On the impact of middle-atmosphere thermal tides on the propagation and dissipation of gravity waves

F. Senf[1] and U. Achatz[2]

U. Achatz, Institute for Atmospheric and Environmental Sciences, Goethe University, Frankfurt (Main), Germany.

F. Senf, Leibniz Institute of Atmospheric Physics at the Rostock University, Kühlungsborn, Germany.

[1] Leibniz Institute of Atmospheric Physics
at the Rostock University, Kühlungsborn,
Germany.

[2]Institute for Atmospheric and
Environmental Sciences, Goethe University,
Frankfurt (Main), Germany.

Abstract. In the middle atmosphere, solar thermal tides cause large variations in the background conditions for gravity-wave propagation. The induced modulation of gravity-wave pseudo-momentum fluxes is responsible for a diurnal force. In past studies, this forcing was derived from gravity-wave parameterizations which neglect time-dependence and horizontal inhomogeneities of the background flow. In our study, we evaluate these assumptions using a highly simplified gravity-wave ensemble. With the help of a global ray-tracing model, a small number of different gravity-wave fields is transported through a time-changing background which is composed of a climatological mean and tidal fields from a general circulation model. Within three off-line experiments, assumptions on horizontal and temporal dependence of the background conditions have been successively omitted. Time-dependence leads to a modulation of gravity-wave observed frequencies and its phase velocities. Transient critical layers disappear. The amplitude of the diurnal forcing is reduced. Horizontal inhomogeneities induce a refraction of the gravity waves into the jet-stream cores. Horizontal propagation can lead to large meridional displacements and an inter-hemispheric exchange of gravity-wave energy. With equivalent Rayleigh friction coefficients, it is shown that for the gravity-wave ensemble in use the damping of tidal amplitudes is reduced when horizontal and time dependence of tidal background conditions are taken into account.

1. Introduction

Upward propagating gravity waves (GWs) transport a significant amount of momentum and energy from the lower to the middle atmosphere [*Fritts and Alexander*, 2003]. In the mesosphere / lower thermosphere (MLT) region GW breaking causes a mean force which is approximately balanced by a mean Coriolis torque and drives the large-scale meridional circulation. Main mechanisms which lead to the interaction of GWs and temporally averaged flow are well established [*Lindzen*, 1981; *Holton*, 1982; *Dunkerton*, 1982], but there is still some uncertainty concerning the interaction of GWs with middle-atmosphere variability patterns. One of them are solar thermal tides. These are excited to the largest part by large-scale solar heating of water vapor in the upper troposphere and ozone in the stratosphere as well as latent heat release in tropical convection regions [*Chapman and Lindzen*, 1970; *Grieger et al.*, 2004; *Achatz et al.*, 2008].

For the GW-tide interaction it is believed that the periodic modulation of GW breaking into small turbulent structures is responsible for the diurnal GW forcing. Hence, a detailed description of the GW-tide interaction process should incorporate a huge range of scales, from global structures to small-scale eddies. But, this is beyond the current computer capabilities. Most of former investigations used parameterizations of turbulent and GW forces and heating rates. Especially, GW parameterizations are not well constrained in their choice of GW source parameters as well as diffusion mechanisms [*Alexander et al.*, 2010]. This seems to be the origin of an ongoing controversy about the effect of GWs on tidal amplitudes [*Ortland and Alexander*, 2006, and discussion therein].

Previous investigations of the GW-tidal interaction may be sorted into two groups: (1) global modeling applying a linear tidal model [*Miyahara and Forbes*, 1991; *Forbes et al.*, 1991; *Meyer*, 1999; *Ortland and Alexander*, 2006] or a non-linear general circulation model (GCM) [*Mayr et al.*, 1999, 2001; *Akmaev*, 2001; *McLandress*, 2002] with a simplified GW parameterization and (2) local ray-tracing studies focusing on the interaction between large-scale and groups of small-scale waves [*Broutman*, 1984; *Broutman and Young*, 1986; *Zhong et al.*, 1995; *Eckermann and Marks*, 1996; *Sonmor and Klaassen*, 2000; *Walterscheid*, 2000; *Sartelet*, 2003]. Additionally, some new input to the field comes from non-linear GCM studies with resolved hydrostatic GWs [*Watanabe and Miyahara*, 2009].

Although simulations of the first group reproduce many features of both the atmospheric mean circulation and the solar tides, they possibly suffer from one major disadvantage hidden in the GW parameterization. In these, strong assumptions about the propagation and time dependence of GW fields have been imposed. Conventional GW parameterizations work in vertical columns which are assumed to be independent from each other, ignoring horizontal inhomogeneities in the large-scale flow [*McLandress*, 1998]. Furthermore, time-dependence of the background (BG) conditions is neglected. It is supposed that GW fields just see a quasi-stationary mean flow and adjust instantaneously to its changes. This assumption has originally been introduced for the representation of the interaction between GWs and a very slowly developing mean flow, but might be less appropriate for the interaction of GWs with solar tides.

In the second group, detailed studies of GW propagation in more or less extremely simplified BG situations have been performed. For instance, *Eckermann and Marks* [1996]

investigated a set of GW rays within a monochromatic and an amplitude-modulated tidal wave. The time-dependence of their chosen large-scale waves caused (1) a modulation of the GW observed frequency and thus of the horizontal phase velocity and (2) a local temporal change in the GW amplitude. From the latter, a non-dissipative GW force resulted induced by transient Eliassen-Palm (EP) flux effects. In the saturation region, lower diurnal GW forces were found compared to the conventional Lindzen GW parameterization [*Lindzen*, 1981].

The aim of our study is to extend the results by *Eckermann and Marks* [1996] to more realistic tidal motion and investigate the effect of propagation and dissipation of GWs in realistic tidal fields with the help of global ray-tracing simulations. We successively relax assumptions on horizontal inhomogeneity and time-dependence of the BG conditions and directly compare different results of each approximation. Special focus is on the diurnal GW force which acts back on the tide.

We like to emphasize that the current study is restricted by the use of an extremely simplified GW ensemble. For the sake of simplicity, a small number of horizontally homogeneous and continuously emitting GW sources have been considered. This has the advantage that all resulting temporal variability and horizontal inhomogeneity in the GW fields can be uniquely attributed to the impact of the BG conditions. The investigation of more realistic source configurations is left to future research.

The paper is structured as follows: In Sect. 2, the global ray-tracing model, the GW ensemble and the background-flow data for the different simulation setups are described. In the following sections 3 and 4 effects of GW frequency modulation and refraction of the horizontal GW vector are discussed, respectively. In Sect. 5, the periodic forces due

to GW stresses are presented. Possible effects of the GW forcing on tidal structures are discussed on the basis of equivalent Rayleigh-friction coefficients. A summary is given in Sect. 6 and supplementary material is provided in the appendix.

2. Model description

2.1. Basics

Under the assumption of a clear scale separation between background and small-scale gravity wave structures, an approximate WKB theory of locally monochromatic GWs can be established. With the help of multi-scale analysis, a hierarchy of equations can be derived [*Grimshaw*, 1975; *Achatz et al.*, 2010]. To leading order, a local dispersion relation and polarization relations between GW amplitudes are obtained.

For our study, the dispersion relation

$$\hat{\omega}^2 = (\omega - \boldsymbol{u}\cdot\boldsymbol{k})^2 = \frac{N^2k_h^2 + f^2m^2}{|\boldsymbol{k}|^2} \tag{1}$$

has been employed where $\boldsymbol{k} = k\,\boldsymbol{e}_\lambda + l\,\boldsymbol{e}_\varphi + m\boldsymbol{e}_z$, $k_h = \sqrt{k^2+l^2}$, $\hat{\omega}$ and ω denote wave vector, the horizontal wave number, intrinsic frequency and observed frequency, respectively, with the set of unit vectors $\{\boldsymbol{e}_\lambda, \boldsymbol{e}_\varphi, \boldsymbol{e}_z\}$ of the spherical coordinate system. The horizontal BG wind $\boldsymbol{u} = \boldsymbol{u}(\lambda,\varphi,z,t)$, the reference buoyancy frequency $N(z)$ and the Coriolis parameter $f(\varphi)$ are allowed to vary slowly in λ, φ, z, and t which are geographic longitude, latitude, geometric altitude and time, respectively. Thermodynamic reference profiles have been calculated via horizontal averaging of the 3D mean flow. Compared to previous ray-tracing studies [*Marks and Eckermann*, 1995; *Hasha et al.*, 2008], temporal and horizontal variations of the buoyancy frequency N and the scale height factor $1/4H_\rho^2$ are neglected. Detailed investigations showed that these terms do not significantly con-

tribute to the diurnal forces which is in line with *Zhong et al.* [1995]. Furthermore, the Doppler shift by the vertical BG wind wm as investigated by *Walterscheid* [2000] was a priory neglected in our study. The quantification of the impact of this term on the diurnal force is needed in future research.

In ray tracing, an initial, locally monochromatic GW field is divided into small parts in which local values of ω, $\boldsymbol{k}$ and an appropriate amplitude measure can be defined. Each part of the GW field is called wave parcel and is followed along its group velocity $\boldsymbol{c}_g = c_{g\lambda}\boldsymbol{e}_\lambda + c_{g\varphi}\boldsymbol{e}_\varphi + c_{gz}\boldsymbol{e}_z$ given in (A13)-(A15). The geometric position $\boldsymbol{x}$ of the wave parcel is determined by its initial position and the solution of $d_t\boldsymbol{x} = \boldsymbol{c}_g$ where d_t is the derivative along the group ray.

As shown in appendix A1, the ray tracing equations in a shallow atmosphere are

$$d_t\omega = \boldsymbol{k}\cdot\partial_t\boldsymbol{u}\,, \tag{2}$$

$$d_t k = -\,\boldsymbol{k}\cdot\frac{\partial_\lambda\boldsymbol{u}}{a_E\cos\varphi} + \frac{k\tan\varphi}{a_E}\,\hat{c}_{g\varphi}\,, \tag{3}$$

$$d_t l = -\,\boldsymbol{k}\cdot\frac{\partial_\varphi\boldsymbol{u}}{a_E} - \frac{fm^2}{\hat{\omega}|\boldsymbol{k}|^2}\,\frac{\partial_\varphi f}{a_E} - \frac{k\tan\varphi}{a_E}\,\hat{c}_{g\lambda}\,, \tag{4}$$

$$d_t m = -\,\boldsymbol{k}\cdot\partial_z\boldsymbol{u} - \frac{Nk_h^2}{\hat{\omega}|\boldsymbol{k}|^2}\,\partial_z N \tag{5}$$

where a_E is the earth radius, and $\hat{c}_{g\lambda}$, $\hat{c}_{g\varphi}$ denote the intrinsic zonal and meridional group velocity, respectively. The time-dependence of the BG wind, in our case the effect of the diurnal tide, induces a modulation of GW observed frequency ω along the ray. The horizontal gradients in the BG conditions lead to changes in the horizontal GW numbers. Furthermore, the convergence of the meridians due to the curvature of earth lead to turning of the horizontal wave vector $\boldsymbol{k}_h$ as indicated by the last terms of eq. (3) and (4).

Several aspects of the numerical implementation of the global ray tracing are discussed in appendix A2.

Following *Grimshaw* [1975], the GW amplitude equation arises in next order of WKB expansion. It condenses to the wave action law (see also *Bretherton and Garrett* [1968])

$$d_t A = -A\,\nabla\cdot\boldsymbol{c}_g - \tau^{-1}A \tag{6}$$

with

$$\nabla\cdot\boldsymbol{c}_g = \frac{\partial_\lambda c_{gx} + \partial_\varphi\left(\cos\varphi\, c_{gy}\right)}{a_E\cos\varphi} + \partial_z c_{gz}\,, \tag{7}$$

where A denotes the wave action density and τ^{-1} is the damping rate mainly due to wave breaking processes. The change in the volume of a ray bundle [*Walterscheid*, 2000] is determined by the divergence of the group flow. Wave action conservation is also known in a much more general context [*Andrews and McIntyre*, 1978; *Grimshaw*, 1984].

The damping rate τ^{-1} in the second term of ride-hand side of eq. (6) is estimated via a highly simplified turbulence parameterization based on saturation theory [*Lindzen*, 1981]. In this scheme, the GW amplitudes are forced back to the convective instability threshold if they have the tendency to grow above it. τ^{-1} is calculated in a way to ensure that the saturation condition is fulfilled [*Holton*, 1982]. As we are concerned with GW forces only, the explicit dependence of τ on the diffusion coefficient and Prandtl number can remain unspecified (for a sophisticated approach see *Marks and Eckermann* [1995]). Additionally, in the MLT region molecular viscosity and thermal diffusivity become more important and are included into the damping process. Note however that in the middle and upper thermosphere, also the dispersion of GW fields would be strongly affected by molecular motion [*Vadas and Fritts*, 2005].

