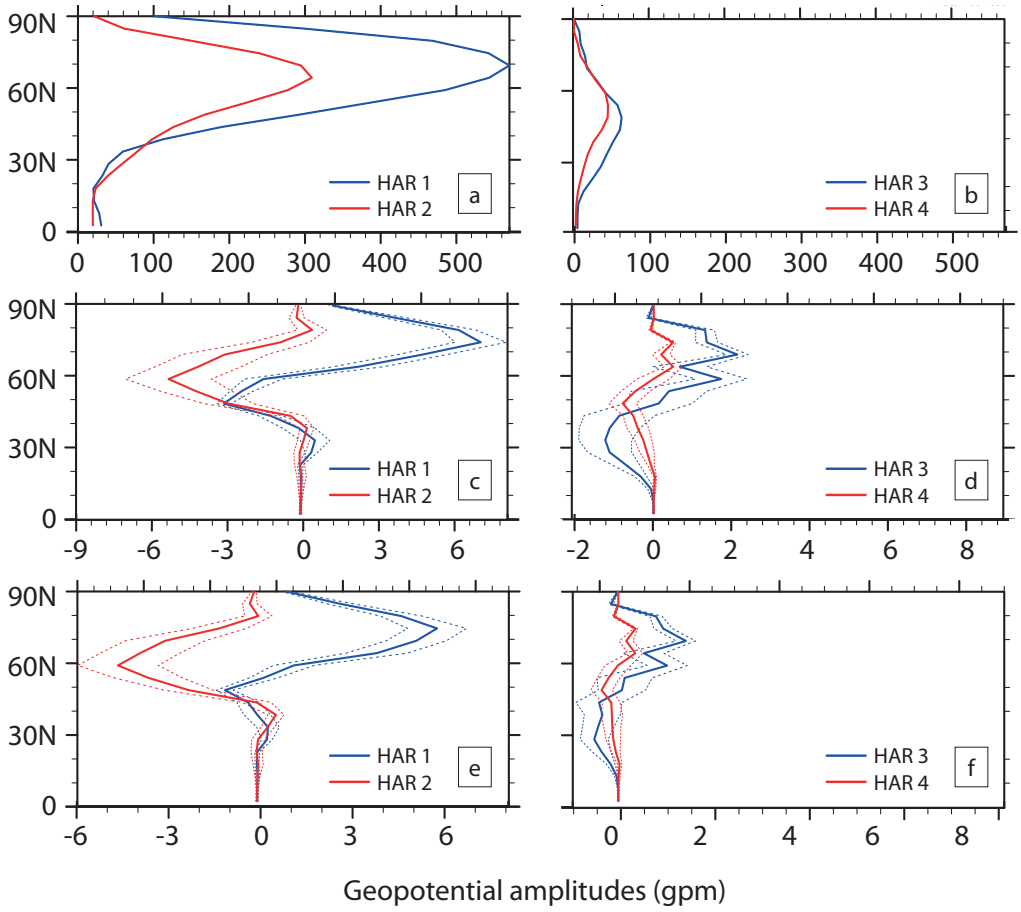

Figure 4.9: Mean (7th to 30th day) latitudinal structure of the amplitude of selected harmonics for the Box0.1 simulation. From top left to bottom right: a) harmonics 1 and 2 for Box0.1, b) harmonics 3 and 4 for Box0.1, c) differences of a) from the reference simulation, d) differences of b) from reference simulation, e) differences of a) from Zon0.1, f) differences of b) from Zon0.1. At approx. 35 km log-pressure height. Units are given in *gpm*. Dotted lines show the standard deviation differences.

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. 4.9, sec. a and b. Anomalous amplitudes (Box0.1-Ref simulation) are presented in Fig. 4.9, sec c, d and differences from the Zon0.1 simulation are shown in Fig. 4.9, sec. e and 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-Ref; Fig. 4.9, sec. c). The amplitude anomaly is positive starting at the northern flank of the artificial GWD (37.5–62.5^{o}N) and further poleward. The maximum is gained between 70–75^{o}N. Another smaller, but still significant, region of positive wave-1 amplitude anomaly is located around 30^{o}N south of the GWD. Smaller negative wave-1 amplitude anomaly lies inside the GWD area. In the Box0.1 simulation, wave-2 (Fig. 4.9, sec. c) has a pronounced negative amplitude anomaly inside the latitudinal belt encompassing the enhancement region. For wave-3 (Fig. 4.9, sec. d), we find positive anomalous amplitudes starting from central latitudes of the GWD region and ending around 80^{o}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^{o}N. The effect on wave-4 amplitudes is almost negligible (Fig. 4.9, sec. d). 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. 4.9, sec. c and d) and differences with Zon0.1 simulation (Fig. 4.9, sec. e and 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.

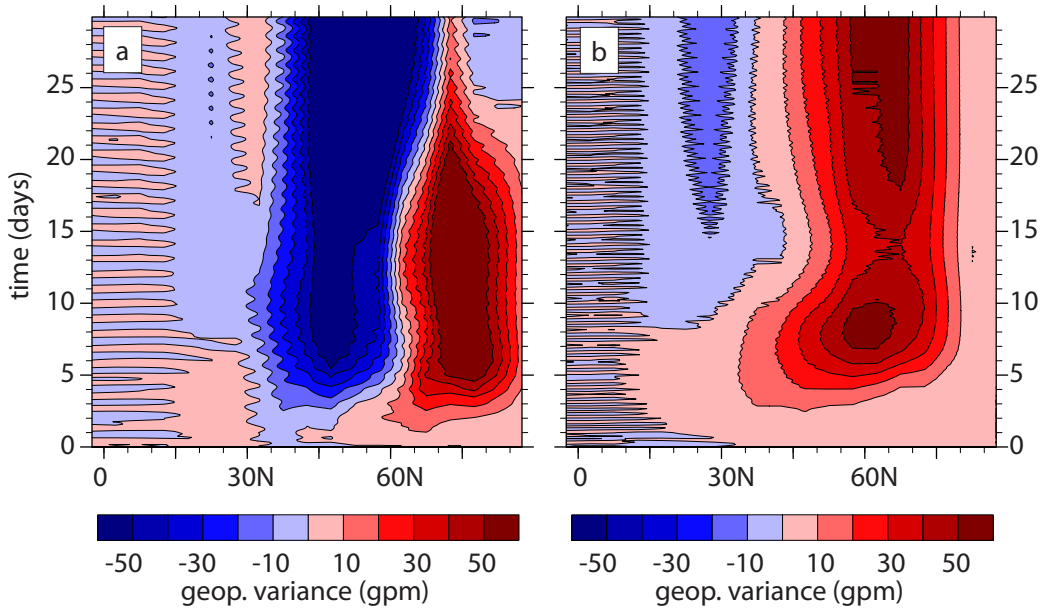

Figure 4.10: Time evolution of a wave-1 (a) and wave-3 (b) amplitude difference with respect to the reference run, as given by the FT of geopotential height at approximately 35 km log pressure height for Box0.1 simulation. Units are given in *gpm*.

Another indication that the PWs are indeed generated by the GWD box enhancement is given in Fig. 4.10, where the time evolutions of the anomalous wave-1 and wave-3 amplitudes are presented. Especially in the first approximately seven days from the GWD injection, we can observe a slow propagation of anomalous wave-1 (Fig. 4.10, sec. a) and wave-3 (Fig. 4.10, sec. b) 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. 4.10

most likely originate from a non-linear interaction between anomalously generated inertia IGWs and solar tides [see e.g. Walterscheid, 1981]. Those inertia IGWs 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. 4.10).

4.2.3 Residual circulation response

The first row of Fig. 4.11 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. Firstly, 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. 4.11, sec. e). For the 10box simulation the differences reach about 40 % and create a similar pattern as for the conservative enhancement (Fig. 4.11, sec. f). 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 constitute together a butterfly like pattern in the box-ring differences centered at approximately 45^oN (the center of the enhancement region) and influencing a shallow BDC branch. Taking into account the reference field (see Fig. 4.11, sec. a) we can explain this pattern by a faster northward advection starting at approx. 45^oN and stronger subsidence northward of 60^oN. 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^oN.

We observe a similar but stronger pattern in the anomalies (Fig. 4.11, sec. c and d), with the mean residual circulation mass flux anomaly reaching up to 5 % for the Box0.1 simulation and more than 60 % for the 10Box simulation. The position of the anomalous residual circulation patterns corresponds with the E-P flux divergence anomalies (Fig. 4.7, sec. b and f), where, for the box simulations, we observed anomalous E-P flux divergence at the southern flank and convergence north of the GWD region. In Fig. 4.11, sec. c and d the butterfly like pattern is centered more southward (35^oN) than in box-ring differences and the anomalous pattern on the south of the GWD region is not as well pronounced and appears to be shifted above the GWD region for the 10box simulation (Fig. 4.11, sec. d).

In the upper stratosphere there are anomalies up to 2 % only for the Box0.1 (Fig. 4.11, sec. c) and locally around 25 % for the 10box simulation (Fig. 4.11, sec. d). The box simulations (not significant for Box0.1) show weaker subsidence towards the polar vortex center than the Ref simulation in the upper stratosphere and there is also anomalously low mass flux poleward and downward between

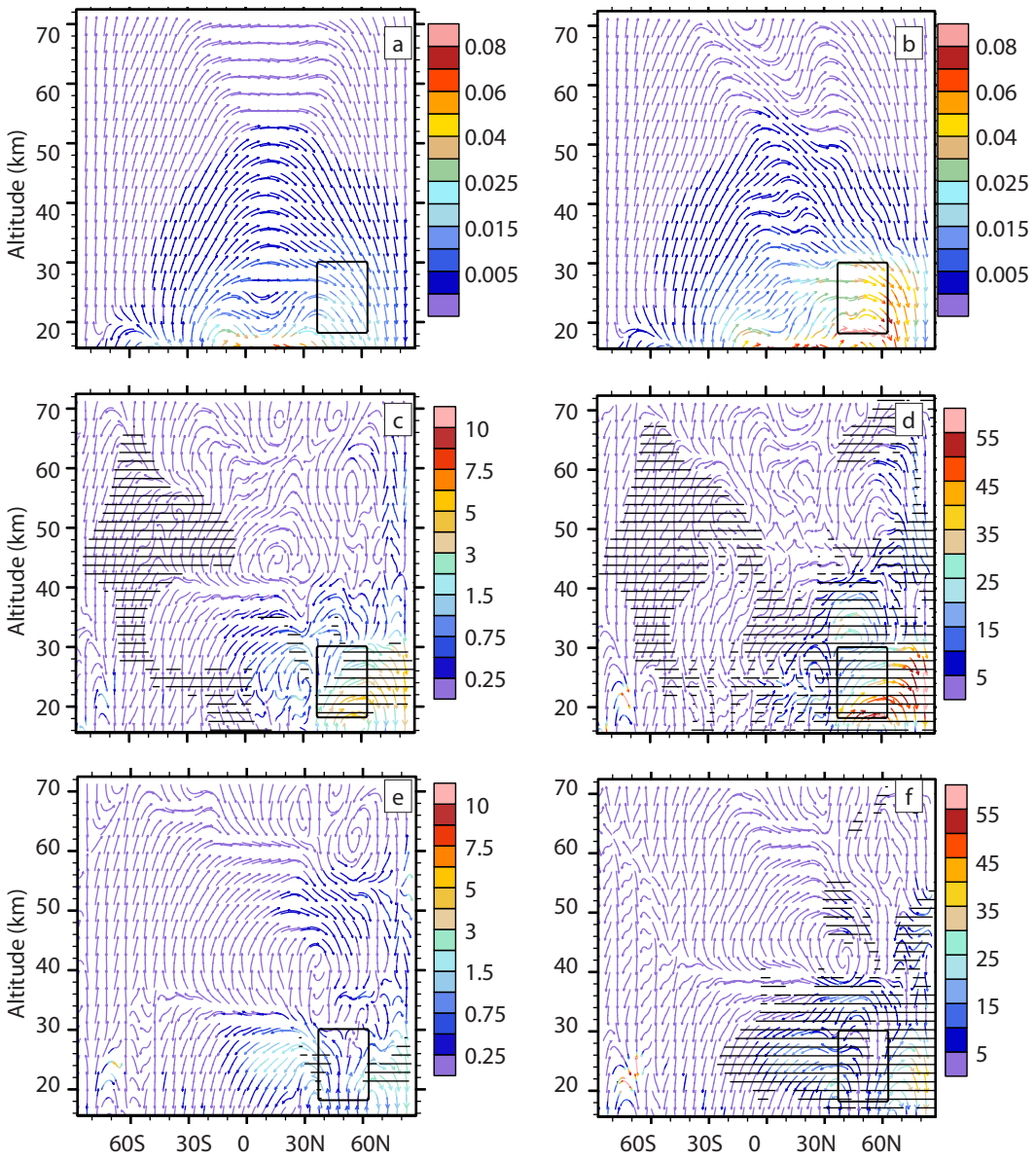

Figure 4.11: Mean January zonal mean residual circulation (stream lines for illustration of direction only) and its mass flux (colors, in $kg \cdot m^2 s^{-2}$) on the left (from top to bottom: Ref simulation (a), relative Box0.1-Ref simulation anomaly (c), relative Box0.1−Zon0.1 simulation difference (e)) and on the right (from top to bottom: snapshot of the SSWbox simulation at 5 days after the GWD injection (b), 10box-ref 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.

30 and 40 km of height above the GWD enhancement region. For both box enhancements, there is a large area of statistical significant anomalies giving a weak hint of lesser upwelling in the SH stratosphere (Fig. 4.11, sec. c and d). The differences between the two sets of box and ring GWD configuration are not significant in the SH (Fig. 4.11, sec. e and f).

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. 4.11, sec. b, snapshot for a SSWbox run) is constantly flown around inducing a significant mean anomaly (Fig. 4.11, sec. c and d) with anomalous upwelling in its southern part and downwelling on the northern flank. But, the GWD region (obstacle) creates also 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 gets affected by the NH anomalous forcing. 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 Pole to the lowest analyzed 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 IGW hotspot for the longitudinal variability of the BDC were studied by means of 3D residual circulation analysis according to Kinoshita and Sato [2013]. 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 penetrating to lower levels than elsewhere. This is in agreement with Fig. 3 in Demirhan Bari et al. [2013]. Supporting results highlighting the importance of future research in this region are given in the supplement. In Fig. S.18 in the supplement we present a thirty-year average January MSR total ozone column field with a total ozone column maximum located in the EA/NP region. In Fig. S.19 in the supplement, longitudinal cross-sections of MIPAS CH4 volume mixing ratios show a peak of subsidence around 15 km in the EA/NP region (at 140^{o}E) and the interesting massive upwelling branch east of it.

To evaluate the possible role of the IGW activity in the longitudinal variability of the BDC, we present longitudinal cross-sections of the reference 3D vertical residual velocity and Box0.1 anomalies going from the northern to southern part of the artificial GWD (Fig. 4.12). From longitudinal cross-sections of the reference vertical residual velocity (left side of Fig. 4.12), 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.

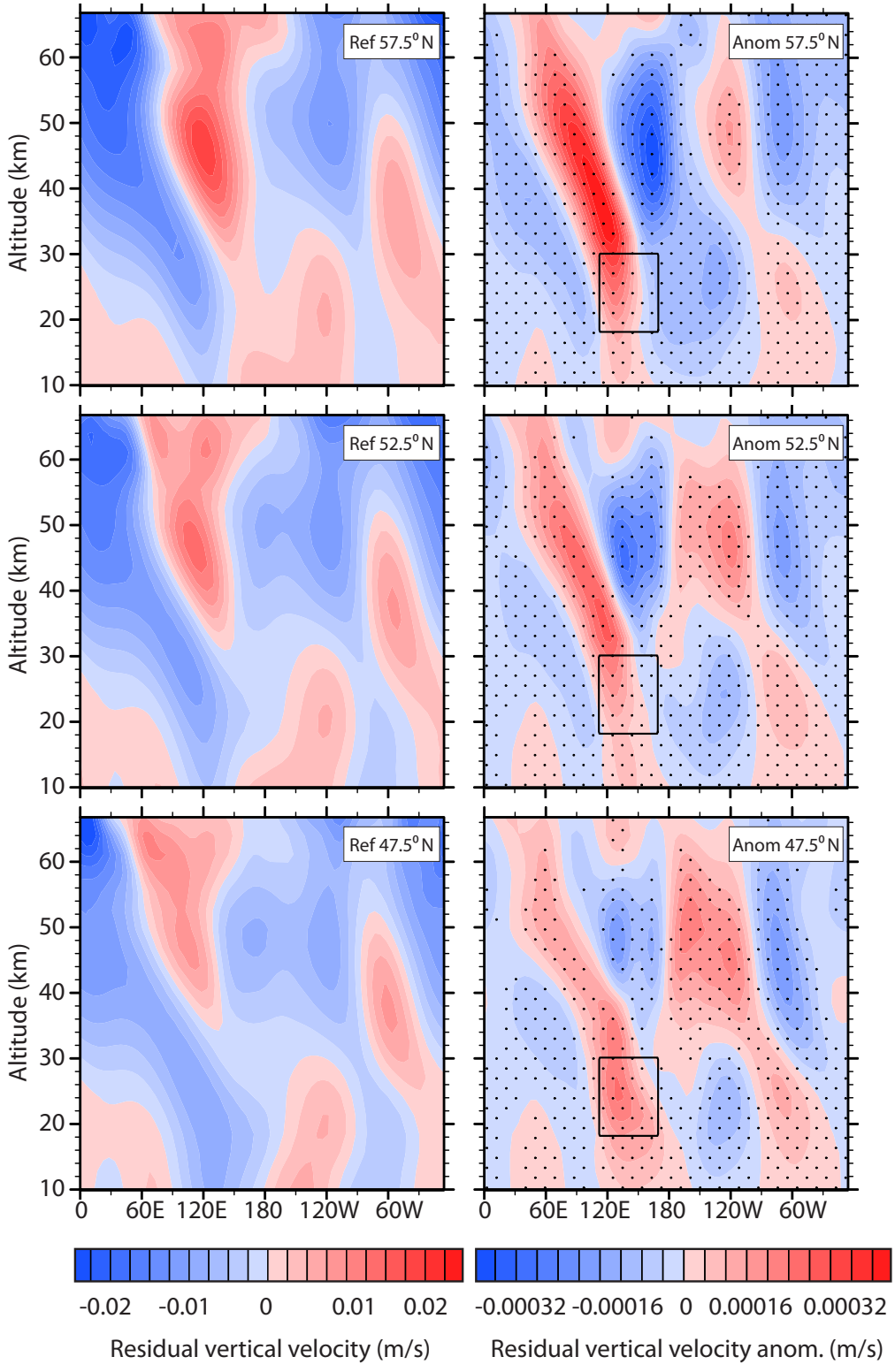

Figure 4.12: Mean (7th to 30th day) January longitudinal cross-sections of reference residual vertical velocity ($m \cdot s^{-1}$; on the left) and Box0.1 simulation anomalies (on the right) at selected latitudes. The contours illustrate the position of the artificial GWD. The statistical significance of the mean anomalies was computed by a t-test and regions with p values < 0.05 are stippled.