Using eq. (6), a ray equation for the vertical flux of wave action $F_A = c_{gz}A$ is obtained, i.e. $d_t F_A = d_t c_{gz} A + c_{gz} d_t A$, and can be written as

$$d_t F_A = -\left(\tau^{-1} - \tau_{non}^{-1}\right) F_A \,, \tag{8}$$

where all non-dissipative effects have been collected into the rate

$$\tau_{non}^{-1} = c_{gz}^{-1}\Bigg(\partial_t c_{gz} + \frac{c_{gx}\partial_\lambda c_{gz} - c_{gz}\partial_\lambda c_{gx}}{a_E \cos\varphi} + \frac{(\cos\varphi\, c_{gy})\,\partial_\varphi c_{gz} - c_{gz}\partial_\varphi\left(\cos\varphi\, c_{gy}\right)}{a_E \cos\varphi}\Bigg) \tag{9}$$

which can be either positive or negative. τ_{non}^{-1} is derived by expanding and rewriting the terms $d_t c_{gz} A$ and $-A\nabla\cdot\boldsymbol{c}_g$ via $d_t = \partial_t + c_{gx}/(a_E\cos\varphi)\,\partial_\lambda + c_{gy}/a_E\,\partial_\varphi + c_{gz}\partial_z$ and eq. (7). Eq. (8) extends the relation given by *Marks and Eckermann* [1995] to time-dependent flows in spherical geometry. Changes in F_A result from dissipation via $-\tau^{-1}F_A$ and from temporal and horizontal variations of group velocity via $\tau_{non}^{-1}F_A$. The latter are connected to a local change of the volume which neighboring GW rays occupy [*Broutman et al.*, 2004]. In our simulations, the turbulent damping is the major contribution and changes in GW properties, e.g. ω and k_h, modify the GW breakdown, in our formulation, the damping rate τ^{-1}. Hence, time- and horizontal dependence of the background flow have mainly an indirect impact on the diurnal GW force in changing the turbulence parameterization. This is in contrast to direct non-dissipative forces due to transience and horizontal refraction, i.e. from $\tau_{non}F_A$, as discussed by e.g. *Dunkerton* [1981]; *Eckermann and Marks* [1996]; *Bühler* [2009].

2.2. Gravity-wave ensemble

In the present simulations, a small and highly idealized GW ensemble of *Becker and Schmitz* [2003], listed in tab. 1, was used. GWs with horizontal wave lengths between

about 400 km to 600 km and random initial phases are globally homogeneously and continuously emitted at the lower boundary, $\hat{z}_B = 20$ km ($\hat{z}$ denotes the average geo-potential height, see appendix A2). Each of the 14 individual and independent GW components are integrated forward separately. The individual GWs have initial horizontal phase velocities c_h between 7 and 30 m/s and are directed into 8 equi-distant azimuth directions with an increment of 45° beginning at east and increasing counter-clockwise. Furthermore, the GW ensemble is non-isotropic with largest k_h directed to east, largest c_h in zonal directions and largest momentum flux to the west as given in tab. 1.

It was shown by *Becker and Schmitz* [2003] that the mean residual circulation of middle atmosphere is well reproduced in a large-scale GCM when their GW ensemble is used in a Lindzen GW parameterization. Note however that, as that mostly resulted from tuning the GW parameters, this GW ensemble is just one of many possibilities. Therefore, the simple GW ensemble is viewed as a toy configuration in which the effect of temporal and horizontal variation of the BG conditions is investigated by way of a reasonably-motivated example. Beside its shortcomings, we did not indent to retune the given GW ensemble for the present study.

2.3. Background data

In the ray simulations, the background-field has been taken from the coupled chemistry climate model HAMMONIA which is explained by *Schmidt et al.* [2006] in detail. It was shown by several studies that simulation results from HAMMONIA compare quite well with recent observations [e.g *Achatz et al.*, 2008; *Yuan et al.*, 2008]. Global horizontal wind, temperature and geo-potential height data have been provided from a twenty year time slice experiment from 1980 to 1999 in typical solar maximum conditions with

a spectral truncation at T31 and 67 vertical levels. Monthly averaged values at eight different times a day within an interval of 3 hours have been used to calculate a monthly-mean diurnal cycle. Mean January values have been chosen. By a Fourier analysis in time, the latter has been analyzed for the monthly average and the diurnal tide. E.g. the zonal wind is represented by

$$u(\lambda,\, \varphi,\, \eta,\, t) = \bar{u}(\lambda,\, \varphi,\, \eta) + u_T(\lambda,\, \varphi,\, \eta,\, t)\,, \tag{10}$$

where the tidal wind component is

$$u_T = u_R\, \cos(\Omega t) + u_I\, \sin(\Omega t)\,. \tag{11}$$

Here, $\Omega = 2\pi\,(24\,\mathrm{h})^{-1}$ is the diurnal frequency, $u_R(\lambda,\, \varphi,\, \eta)$ and $u_I(\lambda,\, \varphi,\, \eta)$ are the corresponding Fourier coefficients and η is HAMMONIA's generalized vertical coordinate. For simplicity, semi-diurnal and shorter-period tidal variations have been excluded.

The zonally and temporally averaged zonal wind $[\bar{u}]$ and temperature $[\bar{T}]$ are plotted in fig. 1 for reference. Here, brackets and overbars denote zonal and temporal averaging over one latitude circle and one period, respectively. The climatological flow is reasonably well represented.

The total amplitude of the diurnal variation is defined in analogy to a zonal and temporal root-mean square, e.g. as

$$U = \sqrt{[u_R^2 + u_I^2]}\,, \tag{12}$$

which corresponds to the amplitude definition of *Ortland and Alexander* [2006]. The zonal wind amplitude U and meridional wind amplitude V are given in fig. 2 and have a double maximum structure with peaks at about 20°N and 20°S and between 100 km and 110 km.

The migrating parts of diurnal variations have zonal wave number one and follow the apparent motion of sun. Their corresponding amplitudes, also plotted in fig. 2, share the same double maximum structure, but are less in strength than the total amplitudes. The non-migrating tides, investigated by several authors [*Achatz et al.*, 2008, and references therein], are excited by the heating in equatorial convection zones and by the interaction of the migrating tide with planetary waves. The amplitudes of non-migrating tides show a single maximum in the tropics and are comparable to or even larger than their migrating counterparts. Migrating as well as non-migrating tides are considered as background for GW propagation.

2.4. Experimental setup

A hierarchy of three different experiments, named "full", "noREF" and "TS", with decreasing complexity has been used, as listed in tab. 2.

The "full" experiment refers to a full ray-tracing simulation without any approximations for horizontal and time dependence, i.e. the unmodified eqns. (2) - (5), (8) are integrated along the ray path given by eqns. (A10), (A11), (A12). Thus, changes in ω, $\boldsymbol{k}_h$ and m appear and are induced by mean flow changes. Also, the geographical distribution of the GW fields is altered.

"noREF" (no refraction) is a simplified ray-tracing experiment in which neither horizontal refraction nor horizontal propagation are allowed. In the simulation, the right-hand sides of eqns. (A10), (A11) and (3), (4) have been set to zero. Additionally, horizontal derivatives and curvature terms in eq. (8) have been ignored. Hence in experiment "noREF", ray points are only allowed to propagate vertically, but have a finite group velocity and feel the transience of the BG wind. The horizontal wave vector $\boldsymbol{k}_h$ is constant

along each group ray, but ω and m vary to compensate temporal and vertical changes in the BG conditions, respectively.

The third experiment is denoted by "TS" (time slicing) and is equivalent to a Lindzen-type vertical column parameterization with temporally fixed BG fields at each time step [*McLandress*, 1998]. Only vertical variations of the background are taken into account. A set of simulations was performed in a stationary background containing the temporal mean state and the diurnal tide fixed at one particular phase, i.e.

$$u_{\mathrm{TS},n} = \bar{u} + u_R \cos(2\pi n/12) + u_I \sin(2\pi n/12) \ , \tag{13}$$

where $n = 1, \ldots, 12$ was chosen for a good sampling of the diurnal cycle. At the end, all results for different tidal phases have been combined together.

With the three experiments, effects of frequency modulation and the refraction of horizontal wave vector can be extracted. Differences between "TS" and "noREF" are attributed to the first, whereas differences between "noREF" and "full" to the latter. As the simpler simulations "TS" and "noREF" are obtained by successively simplifying the "full" ray-tracing system, a consistent comparison of the results is possible while keeping numerical and implementation aspects the same.

3. Gravity-wave frequency and phase-speed modulation

3.1. The mechanism of frequency modulation

We consider a simple background flow which only consists of a zonal wind component $u(z,t) = U \sin(Mz - \Omega t)$ with the large-scale vertical wave number $M = -2\pi/L_z$ and vertical wave length L_z [*Broutman*, 1984; *Broutman and Young*, 1986; *Eckermann and Marks*, 1996]. For diurnal tides with $M = -2\pi\,(30\,\mathrm{km})^{-1}$ and $\Omega = 2\pi\,(1\,\mathrm{day})^{-1}$, the phase

progression $C = \Omega/M$ is downward and in the order of 0.3 m/s. We assume that the wind amplitude U is zero at the ground and slowly increases to a constant value above a certain altitude z_0. Additionally, the thermodynamic BG state is set to isothermal.

Two counter-propagating GW trains, i.e. propagating in positive and negative zonal direction with zonal wave number $k = \pm k_0$, are continuously emitted at ground with the phase velocities $c = \pm c_0$. Then the ray-tracing eqns. (2), (3) reduce to

$$d_t \omega = k \, \partial_t u \,, \tag{14}$$

$$d_t k = 0 \,, \tag{15}$$

which can be combined to the ray equation for the zonal phase velocity $c = \omega/k$

$$d_t c = \partial_t u \,. \tag{16}$$

From eq. (16), we infer that a local tendency of BG wind is connected to a change of zonal phase velocity c along the ray. But as also shown by *Eckermann and Marks* [1996] and *Walterscheid* [2000], phase velocity changes arise only due to frequency changes. Above z_0, u is monochromatic and a solution of the form $c(z - Ct)$ can be found for which eq. (16) changes to

$$\left(1 + \frac{c_{gz}}{|C|}\right) \partial_t c = \partial_t u \,. \tag{17}$$

Assuming small U, the ansatz

$$c \approx \pm c_0 + \delta c \, \sin(Mz - \Omega t) \,. \tag{18}$$

gives the phase velocity variation δc to the lowest order

$$\delta c = \frac{U}{1 + \frac{c_{gz,0}}{|C|}} \,, \tag{19}$$

where the initial vertical group velocity $c_{gz,0}$ was included. For upward propagating GWs, the frequency modulation acts so that c follows u as shown in fig. 3. Approximations (18) and (19) perform quite well even in the case of large U. Only peak values of c in fig. 3 are under-estimated in regions where vertical GW motion is slowed down. As pointed out e.g. by *Walterscheid* [2000], the slower the GW propagates in the vertical, the more pronounced is the effect of frequency modulation. In the two limits, we obtain for $c_{gz,0} \ll |C| : \delta c \to U$ and for $c_{gz,0} \gg |C| : \delta c \to 0$. Hence, the effect is important for slow GWs and seems to be negligible for fast GWs. For typical values of $N = 2\pi\,(300\,\mathrm{s})^{-1}$, $k_h = 2\pi\,(300\,\mathrm{km})^{-1}$ and $c_0 = 20\,\mathrm{m/s}$ the vertical group velocity of GWs is nearly equal to the tidal phase progression, i.e. $c_{gz,0} \approx |C|$, therefore the variation of the phase velocity is about $\delta c \approx U/2$.

One might conclude that the effect of frequency modulation is restricted to only the small part of very slowly vertical propagating GWs in the spectrum. However, it is believed that critical layer-type interactions with the mean wind are very important to induce breakdown and dissipation of GWs in the middle atmosphere and especially in the mesopause region [*Fritts and Alexander*, 2003]. When a spectrum of GWs approaches a critical region, a large part of the spectrum is slowed down and becomes aware of the time-dependence of the BG conditions. This makes us believe that the effect of frequency modulation is of overall importance in a realistic middle-atmosphere including temporal variation of solar-thermal tides.

3.2. The impact on saturated gravity-wave trains

In the following, we compare forces induced by saturated GW trains in the simple example from above: for a conventional vertical-column parameterization and for the

consistent time-dependent solution of eq. (16), respectively. For mid-frequency GWs, the zonal wind amplitude $|u'|$ is set back to the saturation threshold $|u'_s| = |c-u|$ above the level of convective instability [*Fritts*, 1984]. This also applies for time-dependent flows. Saturation leads to a flux of zonal GW pseudo-momentum

$$F = \frac{\rho_r}{2}\frac{k}{N}\,\hat{c}_h^3 = \frac{\rho_r}{2}\frac{k_0}{N}\,(c-u)^3\ , \tag{20}$$

where the horizontal intrinsic phase velocity $\hat{c}_h = \pm(c-u)$, the zonal wave number $k = \pm k_0$ and the reference density ρ_r were used. Restricting to the case of small U and $c_0 \gg U$, where in the conventional approach no critical levels are encountered and saturation is not disrupted due to strong wind shears which may overcome the effect of density decrease, the resulting zonal force $-\partial_z F/\rho_r$.

In the conventional approach, the GW phase velocity is assumed is to be constant, i.e. $c = \pm c_0$, and thus, the saturation flux becomes

$$F^{\pm}_{conv} = \frac{\rho_r}{2}\frac{k_0}{N}\left(\pm c_0^3 - 3c_0^2 U\sin(Mz-\Omega t) + \ldots\right)\ , \tag{21}$$

where terms nonlinear in U are not given explicitly. The diurnal force exerted on the mean flow due to the damping of counter-propagating GWs is

$$f_{conv,T} = -\frac{1}{\rho_r}\frac{\partial}{\partial z}\left(F^+_{conv} + F^-_{conv}\right) \tag{22}$$

$$= -\frac{3c_0^2 U k_0}{N}\left(\frac{1}{H_\rho}\sin\Phi_T - M\,\cos\Phi_T\right) \tag{23}$$

with the tidal phase $\Phi_T = Mz - \Omega t$ and the density scale height $H_\rho = -(\partial_z \ln\rho_r)^{-1}$. Terms nonlinear in U/c_0 have been neglected.

But, by taking realistic GW propagation into account, the periodic change in the BG wind induces a modulation of frequency and hence zonal phase velocity (see eq. (19) and (18)). This effect reduces the variation of the intrinsic horizontal phase velocity, saturation

pseudo-momentum flux and hence the diurnal force due to GWs. Utilizing eq. (18), we obtain for the diurnal force

$$f_T = f_{conv,T}\left(1 - \frac{\delta c}{U}\right) \tag{24}$$

and recall that $\delta c < U$, which is ensured by eq. (19). Therefore, the diurnal GW force f_T is reduced due to phase velocity variations compared to the conventional approach. Note also, that no critical layer is encountered for GWs in the time-dependent approach. The localized deposition of GW pseudo-momentum at the conventional critical layer is smoothed out by the effects of frequency modulation.

3.3. Vertical-column thinking and phase-velocity modulation in realistic flows

As mentioned before, large-scale circulation models need to apply GW parameterizations [*McLandress*, 1998]. Horizontal gradients of the BG medium are neglected which leads via eq. (3) and (4) to a conserved horizontal wave number k_h. Possibly of graver consequence, time-dependence of the transient large-scale motion is neglected in the vertical column. GW trains are assumed to feel a stationary background and adjust instantaneously to a given wind field. In this sense, perturbations in the GW field propagate infinitely fast to the levels above. The advective time scale, however, connected to the time which a part of a GW field vertically propagates can be in the order of a day and longer. But, the scale-separation assumption is still fulfilled if the GW times scale $\sim \hat{\omega}^{-1}$ is significantly smaller than a day. In eq. (18), the ratio $c_{gz,0}/C$ can be interpreted as ratio between BG time scale and GW advective time scale and directly affects the variation of GW phase velocities and diurnal forces.

The experience obtained from the vertical column model has guided the conventional thinking of gravity wave - mean flow interaction. There, the horizontal phase velocity of the GWs, c_h, is assumed to be constant and compared to the horizontal BG wind in GW direction, $u_h = \boldsymbol{u} \cdot \boldsymbol{k}_h / k_h$. The difference between both, i.e. the intrinsic horizontal phase velocity $\hat{c}_h = c_h - u_h$, is to a good approximation direct proportional to the vertical GW length. When $\hat{c}_h$ approaches its minimum, the vertical structure of the GW shrinks and turbulent diffusion becomes much more effective. The saturation momentum flux (20) is $\propto \hat{c}_h^3$ whose vertical variations determine the GW force on the BG medium. Hence, a consistent estimate of $\hat{c}_h$ is of major importance.

Fig. 4 shows the zonal phase velocity for GW ensemble member 12 (see tab. 1) at $\lambda = 0$ and $\varphi = 15°$N from the "noREF" experiment at four different times. The zonal wind jet favors the passages of westward GWs. Beyond 80 km altitude, phase velocity variations δc grow in amplitude up to 15 m/s. Note that this is only due to ω-modulation since horizontal refraction and propagation is switched off in the "noREF" experiment.

For the eastward propagating GW member 9, a temporal snapshot of c_h and u_h at time $t = 0$ and day 16 is given in fig. 5 at $\varphi = 15°$S for the "full" experiment. This experiment also includes effects of horizontal propagation and refraction. But, variations in c_h match surprisingly good to the variations in u_h in the mesopause region. This suggests a cooperation between frequency modulation and horizontal refraction due to tides. The amplitude of the c_h-modulation becomes with more than 30 m/s in the lower thermosphere larger than the initial phase velocity. Hence, there is no single (constant) phase velocity which can be attributed to the GW field when the temporal variation of

the thermal tide is present. Furthermore, due to the large ω-variations, negative values of c_h occur in the minima of the tidal winds.

4. Horizontal propagation and refraction of gravity-wave fields

4.1. Mechanisms of mean horizontal refraction

The temporally average horizontal refraction leads to changes in the GW fields which indirectly affect the diurnal tides. Additionally, the average horizontal GW propagation is responsible for a redistribution of the mean, but also of the diurnal GW force. In the following, we use simple examples to discuss the mean refraction effects.

Dunkerton [1984] and *Eckermann* [1992] showed that mainly the meridional gradients of the mean zonal wind $[\bar{u}]$ cause horizontal refraction of GW fields in the middle atmosphere. GWs propagating against the jet are refracted into its maximum. This is easily shown using the ray tracing eq. (4). We assume typical northern winter conditions for the upper stratosphere. A change in l due to the zonal mean wind is

$$d_t l\big|_{mean} = -k\,\partial_\varphi[\bar{u}]/a_E\,, \tag{25}$$

where for simplicity curvature effects have been excluded. In the winter hemisphere, GWs propagating against the jet, e.g. $k < 0$, have best propagation conditions. On the northern flank of $[\bar{u}]$, the wind increases with decreasing φ, thus $\partial_\varphi[\bar{u}] < 0$ and vice versa on the southern flank. Hence, $d_t l|_{mean} < 0$ north-ward the jet and > 0 south-ward the jet. An initially zonally aligned GW packet is refracted into the winter west-wind jet. On the summer hemisphere, the east-wind jet, $[\bar{u}] < 0$, supports east-ward GW motion with $k > 0$. But, as the Doppler shift $k[\bar{u}]$ remains negative as in the winter hemisphere, the same arguments apply here. The effect is illustrated in fig. 6(a).

A rough estimate of the change in meridional wave number $|\Delta l|$ of an initially zonally directed GW field is presented in the following. Within a latitude interval of around 20°, equivalent to a meridional distance of $\Delta y \approx 2000$ km, the mean zonal wind $[\bar{u}]$ increases (decreases) about $\Delta u \approx 60$ m/s from high- to mid-latitudes in the northern (southern) upper stratosphere (see fig. 1). Hence, for the meridional gradient one obtains $|\partial_\varphi[\bar{u}]/a_E| \approx |\Delta u/\Delta y| \approx 3 \times 10^{-5}$ s^{-1}. A GW with initial values of $k = 2\pi\,(300\,\mathrm{km})^{-1}$, $c = 20$ m/s, and thus $c_{gz} \approx 0.4$ m/s excited in the lower atmosphere propagates into the jet region. Due to the meridional wind gradients, it is refracted into the jet core, whereas the vertical wind shear leads to a decrease of its vertical wave number $|m|$. The GW accelerates up to a maximum vertical group velocity of ≈ 6 m/s. The wave field goes along a path of minimal vertical travel time (analogous to Fermat's principle in geometric optics). Between 30 km and 70 km altitude, i.e. $\Delta z \approx 40$ km, the average group velocity is about $\tilde{c}_{gz} \approx 3$ m/s. Thus, the GW stays there $\Delta t \approx \Delta z/\tilde{c}_{gz} \approx 1.3 \times 10^4 s$ which is about 4 hours. In this time interval, meridional refraction is most effective and induces a cumulative change of $|\Delta l|/k \approx |\partial_\varphi[\bar{u}]/a_E|\Delta t \approx 40\%$ above the wind jet. Note, as the zonal wave number remains constant, as $\partial_\varphi[\bar{u}] = 0$ and ignoring metric correction in eq. (3), the horizontal wave number $k_h = |k|\sqrt{1 + \Delta l^2/k^2}$ increases of about $|\Delta k_h/k| \approx \Delta l^2/(2k^2) \approx 8\%$. In addition to the effect on k_h, a horizontal redistribution of the GW field happens. Especially for slow GWs, large meridional displacements appear in the simulation and lead in some circumstances to the formation of caustics [*Dunkerton*, 1984].

Planetary Rossby waves in the winter stratosphere also affect the horizontal refraction of GWs [*Dunkerton and Butchart*, 1984]. We refer to a simple wave field in a channel

with Cartesian geometry shown in fig 6(b). The planetary wave is described by a simple stream function $\psi = -\Psi \sin(x) \sin(y)$, where x and y have been scaled by the channel size and Ψ is arbitrary. For $u = -\partial_y \psi$ and $v = \partial_x \psi$, divergence and shear deformation are zero. Hence, the GWs in this simple planetary wave are only affected by the vorticity $\zeta = 2\,\partial_x v$ and the strain deformation $\vartheta = 2\,\partial_x u$. Following *Bühler* [2009], the tendencies of k and l are

$$d_t \boldsymbol{k}_h \big|_{pw} = -\mathsf{S} \cdot \boldsymbol{k}_h \tag{26}$$

with the wind-shear tensor

$$\mathsf{S} = \begin{pmatrix} \partial_x u & \partial_x v \\ \partial_y u & \partial_y v \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \vartheta & \zeta \\ -\zeta & -\vartheta \end{pmatrix} . \tag{27}$$

Thus, the planetary wave vorticity leads to a rotation of $\boldsymbol{k}_h$ via $d_t k|_\zeta = -\zeta l$ and $d_t l|_\zeta = \zeta k$ in the sense of background vorticity. Cyclonic vorticity leads to anti-clock-wise turning of $\boldsymbol{k}_h$ and vice versa (on the northern hemisphere). The strain deformation induces via $d_t k|_\vartheta = -\vartheta k$ and $d_t l|_\vartheta = \vartheta l$ a change in the magnitude of the corresponding wave numbers. For instance, positive strain increases the magnitude of l. Both effects are summarized in fig. 6(b). Initially west-ward propagating GWs crossing the planetary wave trough in high latitudes are refracted to the south downstream the ridge. GWs from lower latitudes are refracted north-ward upstream the ridge, (down- and upstream with respect to $[u] > 0$).

The beta-effect (the second term on the right-hand side of eq. (4)) is usually small and causes a decrease of l in the northern and a increase of l in the southern mid-latitudes. At last, geometric effects due the to spherical shape of earth induce additional GW refraction (last terms in eq. (3) and (4)). In a hypothetical isothermal earth at rest, the angular momentum of a GW packet $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{k} A$ is constant along its path. The wave packet is

forced to move on a great circle [*Dunkerton and Butchart*, 1984; *Hasha et al.*, 2008]. For an initially zonally aligned GW at middle or higher latitudes, the geometric refraction cause equator-ward motion.

4.2. The impact on saturated gravity-wave trains

Analogously to Sect. 3.2, we estimate the impact of horizontal (in our example meridional) refraction on saturated GW trains. Again, the conventional vertical-column approach is compared to a solution which take horizontal gradients of the background flow into account. In the following, we refer to the example of the last paragraph, a GW with initial values of $k = 2\pi\,(300\,\mathrm{km})^{-1}$ and $c = 20$ m/s moving through a typical January zonal mean wind $[\bar{u}]$. If the influence of the mean meridional wind on the Doppler shift is neglected, i.e. $|l[\bar{v}]| \ll |k[\bar{u}]|$, than the intrinsic frequency $\hat{\omega}$ is mostly not affected by meridional refraction. But as the cumulative change in l leads to an increase in the horizontal wave number $\Delta k_h \approx 8\%$, the intrinsic horizontal phase velocity $\hat{c}_h$ is reduced by 8%, too. If the breakdown of the GW is described by saturation, then the corresponding flux of GW pseudo-momentum is given by eq. (20), i.e. $F \propto \hat{c}_h^3$ (again assuming mid-frequency approximation). Compared to the conventional saturation flux F_{conv}, the value of F is reduced due to refraction by a factor of $3\Delta k_h/|k| \approx 24\%$. Therefore, if the vertical dependence of this additional factor is ignored, the real zonal force f_λ is also diminished by horizontal refraction compared to the force $f_{conv,\lambda}$ calculated within the vertical-column approach, i.e.

$$f_\lambda = f_{conv,\lambda}\left(1 - 3\frac{\Delta k_h}{k_h}\right)\,. \tag{28}$$

The force reduction due to horizontal gradients is mainly a temporally average effect, but reduces the diurnal GW force as well.