Šácha et al. [2015] hypothesized that the collocation of the EA/NP IGW hotspot and the enhanced BDC branch can be partly a consequence of the circulation induced by the IGW 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. 4.12, sec. f), the GWD induces predominantly anomalous upward flow. Anomalous subsidence strengthens when going further northward (Fig. 4.12, sec. b and 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^oN. These results show that IGWs can contribute to longitudinal variations in the BDC and not only the downwelling, but also upwelling patterns may be related with IGWs.

The magnitude of the vertical residual velocity anomalies maximizes around 2 % of the reference value for the Box0.1 simulation (Fig. 4.12, sec. b, d and f). For the 10box simulation (Fig. S.20 in the Supplement) 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 north-eastward 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 simulation 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 IGW 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 at least the GWD enhancement to reflect the changing background conditions.

4.3 Discussion and conclusions

We will begin this section by giving a brief summary of results. Then, we will discuss limits of our results stemming from the construction of the sensitivity simulations and afterwards, we will give some conclusions for different research topics in the middle atmosphere.

In this chapter, we presented results of a set of sensitivity simulation to find out the possible role of a localized IGW hotspot and also, generally, to demonstrate the influence of 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. 4.3), all simulations (Table 4.1) have reached a quasi steady state approximately 7 days after the GWD enhancement (Fig. 4.3). The average across this state was considered as a mean response later in the text. Subsection 4.2.1 was concerned with a mean geopotential response at the 6.25 hPa level (Fig. 4.5). 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 a purely constraint (from observations) meridional GWD component, especially for the polar vortex response, was highlighted. Most importantly, for simulations with the strongest GWD enhancement (SSWbox and SSWzon; Table 4.1), 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 localized forcing (Fig. 4.6).

In subsection 4.2.2 we studied the influence of the artificial GWD and of its distribution on the PW activity. We have found (Fig. 4.7) mean E-P flux convergence anomaly centered at the equatorward flank of the GWD enhancement area and an anomalous convergence in a broad area around 60^oN in response to the artificial GWD. The anomalies are bigger for the box enhancements and in the box simulations we have also 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. 4.9), where, for the box simulations we have found especially the wave-1 and also wave-3 mean amplitude anomalously enhanced. Also, the short-term response (Fig. 4.10) showed the origin of the enhanced amplitudes to be in the GWD area.

Subsection 4.2.3 has been concerned with a residual circulation response. It was shown that there are significant differences in a zonal mean residual circulation between different distributions of the same zonal mean GWD (Fig. 4.11). A butterfly like pattern in the box-ring differences was identified centered at approximately 45^oN (the center 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 3D residual circulation analysis (Fig. 4.12). We have found downwelling to the northeast (downwind) and upwelling to the southwest (upwind) of the GWD box showing that IGWs 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 IGWs 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 IGWs are known to be significantly filtered during SSWs [e.g. Holton, 1983; Limpasuvan et al., 2012]. Considering the intermittent nature of IGWs [e.g. Hertzog et al., 2012; Wright and Gille, 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 on the expense of absense of any steady response during the whole simulation. It is also a question, what is a more realistic illustration of the IGW effect on the atmosphere, a sudden GWD injection or smooth increase and decrease with e.g. a 10-day e-folding time to minimize the initial adjustment

noise as proposed by Holton [1983]. Also the spatial distribution of our artificial GWD is highly idealized (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 SSWbox/zon simulation are influencing some minor patterns of the response (e.g. the lee wave pattern in Fig. 4.10, sec. b).

In future work it is therefore necessary to take into account more realistic GWD distributions to address e.g. 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 IGW activity hotspot would favour enhanced wave-2 instead of wave-1 activity, and for comparison a chessboard-like or random distribution of GWD would possibly be more appropriate for comparison. Generally, the fact that the PW activity depends on the longitudinal GWD distribution (Fig. 4.7) 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 IGW hotspots relative to the climatological stationary wave location in the stratosphere and to analyze 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 ridge of the climatological wave-1 and our results show (Fig. 4.8) 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 recognized [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 IGWs (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 sourcing, propagation or breaking conditions for IGWs 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 IGW hotspot can be dependent on the Pacific Decadal Oscillation (PDO) phase and can play a role in the relationship between PDO and SSW occurence 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 common methodology [see e.g. Albers and Birner, 2014] for a review of SSW preconditioning concepts) to estimate e.g. the relative impact of IGWs 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 localized 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 localized 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 SSWs is a well-known correlation [e.g. Andrews et al., 1987; Martius et al., 2009; Nakamura et al., 2014; Albers and Birner, 2014] but the mechanisms standing behind 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 localized GWD (Fig. 4.5) suggesting that IGWs 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 localized 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 been studied yet. Also, horizontal IGW propagation is neglected in most climate model parameterizations [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, IGW 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 hypothesized by Scott and Dritschel [2005]. Traditionally, PWs are thought to be generated in the troposphere and propagate up on the polar vortex edge. But, as Scott and Dritschel [2005] pointed out, when wave amplitudes become large and nonlinear 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 – localized IGW 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 (localized 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 e.g. in the research of future BDC changes from climate models we need to be concerned not merely by 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 IGW activity hotspots. This suggests the need for improvement especially of the nonorographic IGW parameterization (though nonorographic IGW are usually assumed to have significant effect at higher altitudes than in the vertical range analyzed in this paper), since many global climate models use e.g. a globally uniform gravity wave

source function [Geller et al., 2013].

Code availability

MUAM model code is available from the authors upon request.

Data availability

MIPAS CH_4 volume mixing ratio profiles have been provided by Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research - Atmospheric Trace Gases and Remote Sensing through $https://www.imk-asf.kit.edu/english/308.php$. MSR total ozone is 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 have been provided by ECMWF through $http://www.ecmwf.int/en/research/climate-reanalysis/era-interim$.

Appendix – Structure of residual circulation in November

Zonal cross-sections of vertical residual velocity from MUAM simulation for November conditions gives us an interesting insight into the structure of the residual circulation in a month, when the IGW activity in the EA/NP region is indicated by all wave activity proxies (see the chapter 3) as absolutely unique on the NH.

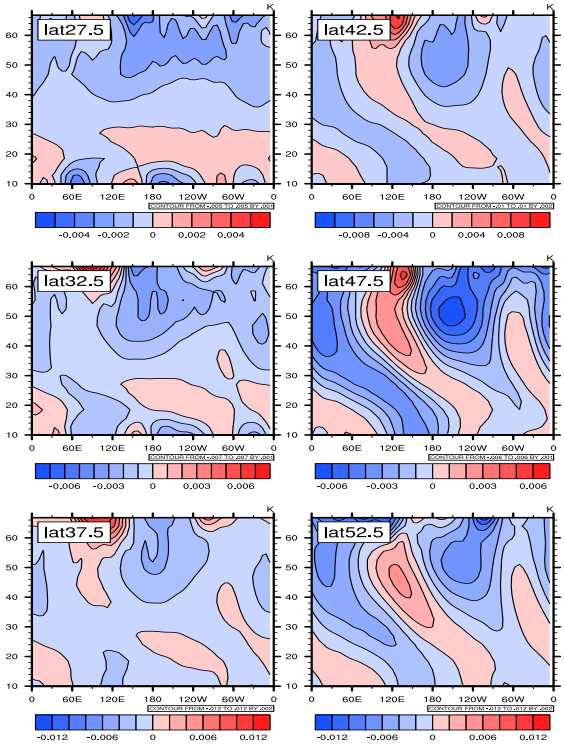

Figure 4.13: Zonal cross-sections at latitudes around 35^{o}N of residual vertical velocity from MUAM November simulations.

In Fig. 4.13 we can see that from roughly 40^{o}N the vertical residual velocity field is dominated by PW2 with the maximum subsidence branch penetrating to the lower stratosphere in the EA/NP region. Ridges and troughs of the wave show a characteristic westward tilt with height, which was already discussed in subsection 4.2.3.

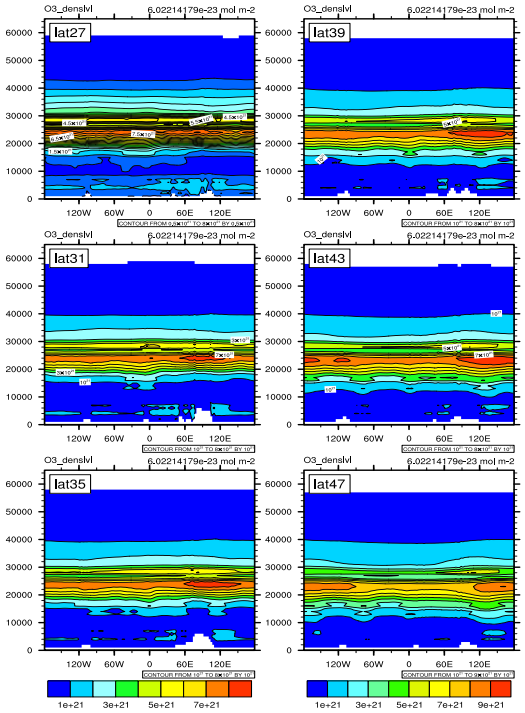

Figure 4.14: As in Fig. 4.13 but for the November 1997–2008 mean of O_3 mole content in an atmospheric layer.

Comparison of MUAM results and trace gases distributions confirms a realistic behavior of the model middle atmospheric circulation. From approximately 35^{o}N we can see evidence (Fig. 4.14, 4.16 and 4.17) of the enhanced branch of BDC above the EA/NP region reaching to the lowest levels around 140^{o}E (compare with the location of the IGW activity hotspot in the chapter 3).

On the other hand around 150^{o}E we can see in the fields of H_2O and N_2O (Fig. 4.16 and 4.17) signatures of upwelling of lower stratospheric air higher into the stratosphere. The other two peaks of PW2 are not as pronounced in tracer distributions.

From the distribution of CO (Fig. 4.15) we find that the tropopause (chemical) is located at lower altitudes above EA/NP region than at other longitudes starting roughly from 35^{o}N. Note also the local maxima of CO concentrations around 15 km altitude between 27.5 and 37.5^{o}N and around 130^{o}E, as they are located above the ocean (no artificial sources) this raises the question whether they can be considered as signatures of STE.

Another interesting feature can be seen in Fig. 4.14 from 31^{o}N to 35^{o}N – the maximum of the ozone mole content is located above the Himalayas. We have no explanation for this feature yet but we take it as a suggestion not to forget the possible role of Himalayas in future studies of EA/NP IGW hotspot and its dynamical implications.

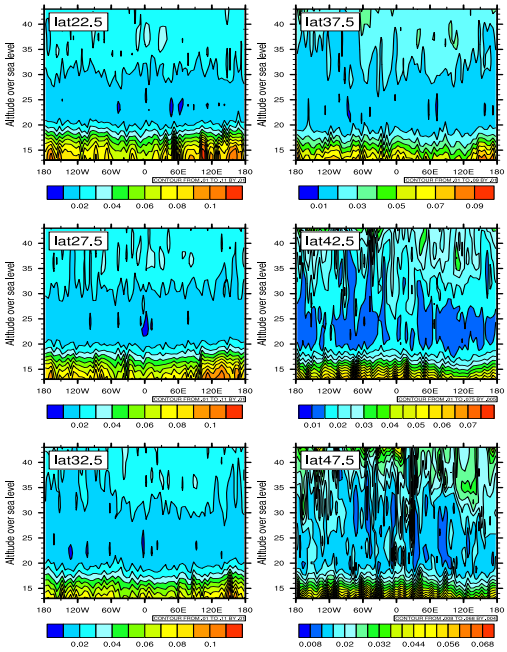

Figure 4.15: As in Fig. 4.13 but for the November 2009 volume mixing ratio of CO.

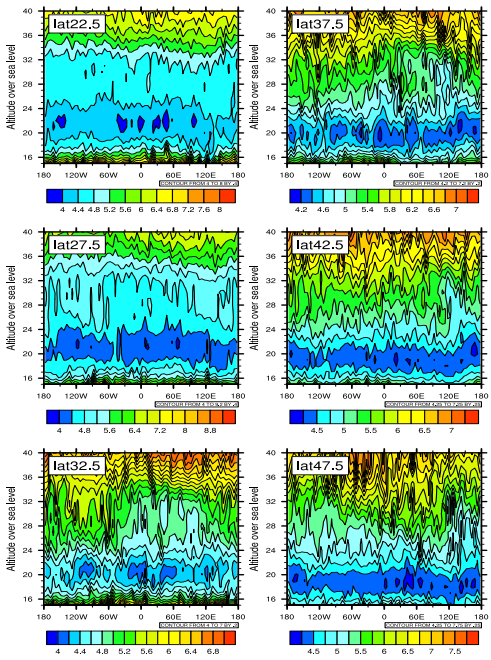

Figure 4.16: As in Fig. 4.13 but for the November 2009 volume mixing ratio of H_2O.

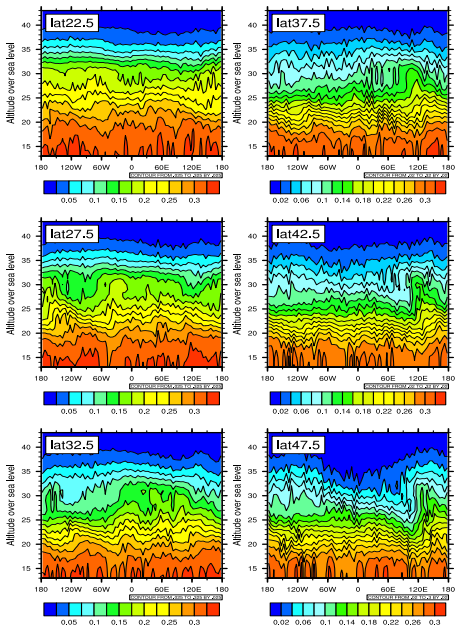

Figure 4.17: As in Fig. 4.13 but for the November 2009 volume mixing ratio of N_2O.

Epilogue

The Epilogue starts by giving a brief summary of each chapter and then general conclusions of this thesis are given. The thesis started with a review of the most classical results of the wave-mean flow theory. In the following of the first chapter, underlying kinematical mechanisms were illustrated and the influence of averaging was discussed. The discussion smoothly passed into an introduction of some of the newly emerging approaches to the wave-mean flow interaction theory.

In the second chapter, GPS RO data were highlighted to be a great tool for atmospheric monitoring and studies, especially for analyses of the IGWs in the upper troposphere and lower stratosphere region. Traditionally, the dry temperature GPS RO data were utilized in the IGW research, but in our study we argued for usage of the dry atmospheric density. It is the first quantity of state gained during the GPS RO retrieval process and is not burdened by any additional assumptions. A novel method for density background separation was introduced and a methodology for the IGW analysis is given using the density profiles. Various background choices were discussed and the correspondence between analytical forms of the density and temperature background profiles was examined. In the stratosphere, a comparison between the power spectrum of normalized density and normalized dry temperature fluctuations confirmed the suitability of the density profiles' utilization. In the height range of 8–40 km, results of the continuous wavelet transform were presented and discussed. Finally, the limits of our approach are discussed and the advantages of the density usage were listed. Those are above all the direct unfiltered information about nonhydrostatic IGWs and a familiar form of the density background.

In the third chapter, analysis has been presented of a stratospheric area of an anomalously low annual cycle amplitude and of the specific dynamics in the stratosphere over the EA/NP region. Using the GPS RO density data and the methodology from the previous chapter, IGW activity and breaking hotspot has been found in this region. Conditions supporting orographic wave sourcing and propagation were found. Other possible sources of IGW activity in this region were listed. The reasons why this particular IGW hotspot was not discovered before as well as why the specific dynamics of this region have not been pointed out were discussed together with the inaccuracy of using the mean potential energy as a wave activity proxy. Possible consequences of the specific dynamics in this region on the middle atmospheric dynamics and transport were outlined. Those were then examined in the fourth chapter.

The fourth chapter has been concerned with an analysis of sensitivity simulations performed with a mechanistic model to study a possible effect of a localized IGW breaking region (EA/NP region) on the large-scale circulation and transport. Also more generally, possible influence of the spatial distribution of gravity wave activity on the middle atmospheric circulation and transport was examined. The results indicated an important role of the spatial distribution of IGW activity for the polar vortex stability, formation of PWs and for the strength and structure of the zonal mean residual circulation. Finally, a possible role of a zonally asymmetric IGW breaking in the longitudinal variability of BDC was analyzed. Consequences of our results for variety of research topics (SSWs, atmospheric

blocking, teleconnections and a compensation mechanism between resolved and unresolved drag) were discussed.

The bottom line of this thesis is to give evidence that IGWs should be acknowledged to play a much bigger role in shaping the stratospheric dynamics than they currently are. Starting with the analytical description of the large scale circulation and dynamics in the stratosphere, which is traditionally described using the QG approximation, where IGWs possess an infinite speed of propagation, we must note that each change between the two consequent states of the analytical QG system can be a product of a sequence of a number of IGW related processes. IGWs are the most immediate and fastest way for communication of information in the vertical (apart from acoustic and acoustic-gravity waves with effects much higher in the middle and in the upper atmosphere) and they are ubiquitous in the atmosphere. So far most studies dealt with IGW impacts higher in the upper stratospheric/mesospheric region and with the modulation of IGWs by PWs [e.g. Cullens et al., 2015], while IGW effects in the lower and middle stratosphere received less attention. One of the reasons is that IGWs induce highest accelerations in the mesosphere, lower thermosphere region. But what is often forgotten is that the imposed drag force is much bigger in the stratosphere.

Šácha et al. [2015] indicated from observations a robust IGW breaking underway in the lower stratosphere of the Eastern Asia/Northern Pacific (EA/NP) region (third chapter of this thesis) and Šácha et al. [2016] demonstrated a possible dynamical impact of such a localized lower stratospheric GWD (Chapter four). Indeed, there are some other modeling studies showing the ability of the GWD in the lower stratosphere to e.g. significantly affect the development of SSWs [Pawson, 1997; Lawrence, 1997] or the large scale flow in the lower stratosphere and troposphere in general [McFarlane, 1987; Sandu et al., 2016]. Moreover, Kruse et al. [2016] showed that the GWD in the lower stratosphere is still underrepresented (e.g. in the MERRA reanalysis).