4.3. Horizontal refraction in realistic flows

A zonally and temporally averaged picture of GW propagation for the four GW members 9, 10, 12 and 13 is shown in fig. 7. The arrows show the mean group velocities $[\bar{c}_{g\varphi}]$ and $100\,[\bar{c}_{gz}]$. The vector field illustrates the mean streaming of GW fields. The shadings indicate the zonally and temporally averaged geographical distribution of the initial latitudinal position the GW field had at the lower boundary, i.e. at $\hat{z}_B = 20$ km, where the ray have been initialized. It visualizes the meridional displacement of the GW field. The last quantity overlaid in the plots is the zonally and temporally averaged BG wind in wave direction $[\bar{u}_h]$. Negative contours indicate GW propagation against the wind whereas positive contours GWs with the wind.

GW members 9 and 10 are east and north-eastward aligned at the lower boundary, i.e. at $\hat{z}_B = 20$ km, respectively. They have favorable propagation conditions in the southern stratosphere. In the jet core at about 30°S, the mean group-velocity vectors are mainly vertically aligned. At the edges, parts of GW fields are refracted into the jet core. In the summer mesopause region, GW fields are slowed down due to the reversing BG winds. The GW fields avoid the positive jet core and are refracted into the mid-latitudes. Especially for GW member 10, this refraction leads to a irreversible growth of the meridional wave number above 90 km and large meridional displacements. In the northern hemisphere, the GW fields can vertically propagate through the minima of planetary wave structures.

GW members 12 and 13 are west and south-westward aligned at the lower boundary, i.e. at $\hat{z}_B = 20$ km, respectively. The westerly-wind vortex of the northern winter hemisphere

provides most favorable propagation conditions. In the jet core, the group velocities are mainly vertical. At the wind reversal, GW fields are refracted into meridional direction. For GW member 13, the mean latitude positions are interchanged in the lower thermosphere. Parts of the GW field initially from the northern mid-latitudes have moved south-ward to the equatorial region and even to the southern hemisphere, whereas parts of the GW field initially from the subtropics have propagated north-wards. As discussed before, due to the modulation of stratospheric winds by planetary waves zonally dependent wave guides develop. The easterly-wind jet in the southern hemisphere mainly prohibits propagation of GW member 12 and 13. Interestingly, some chance exists for parts of the high-latitude GW field to circumvent the jet core (GW member 12). Above the critical jet GW fields are refracted southward and spread over a large horizontal domain. The considerable horizontal expansion of GW fields, as seen for GW member 12 at e.g. 100 km and between 80° and 10°N as well as GW member 13 above 110 km and between 80°S and 30°N, also influences the amplitudes of the GW field via eq. (8). The corresponding change in GW amplitudes is not incorporated in most previous ray-tracing work [*Marks and Eckermann*, 1995; *Hasha et al.*, 2008; *Song and Chun*, 2008] which commonly apply the assumption of a constant vertical flux $F_A = c_{gz}A$ of wave action density A.

The median meridional displacement for GW member 9 and 12 remains around zero in the MLT, but large displacements up to 50° are also possible. For member 10 and 13, median displacement of 26° and −27° occur, respectively, but its distribution is broad with maximum values up to 100°. Hence, some parts of the GW fields are interchanged between both hemispheres.

5. Gravity-wave forces on the tide

5.1. Mean gravity-wave forces

Before investigating the periodic GW forces, which are one major focus of this study, changes in the temporally mean GW force are inspected. As discussed e.g. by *Andrews et al.* [1987], the relevant GW forcing of the mean flow, in our case temporally averaged flow plus diurnal tides, is given by the divergence of the GW pseudo-momentum flux rather than the GW momentum flux itself. The main difference between both arise for slowly vertically propagating, inertia-gravity waves. These waves produce a Stokes drift which is counterbalanced by an Eulerian mean flow locally attached to the waves [*Bühler*, 2009]. Hence, some parts of the force inferred from the divergence of momentum flux are needed to sustain the local Eulerian circulation and do not change the BG conditions. The vertical flux of zonal pseudo-momentum is [*Fritts and Alexander*, 2003]

$$F_{P,\lambda} = \hat{c}_{gz}\, kA = \rho_r \,\langle u'w' \rangle \left(1 - \frac{f^2}{\hat{\omega}^2}\right) , \tag{29}$$

where the prime denotes GW perturbations which are averaged over reasonable GW scales via the bracket operator. Therefore, the wave stress on the (Lagrangian) mean flow is reduced by a factor of f^2 over $\hat{\omega}^2$.

In neglecting horizontal variations in the GW fields, the horizontal force due to GW stresses is expressed as [*Fritts and Alexander*, 2003]

$$\boldsymbol{f}_h \approx -\frac{1}{\rho_r}\partial_z \left(\hat{c}_{gz}\boldsymbol{k}_h A\right) . \tag{30}$$

As we are interested in the effects of horizontal inhomogeneities in the BG conditions on the diurnal GW force, the more complete form [*Grimshaw*, 1975]

$$\boldsymbol{f}_h = -\frac{1}{\rho_r}\nabla \cdot \left(\hat{\boldsymbol{c}}_g \boldsymbol{k}_h A\right) \tag{31}$$

is used which includes the full 3D divergence of the flux tensor of GW pseudo-momentum. In the following, an ensemble mean force is calculated by an arithmetic average over all 14 GW members.

The temporally and zonally averaged zonal GW force is shown in fig. 8. Three experiments "full" (a), "noREF" (b) and "TS" (c) are compared to each other. In the mesopause region at about 80 km to 85 km, the typical dipole structure is visible with a negative forcing peak in winter and a positive one in summer. In the "TS" simulation, the peak values are about -55 and 48 m/s per day which might be a factor 1.5 to 2 smaller than typical GW forcing values in realistic GCM simulations [compare e.g. *Alexander et al.*, 2010; *Richter et al.*, 2010]. Beside this deficit, it is instructive to quantify the impact of tidal time-dependence and mainly mean horizontal gradients on the time-mean GW force using the 3 experiments. There is no significant change between the GW forces stepping from the "TS"-experiment (fig. 8(c)) to the "noREF"- experiment (fig. 8(b)) when focusing on the mesopause region. Hence, the frequency modulation has no impact on the mean force there. On the other hand, the GW force is diminished in the "full"-simulation with values of about -40 and 36 m/s per day due to the impact of horizontal refraction. This is a reduction of about 17% and 35% in the southern and northern hemisphere, respectively, and, as explained in Sect. 4.2, a temporally mean effect resulting from a cumulative change in the horizontal wave number k_h. For the "full"-simulation, a force-weighted hemispheric average [*Preusse et al.*, 2009] of $|\Delta k|/k_h$, $|\Delta l|/k_h$ and $|\Delta k_h|/k_h$ was calculated with for instance northern hemispheric values of 47%, 11% and 12% respectively. There, changes in l seem to dominate the changes in k_h. Using eq. (28), the increase in k_h can explain a reduction of the temporally mean force of about 36%.

Different ray simulations sequentially excluding tides and planetary waves were also performed (not shown). For the reduction of the temporally mean force, equivalent values were found even if temporal and horizontal variations of the background flow are excluded. Furthermore, runs with tides only were made but no clear impact of the tidal gradients on the GW saturation was found.

5.2. Periodic forces due to wave stresses

Since GW fields in the MLT region are periodically modulated by tidal winds, they produce a periodic force acting back on the diurnal tides. The diurnal amplitude of the zonal force $f_\lambda = \boldsymbol{f}_h \cdot \boldsymbol{e}_\lambda$, calculated analogously to eq. (12), is shown in fig. 9 for the three experiments "full", "noREF" and "TS". The results from the "TS" experiment, fig. 9(c), which mimics the effect of a conventional Lindzen GW parameterization, are chosen as basis for comparison.In the northern hemisphere (winter) a pronounced subtropical maximum with a peak force of 40 m/s per day can be seen whereas in the southern hemisphere (summer) and high latitudes the forcing peaks at about 80 m/s per day. The overall structure of the forcing amplitudes of the "TS" experiment compares quite well to past investigations of several authors with very different assumptions on GW source parameters [*Miyahara and Forbes*, 1991; *Meyer*, 1999; *Ortland and Alexander*, 2006; *Watanabe and Miyahara*, 2009], even though a highly simplified GW ensemble is used here.

For the "noREF" experiment in fig. 9(b), the total forcing amplitude is decreased. The northern hemispheric maximum is reduced to 28 m/s per day which is about 30% less than the "TS" value. In the southern hemisphere, dramatic differences arise in the high-latitudinal maximum with a reduction up to 70%. Now, relations (17) and (24)

are used to roughly estimate this frequency-induced reduction: When averaged globally over the altitude range of 60 km to 120 km, the mean tidal vertical phase velocity is $C \approx -0.34$ m/s, and when additionally averaged over all GW ensemble members, the mean GW vertical group velocity is $c_{gz,0} \approx 0.52$ m/s. With eq. (17), a factor of 0.4 is obtained which gives with eq. (24) a reduction of 40% of the mean diurnal force. For the crude assumption made, this estimate is quite good. The comparison also shows, that the frequency modulation mainly indirectly affects the GW force as the saturation strength is altered. The direct "transient EP-flux effects" [*Eckermann and Marks*, 1996] are at least one order of magnitude smaller, here.

The forcing amplitudes are further reduced in the "full" experiment in fig. 9(a). The asymmetry between subtropical winter and summer maximum has disappeared. Peak values nearly reach 20 m/s per day which is only 50% of the winter peak value of the "TS" experiment. In addition, the mentioned high-latitude peak at about 110 km has moved to 30°S. Its value is reduced to about 90% of the conventional one. As discussed in Sect. 5.1, the reduction of the diurnal forcing amplitude is caused by an overall decrease of the total GW force. The globally, vertically and spectrally averaged reduction of the diurnal, zonal GW force is 65% for the "full" experiment compared to the "TS" experiment.

The amplitude of the meridional diurnal force is shown in fig. 10. Its structure is similar to the zonal forcing given in fig. 9, but its maxima reach only about 40% of f_λ which is likely the result of the anisotropic GW ensemble. With reference to satellite observations, it was discussed by *Lieberman et al.* [2010] that the meridional GW force can be as twice as large than the zonal GW force. We are not able to reproduce this feature with our

toy GW ensemble. In the "noREF" and "full" experiments (fig. 10(a) and 10(b)) the winter peak is reduced to 30%. The subtropical summer forcing changes slightly between the given experiments.

We also investigated the direct effects of horizontal GW propagation and refraction on the diurnal forcing. We found that in the main forcing region between 80 and 95 km, the diurnal force amplitudes due to horizontal divergence of the pseudo-momentum flux-tensor, horizontal turning and stretching of the wave field are in the order of 10% to 20% for our chosen GW ensemble. Hence, the force due to dissipation of wave action is the dominant contribution to the total force.

5.3. Equivalent Rayleigh frictions coefficients

Equivalent Rayleigh friction coefficients (ERFs) have been introduced in the context of GW-tidal interaction by *Miyahara and Forbes* [1991]; *Forbes et al.* [1991] and further discussed e.g. by *McLandress* [2002]. With the help of ERFs, the effects of GWs can be incorporated into a linear tidal model [*Miyahara and Forbes*, 1991; *Forbes et al.*, 1991; *Ortland*, 2005a]. However, the concept has also diagnostic value for non-linear simulations with parametrized GWs [*McLandress*, 2002] as well as resolved GWs [*Watanabe and Miyahara*, 2009].

Previous studies, mainly focused on the GW effect on the migrating tidal components whereas non-migrating parts have been ignored. Here, we discuss both effects in a zonally averaged manner. With the definition of the real part γ_R and imaginary part γ_I of ERFs

$$\gamma_R = -U^{-2} \left[\overline{f_\lambda \, u_T}\right] \,, \tag{32}$$

$$\gamma_I = -\Omega^{-1} U^{-2} \left[\overline{f_\lambda \, \partial_t u_T}\right] \,, \tag{33}$$

569 the diurnal force is approximated by

$$f_\lambda \approx -\gamma_R \, u_T - \frac{\gamma_I}{\Omega} \, \partial_t u_T \,, \tag{34}$$

570 where again bracket and overbar denote zonal and temporal average, respectively.

571 Note that $\overline{[f_\lambda \, u_T]}$ is the average tidal kinetic energy tendency induced by zonal GW
572 forcing. Since $\overline{[f_\lambda \, u_T]} < 0$ is equivalent to $\gamma_R > 0$, positive real parts of the ERFs indicate
573 regions of decrease in tidal kinetic energy and therefore damping of the tides and vice
574 versa. The imaginary part of ERF acts on the tidal phase structure. For $\gamma_I < 0$ decrease
575 in tidal vertical wave length is observed and vice versa (see discussion by *McLandress*
576 [2002]; *Ortland* [2005b]). A reduction of tidal vertical wave length is a very robust result
577 in previous investigations, whereas the GW effect on tidal amplitudes is controversial
578 [*Ortland and Alexander*, 2006, and references therein].

579 The real parts of ERFs are shown in fig. 11 for the "full", "noREF" and "TS" simu-
580 lations.For the reference simulation "TS" in fig. 11(c), large positive peaks up to 60 in
581 $10^{-6}s^{-1}$ occur. The maxima correspond to values of 2 to 5 per day which are a factor of
582 3 to 5 larger than values reported by *Forbes et al.* [1991] and *McLandress* [2002], but in
583 the line with *Miyahara and Forbes* [1991].

584 In fig. 11(c), the typical structures of the real part of ERF for a Lindzen-type saturation
585 parameterization are shown. Consider, e.g. a vertical profile at 45°S. Negative values of
586 γ_R are encountered below about 78 km and positive ones above that altitude. The mean
587 onset of GW breaking is around 75 km where for this profile the negative peak appears.
588 Negative γ_R can lead to an increase in tidal amplitudes which is a typical result for GWs
589 which approach their critical levels or even if the onset of instability is dominated by tidal
590 winds as discussed by *Lu and Fritts* [1993] and *Mayr et al.* [1999]. In altitudes above

591 the onset of breaking, the saturation is controlled rather by the density decrease than
592 the increase in tidal wind amplitudes. We refer to the simple wave example in the limit
593 $c_0 \gg U$ and eq. (23). Assuming additionally a slow vertical increase in tidal amplitudes
594 with $H_U = (\partial_z \ln U)^{-1} > 0$, the real part of the ERF is [*Lu and Fritts*, 1993]

$$\gamma_{R,conv} = \frac{3c_0^2 k_0}{N}\left(\frac{1}{H_\rho} - \frac{1}{H_U}\right). \tag{35}$$

595 Clearly, density and wind changes act in complementary ways on the gravity wave satura-
596 tion. Positive values of γ_R are consistently obtained above the onset of GW breaking. This
597 is in line with sensitivity studies reported by *McLandress* [1997] and is supplementary to
598 the discussion raised by *Akmaev* [2001] on the effect of Lindzen-type saturations.

599 For the "noREF" experiment, in fig. 11(b), the magnitude of γ_R is reduced. The
600 latitude-altitude structure is wave-like with a vertical wave length comparable to the tidal
601 wave length. In fig. 11(a), the magnitude of the ERF is further reduced. Compared to
602 the "noREF" experiment, the influence of γ_R is drastically lowered in high-latitudes and
603 in the thermosphere. Fig. 11(a) corresponds surprisingly well to the non-linear simulation
604 with resolved GWs by *Watanabe and Miyahara* [2009], even though here an extremely
605 simple GW ensemble was used and no feedback between GWs and tides was taken into
606 account.

607 The imaginary parts γ_I of the ERFs are shown in fig. 12. For the "TS" experiment
608 in fig. 12(c), two distinct negative peaks appear at 45°N, 85 km and 60°S, 100 km with
609 magnitudes around -27 and -60 in $10^{-6}\mathrm{s}^{-1}$, respectively. In the "noREF" experiment,
610 the mid-latitude winter maximum is reduced to 40%. For γ_I from the "full" experiment,
611 any significant impact on lower thermosphere has disappeared. Two negative maxima in

the mid-latitude mesopause region are present with peaks around -20 and -15 in $10^{-6}\mathrm{s}^{-1}$ in the southern and northern hemisphere, respectively.

Horizontal averages of γ_R and γ_I are comparable in magnitude to the values published by *McLandress* [2002]. For the "full" simulation, one distinct peak appears between 80 and 85 km with values of 3 and -7 in $10^{-6}\mathrm{s}^{-1}$ for γ_R and γ_I, respectively.

6. Conclusions and summary

A global ray-tracing analysis of GW fields has been performed in which a zonally dependent climatological mean flow and diurnal tides from a GCM have been used. Our main objective is an evaluation of the assumptions usually made by single-column GW parameterizations. To quantify the impact of the temporal and horizontal variability, a small and highly simplified GW ensemble is used as a toy configuration for the investigation of GW propagation. For this, different ray-tracing experiments with increasing complexity have been performed in which the background conditions were unaffected by GW forces. The successive reduction of imposed assumptions give us the opportunity to consistently compare the ray tracing results with the conventional approach.

First, when the time-dependence of the thermal tides is included in the description of GW propagation, GW observed frequencies are modulated [*Eckermann and Marks*, 1996; *Walterscheid*, 2000]. Also, the GW phase velocity c_h is periodically changed such that GWs avoid their conventional critical level. As c_h follows the shape of the background wind, the diurnal GW forcing is reduced. Second, especially meridional gradients of the zonal-mean flow refract GW fields into the wind jets. Permanent changes in the horizontal wave number k_h reduce the value c_h which is responsible for a reduction of the temporally mean GW force, but also for a further decrease of the diurnal GW force.

Frequency modulation and refraction of horizontal wave numbers mainly have an indirect impact on diurnal GW forces in changing the saturation strength. Cumulative changes in ω and $\boldsymbol{k}_h$, which each part of the gravity-wave fields undergoes during its propagation, are important and sum up significantly. Additionally, the inclusion of horizontal propagation decreases the forcing in polar regions and in the lower thermosphere. Direct forces due to horizontal refraction are not dominant, but may lead to changes of 10% to 20% with respect to the total diurnal force. Note, however, that no feedback between GWs and tides is included in this study. We do expect, that the non-linear coupling will further modify the GW impact on the tides.

With the help of Rayleigh friction coefficients, the possible effect of the diurnal GW force on the diurnal tides is estimated. For the conventional GW parameterization with the Lindzen saturation assumption, the real part of the equivalent Rayleigh friction coefficient is mainly positive and expected to produce a damping of tidal amplitudes. This situation changes when temporal and horizontal dependence of the background conditions are taken into account. In the more complex ray tracing simulations, which also use the simple saturation approach, the forcing is more restricted to the mesopause region with much smaller coefficients. Furthermore, alternating areas of positive and negative influence on the tide exist. For the imaginary part of the Rayleigh friction coefficient, predictions from previous investigations are confirmed [*Ortland and Alexander*, 2006, and reference therein]. Two negative peaks are found in the mesopause region. As shown by those authors, this will decrease the vertical wave length of the thermal tides.

The extreme simplifications, i.e. the use of the toy GW ensemble, the saturation approach for the turbulence parameterizations and the non-interactive calculations, provided

us the starting point for a simple attribution and quantification of the discussed GW effects. These idealizations undoubtedly limit the generality of the results in this study on the realistic impact of GWs on the tide. Nevertheless, it seems to support that GW parameterization should not be blindly used. All corresponding assumptions have to be tested for each target problem. The investigation of more realistic GW fields, more sophisticated turbulence parameterization and especially the feedback between GW forces and diurnal tides is left to future research.

Appendix A: Ray tracing in a shallow, spherical atmosphere

A1. Basic ray equations

The basic ray tracing equations are derived here with special emphasis on metric corrections appearing in a shallow, spherical atmosphere [*Hasha et al.*, 2008].

Following *Hayes* [1970], GW observed frequency ω and GW vector $\boldsymbol{k}$ are connected to local variations of the GW phase Θ by

$$\omega = -\partial_t \Theta \quad \text{and} \quad \boldsymbol{k} = \nabla\Theta\,, \tag{A1}$$

where according to the shallow atmosphere approximation the radial distance r has been replaced by the mean radius of earth a_E in $\nabla = \boldsymbol{e}_\lambda\,\partial_\lambda/(a_E\cos\varphi) + \boldsymbol{e}_\varphi\,\partial_\varphi/a_E + \boldsymbol{e}_z\,\partial_z$. Note that $\boldsymbol{k}$ is defined as local Cartesian quantity, but its projection on the set of spherical unit vectors $\{\boldsymbol{e}_\lambda, \boldsymbol{e}_\varphi, \boldsymbol{e}_z\}$ changes during its evolution.

With this in mind, the dispersion relation (1) gives a Hamilton-Jacobi equation for Θ and is used to derive the evolution equations for ω and $\boldsymbol{k}$. In the following, we write $\omega = \omega(\boldsymbol{k}, \boldsymbol{\Lambda})$, where the local properties of the background medium are summarized in the vector $\boldsymbol{\Lambda}$ [*Bretherton and Garrett*, 1968]. For a local change in observed frequency,

676 we have

$$\partial_t \omega = \frac{\partial \omega}{\partial k_i}\, \partial_t k_i + \frac{\partial \omega}{\partial \Lambda_n}\, \partial_t \Lambda_n \tag{A2}$$

677 where $k_i = \boldsymbol{k} \cdot \boldsymbol{e}_i$ and i and n count coordinate directions and the number of background
678 quantities, respectively. As the unit vectors do not depend on time, we obtain

$$\partial_t \boldsymbol{k} = -\nabla \omega \tag{A3}$$

679 and

$$d_t \omega = \frac{\partial \omega}{\partial \Lambda_n}\, \partial_t \Lambda_n \,, \tag{A4}$$

680 where the group velocity is $\boldsymbol{c}_g = (\partial_{k_i} \omega)\, \boldsymbol{e}_i$, and advective derivative along a ray $d_t =$
681 $\partial_t + \boldsymbol{c}_g \cdot \nabla$.

682 For the GW vector $\boldsymbol{k}$, the same procedure applies and using eq. (A3) gives

$$\partial_t \boldsymbol{k} = -\frac{\partial \omega}{\partial k_i}\, \nabla k_i - \frac{\partial \omega}{\partial \Lambda_n}\, \nabla \Lambda_n \,. \tag{A5}$$

683 Next we show that the term $c_{gi} \nabla k_i$ can be rewritten as an advective derivative supple-
684 mented by metric corrections. We get

$$c_{gi} \nabla k_i = c_{gi} \nabla \left(\boldsymbol{k} \cdot \boldsymbol{e}_i \right) = \nabla \boldsymbol{k} \cdot \boldsymbol{c}_g + c_{gi} \nabla \boldsymbol{e}_i \cdot \boldsymbol{k} \,.$$

685 Since $\nabla \boldsymbol{k} = \nabla \nabla \Theta$ is a symmetric tensor of second order, i.e. $\nabla \boldsymbol{k} = (\nabla \boldsymbol{k})^T$, we obtain
686 $\nabla \boldsymbol{k} \cdot \boldsymbol{c}_g = \boldsymbol{c}_g \cdot \nabla \boldsymbol{k}$. Applying the projection $\boldsymbol{k} = k_i \boldsymbol{e}_i$ once again, we arrive at

$$\begin{aligned} c_{gi} \nabla k_i &= \left(\boldsymbol{c}_g \cdot \nabla k_i \right)\, \boldsymbol{e}_i + \boldsymbol{c}_g \cdot \nabla \boldsymbol{e}_i \, k_i + c_{gi} \nabla \boldsymbol{e}_i \cdot \boldsymbol{k} \\ &= \left(\boldsymbol{c}_g \cdot \nabla k_i \right)\, \boldsymbol{e}_i + \boldsymbol{c}_g \cdot \Big(\nabla \boldsymbol{e}_i - (\nabla \boldsymbol{e}_i)^T \Big)\, k_i \,, \end{aligned}$$

where in the last line $\nabla\left(\boldsymbol{e}_i \cdot \boldsymbol{e}_j\right) = 0$ was used. Hence, the ray equations for the wave numbers k_i are

$$(d_t k_i)\ \boldsymbol{e}_i = -\frac{\partial\omega}{\partial\Lambda_n}\,\nabla\Lambda_n - \boldsymbol{c}_g \cdot \left(\nabla\boldsymbol{e}_i - (\nabla\boldsymbol{e}_i)^T\right) k_i\,, \tag{A6}$$

which are valid for quite general coordinate systems [*Hasha et al.*, 2008].