This thesis has brought to attention the aspect of spatial distribution, especially longitudinal variability of the GWD. In contrast to the PW activity and breaking, where there has been absolutely no emphasis in the literature about the geometry and longitudinal variability of the imposed drag force, the distribution of the IGW activity in the atmosphere is quite well known from observations. Also, from comprehensive IGW parametrizations and IGW resolving models we have information about the 3D distribution of the GWD. This distribution (especially for orographic IGWs) is highly zonally asymmetric and the sign of the drag force can have not only negative but also a positive sign locally. Therefore it has a little sense to compare the impact of PWs and IGWs based on the magnitude of their respective zonal mean drag forces, as it is a common methodology in the middle atmospheric research [see e.g. Albers and Birner, 2014]. In this thesis, the evidence has been presented that the localized GWD has larger dynamical impact than a corresponding zonally symmetric distribution.

It would be certainly correct to argue using a scale analysis that IGWs are a small scale phenomenon and so their effects should be local only, but we must be conscious that in the atmosphere IGW activity is often organized in large-scale hotspots (significant orography, monsoon, etc.) constituting a large-scale forcing. Although the IGW sourcing (adjustment processes), propagation and breaking is governed to some extent by processes in the stratosphere, there is

a significant portion of the IGW spectra created in the troposphere (orography, convection), with the highest amplitude modes breaking already in the lower or middle stratosphere. This opens new horizons for research of teleconnections between tropospheric (e.g. El Niño Southern Oscillation, North Atlantic Oscillation, Pacific Decadal Oscillation (PDO)) and stratospheric (polar vortex stability) phenomena, when the tropospheric variability almost immediately affects the GWD distribution in the stratosphere.

For a better illustration, one hypothetical example concerning the IGW hotspot in the EA/NP region: Kren et al. [2015] gave evidence that the Pacific Decadal Oscillation influences the frequency of SSWs confirming the impact of North Pacific anomalous circulation on the wintertime polar vortex [e.g. Woo et al., 2015; Kidston et al., 2015]. Woo et al. [2015] attributed the differences between PDO phase effects on the polar vortex strength to the PDO-induced tropospheric circulation anomalies over the North Pacific and consequent interference (constructive or destructive) of the anomalous wave-1 component with climatological PWs. Hypothetically, PDO connected sea surface temperature anomalies will lead to anomalous surface winds or convection in this region with consequent anomalous orographic or convective IGW creation. Those anomalous IGWs will immediately influence the occurrence and strength of the EA/NP hotspot leading to anomalous in-situ generation of PWs propagating horizontally in the lower stratosphere and influencing the polar vortex as was demonstrated in the fourth chapter. This does not need to work for the PDO only, a similar mechanism can also be behind e.g. the West Pacific teleconnection pattern or others. Of course, this is a very simplified view not taking into account the inner variability of the stratosphere, but this part of mutual interactions between the troposphere and stratosphere is the quickest and almost constantly present.

Although nonorographic IGWs are usually believed to have their effects higher in the middle atmosphere, de la Cámara et al. [2016] argued that they contribute to the IGW forcing in the stratosphere evenly as the orographic IGWs. Moreover, their distribution is also highly variable. This makes our results and conclusions applicable to the nonorographic IGWs as well. Since many global climate models use e.g. a globally uniform gravity wave source function [Geller et al., 2013], the inaccurate GWD distribution can introduce additional uncertainty espeicially in estimating the future circulation change. For example, He et al. [2015] found projected changes under the RCP4.5 and RCP8.5 scenarios at 850 hPa in the western North Pacific subtropical high region to be around 1 m/s, which is comparable in magnitude with the horizontal wind response in this region to a conservative GWD enhancement in the stratosphere (fourth chapter). To sum up, our results suggest that the global climate models should be at least able to mimic the observed IGW activity distribution, since it has direct implications e.g. for the strength and structure of the residual circulation, when the matching between observations and the robust model stratospheric change projections determined to a large extent by a speeding-up of the BDCis one of the most pressing issues of todays climate research.

Bibliography

Albers, J. R. and Birner, T.: Vortex preconditioning due to planetary and gravity waves prior to sudden stratospheric warmings, Journal of the Atmospheric Sciences, 71, 4028–4054, doi:10.1175/JAS-D-14-0026.1, 2014.

Alexander, M., Geller, M., McLandress, C., Polavarapu, S., Preusse, P., Sassi, F., Sato, K., Eckermann, S., Ern, M., Hertzog, A., et al.: Recent developments in gravity-wave effects in climate models and the global distribution of gravity-wave momentum flux from observations and models, Quarterly Journal of the Royal Meteorological Society, 136, 1103–1124, 2010.

Alexander, M. J. and Sato, K.: Gravity Wave Dynamics and Climate: An Update from the SPARC Gravity Wave Activity, Report on the 35th Session of the Joint Scientific Committee of the World Climate Research Programme 29 June–4 July 2014, Heidelberg, Germany, 9, 2015.

Alexander, P., de la Torre, A., Llamedo, P., Hierro, R., Schmidt, T., Haser, A., and Wickert, J.: A method to improve the determination of wave perturbations close to the tropopause by using a digital filter, Atmospheric Measurement Techniques, 4, 1777–1784, doi:10.5194/amt-4-1777-2011, URL `http://www.atmos-meas-tech.net/4/1777/2011/`, 2011.

Alexander, S., Klekociuk, A., and Tsuda, T.: Gravity wave and orographic wave activity observed around the Antarctic and Arctic stratospheric vortices by the COSMIC GPS-RO satellite constellation, Journal of Geophysical Research: Atmospheres (1984–2012), 114, 2009.

Alexander, S. P. and Shepherd, M. G.: Planetary wave activity in the polar lower stratosphere, Atmos. Chem. Phys., 10, 707–718, doi:10.5194/acp-10-707-2010, 2010.

Alexander, S. P., Tsuda, T., and Kawatani, Y.: COSMIC GPS Observations of Northern Hemisphere winter stratospheric gravity waves and comparisons with an atmospheric general circulation model, Geophysical Research Letters, 35, L10 808, doi:10.1029/2008GL033174, URL `http://doi.wiley.com/10.1029/2008GL033174`, 2008.

Allen, S. J. and Vincent, R. A.: Gravity wave activity in the lower atmosphere: Seasonal and latitudinal variations, Journal of Geophysical Research: Atmospheres (1984–2012), 100, 1327–1350, 1995.

Andrews, D. and McIntyre, M.: Generalized Eliassen-Palm and Charney-Drazin theorems for waves on axismmetric mean flows in compressible atmospheres, Journal of the Atmospheric Sciences, 35, 175–185, 1978.

Andrews, D. G., Holton, J. R., and Leovy, C. B.: Middle atmosphere dynamics, vol. 40, Academic press, 1987.

Anthes, R. A., Ector, D., Hunt, D. C., Kuo, Y.-H., Rocken, C., Schreiner, W. S., Sokolovskiy, S. V., Syndergaard, S., Wee, T.-K., Zeng, Z., Bernhardt,

P. A., Dymond, K. F., Chen, Y., Liu, H., Manning, K., Randel, W. J., Trenberth, K. E., Cucurull, L., Healy, S. B., Ho, S.-P., McCormick, C., Meehan, T. K., Thompson, D. C., and Yen, N. L.: The COSMIC/FORMOSAT-3 Mission: Early Results, Bulletin of the American Meteorological Society, 89, 313–333, doi:10.1175/BAMS-89-3-313, URL http://dx.doi.org/10.1175/BAMS-89-3-313, 2008.

Ayarzaguena, B., Langematz, U., and Serrano, E.: Tropo- spheric forcing of the stratosphere: A comparative study of the two different major stratospheric warmings in 2009 and 2010, J. Geophys. Res, 116, doi:10.1029/2010JD015023, 2011.

Baumgaertner, A. and McDonald, A.: A gravity wave climatology for Antarctica compiled from Challenging Minisatellite Payload/Global Positioning System (CHAMP/GPS) radio occultations, Journal of Geophysical Research: Atmospheres (1984–2012), 112, 2007.

Berthet, G., Esler, J. G., and Haynes, P. H.: A Lagrangian perspective of the tropopause and the ventilation of the lowermost stratosphere, Journal of Geophysical Research, 112, D18 102, doi:10.1029/2006JD008295, URL http://doi.wiley.com/10.1029/2006JD008295, 2007.

Bishop, C. H. and Thorpe, A. J.: Potential vorticity and the electrostatics analogy: Quasi-geostrophic theory, Quarterly Journal of the Royal Meteorological Society, 120, 713–731, 1994.

Boyd, J. P.: The Noninteraction of Waves with the Zonally Averaged Flow on a Spherical Earth and the Interrelaionships of Eddy Fluxes of Energy, Heat, and Momentum, doi:10.1175/1520-0469(1976)033⟨2285:TNOWWT⟩2.0.CO;2, 1976.

Brewer, A. W.: Evidence for a world circulation provided by the measurements of helium and water vapor distribution in the stratosphere, Quart. J. Roy. Meteor. Soc., 75, 351–363, 1949.

Bühler, O.: Waves and mean flows, Cambridge University Press, 2014.

Calvo, N., Polvani, L. M., and Solomon, S.: On the surface impact of Arctic stratospheric ozone extremes, Environmental Research Letters, 10, 094 003, 2015.

CCMVal, S.: SPARC Report on the Evaluation of Chemistry-Climate Models, edited by: Eyring, V., Shepherd, TG, and Waugh, DW, Tech. rep., SPARC Report, 2010.

Chandran, A., Collins, R. L., Garcia, R. R., Marsh, D. R., Harvey, V. L., Yue, J., and de la Torre, L.: A climatology of elevated stratopause events in the whole atmosphere community climate model, J. Geophys. Res. Atmos., 118, 1234–1246, doi:10.1002/jgrd.50123, 2013.

Chane-Ming, F., Molinaro, F., Leveau, J., Keckhut, P., and Hauchecorne, A.: Analysis of gravity waves in the tropical middle atmosphere over La Reunion Island (21 S, 55 E) with lidar using wavelet techniques, Annales Geophysicae, 18, 485–498, doi:10.1007/s00585-000-0485-0, URL http://dx.doi.org/10.1007/s00585-000-0485-0, 2000.

Chiu, Y. T. and Ching, B. K.: The response of atmospheric and lower ionospheric layer structures to gravity waves, Geophysical Research Letters, 5, 539–542, doi:10.1029/GL005i006p00539, URL http://dx.doi.org/10.1029/GL005i006p00539, 1978.

Cohen, N. Y., Gerber, E. P., and Bühler, O.: Compensation between resolved and unresolved wave driving in the stratosphere: Implications for downward control, J. Atmos. Sci, 70, 3780–3798, doi:10.1175/JAS-D-12-0346.1, 2013.

Cohen, N. Y., Gerber, E. P., and Bühler, O.: What drives the Brewer- Dobson circulation?, Journal of the Atmospheric sciences, 71, 3837–3855, doi:10.1175/JAS-D-14-0021.1, 2014.

Cook, K., Wenkel, M. J., Fong, C.-J., Yen, N., and Chang, G.: From paper to production: Status update for the COSMIC-2/FORMOSAT-7 program, in: Aerospace Conference, 2016 IEEE, pp. 1–10, IEEE, 2016.

Craik, A. D.: Wave interactions and fluid flows, Cambridge University Press, 1988.

Cullens, C. Y., England, S. L., and Immel, T. J.: Global responses of gravity waves to planetary waves during stratospheric sudden warming observed by SABER, J. Geophys. Res. Atmos., 120, 12,018–12,026, doi:10.1002/2015JD023966, 2015.

Cushman-Roisin, B.: Introduction to geophysical fluid dynamics, 1994.

de la Cámara, A., Lott, F., Jewtoukoff, V., Plougonven, R., and Hertzog, A.: On the gravity wave forcing during the southern stratospheric final warming in LMDz, Journal of the Atmospheric Sciences, 2016.

De la Torre, A., Schmidt, T., and Wickert, J.: A global analysis of wave potential energy in the lower stratosphere derived from 5 years of GPS radio occultation data with CHAMP, Geophysical research letters, 33, 2006.

Demirhan Bari, D., Gabriel, A., Körnich, H., and Peters, D. W. H.: The effect of zonal asymmetries in the Brewer-Dobson circulation on ozone and water vapor distributions in the northern middle atmosphere, Journal of Geophysical Research: Atmospheres, 118, 3447–3466, doi:10.1029/2012JD017709, URL http://doi.wiley.com/10.1029/2012JD017709, 2013.

Dobson, G. M. B.: Origin and distribution of the polyatomic molecules in the atmosphere, in: Proceedings of the Royal Society of . Series A, Mathematical and Physical Sciences, 236(1205, pp. 187–193, 1956.

Dunkerton, T.: On the mean meridional mass motions of the stratosphere and mesosphere, J, Atmos. Sci, 35, 2325–2333, 1978.

Dunkerton, T. J. and Butchart, N.: Propagation and selective transmission of internal gravity waves in a sudden warming, J. Atmos. Sci, 41, 1520–0469, doi:10.1175/1520-0469(1984)041,1443:PASTOI.2.0.CO;2, 1984.

Durran, D. R.: Lee Waves and Mountain Waves, Encyclopedia of Atmospheric Sciences, p. 13, doi:10.1175/1520-0469(1982)039⟨2490:TEOMOT⟩2.0.CO;2, URL http://www.atmos.washington.edu/2010Q1/536/2003AP_lee_waves.pdf, 2003.

Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., Harada, Y., Yasui, S., Miyaoka, K., Takahashi, K., et al.: The Japanese 55-year Reanalysis (JRA-55): an interim report, Sola, 7, 149–152, 2011.

ECMWF: ERA-Interim, URL http://www.ecmwf.int/en/research/climate-reanalysis/era-interim, 2016.

Ern, M. and Preusse, P.: Gravity wave momentum flux spectra observed from satellite in the summertime subtropics: Implications for global modeling, Geophysical Research Letters, 39, 1–5, doi:10.1029/2012GL052659, URL http://www.agu.org/pubs/crossref/2012/2012GL052659.shtml, 2012.

Ern, M., Preusse, P., Alexander, M. J., and Warner, C. D.: Absolute values of gravity wave momentum flux derived from satellite data, Journal of Geophysical Research D: Atmospheres, 109, 1–17, doi:10.1029/2004JD004752, 2004.

Ern, M., Preusse, P., Gille, J., Hepplewhite, C., Mlynczak, M., Russell, J., and Riese, M.: Implications for atmospheric dynamics derived from global observations of gravity wave momentum flux in stratosphere and mesosphere, Journal of Geophysical Research: Atmospheres (1984–2012), 116, 2011.

Ern, M., Arras, C., Faber, A., and Fröhlich, K.: Observations and Ray Tracing of Gravity Waves: Implications for Global Modeling, Climate and Weather of . . . , pp. 383–408, doi:10.1007/978-94-007-4348-9, URL http://link.springer.com/chapter/10.1007/978-94-007-4348-9_21, 2013.

Ern, M., Ploeger, F., Preusse, P., Gille, J. C., Gray, L. J., Kalisch, S., Mlynczak, M. G., Russell, J. M., and Riese, M.: Interaction of gravity waves with the QBO: A satellite perspective, Journal of Geophysical Research: Atmospheres, 119, 2329–2355, doi:10.1002/2013JD020731, 2014.

Faber, A., Llamedo, P., Schmidt, T., de la Torre, A., and Wickert, J.: On the determination of gravity wave momentum flux from GPS radio occultation data, Atmospheric Measurement Techniques, 6, 3169–3180, doi:10.5194/amt-6-3169-2013, URL http://www.atmos-meas-tech.net/6/3169/2013/, 2013.

Foelsche, U., Borsche, M., and Steiner, A.: Observing upper troposphere–lower stratosphere climate with radio occultation data from the CHAMP satellite, Climate Dynamics, pp. 49–65, doi:10.1007/s00382-007-0337-7, URL http://link.springer.com/article/10.1007/s00382-007-0337-7, 2008.

France, J. A. and Harvey, V. L.: A climatology of the stratopause in WACCM and the zonally asymmetric elevated stratopause, J. Geophys. Res. Atmos., 118, 2241–2254, doi:10.1002/jgrd.50218, 2013.

Fritts, D. C.: Gravity wave saturation in the middle atmosphere: A review of theory and observations, Reviews of Geophysics, 22, 275, doi:10.1029/RG022i003p00275, URL http://doi.wiley.com/10.1029/RG022i003p00275, 1984.

Fritts, D. C. and Alexander, M. J.: Gravity wave dynamics and effects in the middle atmosphere, Reviews of Geophysics, 41, n/a–n/a, doi:10.1029/2001RG000106, URL http://dx.doi.org/10.1029/2001RG000106, 2003.

Fröhlich, K., Pogoreltsev, A., and Jacobi, C.: Numerical simulation of tides, Rossby and Kelvin waves with the COMMA-LIM model, Adv. Space Res, 32, 863–868, doi:10.1016/S0273-1177(03)00416-2, 2003.

Fukao, S., Yamanaka, M. D., Ao, N., Hocking, W. K., Sato, T., Yamamoto, M., Nakamura, T., Tsuda, T., and Kato, S.: Seasonal variability of vertical eddy diffusivity in the middle atmosphere: 1. Three-year observations by the middle and upper atmosphere radar, Journal of Geophysical Research: Atmospheres, 99, 18 973–18 987, doi:10.1029/94JD00911, URL http://dx.doi.org/10.1029/94JD00911, 1994.

Garcia, R. R. and Randel, W. J.: Acceleration of the Brewer-Dobson circulation due to increases in greenhouse gases, Journal of the Atmospheric Sciences, 65, 2731–2739, 2008.

Gardner, C. S. and Shelton, J. D.: Density response of neutral atmospheric layers to gravity wave perturbations, Journal of Geophysical Research: Space Physics, 90, 1745–1754, doi:10.1029/JA090iA02p01745, URL http://dx.doi.org/10.1029/JA090iA02p01745, 1985.

Gardner, C. S., Miller, M. S., and Liu, C. H.: Rayleigh Lidar Observations of Gravity Wave Activity in the Upper Stratosphere at Urbana, Illinois, Journal of the Atmospheric Sciences, 46, 1838–1854, doi:10.1175/1520-0469(1989)046⟨1838:RLOOGW⟩2.0.CO;2, URL http://dx.doi.org/10.1175/1520-0469(1989)046<1838:RLOOGW>2.0.CO;2, 1989.

Garny, H., Birner, T., Bönisch, H., and Bunzel, F.: The effects of mixing on age of air, J. Geophys. Res. Atmos., 119, 7015–7034, doi:10.1002/2013JD021417, 2014.

Gaßmann, A.: Deviations from a general nonlinear wind balance: Local and zonal-mean perspectives, Meteorologische Zeitschrift, pp. 467–481, 2014.

Geller, M. A., Alexander, M. J., Love, P. T., Bacmeister, J., Ern, M., Hertzog, A., Manzini, E., Preusse, P., Sato, K., Scaife, A. A., and Zhou, T.: A comparison between gravity wave momentum fluxes in observations and climate models, J. Climate, 26, 6383–6405, doi:10.1175/JCLI-D-12-00545.1, 2013.

Gerber, E. P., Butler, A., Calvo, N., Charlton-Perez, A., Giorgetta, M., Manzini, E., Perlwitz, J., Polvani, L. M., Sassi, F., Scaife, A. A., Shaw, T. A., w. Son, S., and Watanabe, S.: Assessing and understanding the impact of stratospheric dynamics and variability on the Earth system, B. Amer. Meteorol. Soc., 93, 845–859, 2012.

Gubenko, V. N., Pavelyev, A. G., and Andreev, V. E.: Determination of the intrinsic frequency and other wave parameters from a single vertical temperature or density profile measurement, Journal of Geophysical Research: Atmospheres, 113, n/a–n/a, doi:10.1029/2007JD008920, URL http://dx.doi.org/10.1029/2007JD008920, d08109, 2008.

Gubenko, V. N., Pavelyev, a. G., Salimzyanov, R. R., and Pavelyev, a. a.: Reconstruction of internal gravity wave parameters from radio occultation retrievals of vertical temperature profiles in the Earth's atmosphere, Atmospheric Measurement Techniques, 4, 2153–2162, doi:10.5194/amt-4-2153-2011, URL http://www.atmos-meas-tech.net/4/2153/2011/, 2011.

Gubenko, V. N., Pavelyev, A. G., Salimzyanov, R. R., and Andreev, V. E.: A method for determination of internal gravity wave parameters from a vertical temperature or density profile measurement in the Earth's atmosphere, Cosmic Research, 50, 21–31, doi:10.1134/S0010952512010029, URL http://dx.doi.org/10.1134/S0010952512010029, 2012.

Hardiman, S. C. and Haynes, P. H.: Dynamical sensitivity of the stratospheric circulation and downward influence of upper level perturbations, Journal of Geophysical Research, 113, D23 103, doi:10.1029/2008JD010168, URL http://doi.wiley.com/10.1029/2008JD010168, 2008.

Hardiman, S. C., Andrews, D. G., White, A. A., Butchart, N., and Edmond, I.: Using Different Formulations of the1065 Transformed Eulerian Mean Equations and EliassenPalm Diagnostics in General Circulation Models, J. Atmos. Sci, 67, 1983, doi:doi:10.1175/2010JAS3355.1, URL http://dx.doi.org/10.1175/2010JAS3355.1, 2010.

Hartley, D., Villarin, J., Black, R., and Davis, C.: A new perspective on the dynamical link between the stratosphere and troposphere, Nature, 8311, 1996–1999, URL http://www.nature.com/nature/journal/v391/n6666/abs/391471a0.html, 1998.

Haynes, P.: Stratospheric Dynamics, Annual Review of Fluid Mechanics, 37, 263–293, doi:10.1146/annurev.fluid.37.061903.175710, URL http://www.annualreviews.org/doi/abs/10.1146/annurev.fluid.37.061903.175710, 2005.

Haynes, P., McIntyre, M., Shepherd, T., Marks, C., and Shine, K. P.: On the "downward control" of extratropical diabatic circulations by eddy-induced mean zonal forces, Journal of the Atmospheric Sciences, 48, 651–678, 1991.

He, C., Zhou, T., Lin, A., Wu, B., Gu, D., Li, C., and Zheng, B.: Enhanced or Weakened Western North Pacific Subtropical High under Global Warming?, Scientific reports, 5, 2015.

Hei, H., Tsuda, T., and Hirooka, T.: Characteristics of atmospheric gravity wave activity in the polar regions revealed by GPS radio occultation data with CHAMP, Journal of Geophysical Research: Atmospheres (1984–2012), 113, 2008.

Hertzog, A., Alexander, M. J., and Plougonven, R.: On the intermittency of gravity wave momentum flux in the stratosphere, Journal of the Atmospheric Sciences, 69, 3433–3448, 2012.

Hindley, N. P., Wright, C. J., Smith, N. D., and Mitchell, N. J.: The southern stratospheric gravity wave hot spot: individual waves and their momentum fluxes measured by COSMIC GPS-RO, Atmospheric Chemistry and Physics, 15, 7797–7818, doi:10.5194/acp-15-7797-2015, URL http://www.atmos-chem-phys.net/15/7797/2015/, 2015.

Hines, C. O.: INTERNAL ATMOSPHERIC GRAVITY WAVES AT IONOSPHERIC HEIGHTS, Canadian Journal of Physics, 38, 1441–1481, doi:10.1139/p60-150, URL http://dx.doi.org/10.1139/p60-150, 1960.

Hirota, I. and Niki, T.: A Statistical Study of Inertia-Gravity Waves in the Middle Atmosphere, J. Meteor. Soc. Japan., 63, 1055–1066, 1985.

Hoffmann, P., Singer, W., Keuer, D., Hocking, W. K., Kunze, M., and Murayama, Y.: Latitudinal and longitudinal variability of mesospheric winds and temperatures during stratospheric warming events, J. Atmos. Sol. Terr. Phys., 69, 2355–2366, 2007.

Holton, J.: An introduction to dynamic meteorology (International geophysics series, San Diego, New York, 1992.

Holton, J. R.: The Influence of Gravity Wave Breaking on the General Circulation of the Middle Atmosphere, Journal of the Atmospheric Sciences, 40, 2497–2507, doi:10.1175/1520-0469(1983)040⟨2497:TIOGWB⟩2.0.CO;2, URL http://dx.doi.org/10.1175/1520-0469(1983)040<2497:TIOGWB>2.0.CO;2, 1983.

Holton, J. R.: The generation of mesospheric planetary waves by zonally asymmetric gravity wave breaking, J. Atmos. Sci, 41, 3427–3430, doi:10.1175/1520-0469(1984)041,3427:TGOMPW.2.0.CO;2., 1984.

Horinouchi, T.: Influence of Upper Tropospheric Disturbances on the Synoptic Variability of Precipitation and Moisture Transport over Summertime East Asia and the Northwestern Pacific, Journal of the Meteorological Society of Japan. Ser. II, 92, 519–541, doi:10.2151/jmsj.2014-602, URL https://www.jstage.jst.go.jp/article/jmsj/92/6/92_2014-602/_article, 2014.

Horinouchi, T. and Tsuda, T.: Spatial structures and statistics of atmospheric gravity waves derived using a heuristic vertical cross-section extraction from COSMIC GPS radio occultation data, Journal of Geophysical Research: Atmospheres (1984–2012), 114, 2009.

Jacobi, C., Fröhlich, K., and Pogoreltsev, A.: Quasi two-day-wave modulation of gravity wave flux and consequences for the planetary wave propagation in a simple circulation model, J. Atmos. Sol.-Terr. Phys., 68, 283–292, doi:10.1016/j.jastp.2005.01.017, 2006.

Jacobi, C., Lilienthal, F., Geißler, C., and Krug, A.: Long-term variability of mid-latitude mesosphere-lower thermosphere winds over Collm (51 N, 13 E), Journal of Atmospheric and Solar-Terrestrial Physics, 136, 174–186, 2015.

Jia, J. Y., Preusse, P., Ern, M., Chun, H.-Y., Gille, J. C., Eckermann, S. D., and Riese, M.: Sea surface temperature as a proxy for convective gravity wave excitation: a study based on global gravity wave observations in the middle atmosphere, Annales Geophysicae, 32, 1373–1394, doi:10.5194/angeo-32-1373-2014, URL http://www.ann-geophys.net/32/1373/2014/, 2014.

Jiang, J. H., Eckermann, S. D., Wu, D. L., and Ma, J.: A search for mountain waves in MLS stratospheric limb radiances from the winter Northern Hemisphere: Data analysis and global mountain wave modeling, Journal of Geophysical Research: Atmospheres (1984–2012), 109, 2004.

John, S. R. and Kumar, K. K.: TIMED/SABER observations of global gravity wave climatology and their interannual variability from stratosphere to mesosphere lower thermosphere, Climate dynamics, 39, 1489–1505, 2012.

Kalisch, S., Preusse, P., Ern, M., Eckermann, S. D., and Riese, M.: Differences in gravity wave drag between realistic oblique and assumed vertical propagation, J. Geophys. Res. Atmos., 119, 10,081–10,099, doi:10.1002/2014JD021779, 2014.

Kidston, J., Scaife, A. A., Hardiman, S. C., Mitchell, D. M., Butchart, N., Baldwin, M. P., and Gray, L. J.: Stratospheric influence on tropospheric jet streams, storm tracks and surface weather, Nat, Geosci., 8, 433–440, doi:10.1038/ngeo2424, 2015.

Kinoshita, T. and Sato, K.: A Formulation of Three-Dimensional Residual Mean Flow Applicable Both to Inertia–Gravity Waves and to Rossby Waves, Journal of the Atmospheric Sciences, 70, 1577–1602, doi:10.1175/JAS-D-12-0137.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JAS-D-12-0137.1, 2013.

Kinoshita, T., Tomikawa, Y., and Sato, K.: On the Three-Dimensional Residual Mean Circulation and Wave Activity Flux of the Primitive Equations, Journal of the Meteorological Society of Japan, 88, 373–394, doi:10.2151/jmsj.2010-307, URL http://joi.jlc.jst.go.jp/JST.JSTAGE/jmsj/2010-307?from=CrossRef, 2010.

Kozubek, M., Krizan, P., and Lastovicka, J.: Northern Hemisphere stratospheric winds in higher midlatitudes: longitudinal distribution and long-term trends, Atmospheric Chemistry and Physics, 15, 2203–2213, doi:10.5194/acp-15-2203-2015, URL http://www.atmos-chem-phys.net/15/2203/2015/, 2015.

Kren, A. C., Marsh, D. R., Smith, A. K., and Pilewskie, P.: Wintertime Northern Hemisphere response in the Stratosphere to the Pacific Decadal Oscillation using the Whole Atmosphere Community Climate Model, Journal of Climate, 2015.

Kruse, C. G., Smith, R. B., and Eckermann, S. D.: The Mid-latitude Lower-Stratospheric Mountain Wave Valve Layer, Journal of the Atmospheric Sciences, 2016.

Kuchar, A., Sacha, P., Miksovsky, J., and Pisoft, P.: The 11-year solar cycle in current reanalyses: a (non)linear attribution study of the middle atmosphere, Atmospheric Chemistry and Physics, 15, 6879–6895, doi:10.5194/acp-15-6879-2015, URL http://www.atmos-chem-phys.net/15/6879/2015/, 2015.

Kunkel, D., Hoor, P., and Wirth, V.: Can inertia-gravity waves persistently alter the tropopause inversion layer?, Geophysical Research Letters, 41, 7822–7829, doi:10.1002/2014GL061970, URL http://dx.doi.org/10.1002/2014GL061970, 2014.

Kuroda, Y.: Effect of stratospheric sudden warming and vortex intensification on the tropo- spheric climate, J. Geophys. Res, 113, doi:10.1029/2007JD009550, 2008.

Kuttippurath, J. and Nikulin, G.: A comparative study of the major sudden stratospheric warmings in the Arctic winters 2003/2004–2009/2010, Atmos. Chem. Phys., 12, 8115–8129, doi:10.5194/acp-12-8115-2012, 2012.

Lange, M. and Jacobi, C.: Analysis of gravity waves from radio occultation measurements, in: First CHAMP mission results for gravity, magnetic and atmospheric studies, pp. 479–484, Springer, 2003.

Lawrence, B. N.: The Effect of Parameterized Gravity Wave Drag on Simulations of the Middle Atmosphere During Northern Winter 1991/1992-General Evolution, NATO ASI SERIES I GLOBAL ENVIRONMENTAL CHANGE, 50, 291–308, 1997.

Lehtonen, I. and Karpechko, A. Y.: Observed and modeled tropospheric cold anomalies associated with sudden stratospheric warmings, J. Geophys. Res. Atmos., 121, 1591–1610, doi:10.1002/2015JD023860, 2016.

Lieberman, R. S., Riggin, D. M., and Siskind, D. E.: Stationary waves in the wintertime mesosphere: Evidence for gravity wave filtering by stratospheric planetary waves, J, Geophys. Res, 118, 3139–3149, doi:10.1002/jgrd.50319, 2013.

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

Lindzen, R. S.: Turbulence and stress owing to gravity wave and tidal breakdown, Journal of Geophysical Research, 86, 9709–9714, 1981.

Liou, Y., Pavelyev, A., and Wickert, J.: Observation of the gravity waves from GPS/MET radio occultation data, Journal of Atmospheric and Solar-Terrestrial Physics, 67, 219 – 228, doi:http://dx.doi.org/10.1016/j.jastp.2004.08.001, URL http://www.sciencedirect.com/science/article/pii/S1364682604002214, 2005.

Liou, Y. A., Pavelyev, A. G., Huang, C. Y., Igarashi, K., Hocke, K., and Yan, S. K.: Analytic method for observation of the gravity waves using radio occultation data, Geophysical Research Letters, 30, n/a–n/a, doi:10.1029/2003GL017818, URL http://dx.doi.org/10.1029/2003GL017818, 2021, 2003.

Lübken, F.-J., Berger, U., and Baumgarten, G.: Temperature trends in the mid-latitude summer mesosphere, Journal of Geophysical Research: Atmospheres, 118, 2013.

Luna, D., Alexander, P., and de la Torre, A.: Evaluation of uncertainty in gravity wave potential energy calculations through {GPS} radio occultation measurements, Advances in Space Research, 52, 879 – 882, doi:http://dx.doi.org/10.1016/j.asr.2013.05.015, URL http://www.sciencedirect.com/science/article/pii/S0273117713002883, 2013.

Maddison, J. R. and Marshall, D. P.: The Eliassen–Palm flux tensor, Journal of Fluid Mechanics, 729, 69–102, doi:10.1017/jfm.2013.259, URL http://www.journals.cambridge.org/abstract{_}S0022112013002590, 2013.

Manney, G. L., Schwartz, M. J., Krüger, K., Santee, M. L., Pawson, S., Lee, J. N., Daffer, W. H., Fuller, R. A., and Livesey, N. J.: Aura Microwave Limb Sounder observations of dynamics and transport during the record-breaking 2009 Arctic stratospheric major warming, Geophys. Res. Lett., 36, L12 815, doi:10.1029/2009GL038586, 2009.

Manzini, E.: Northern winter climate change: Assessment of uncertainty in CMIP5 projections related to stratosphere-troposphere coupling, J. Geophys. Res. Atmos., 119, 7979–7998, doi:10.1002/2013JD021403, 2014.

Markwardt, C. B.: Non-linear Least-squares Fitting in IDL with MPFIT, in: Astronomical Data Analysis Software and Systems XVIII, vol. 411 of *Astronomical Society of the Pacific Conference Series*, 2009.

Markwardt, C. B. and Healy, S. B.: Measurement Noise and Stratospheric Gravity Wave Characteristics Obtained from GPS Occultation Data, Journal of the Meteorological Society of Japan, 83(3), 417–428, doi:10.2151/jmsj.83.417, 2005.

Marshall, A. G. and Scaife, A. A.: Impact of the QBO on surface winter climate, Journal of Geophysical Research: Atmospheres, 114, 2009.

Martius, O., Polvani, L., and Davies, H.: Blocking precursors to stratospheric sudden warming events, Geophys. Res. Lett., 36, L14 806, doi:10.1029/2009GL038776, 2009.

Matsuno, T.: Numerical integration of the primitive equations by a simulated backward difference method, Journal of Meteorological Society of Japan, 44, 76–84, 1966.

Matsuno, T.: Lagrangian Motion of Air Parcels in the Stratosphere in the Presence of Planetary Waves, Pure and Applied Geophysics, 118, 189–216, doi: 10.1007/BF01586451, URL http://www.springerlink.com/index/10.1007/BF01586451, 1980.

McDonald, A. J., Tan, B., and Chu, X.: Role of gravity waves in the spatial and temporal variability of stratospheric temperature measured by COSMIC/FORMOSAT-3 and Rayleigh lidar observations, Journal of Geophysical Research, 115, D19 128, doi:10.1029/2009JD013658, URL http://doi.wiley.com/10.1029/2009JD013658, 2010.

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

McIntyre, M. E.: An introduction to the generalized Lagrangian-mean description of wave, mean-flow interaction, Pure and Applied Geophysics PAGEOPH, 118, 152–176, doi:10.1007/BF01586449, URL http://link.springer.com/10.1007/BF01586449, 1980.

McLandress, C., Polavarapu, S., and Beagley, S. R.: Is missing orographic gravity wave drag near 608S the cause of the stratospheric zonal wind biases in chemistry climate models?, J. Atmos. Sci, 69, 802–818, 2012.

Melbourne, W. G.: Radio Occultations Using Earth Satellites: A Wave Theory Treatment, Vol. 5, Wiley-Blackwell, 2005.

Miyakoda, K.: Some characteristic features of winter circulation in the troposphere and lower stratosphere, 14, University of Chicago, Dept. of the Geophysical Sciences, 1963.

Mohri, K.: On the fields of wind and temperature over Japan and adjacent waters during winter of 1950–1951, Tellus, 5, 340–358, 1953.

Nakamura, H., Nishii, K., Wang, L., Orsolini, Y. J., and Takaya, K.: Cold-air outbreaks over East Asia associated with blocking highs: Mechanisms and their interaction with the polar stratosphere, Dynamics and Predictability of Large-Scale High-Impact Weather and Climate Events, 2, 2014.

Nishii, K., Nakamura, H., and Miyasaka, T.: Modulations in the planetary wave field induced by upward propagating Rossby wave packets prior to stratospheric sudden warming events: A case study, Q. J. Roy. Meteor. Soc., 135, 39–52, 2009.

Noda, A.: Generalized Transformed Eulerian Mean (GTEM) Description for Boussinesq Fluids, Journal of the Meteorological Society of Japan. Ser. II, 92, 411–431, doi:10.2151/jmsj.2014-501, URL http://jlc.jst.go.jp/DN/JST.JSTAGE/jmsj/2014-501?lang=en{&}from=CrossRef{&}type=abstract, 2014.

Oberheide, J.: Geostrophic wind fields in the stratosphere and mesosphere from satellite data, Journal of Geophysical Research, 107, 8175, doi:10.1029/2001JD000655, URL http://doi.wiley.com/10.1029/2001JD000655, 2002.

Oberheide, J., l. Liu, H., Gusev, O., and Offermann, D.: Mesospheric surf zone and temperature inversion layers in early November 1994, J, J. Atmos. Sol.-Terr. Phys., 68, 1752–1763, 2006.

Ortland, D. A.: Rossby wave propagation into the tropical stratosphere observed by the High Resolution Doppler Imager, Geophysical research letters, 24, 1999–2002, 1997.