As before, the shallow atmosphere approximation will be used in which vertical derivatives of all unit vectors and all derivatives of the outward pointing unit vector $\boldsymbol{e}_z$ are neglected. Thus, only the convergence of meridians is taken into account via

$$\nabla\boldsymbol{e}_\lambda = \frac{\tan\varphi}{a_E}\,\boldsymbol{e}_\lambda\boldsymbol{e}_\varphi \quad \text{and} \quad \nabla\boldsymbol{e}_\varphi = -\frac{\tan\varphi}{a_E}\,\boldsymbol{e}_\lambda\boldsymbol{e}_\lambda\,. \tag{A7}$$

Using additionally

$$\frac{\partial\omega}{\partial u} = k, \quad \frac{\partial\omega}{\partial v} = l, \quad \frac{\partial\omega}{\partial f} = \frac{f m^2}{\hat{\omega}|\boldsymbol{k}|^2} \quad \text{and} \quad \frac{\partial\omega}{\partial N} = \frac{N k_h^2}{\hat{\omega}|\boldsymbol{k}|^2}\,, \tag{A8}$$

the ray eqns. (2) - (5) are obtained. Furthermore, rewriting $u_i = \boldsymbol{u}\cdot\boldsymbol{e}_i$ in eqns. (2) - (5) led to the change $\boldsymbol{c}_g \to \hat{\boldsymbol{c}}_g$ in the corresponding metric corrections.

A2. RAPAGI: the numerical implementation

The RAy parameterization of Gravity wave Impacts (RAPAGI) is a fast numerical model which allows to solve the ray tracing equations on a spherical globe. For the direct use of GCM data, it is favorable to identify the position $\boldsymbol{x}$ of the wave parcel in spherical coordinates λ, φ and an altitude $\hat{z}$ which will be the globally averaged geo-potential height on surfaces of the vertical hybrid coordinate η. As each change of z along the ray is expressed as

$$d_t z = \partial_t z + (d_t\lambda)\,\partial_\lambda z + (d_t\varphi)\,\partial_\varphi z + (d_t\hat{z})\,\partial_{\hat{z}} z\,, \tag{A9}$$

the evolution of a ray point is given by

$$d_t\lambda = \frac{c_{gx}}{a_E \cos\varphi}\,, \tag{A10}$$

$$d_t\varphi = \frac{c_{gy}}{a_E}\,, \tag{A11}$$

$$d_t\hat{z} = \frac{c_{gz} - \partial_t z - \boldsymbol{c}_g \cdot \nabla_h z}{\partial_{\hat{z}} z}\,, \tag{A12}$$

where the components of group velocity $\boldsymbol{c}_g$ are

$$c_{g\lambda} = u + \frac{k}{|\boldsymbol{k}|^2}\frac{N^2 - \hat{\omega}^2}{\hat{\omega}}\,, \tag{A13}$$

$$c_{g\varphi} = v + \frac{l}{|\boldsymbol{k}|^2}\frac{N^2 - \hat{\omega}^2}{\hat{\omega}}\,, \tag{A14}$$

$$c_{gz} = -\frac{m}{|\boldsymbol{k}|^2}\frac{\hat{\omega}^2 - f^2}{\hat{\omega}}\,. \tag{A15}$$

This facilitates inter-model communication. The partial derivatives in eqns. (2)- (5) are given in a coordinate system with geometric altitude z, while these quantities are usually calculated from the large-scale flow in generalized coordinates $\{\lambda, \varphi, \hat{z}(\eta)\}$. The transformation between both was taken into account in our ray tracing simulations.

The time-integration of eqns. (2) - (5) is done in two stages. First, an integration estimate $\{\omega^*_{n+1}, \boldsymbol{k}^*_{n+1}\}$ for time $(n+1)\Delta t$ is obtained using the Heun scheme with a fixed time step of $\Delta t = 5$ min for which convergence has been verified. Second, an optimization technique is used to adaptively change all ray properties till the dispersion relation is retained. For $\beta_i \ll 1$, the corrected estimates $\omega_{n+1} = \omega_n + \Delta\omega\,(1+\beta_0)$ with $\Delta\omega = \omega^*_{n+1} - \omega_n$ and $k_{i,n+1} = k_{i,n} + \Delta k_i\,(1+\beta_i)$ with $\Delta k_i = k^*_{i,n+1} - k_{i,n}$ fulfill dispersion relation (1). In the optimization progress, the functional

$$\mathcal{G} = \frac{1}{2}\sum_{i=0}^{3}\beta_i^2 + \hat{\beta}\,\left(\omega(\boldsymbol{k}_{n+1}, \boldsymbol{\Lambda}_{n+1}) - \omega_{n+1}\right) \tag{A16}$$

is minimized. The variation of $\mathcal{G}$ with respect to β_i gives $\beta_0 = \hat{\beta}\Delta\omega$ and $\beta_i = -\hat{\beta}c_{gi,n+1}\Delta k_i$ for $i > 0$. Inserted in the dispersion relation, a non-linear equation for the Lagrangian

multiplier $\hat{\beta}$ results which is solved numerically via the Newton method. Therefore, in the two-stage scheme, the additional information gained by the ω-equation (2) is used to correct numerical errors and stabilize the implemented method.

Each time step, new ray points are injected at $\hat{z}_B = 20$ km and after a warming time of one day most of the model domain, in which GW propagation is possible, is filled with ray points. Furthermore, ray points are randomly removed when their number exceeds 32 in a grid box of the large-scale model.

All BG quantities are interpolated to the ray position via a linear polygonal interpolation. Furthermore, a distance-weighted interpolation and running median average is used to obtain smooth GW properties on the large-scale mesh. Especially, for the forcing term (9), the group velocity $\boldsymbol{c}_g$ is smoothly interpolated to the large-scale mesh. Derivatives of $\boldsymbol{c}_g$ were calculated using centered differences and τ_{non} is obtained via eq.(9). In the last step, τ_{non} is interpolated back to the ray position. Within this pragmatic approach, caustics appearing at ray crossings, are smoothed out and the corresponding wave action density remains finite, there. This might be interpreted as caustic correction, for which, to our current knowledge, no efficient method for full time-dependent 3D flows exists.

For ray integrations, no explicit test of WKB validity is performed. Only rays which cross the extreme thresholds of 100 km vertical wave length and 10 days intrinsic period are removed from the model run. As noted by *Sartelet* [2003], ray theory performs remarkably good even if the scale separation assumption is not fulfilled.

Acknowledgments. The authors like to thank Erich Becker for fruitful and inspiring discussions and Hauke Schmidt from MPI Hamburg for providing the set of HAMMONIA data. Furthermore, we thank three anonymous reviewers whose suggestions led to con-

siderable improvements. U.A. thanks Deutsche Forschungsgemeinschaft for partial support through the MetStröm Priority Research Program (SPP 1276), and through Grant Ac 71/4-1. U.A. and F.S. thank Deutsche Forschungsgemeinschaft for partial support through the CAWSES Priority Research Program (SPP 1176), and through Grant Ac 71/2-1.

References

Achatz, U., N. Grieger, and H. Schmidt (2008), Mechanisms controlling the diurnal solar tide: Analysis using a GCM and a linear model, *J. Geophys. Res.*, *113*, A08,303, doi: 10.1029/2007JA012967.

Achatz, U., R. Klein, and F. Senf (2010), Gravity waves, scale asymptotics and the pseudo-incompressible equations, *J. Fluid Mech.*, *663*, 120–147, doi: 10.1017/S0022112010003411.

Akmaev, R. A. (2001), Simulation of large-scale dynamics in the mesosphere and lower thermosphere with the Doppler-spread parameterization of gravity waves 2. Eddy mixing and the diurnal tide, *J. Geophys. Res.*, *106*, 1205–1213, doi:10.1029/2000JD900519.

Alexander, M., et al. (2010), Recent developments in gravity-wave effects in climate models and the global distribution of gravity-wave momentum flux from observations and models, *Quart. J. R. Met. Soc.*, *136*(650), 1103–1124.

Andrews, D. G., and M. E. McIntyre (1978), On wave-action and its relatives, *J. Fluid Mech.*, *89*, 647–664, doi:10.1017/S0022112078002785.

Andrews, D. G., J. R. Holton, and C. B. Leovy (1987), *Middle atmosphere dynamics.*, Academic Press, New York.

761 Becker, E., and G. Schmitz (2003), Climatological Effects of Orography and Land-Sea
762 Heating Contrasts on the Gravity Wave-Driven Circulation of the Mesosphere., *J. At-*
763 *mos. Sci.*, *60*, 103–118, doi:10.1175/1520-0469(2003)060.

764 Bretherton, F. P., and C. J. R. Garrett (1968), Wavetrains in Inhomogeneous Moving
765 Media, *Proc. R. Soc., London A*, *302*, 529–554, doi:10.1098/rspa.1968.0034.

766 Broutman, D. (1984), The focusing of short internal waves by an inertial wave, *Geophys.*
767 *Astrophys. Fluid Dyn.*, *30*, 199–225, doi:10.1080/03091928408222850.

768 Broutman, D., and W. R. Young (1986), On the interaction of small-scale oceanic
769 internal waves with near-inertial waves, *J. Fluid Mech.*, *166*, 341–358, doi:
770 10.1017/S0022112086000186.

771 Broutman, D., J. W. Rottman, and S. D. Eckermann (2004), Ray Methods for Inter-
772 nal Waves in the Atmosphere and Ocean, *Annu. Rev. Fluid Mech.*, *36*, 233–253, doi:
773 10.1146/annurev.fluid.36.050802.122022.

774 Bühler, O. (2009), *Waves and Mean Flows*, Cambrigde University Press, Cambridge, UK
775 ; New York.

776 Chapman, S., and R. Lindzen (1970), *Atmospheric tides. Thermal and gravitational*, D.
777 Reidel Publishing Company, Dordrecht, Holland.

778 Dunkerton, T. J. (1981), Wave Transience in a Compressible Atmosphere. Part I:
779 Transient Internal Wave, Mean-Flow Interaction., *J. Atmos. Sci.*, *38*, 281–297, doi:
780 10.1175/1520-0469(1981)038.

781 Dunkerton, T. J. (1982), Stochastic parameterization of gravity wave stresses, *J. Atmos.*
782 *Sci.*, *39*, 1711–1725, doi:10.1175/1520-0469(1982)039.

783 Dunkerton, T. J. (1984), Inertia-gravity waves in the stratosphere, *J. Atmos. Sci.*, *41*,
784 3396–3404, doi:10.1175/1520-0469(1984)041.
785 Dunkerton, T. J., and N. Butchart (1984), Propagation and selective transmission of inter-
786 nal gravity waves in a sudden warming, *J. Atmos. Sci.*, *41*, 1443–1460, doi:10.1175/1520-
787 0469(1984)041.
788 Eckermann, S. D. (1992), Ray-Tracing Simulation of the Global Propagation of Inertia
789 Gravity Waves Through the Zonally Averaged Middle Atmosphere, *J. Geophys. Res.*,
790 *97*(D14), 15,849–15,866, doi:10.1029/92JD01410.
791 Eckermann, S. D., and C. J. Marks (1996), An idealized ray model of gravity wave-tidal
792 interactions, *J. Geophys. Res.*, *101*(D16), 21,195–21,212, doi:10.1029/96JD01660.
793 Forbes, J. M., J. Gu, and S. Miyahara (1991), On the interactions between gravity waves
794 and the diurnal propagating tide, *Planet. Space Sci.*, *39*, 1249–1257, doi:10.1016/0032-
795 0633(91)90038-C.
796 Fritts, D. C. (1984), Gravity wave saturation in the middle atmosphere - A review of
797 theory and observations, *Rev. Geophys.*, *22*, 275–308, doi:10.1029/RG022i003p00275.
798 Fritts, D. C., and M. J. Alexander (2003), Gravity wave dynamics and effects in the
799 middle atmosphere, *Rev. Geophys.*, *41*, 1003, doi:10.1029/2001RG000106.
800 Grieger, N., G. Schmitz, and U. Achatz (2004), The dependence of the nonmigrating
801 diurnal tide in the mesosphere and lower thermosphere on stationary planetary waves,
802 *J. Atmos. Solar-Terr. Phys.*, *66*, 733–754, doi:10.1016/j.jastp.2004.01.022.
803 Grimshaw, R. (1975), Nonlinear internal gravity waves in a rotating fluid, *J. Fluid Mech.*,
804 *71*, 497–512, doi:10.1017/S0022112075002704.