Pavelyev, A. G., Wickert, J., Liou, Y. A., Pavelyev, A. A., and Jacobi, J.: Analysis of atmospheric and ionospheric wave structures using the CHAMP and GPS/MET radio occultation database, in: Atmosphere and Climate Studies by Occultation Methods, Springer Verlag, Springer, Berlin Heidelberg, doi:10.1007/3-540-34121-8_19, 2006.

Pavelyev, A. G., Liou, Y. A., Wickert, J., Gubenko, V. N., Pavelyev, A. A., and Matyugov, S. S.: New Applications and Advances of the GPS Radio Occultation Technology as Recovered by Analysis of the FORMOSAT-3/COSMIC and CHAMP Data-Base, in: New Horizons in Occultation Research: Studies in Atmosphere and Climate, Springer Verlag, Springer, Berlin Heidelbergl, doi:10.1007/3-540-34121-8_19, 2009.

Pawson, S.: Effects of gravity wave drag in the Berlin troposphere-stratosphere-mesosphere GCM, in: Gravity Wave Processes, pp. 327–336, Springer Berlin Heidelberg, Berlin, 1997.

Percival, D. B. and Walden, A. T.: Wavelet methods for time series analysis, vol. 4, Cambridge University Press, 2006.

Pišoft, P., Mikšovskỳ, J., and Žák, M.: An analysis of the spatial distribution of approximate 8 years periodicity in NCEP/NCAR and ERA-40 temperature fields, The European Physical Journal-Special Topics, 174, 147–155, 2009.

Pisoft, P., Miksovsky, J., Kalvova, J., Raidl, A., and Zak, M.: Areal analysis of oscillations in 500-hPa temperature field: a pseudo-2D wavelet transform approach, International Journal of Climatology, 31, 1545–1553, 2011.

Pisoft, P., Holtanova, E., Huszar, P., Kalvova, J., Miksovsky, J., Raidl, A., Zemankova, K., and Zak, M.: Manifestation of reanalyzed QBO and SSC signals, Theoretical and Applied Climatology, pp. 1–10, 2013.

Plieninger, J., von Clarmann, T., Stiller, G. P., Grabowski, U., Glatthor, N., Kellmann, S., Linden, A., Haenel, F., Kiefer, M., Höpfner, M., Laeng, A., and Lossow, S.: Methane and nitrous oxide retrievals from MIPAS-ENVISAT, Atmos. Meas. Tech, 8, 4657–4670, doi:10.5194/amt-8-4657-2015, 2015.

Plougonven, R. and Zhang, F.: Internal gravity waves from atmospheric jets and fronts, Reviews of Geophysics, 52, 33–76, doi:10.1002/2012RG000419, URL http://dx.doi.org/10.1002/2012RG000419, 2014.

Plumb, R. A.: Three-Dimensional Propagation of Transient Quasi-Geostrophic Eddies and Its Relationship with the Eddy Forcing of the Time–Mean Flow, Journal of the Atmospheric Sciences, 43, 1657–1678, doi:10.1175/1520-0469(1986)043⟨1657:TDPOTQ⟩2.0.CO;2, URL http://dx.doi.org/10.1175/1520-0469(1986)043<1657:TDPOTQ>2.0.CO;2, 1986.

Pogoreltsev, A.I., A. V. K. F. and Jacobi, C.: Planetary waves in coupling the lower and upper atmosphere, J. Atmos. Sol.-Terr. Phys., 69, 2083–2101, doi:10.1016/j.jastp.2007.05.014, 2007.

Preusse, P. and Ern, M.: Indication of convectively generated gravity waves observed by CLAES, Advances in Space Research, 35, 1987–1991, 2005.

Preusse, P., Eidmann, G., Eckermann, S. D., Schaeler, B., Spang, R., and Offermann, D.: Indications of convectively generated gravity waves in crista temperatures, Advances in Space Research, 27, 1653–1658, doi:10.1016/S0273-1177(01)00231-9, 2001.

Preusse, P., Eckermann, S. D., and Ern, M.: Transparency of the atmosphere to short horizontal wavelength gravity waves, Journal of Geophysical Research: Atmospheres, 113, n/a–n/a, doi:10.1029/2007JD009682, URL http://dx.doi.org/10.1029/2007JD009682, d24104, 2008.

Preusse, P., Eckermann, S. D., Ern, M., Oberheide, J., Picard, R. H., Roble, R. G., Riese, M., Iii, J. M. R., and Mlynczak, M. G.: Global ray tracing simulations of the SABER gravity wave climatology, J. Geophys. Res, 114, D08 126, doi:10.1029/2008JD011214, 2009.

Ratnam, M., Tetzlaff, G., and Jacobi, C.: Global and Seasonal Variations of Stratospheric Gravity Wave Activity Deduced from the CHAMP/GPS Satellite, Journal of the Atmospheric ..., URL http://journals.ametsoc.org/doi/full/10.1175/1520-0469(2004)061%3C1610:GASVOS%3E2.0.CO%3B2, 2004.

Richter, J. H., Sassi, F., and Garcia, R. R.: Toward a physically based gravity wave source parameterization in a general circulation model, Journal of the Atmospheric Sciences, 67, 136–156, 2010.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G. K., et al.: MERRA: NASA's modern-era retrospective analysis for research and applications, Journal of Climate, 24, 3624–3648, 2011.

Riese, M., Oelhaf, H., Preusse, P., Blank, J., Ern, M., Friedl-Vallon, F., Fischer, H., Guggenmoser, T., Höpfner, M., Hoor, P., and Others: Gimballed Limb Observer for Radiance Imaging of the Atmosphere (GLORIA) scientific objectives, Atmospheric Measurement Techniques, 7, 1915–1928, 2014.

Sandu, I., Bechtold, P., Beljaars, A., Bozzo, A., Pithan, F., Shepherd, T. G., and Zadra, A.: Impacts of parameterized orographic drag on the Northern Hemisphere winter circulation, J. Adv. Model. Earth Syst., 7, doi:10.1002/2015MS000564, 2016.

Sato, K., Kinoshita, T., and Okamoto, K.: A New Method to Estimate Three-Dimensional Residual-Mean Circulation in the Middle Atmosphere and Its Application to Gravity Wave–Resolving General Circulation Model Data, Journal of the Atmospheric Sciences, 70, 3756–3779, doi:10.1175/JAS-D-12-0352.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JAS-D-12-0352.1, 2013.

Scherllin-Pirscher, B., Steiner, A. K., and Kirchengast, G.: Deriving dynamics from GPS radio occultation: Three-dimensional wind fields for monitoring the climate, Geophysical Research Letters, p. 2014GL061524, doi:10.1002/2014GL061524, URL http://dx.doi.org/10.1002/2014GL061524, 2014.

Schmidt, T., de la Torre, A., and Wickert, J.: Global gravity wave activity in the tropopause region from CHAMP radio occultation data, Geophysical Research Letters, 35, 2008.

Schmidt, T., Wickert, J., De la Torre, A., Alexander, P., Faber, A., Llamedo, P., and Heise, S.: The effect of different background separation methods on gravity wave parameters in the upper troposphere and lower stratosphere region derived from GPS radio occultation data., In 39th COSPAR Scientific Assembly, 39, 1721, 2012.

Schreiner, W., Rocken, C., Sokolovskiy, S., Syndergaard, S., and Hunt, D.: Estimates of the precision of GPS radio occultations from the COSMIC/FORMOSAT-3 mission, Geophysical Research Letters, 34, n/a–n/a, doi:10.1029/2006GL027557, URL http://dx.doi.org/10.1029/2006GL027557, l04808, 2007.

Schroeder, S., Preusse, P., Ern, M., and Riese, M.: Gravity waves resolved in ECMWF and measured by SABER, Geophys. Res. Lett., 36, L10 805, doi:10.1029/2008GL037054, 2009.

Scott, R. K. and Dritschel, D. G.: Downward Wave Propagation on the Polar Vortex, J. Atmos. Sci, 62, 3382–3395, 2005.

Senft, D. C. and Gardner, C. S.: Seasonal variability of gravity wave activity and spectra in the mesopause region at Urbana, Journal of Geophysical Research, 96, 17 229, doi:10.1029/91JD01662, URL http://doi.wiley.com/10.1029/91JD01662, 1991.

Seviour, W. J. M., Gray, L. J., and Mitchell, D. M.: Stratospheric polar vortex splits and displacements in the high-top CMIP5 climate models, J. Geophys. Res. Atmos., 121, 1400–1413, doi:10.1002/2015JD024178, 2016.

Shaw, T. A. and Boos, W. R.: The tropospheric response to tropical and subtropical zonally-asymmetric torques: Analytical and idealized numerical model results, J. Atmos. Sci, 69, 214–234, 2012.

Sica, R. J. and Russell, A. T.: Measurements of the Effects of Gravity Waves in the Middle Atmosphere Using Parametric Models of Density Fluctuations. Part I: Vertical Wavenumber and Temporal Spectra, Journal of the Atmospheric Sciences, 56, 1308–1329, doi:10.1175/1520-0469(1999)056⟨1308:

MOTEOG⟩2.0.CO;2, URL http://dx.doi.org/10.1175/1520-0469(1999)056<1308:MOTEOG>2.0.CO;2, 1999.

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

Smith, A. K.: The Origin of Stationary Planetary Waves in the Upper Mesosphere, Journal of the Atmospheric Sciences, 60, 3033–3041, doi:10.1175/1520-0469(2003)060⟨3033:TOOSPW⟩2.0.CO;2, 2003.

Smith, A. K.: Global dynamics of the MLT, Surv. Geophys., 33, 1177–1230, doi:10.1007/s10712-012-9196-9, 2012.

Smith, S. A., Fritts, D. C., and Vanzandt, T. E.: Evidence for a saturated spectrum of atmospheric gravity waves, Journal of the Atmospheric Sciences, 44, 1404–1410, 1987.

Steiner, A. K. and Kirchengast, G.: Gravity Wave Spectra from GPS/MET Occultation Observations, Journal of Atmospheric and Oceanic Technology, 17, 495–503, doi:10.1175/1520-0426(2000)017⟨0495:GWSFGM⟩2.0.CO;2, URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0426%282000%29017%3C0495%3AGWSFGM%3E2.0.CO%3B2, 2000.

Sutherland, B. R.: Internal Gravity Waves, Cambridge University Press, 2010.

Temis, T. E. M. I. S.: Monthly mean ozone from the multi-sensor reanalysis (MSR), p. 2016, URL http://www.temis.nl/protocols/o3field/o3mean_msr.php.

Torrence, C. and Compo, G. P.: A practical guide to wavelet analysis, Bulletin of the American Meteorological society, 79, 61–78, 1998.

Tsuda, T. and Hocke, K.: Vertical Wave Number Spectrum of Temperature Fluctuations in the Stratosphere using GPS Occultation Data, Journal of the Meteorological Society of Japan. Ser. II, 80, 925–938, doi:10.2151/jmsj.80.925, 2002.

Tsuda, T., Murayama, Y., Nakamura, T., Vincent, R., a.H. Manson, Meek, C., and Wilson, R.: Variations of the gravity wave characteristics with height, season and latitude revealed by comparative observations, Journal of Atmospheric and Terrestrial Physics, 56, 555–568, doi:10.1016/0021-9169(94)90097-3, URL http://linkinghub.elsevier.com/retrieve/pii/0021916994900973, 1994.

Tsuda, T., Nishida, M., Rocken, C., and Ware, R. H.: A global morphology of gravity wave activity in the stratosphere revealed by the GPS occultation data (GPS/MET), Journal of Geophysical Research: Atmospheres (1984–2012), 105, 7257–7273, 2000.

Tsuda, T., Lin, X., Hayashi, H., and Noersomadi: Analysis of vertical wave number spectrum of atmospheric gravity waves in the stratosphere using COSMIC GPS radio occultation data, Atmospheric Measurement Techniques, 4, 1627–1636, doi:10.5194/amt-4-1627-2011, URL http://www.atmos-meas-tech.net/4/1627/2011/, 2011.

van der A., R. J., Allaart, M. A. F., and Eskes, H. J.: Extended and refined multi sensor reanalysis of total ozone for the period 1970Ű20132012, Atmos. Meas. Tech, 8, 3021–3035, doi:10.5194/amt-8-3021-2015, 2015.

VanZandt, T. E.: A model for gravity wave spectra observed by Doppler sounding systems, Radio Science, 20, 1323–1330, 1985.

Verkhoglyadova, O. P., Leroy, S. S., and Ao, C. O.: Estimation of Winds from GPS Radio Occultations, Journal of Atmospheric and Oceanic Technology, 31, 2451–2461, doi:10.1175/JTECH-D-14-00061.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JTECH-D-14-00061.1, 2014.

Vincent, R. A., Allen, S. J., and Eckermann, S. D.: Gravity-wave parameters in the lower stratosphere, in: GravityWave Processes, Springer, Berlin Heidelberg, 1997.

von Clarmann, T., Höpfner, M., Kellmann, S., Linden, A., Chauhan, S., Funke, B., Grabowski, U., Glatthor, N., Kiefer, M., Schieferdecker, T., Stiller, G. P., and Versick, S. R. o. t.: H2O, O3, HNO3, CH4, N2O, ClONO2 and ClO from MIPAS reduced resolution nominal mode limb emission measurements, Atmos. Meas. Tech, 2, 159–175, doi:10.5194/amt-2-159-2009, 2009.

Šácha, P., Foelsche, U., and Pišoft, P.: Analysis of internal gravity waves with GPS RO density profiles, Atmospheric Measurement Techniques, 7, 4123–4132, doi:10.5194/amt-7-4123-2014, URL http://www.atmos-meas-tech.net/7/4123/2014/, 2014.

Šácha, P., Kuchař, A., Jacobi, C., and Pišoft, P.: Enhanced internal gravity wave activity and breaking over the northeastern Pacific–eastern Asian region, Atmospheric Chemistry and Physics, 15, 13 097–13 112, doi:10.5194/acp-15-13097-2015, URL http://www.atmos-chem-phys.net/15/13097/2015/, 2015.

Šácha, P., Lilienthal, F., Jacobi, C., and Pišoft, P.: Influence of the spatial distribution of gravity wave activity on the middle atmospheric circulation and transport, Atmospheric Chemistry and Physics Discussions, 2016, 1–35, doi:10.5194/acp-2016-548, URL http://www.atmos-chem-phys-discuss.net/acp-2016-548/, 2016.

Škerlak, B., Sprenger, M., Pfahl, S., Tyrlis, E., and Wernli, H.: Tropopause folds in ERA-Interim: Global climatology and relation to extreme weather events, Journal of Geophysical Research: Atmospheres, pp. n/a–n/a, doi:10.1002/2014JD022787, URL http://dx.doi.org/10.1002/2014JD022787, 2014JD022787, 2015.

Walterscheid, R. L.: Inertio-gravity wave induced accelerations of mean flow having an imposed periodic component: Implications for tidal observations in the meteor region, Journal of Geophysical Research: Oceans, 86, 9698–9706, 1981.

Wang, L. and Alexander, M.: Global estimates of gravity wave parameters from GPS radio occultation temperature data, Journal of Geophysical Research: Atmospheres (1984–2012), 115, 2010.

Wang, L. and Alexander, M. J.: Gravity wave activity during stratospheric sudden warmings in the 2007–2008 Northern Hemisphere winter, Journal of Geophysical Research: Atmospheres (1984–2012), 114, 2009.

Watt-Meyer, O. and Kushner, P. J.: The Role of Standing Waves in Driving Persistent Anomalies of Upward Wave Activity Flux, Journal of Climate, 28, 9941–9954, 2015.

White, A. A., Hoskins, B. J., Roulstone, I., and Staniforth, A.: Consistent approximate models of the global atmosphere: shallow, deep, hydrostatic, quasi-hydrostatic and non-hydrostatic, Quarterly Journal of the Royal Meteorological Society, 131, 2081–2107, doi:10.1256/qj.04.49, URL http://doi.wiley.com/10.1256/qj.04.49, 2005.

Wickert, J., Schmidt, T., Michalak, G., Heise, S., Arras, C., Beyerle, G., Falck, C., König, R., Pingel, D., and Rothacher, M.: GPS Radio Occultation with CHAMP, GRACE-A, SAC-C, TerraSAR-X, and FORMOSAT-3/COSMIC: Brief Review of Results from GFZ, pp. 3–15, Springer Berlin Heidelberg, Berlin, Heidelberg, doi:10.1007/978-3-642-00321-9_1, URL http://dx.doi.org/10.1007/978-3-642-00321-9_1, 2009.

Wilson, R., Chanin, M., and Hauchecorne, A.: Gravity waves in the middle atmosphere observed by Rayleigh lidar: 2. Climatology, Journal of Geophysical Research: Atmospheres (1984–2012), 96, 5169–5183, 1991.

Woo, S.-h., k. Sung, M., w. Son, S., and s. Kug, J.: Connection between weak stratospheric vortex events and the Pacific Decadal Oscillation, Climate Dynamics, pp. 1–12, doi:10.1007/s00382-015-2551-z, 2015.

Wright, C. J. and Gille, J. C.: Detecting overlapping gravity waves using the S-Transform, Geophysical Research Letters, 40, 1850–1855, doi:10.1002/grl.50378, URL http://dx.doi.org/10.1002/grl.50378, 2013.

Wright, C. J., Rivas, M. B., and Gille, J. C.: Intercomparisons of HIRDLS, COSMIC and SABER for the detection of stratospheric gravity waves, Atmospheric Measurement Techniques, 4, 1581–1591, doi:10.5194/amt-4-1581-2011, URL http://www.atmos-meas-tech.net/4/1581/2011/, 2011.

Wu, D. L., Preusse, P., Eckermann, S. D., Jiang, J. H., de la Torre Juarez, M., Coy, L., and Wang, D. Y.: Remote sounding of atmospheric gravity waves with satellite limb and nadir techniques, Advances in Space Research, 37, 2269 – 2277, doi:http://dx.doi.org/10.1016/j.asr.2005.07.031, URL http:

//www.sciencedirect.com/science/article/pii/S0273117705009257, atmospheric Remote Sensing: Earth's Surface, Troposphere, Stratosphere and Mesosphere - {II}, 2006.

Yamashita, C., l. Liu, H., and Chu, X.: Gravity wave variations during the 2009 stratospheric sudden warming as revealed by ECMWF-T799 and observations, Geophys. Res. Lett., 37, L22 806, doi:10.1029/2010GL045437, 2010.

Yamashita, Y., Akiyoshi, H., Shepherd, T. G., and Takahashi, M.: The Combined Influences of Westerly Phase of the Quasi-Biennial Oscillation and 11-year Solar Maximum Conditions on the Northern Hemisphere Extratropical Winter Circulation, Journal of the Meteorological Society of Japan. Ser. II, 93, 629–644, doi:10.2151/jmsj.2015-054, 2015.