Grimshaw, R. (1984), Wave action and wave-mean flow interaction, with application to stratified shear flows, *Annu. Rev. Fluid Mech.*, *16*, 11–44, doi: 10.1146/annurev.fl.16.010184.000303.

Hasha, A., O. Bühler, and J. Scinocca (2008), Gravity Wave Refraction by Three-Dimensionally Varying Winds and the Global Transport of Angular Momentum, *J. Atmos. Sci.*, *65*, 2892–2906, doi:10.1175/2007JAS2561.1.

Hayes, W. D. (1970), Kinematic Wave Theory, *Proc. R. Soc., London A*, *320*, 209–226, doi:10.1098/rspa.1970.0206.

Holton, J. R. (1982), The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere, *J. Atmos. Sci.*, *39*, 791–799, doi:10.1175/1520-0469(1982)039.

Lieberman, R. S., D. A. Ortland, D. M. Riggin, Q. Wu, and C. Jacobi (2010), Momentum budget of the migrating diurnal tide in the mesosphere and lower thermosphere, *J. Geophys. Res.*, *115*, D20,105, doi:10.1029/2009JD013684.

Lindzen, R. S. (1981), Turbulence and stress owing to gravity wave and tidal breakdown, *J. Geophys. Res.*, *86*, 9707–9714, doi:10.1029/JC086iC10p09707.

Lu, W., and D. C. Fritts (1993), Spectral estimates of gravity wave energy and momentum fluxes. Part 3: Gravity wave-tidal interactions, *J. Atmos. Sci.*, *50*, 3714–3727, doi: 10.1175/1520-0469(1993)050.

Marks, C. J., and S. D. Eckermann (1995), A Three-Dimensional Nonhydrostatic Ray-Tracing Model for Gravity Waves: Formulation and Preliminary Results for the Middle Atmosphere., *J. Atmos. Sci.*, *52*, 1959–1984, doi:10.1175/1520-0469(1995)052.

Mayr, H. G., J. G. Mengel, K. L. Chan, and H. S. Porter (1999), Seasonal variations and planetary wave modulation of diurnal tides influenced by gravity waves, *Adv. Space*

828 *Res.*, *24*, 1541–1544, doi:10.1016/S0273-1177(99)00877-7.

829 Mayr, H. G., J. G. Mengel, K. L. Chan, and H. S. Porter (2001), Mesosphere dynamics
830 with gravity wave forcing: Part I. Diurnal and semi-diurnal tides, *J. Atmos. Solar-Terr.*
831 *Phys.*, *63*, 1851–1864, doi:10.1016/S1364-6826(01)00056-6.

832 McLandress, C. (1997), Sensitivity studies using the Hines and Fritts gravity-wave drag
833 parameterizations, in *Gravity Wave Processes and their Parameterization in Global*
834 *Climate Models*, edited by K. Hamilton, pp. 245–256, Springer-Verlag, Berlin.

835 McLandress, C. (1998), On the importance of gravity waves in the middle atmosphere
836 and their parameterization in general circulation models., *J. Atmos. Solar-Terr. Phys.*,
837 *60*, 1357–1383, doi:10.1016/S1364-6826(98)00061-3.

838 McLandress, C. (2002), The Seasonal Variation of the Propagating Diurnal Tide in the
839 Mesosphere and Lower Thermosphere. Part I: The Role of Gravity Waves and Planetary
840 Waves., *J. Atmos. Sci.*, *59*, 893–906, doi:10.1175/1520-0469(2002)059.

841 Meyer, C. K. (1999), Gravity wave interactions with the diurnal propagating tide, *J.*
842 *Geophys. Res.*, *104*, 4223–4240, doi:10.1029/1998JD200089.

843 Miyahara, S., and J. Forbes (1991), Interactions between gravity waves and the diurnal
844 tide in the mesosphere and lower thermosphere, *J. Met. Soc. Japan*, *69*(5), 523–531.

845 Ortland, D. A. (2005a), A Study of the Global Structure of the Migrating Diurnal Tide Us-
846 ing Generalized Hough Modes., *J. Atmos. Sci.*, *62*, 2684–2702, doi:10.1175/JAS3501.1.

847 Ortland, D. A. (2005b), Generalized Hough Modes: The Structure of Damped Global-
848 Scale Waves Propagating on a Mean Flow with Horizontal and Vertical Shear., *J. Atmos.*
849 *Sci.*, *62*, 2674–2683, doi:10.1175/JAS3500.1.

Ortland, D. A., and M. J. Alexander (2006), Gravity wave influence on the global structure of the diurnal tide in the mesosphere and lower thermosphere, *J. Geophys. Res.*, *111*, A10S10, doi:10.1029/2005JA011467.

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

Richter, J. H., F. Sassi, and R. R. Garcia (2010), Toward a Physically Based Gravity Wave Source Parameterization in a General Circulation Model, *J. Atmos. Sci.*, *67*, 136–156, doi:10.1175/2009JAS3112.1.

Sartelet, K. N. (2003), Wave propagation inside an inertia wave. Part I: Role of time dependence and scale separation, *J. Atmos. Sci.*, *60*, 1433–1447, doi:10.1175/1520-0469(2003)060.

Schmidt, H., et al. (2006), The HAMMONIA Chemistry Climate Model: Sensitivity of the Mesopause Region to the 11-Year Solar Cycle and CO2 Doubling, *J. Climate*, *19*, 3903–3931, doi:10.1175/JCLI3829.1.

Song, I., and H. Chun (2008), A Lagrangian Spectral Parameterization of Gravity Wave Drag Induced by Cumulus Convection, *J. Atmos. Sci.*, *65*, 1204–1224, doi:10.1175/2007JAS2369.1.

Sonmor, L. J., and G. P. Klaassen (2000), Mechanisms of gravity wave focusing in the middle atmosphere, *J. Atmos. Sci.*, *57*, 493–510, doi:10.1175/1520-0469(2000)057.

Vadas, S. L., and D. C. Fritts (2005), Thermospheric responses to gravity waves: Influences of increasing viscosity and thermal diffusivity, *J. Geophys. Res.*, *110*, D15,103, doi:10.1029/2004JD005574.

873 Walterscheid, R. L. (2000), Propagation of small-scale gravity waves through large-scale
874 internal wave fields: Eikonal effects at low-frequency approximation critical levels, *J.*
875 *Geophys. Res.*, *105*(D14).

876 Watanabe, S., and S. Miyahara (2009), Quantification of the gravity wave forcing of the
877 migrating diurnal tide in a gravity wave–resolving general circulation model, *J. Geophys.*
878 *Res.*, *114*, D07,110, doi:10.1029/2008JD011218.

879 Yuan, T., C.-Y. She, D. A. Krueger, F. Sassi, R. Garcia, R. G. Roble, H.-L. Liu, and
880 H. Schmidt (2008), Climatology of mesopause region temperature, zonal wind, and
881 meridional wind over Fort Collins, Colorado (41°N, 105°W), and comparison with model
882 simulations, *J. Geophys. Res.*, *113*, D03,105, doi:10.1029/2007JD008697.

883 Zhong, L., L. J. Sonmor, A. H. Manson, and C. E. Meek (1995), The influence of time-
884 dependent wind on gravity-wave propagation in the middle atmosphere, *Ann. Geophys.*,
885 *13*, 375–394, doi:10.1007/s00585-995-0375-6.

#	α [deg]	L_h[km]	c_h[ms^{-1}]	F_h[10^{-3} J m^{-3}]
1	0	385	6.8	0.32
2	45	410	6.8	0.38
3	90	504	10.2	0.35
4	135	570	6.8	0.38
5	180	596	6.8	0.45
6	225	570	6.8	0.38
7	270	504	10.2	0.35
8	315	410	6.8	0.38
9	0	385	32.8	0.32
10	45	410	20.4	0.38
11	135	570	20.4	0.38
12	180	596	32.8	0.45
13	225	570	20.4	0.38
14	315	410	20.4	0.38

Table 1. 14 members of the GW ensemble used in the simulations. α denotes the azimuth angle which is zero towards the east and increases counter-clockwise, L_h and c_h are horizontal wave length and phase velocity in wave direction and F_h vertical flux of horizontal momentum at the lower boundary $\hat{z}_B$.

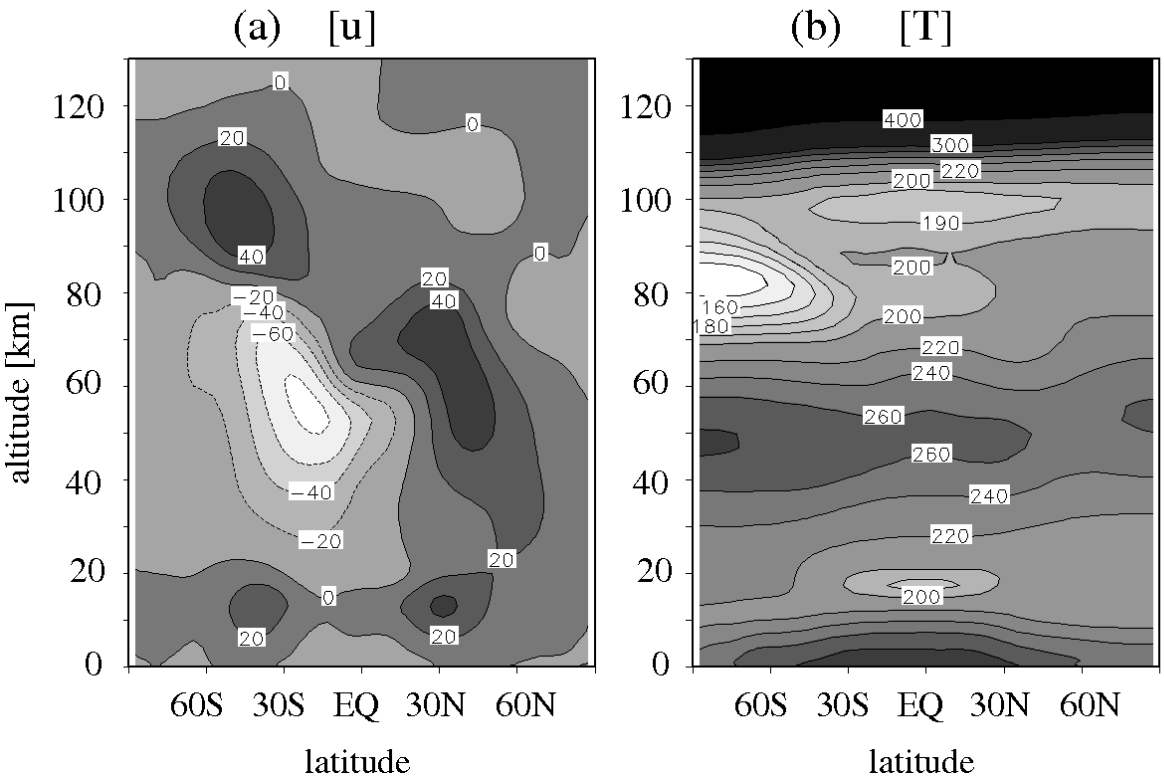

Figure 1. January mean zonally averaged zonal wind $[\bar{u}]$ in $\mathrm{m\,s^{-1}}$ (a) and temperature $[\bar{T}]$ in K (b) from HAMMONIA simulations.