Zhang, Y., Xiong, J., Liu, L., and Wan, W.: A global morphology of gravity wave activity in the stratosphere revealed by the 8-year SABER/TIMED data, Journal of Geophysical Research: Atmospheres (1984–2012), 117, 2012.

List of Figures

List of Tables

List of Abbreviations

Brewer-Dobson Circulation: BDC
Continuous Wavelet Transform: CWT
Eastern Asia/North-western Pacific: EA/NP
Eliassen-Palm: EP
Fourier Transform: FT
Formosat Satellite Mission 3 – Constellation Observing System for Meteorology, Ionosphere, and Climate: FORMOSAT-3-COSMIC
Generalized Transformed Eulerian Mean: GTEM
Geophysical Fluid Dynamics: GFD
Generalized Lagrangian mean: GLM
Geometrical Optics: GO
Global Positioning System: GPS
Gravity Wave Drag: GWD
High Resolution Dynamics Limb Sounder: HIRDLS
Internal Gravity Waves: IGWs
Japanese 55-year Reanalysis: JRA-55
Line of Sight: LOS
Lower Stratosphere: LS
Modern Era Reanalysis for Research and Applications: MERRA
Michelson Interferometer for Passive Atmospheric Sounding: MIPAS
Middle and Upper Atmosphere Model: MUAM
Northern Hemisphere: NH
Pacific Decadal Oscillation: PDO
Power Spectral Density: PSD
Potential Vorticity: PV
Planetary Waves: PWs
Quasi-Biennial Oscillation: QBO
Quasi-Geostrophic: QG
Radio Holography: RH
Radio Occultation: RO
Root Mean Square: RMS
Sounding of the Atmosphere using Broadband Emission Radiometry: SABER
Southern Hemisphere: SH
Stratosphere-Troposphere Exchange: STE
Three-Dimensional: 3D
Transformed Eulerian Mean: TEM
Thermosphere Ionosphere Mesosphere Energetics Dynamics: TIMED
Total Ozone Column: TOC
Wentzel, Kramers and Brillouin: WKB

Attachments

Att. 1: Supplementary figures to the Chapter 3

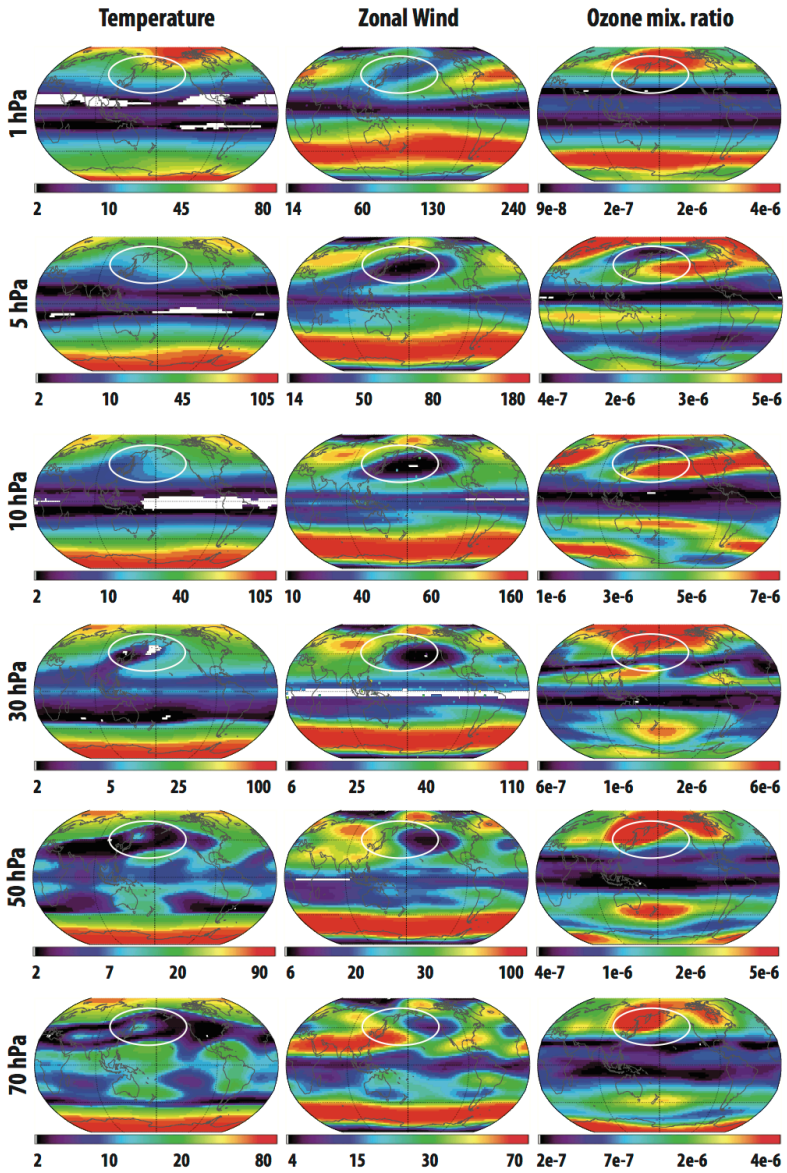

Figure S.1: Annual cycle amplitudes at various geopotential levels in the temperature, zonal winds and ozone mixing ratio series from MERRA. The non-linear color-scale used represents square root of the wavelet power in K, m/s and mg/kg.

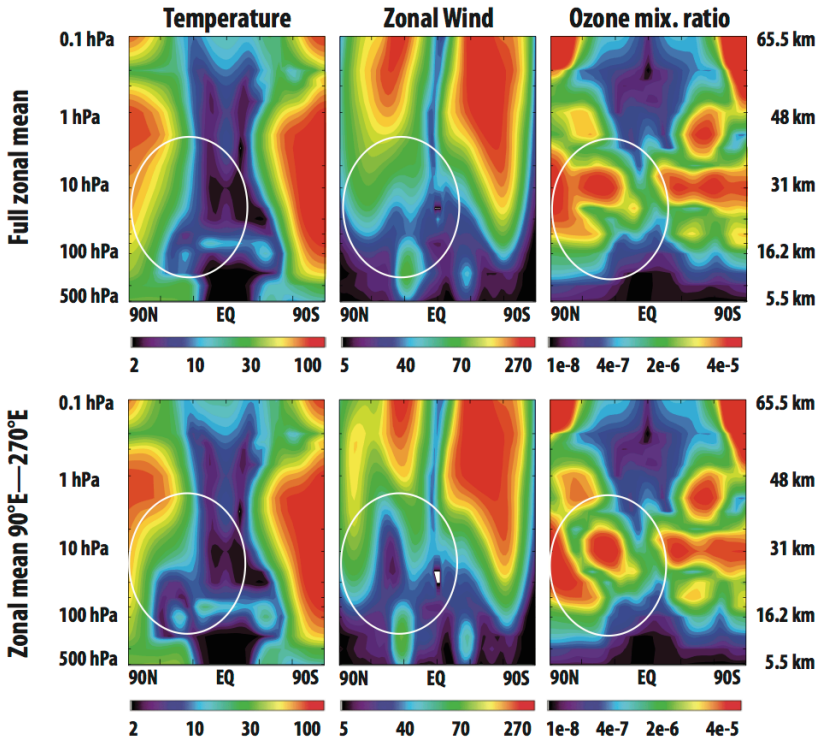

Figure S.2: Vertical profiles of the zonal means of the annual cycle amplitudes in the temperature, zonal winds and ozone mixing ratio series. The upper row shows the zonal means computed for all longitudes, the bottom row presents zonal means computed only for 90^oE–270^oE. The non-linear color scale used represents square root of the wavelet power in K, m/s and mg/kg.

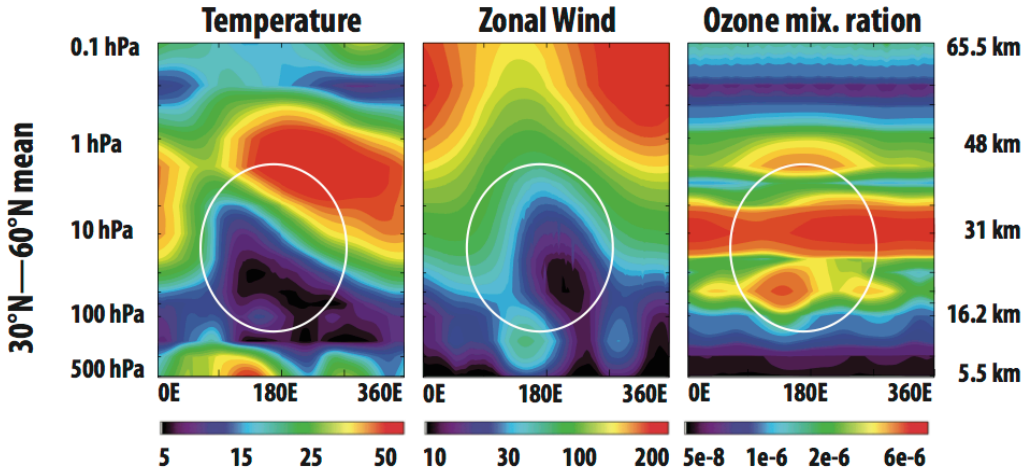

Figure S.3: Vertical profiles of the annual cycle amplitudes in the temperature, zonal winds and ozone mixing ratio series computed for the northern middle latitudes (30^oN–60^oN mean). The non-linear color scale used represents square root of the wavelet power in K, m/s and mg/kg.

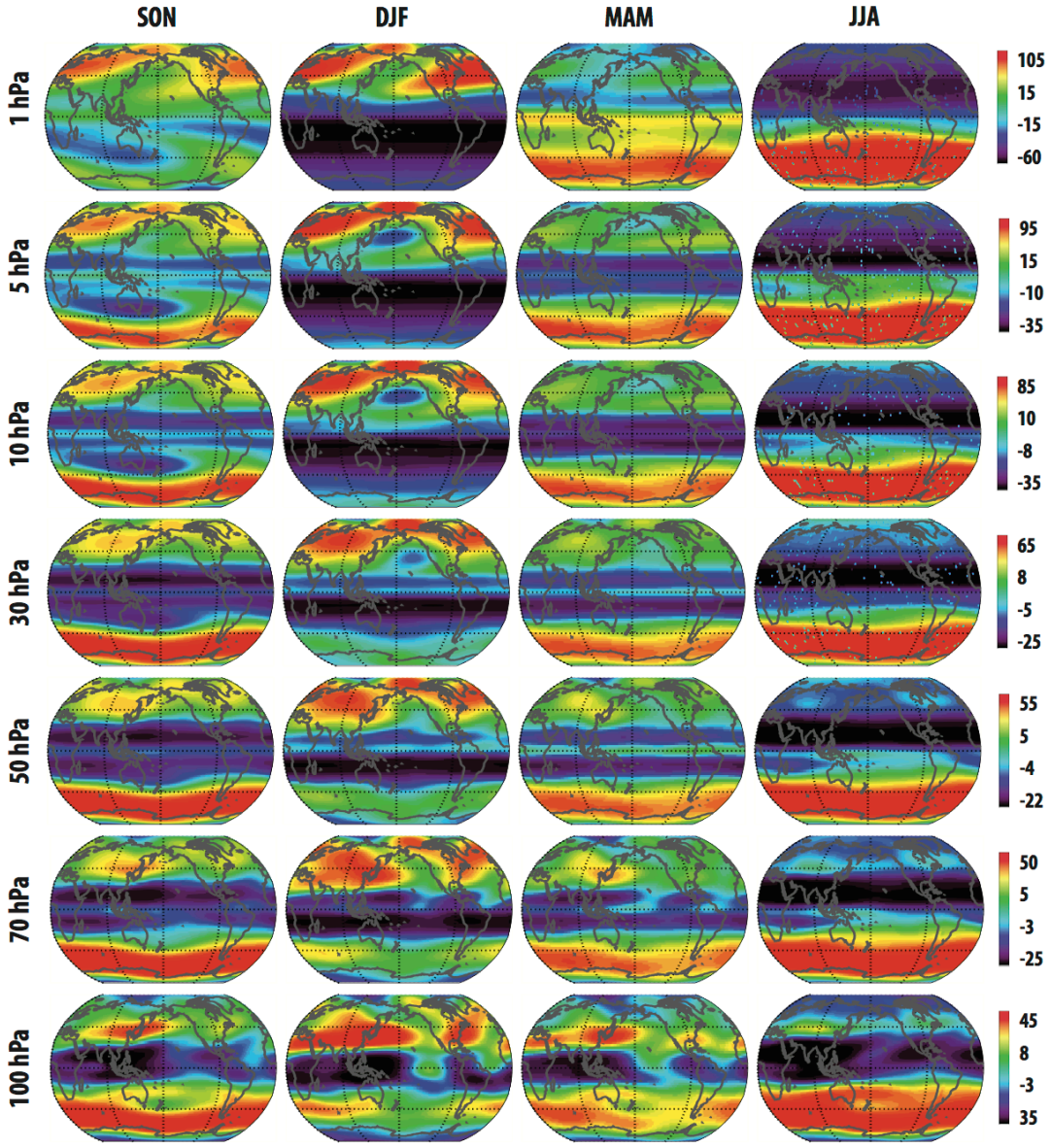

Figure S.4: Seasonal averages in zonal wind in m/s for 1979–2013 time period using MERRA series (non-linear color scale used).

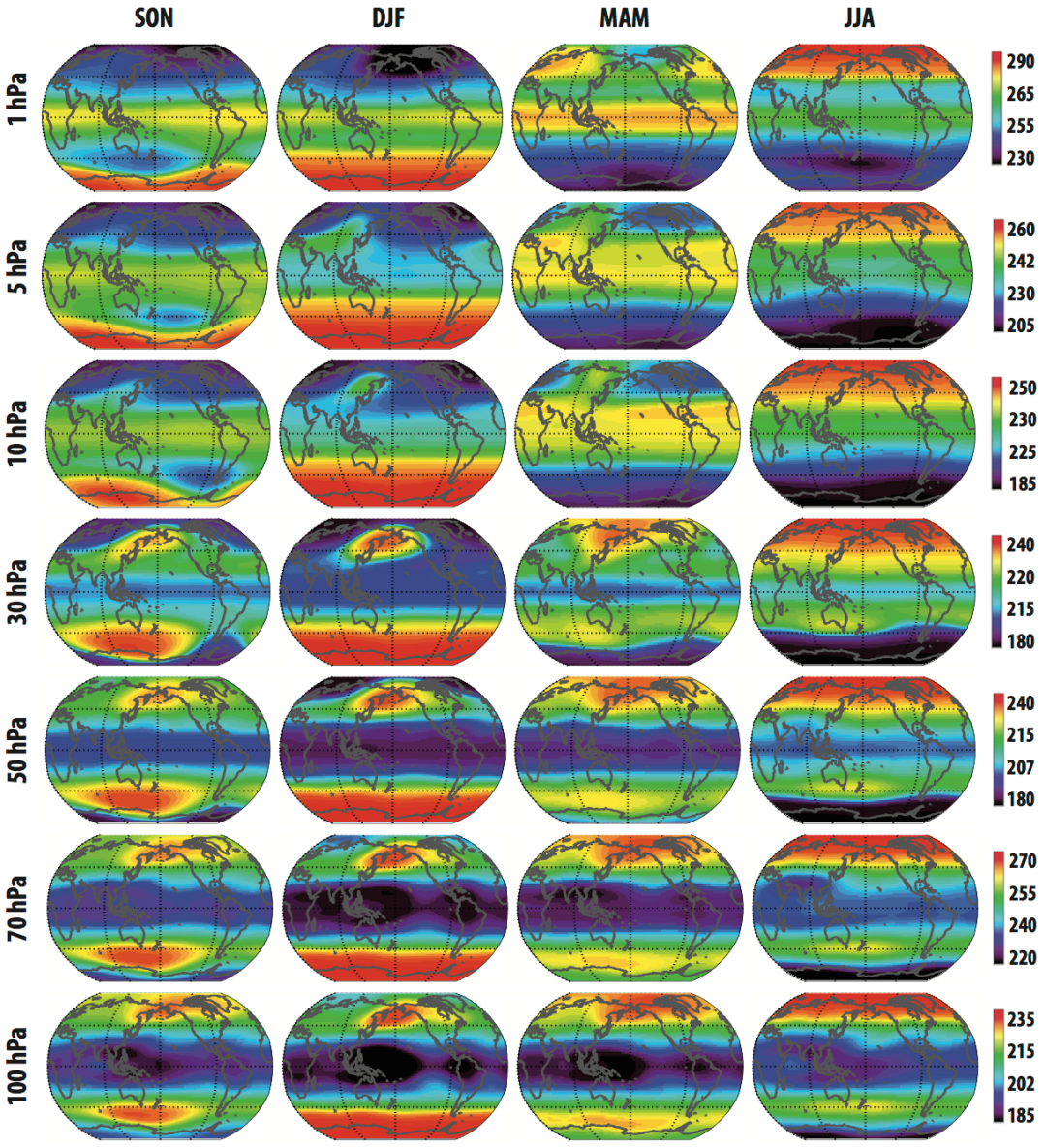

Figure S.5: Seasonal averages in temparature in K for 1979–2013 time period using MERRA series (non-linear color scale used).

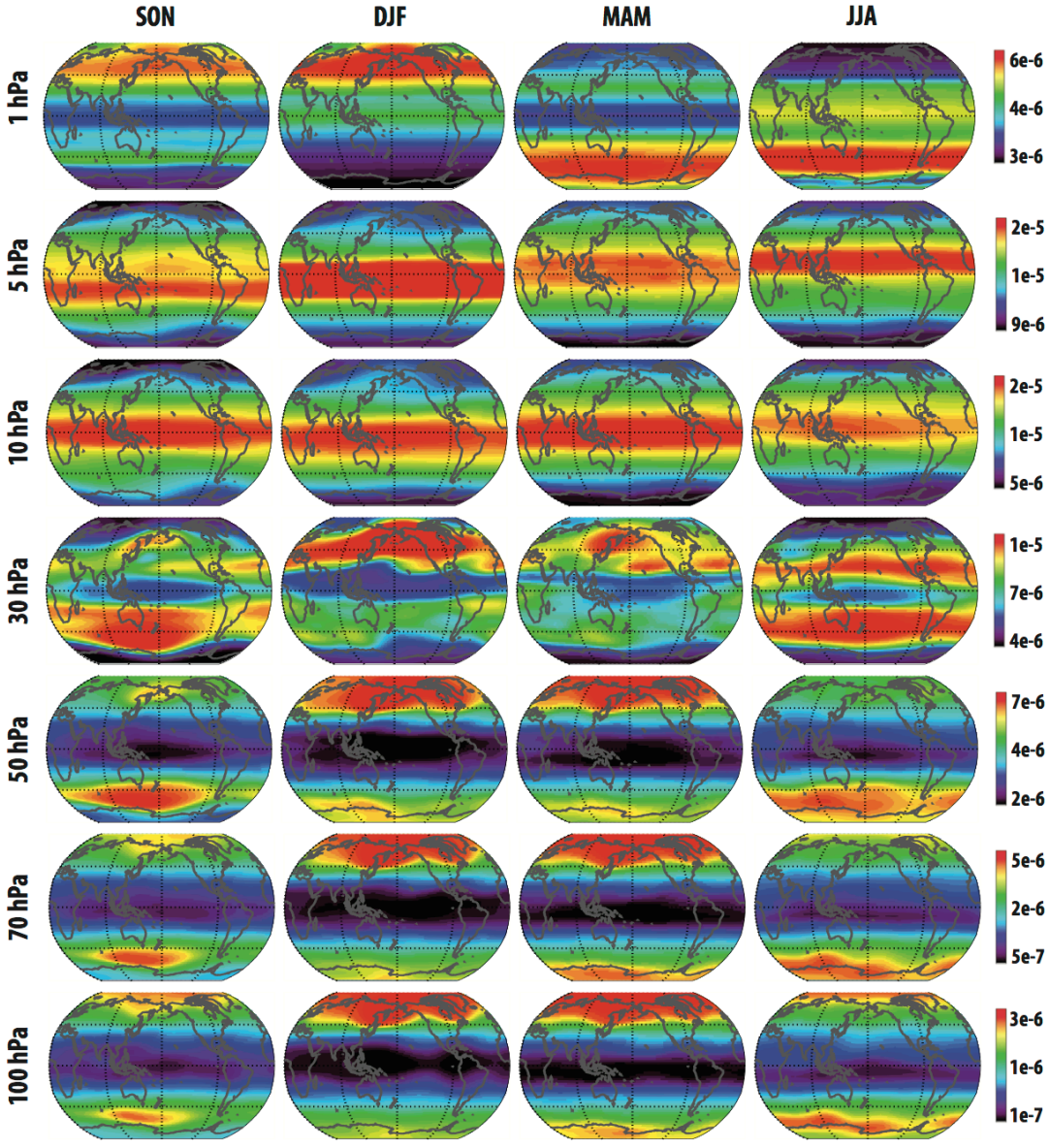

Figure S.6: Seasonal averages in ozone mixing ratio in mg/kg for 1979–2013 time period using MERRA series (non-linear color scale used).

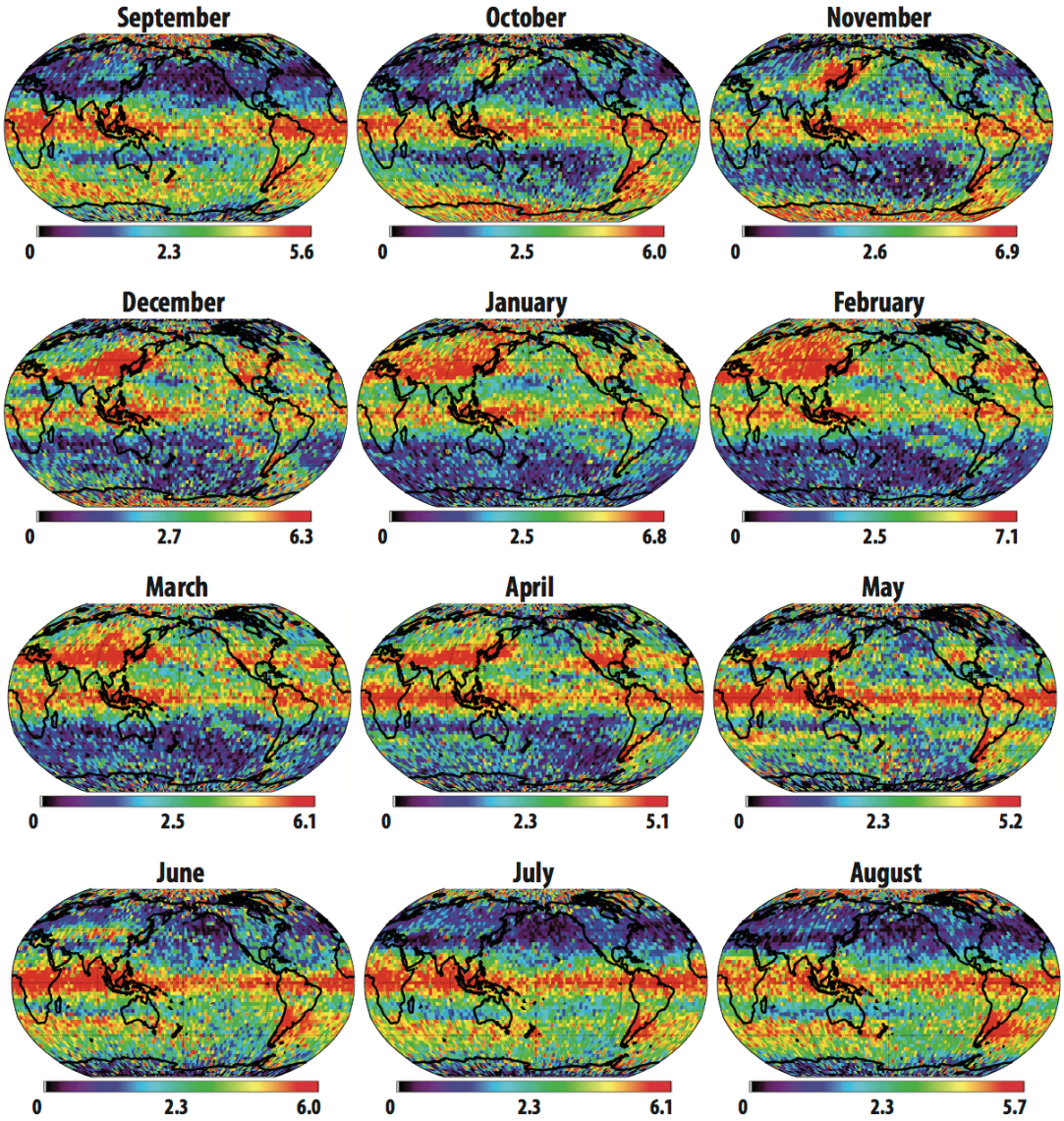

Figure S.7: Monthly means of the potential energy in J/kg averaged across the whole vertical profile for the studied time period 2007–2010 (non-linear color scale used).

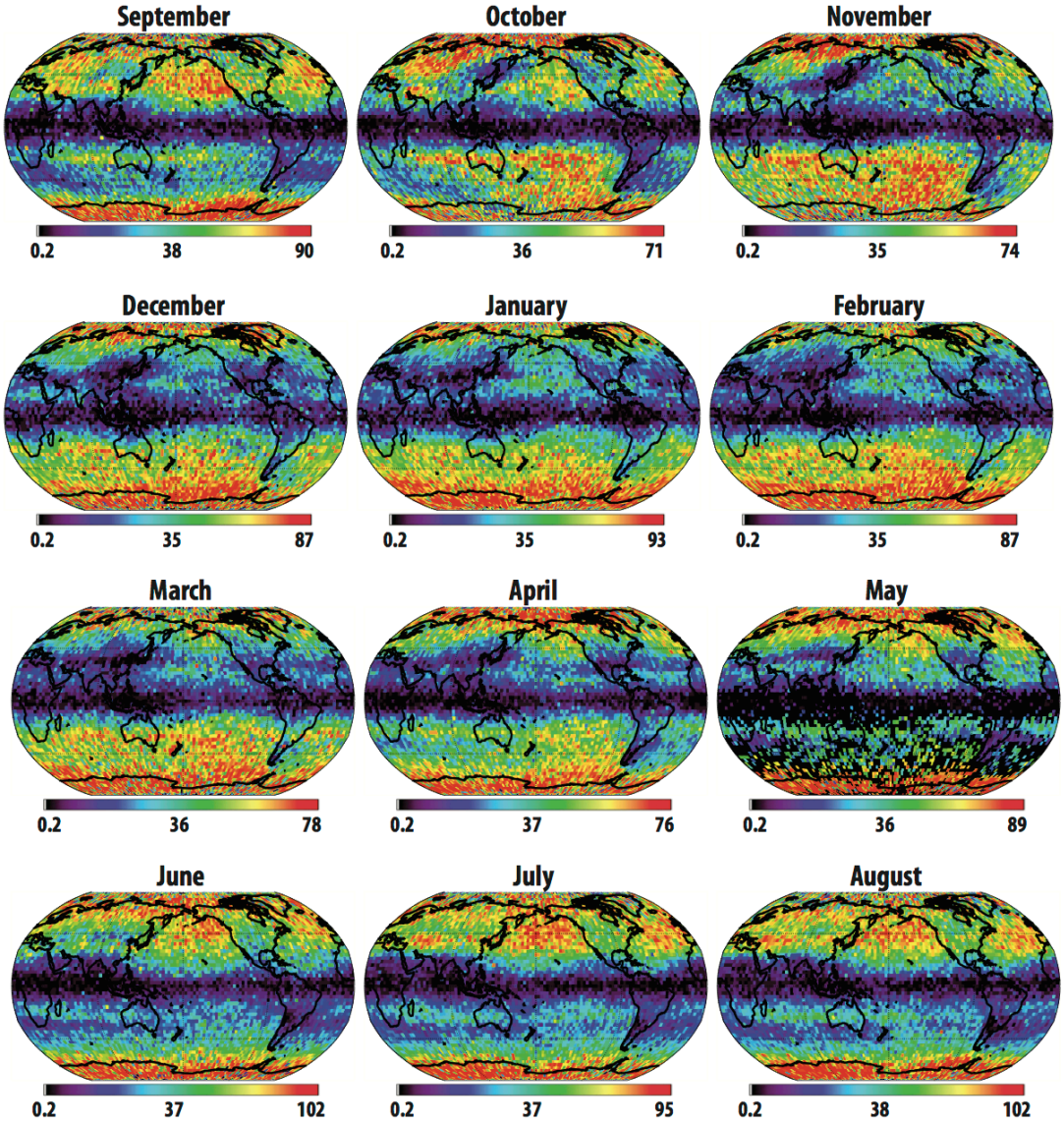

Figure S.8: Monthly means of the gradient Richardson number averaged across the whole vertical profile for the studied time period 2007–2010 (non-linear color scale used).

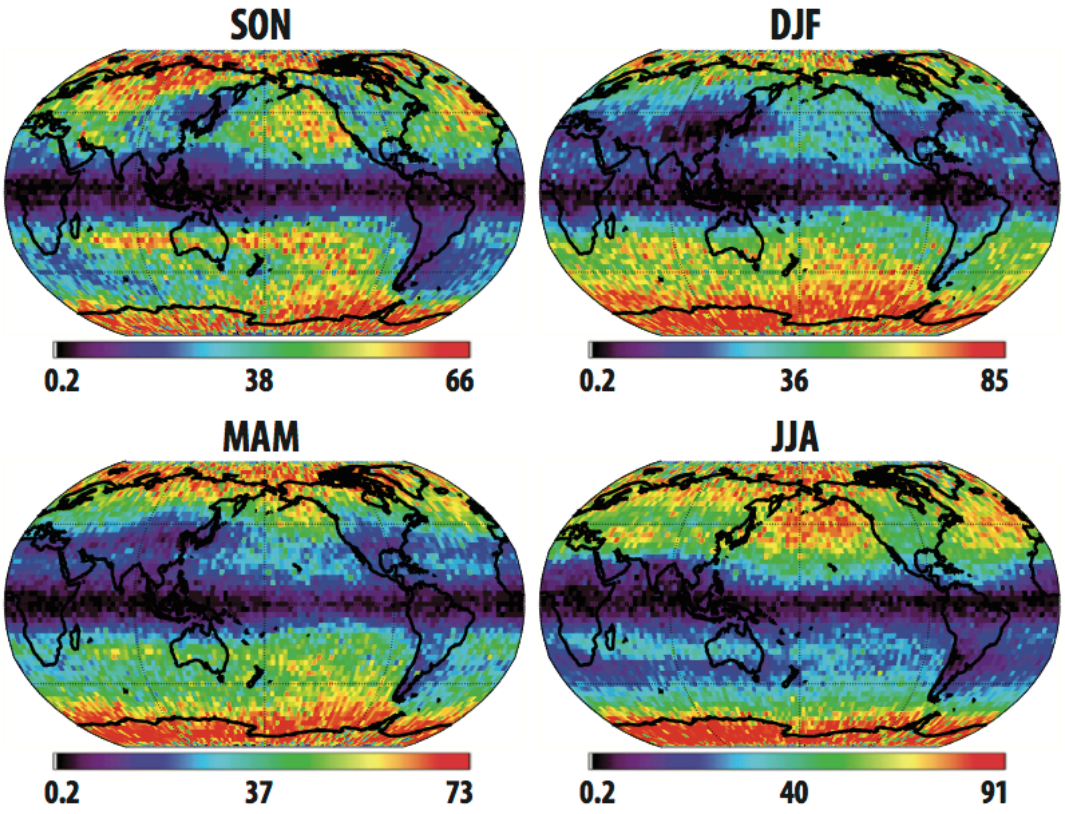

Figure S.9: Seasonal means of the gradient Richardson number averaged across the whole vertical profile for the studied time period 2007–2010 (non-linear color scale used).

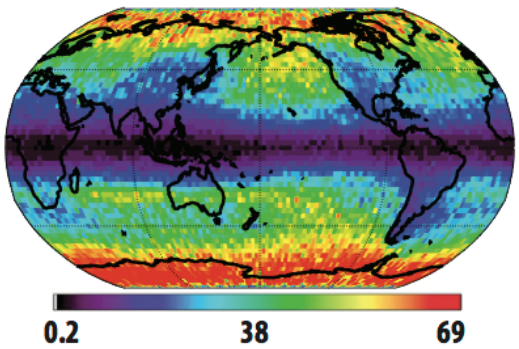

Figure S.10: Annual mean of the gradient Richardson number averaged across the whole vertical profile for the studied time period 2007–2010 (non-linear color scale used).

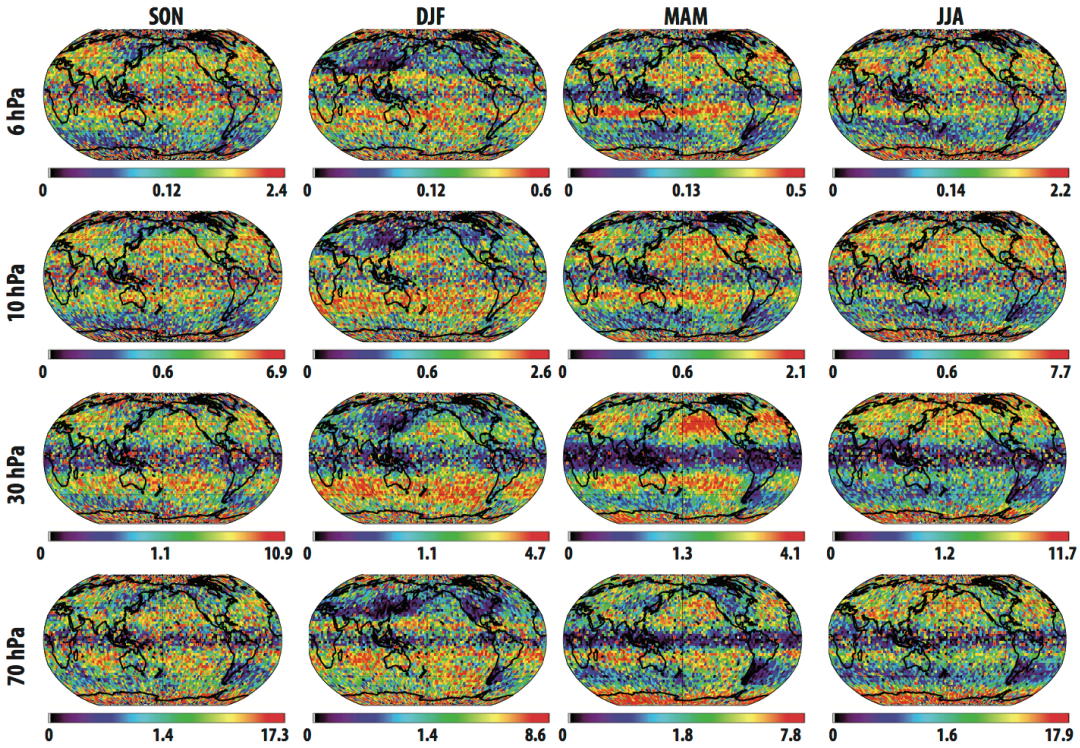

Figure S.11: Seasonal means of the gradient Richardson number at selected pressure levels for the studied time period 2007–2010 (non-linear color scale used).

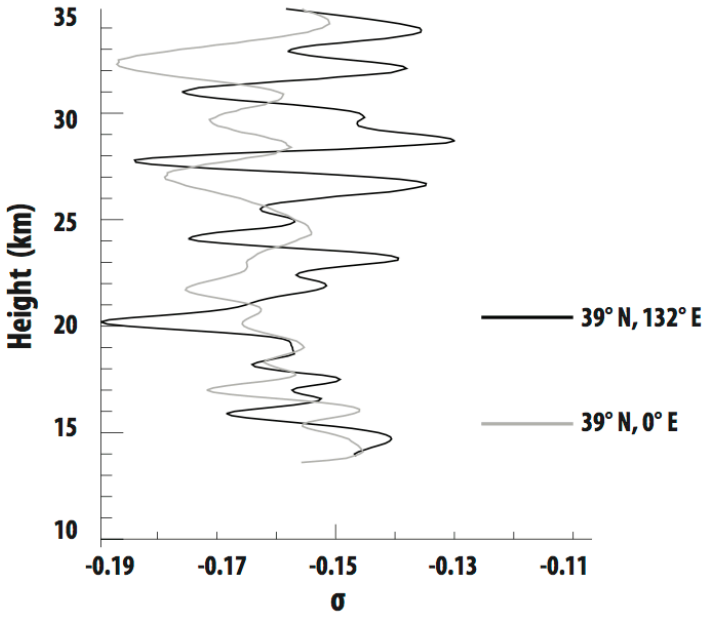

Figure S.12: Vertical profile of the sigma in s^{-2} for January 2008 in 39^{o}N, 132^{o}E (black line) and 39^{o}N, 0^{o}E (grey line).

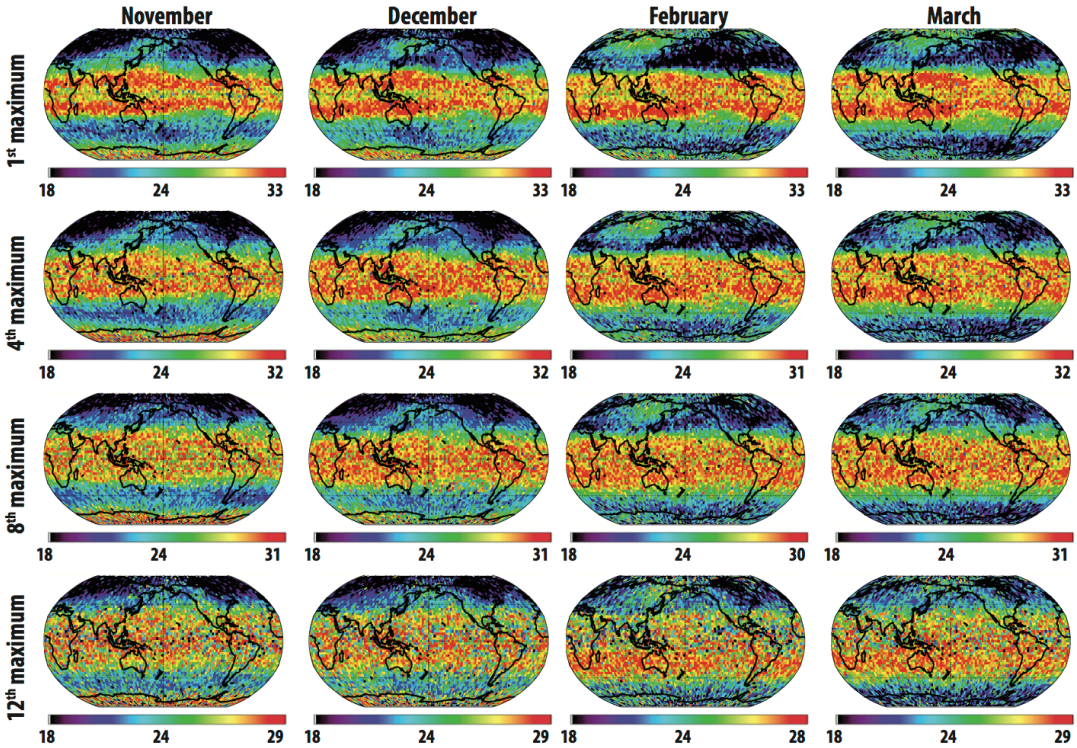

Figure S.13: Selected monthly means of altitudes in km of the primary and selected secondary (i.e., higher order) sigma squared maxima for the studied time period 2007–2010 (non-linear color scale used).

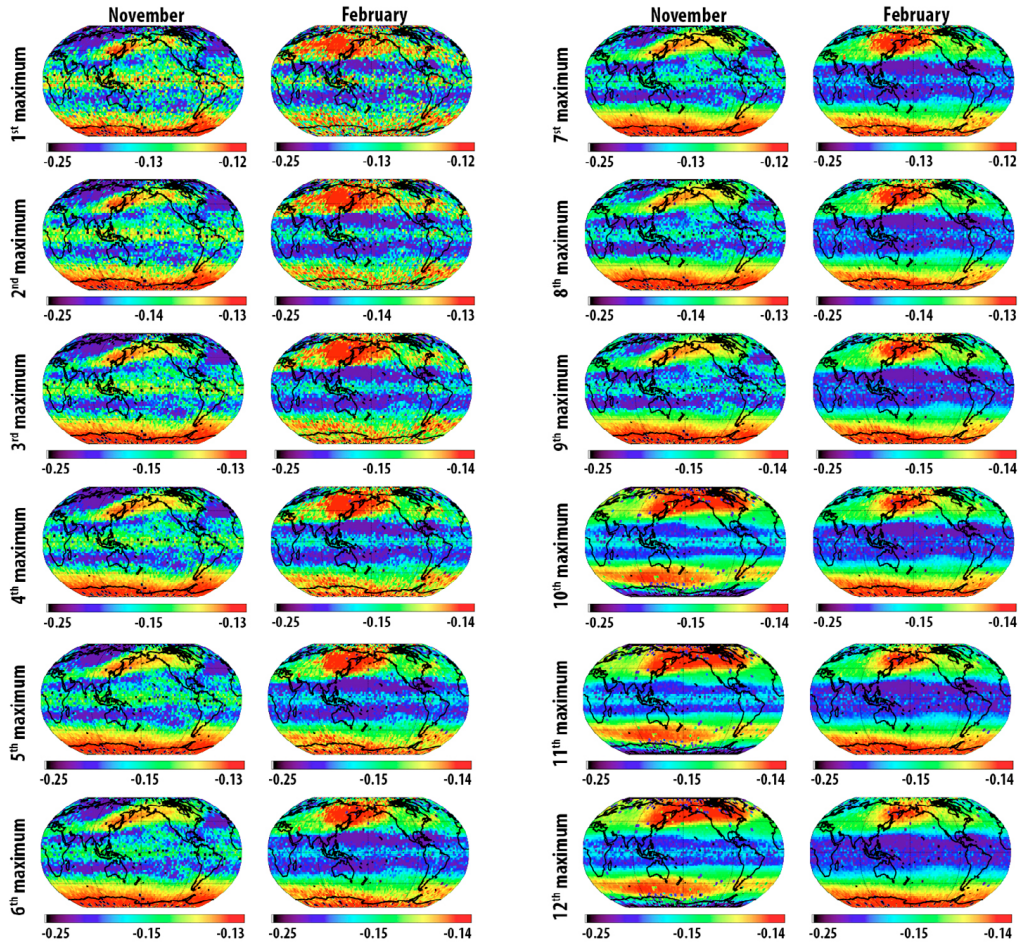

Figure S.14: Selected monthly means of primary and selected secondary (i.e., higher order) sigma squared maxima in s^{-2} for the studied time period 2007–2010 (non-linear color scale used).

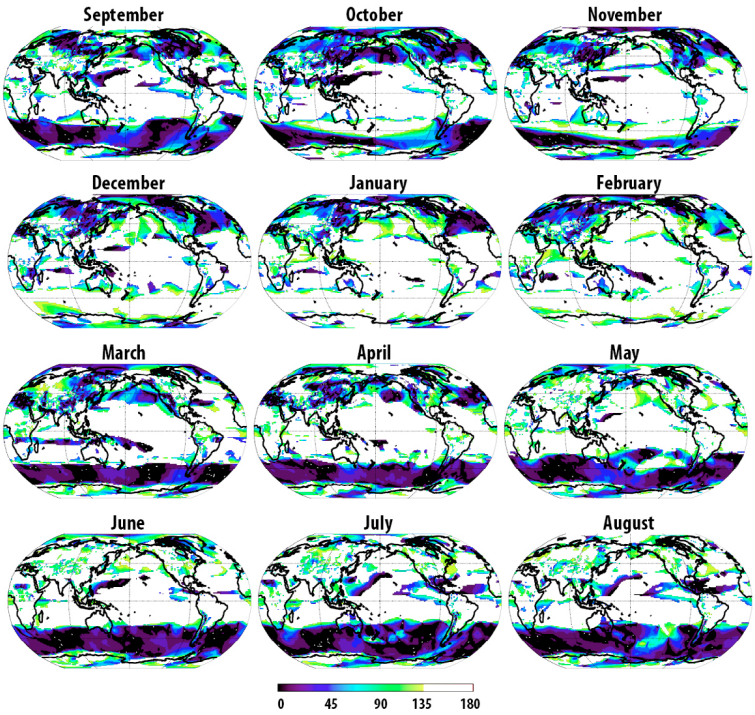

Figure S.15: Cumulative rotation of wind from 975 hPa to 10 hPa. Computed from JRA-55 for 2008.

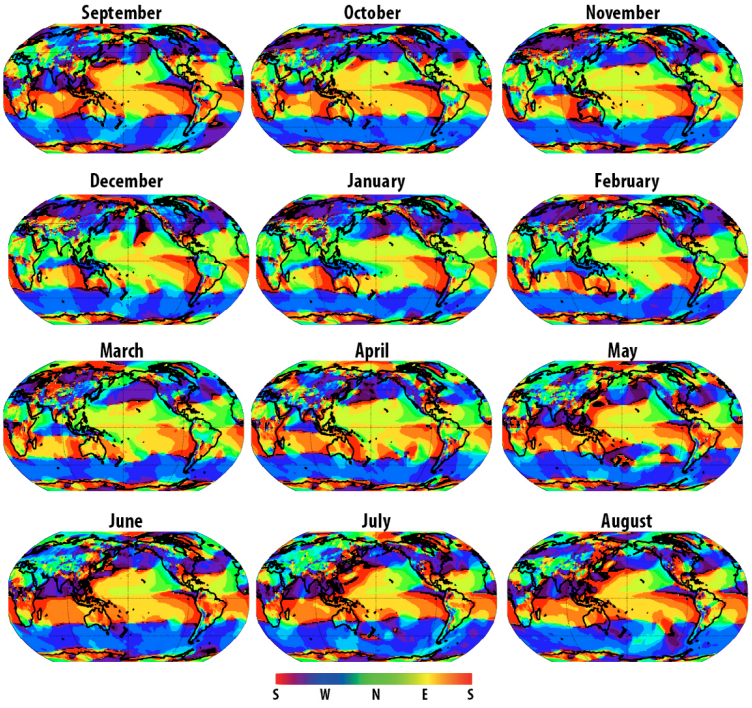

Figure S.16: Prevailing wind direction in the level of 975 hPa. Computed from JRA-55 for 2008.

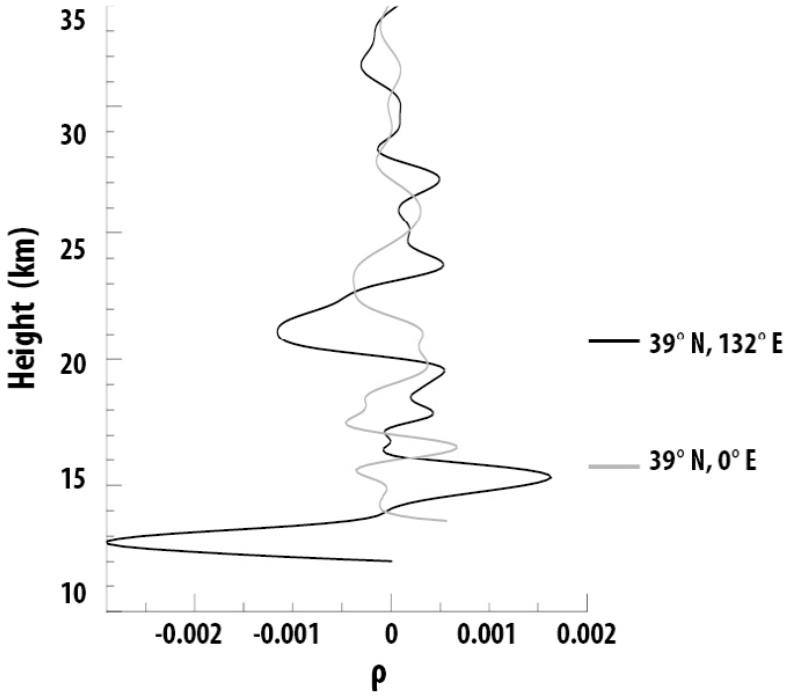

Figure S.17: Vertical profiles of the density perturbation for January 2008 in 39^oN, 132^oE (black line) and 39^oN, 0^oE (gray line).

Att. 2: Supplementary figures to the Chapter 4

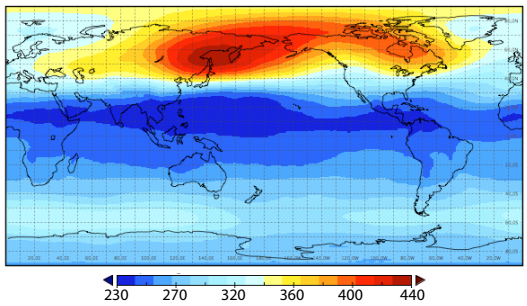

Figure S.18: 1978 to 2008 mean January mean total ozone column field from MSR (units in DU).

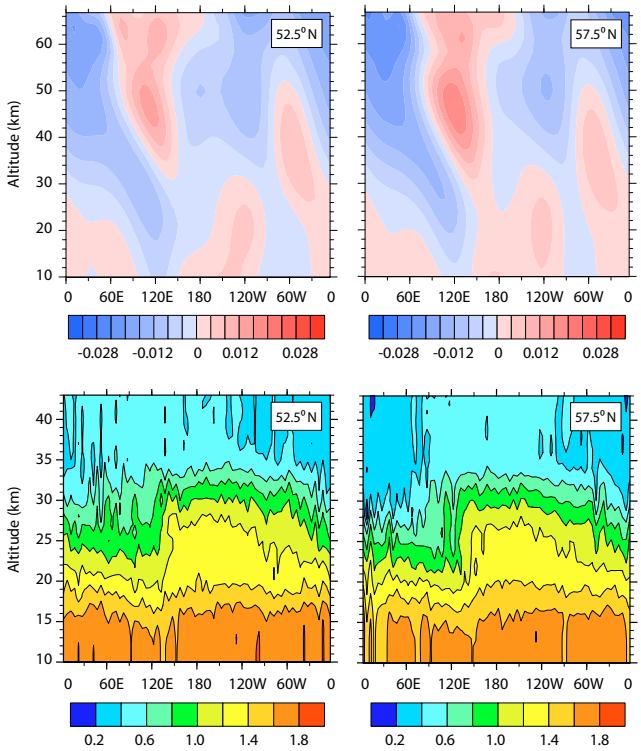

Figure S.19: Mean January zonal cross-sections of MUAM reference simulation residual vertical velocity (upper panels) and MIPAS CH_4 volume mixing ratio in January 2010 (lower panels) at 52.5^{o}N (left panels) and 57.5^{o}N (right panels).

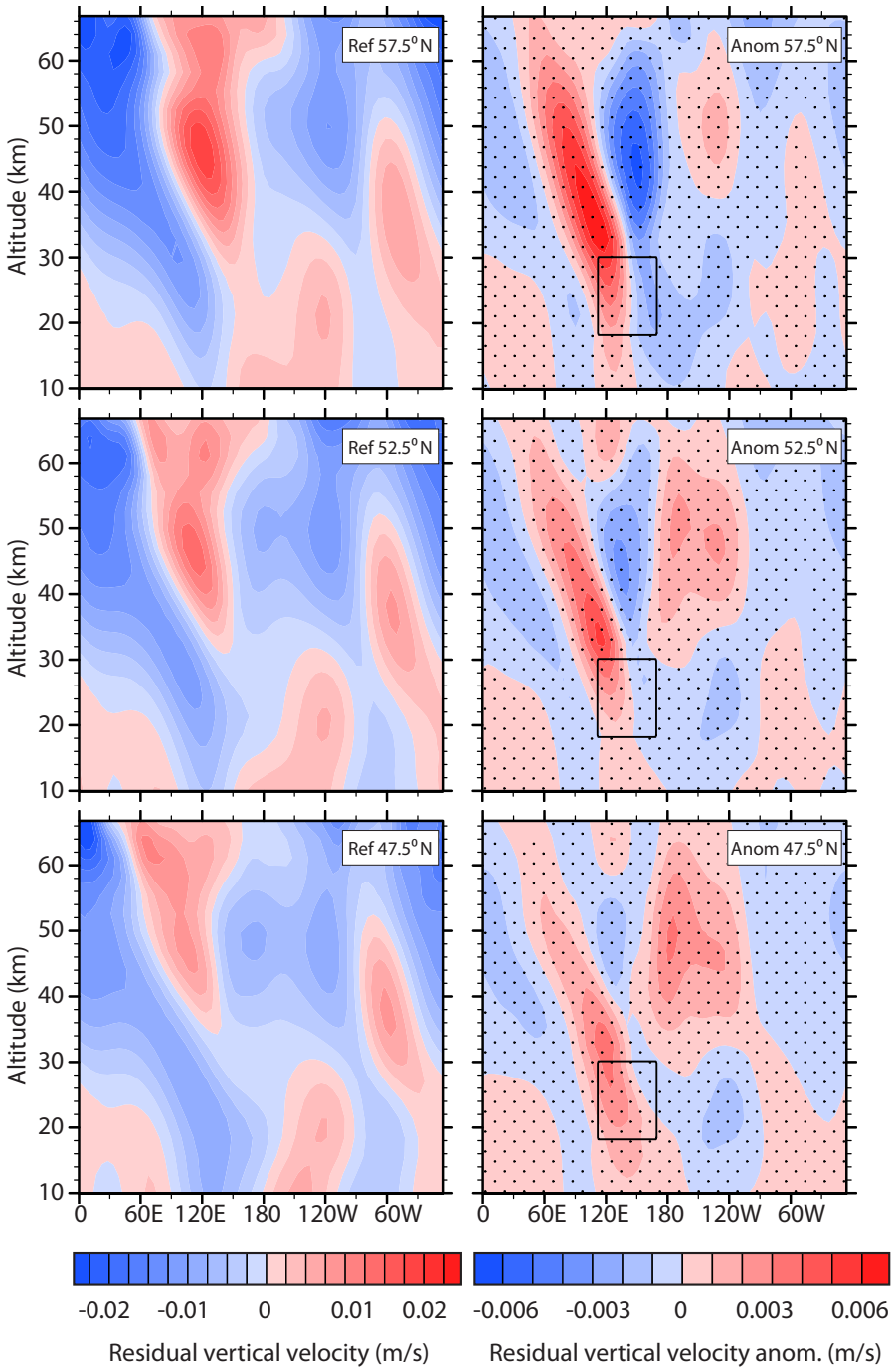

Figure S.20: Mean (7th to 30th day January longitudinal cross-sections of reference residual vertical velocity ($m{\cdot}s^{-2}$; on the left) and 10Box simulation anomalies (on the right) at selected latitudes. The contours illustrate the position of the artificial GWD. The statistical significance of the mean anomalies was computed by a t-test and regions with p values < 0.05 are stippled.

Animation captions:

https://av.tib.eu/media/20495?76
https://av.tib.eu/media/20497?92
Anim. 1: Time evolution of geopotential (colors, given in gpm) and horizontal winds (stream lines, given in $m \cdot s^{-1}$) for the SSWbox run (a) and SSWzon run (b) at the 13th model level (6.25 hPa) at consequent time steps (2h).

https://av.tib.eu/media/20494?108
Anim. 2: Time evolution of E-P flux vectors ($kg \cdot s^{-2}$; arrows are scaled according the relative distances of the plot) and its divergence (colors in $kg \cdot m^{-1} s^{-2}$) for the SSWbox simulation. Location of the artificial GWD is illustrated.

https://av.tib.eu/media/20500?60
Anim. 3: Time evolution of the zonal mean residual circulation (stream lines for illustration of direction only) and its mass flux (colors, in $kg \cdot m^2 s^{-2}$) for the 10box simulation on the left and relative 10Box-Ref simulation difference on the right.

https://av.tib.eu/media/20493?124
Anim. 4: Animation of longitudinal cross-sections of residual vertical velocity for the SSWbox simulation at particular latitudes encompassing the GWD area.