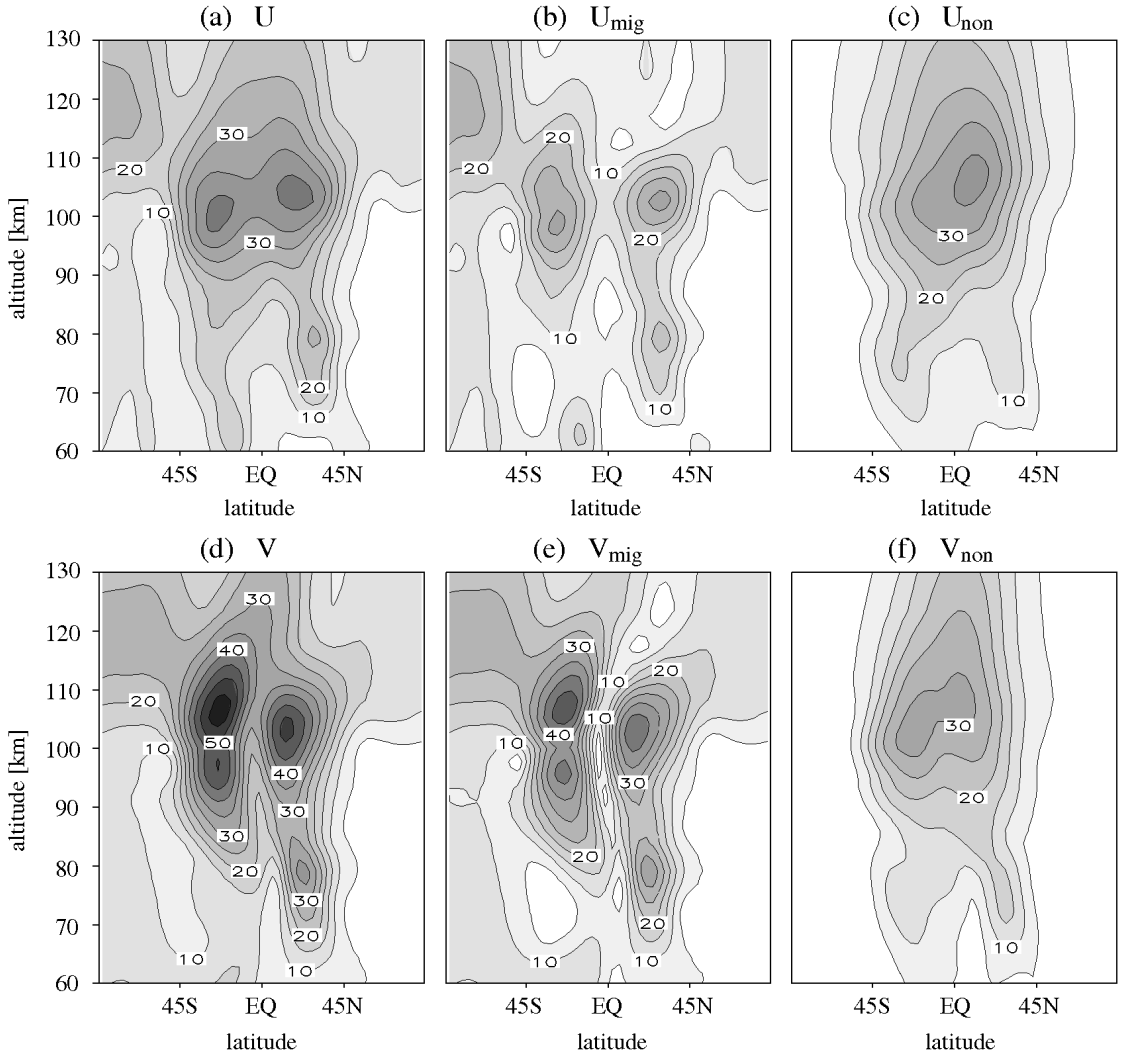

Figure 2. Amplitudes of diurnal variations of zonal wind U (a), (b), (c) and meridional wind V (d), (e), (f) in $\mathrm{m\,s^{-1}}$. Total tidal variations (a), (d) are decomposed in sun-synchronous migrating parts (b), (e) and the residual non-migrating parts (c), (f). Contours are in intervals of 5 $\mathrm{m\,s^{-1}}$.

dependence	"full"	"noREF"	"TS"
time	yes	yes	**no**
horizontal	yes	**no**	**no**

Table 2. Overview of three different experiments which have been performed with the same initial conditions and BG medium. "full", "noREF" and "TS" are the short-cuts of experiments explained in detail in the text.

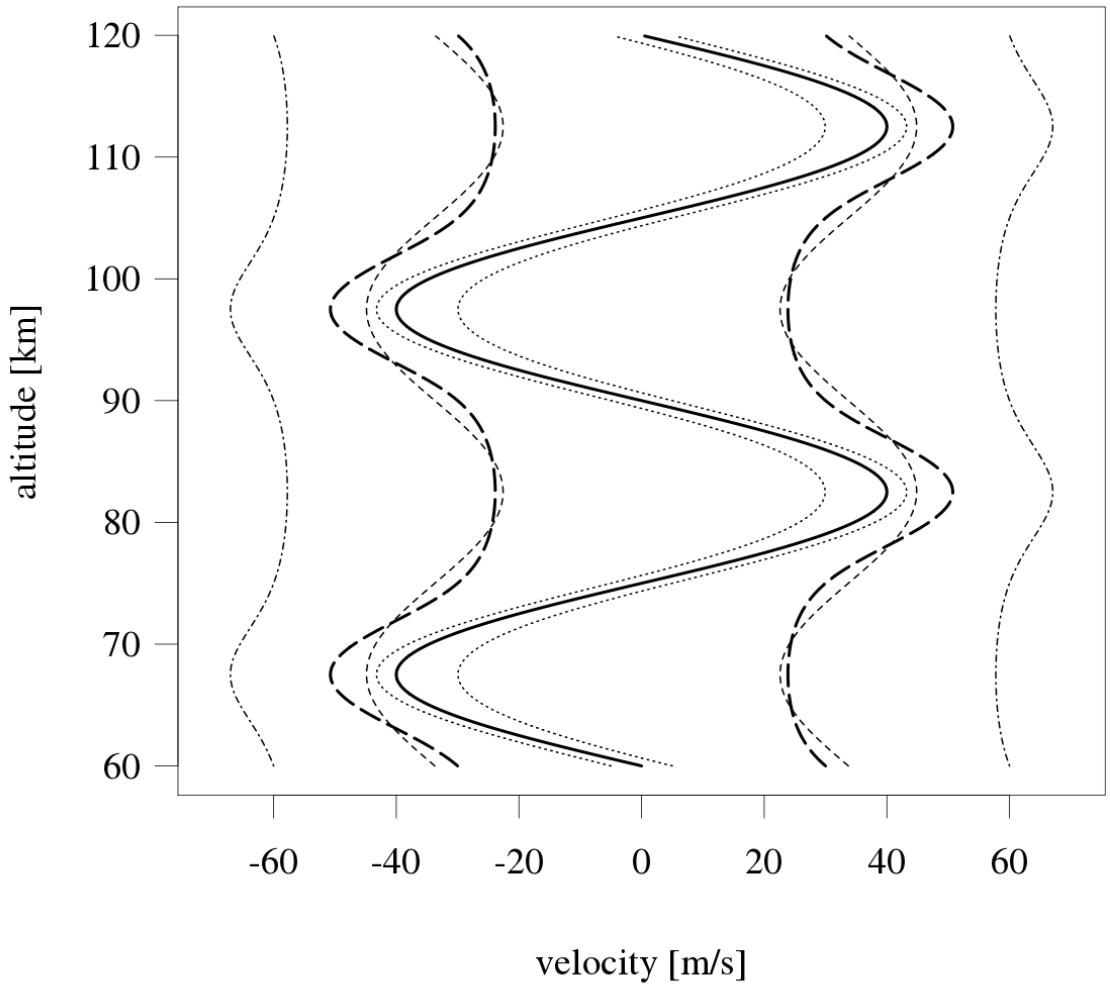

Figure 3. The vertical profiles of zonal wind (thick solid) with $U = 40\,\mathrm{ms}^{-1}$ and the exact solution for c of eq. (17) for 3 counter-propagating GW pairs: $c_0 = 5\,\mathrm{ms}^{-1}$ (dotted), $c_0 = 30\,\mathrm{ms}^{-1}$ (thick dashed) and $c_0 = 60\,\mathrm{ms}^{-1}$ (dot-dashed). The linear approximation (18) for $c_0 = 30\,\mathrm{ms}^{-1}$ (thin dashed) is also shown. In the calculations, the Coriolis effect has been neglected.

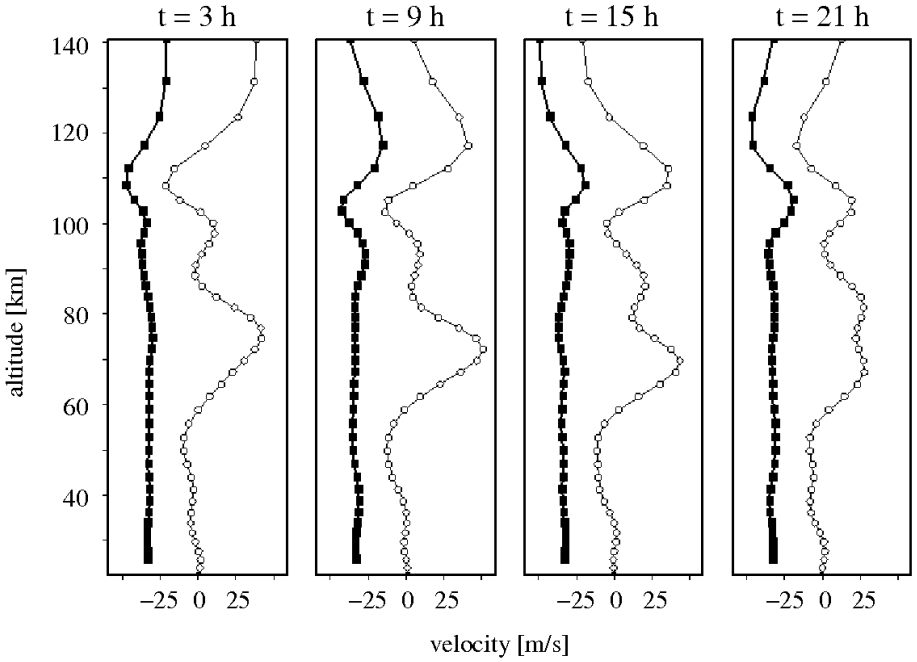

Figure 4. Vertical profiles of the zonal BG wind (open circles) and zonal GW phase velocity (filled squares) of the westward GW ensemble member 12 at $\lambda = 0°$ and $\varphi = 15°$N for the "noREF" experiment at four different times.

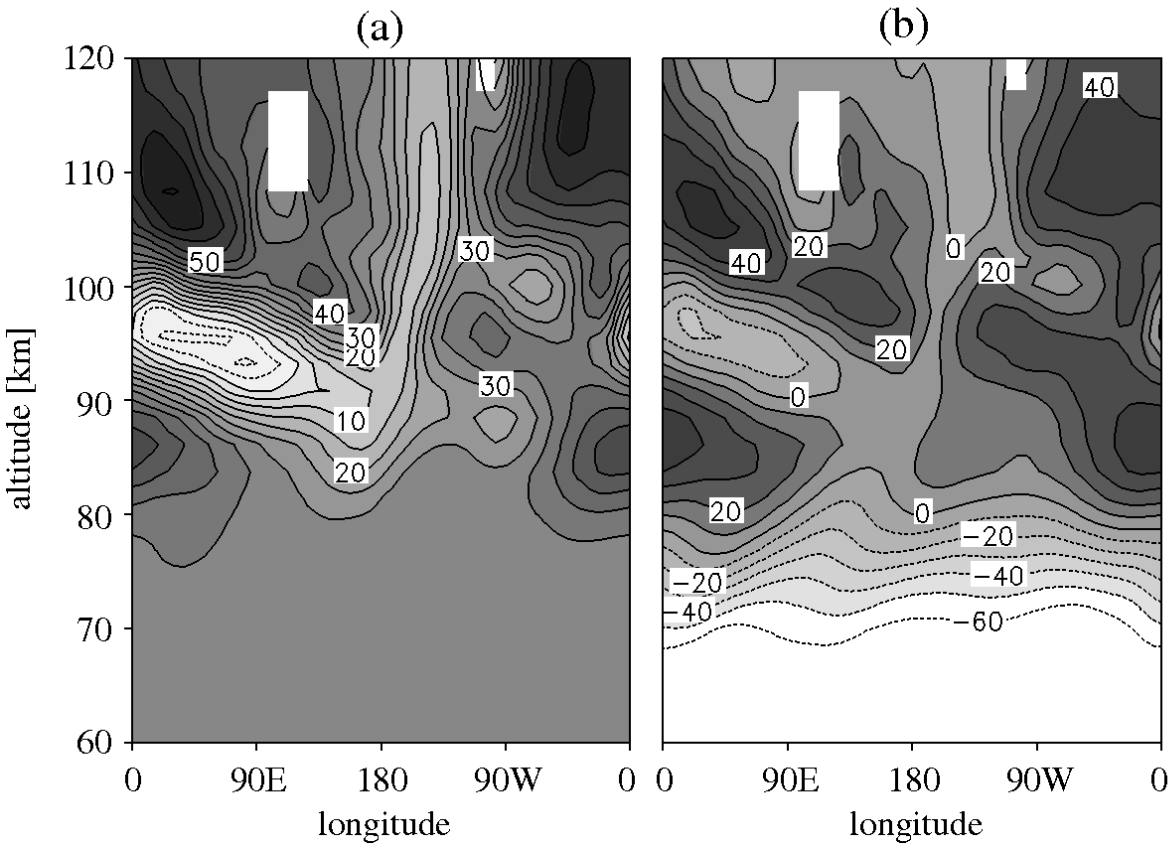

Figure 5. The horizontal phase velocity c_h (a) and the horizontal BG wind u_h in wave direction (b) at $t = 0$ of day 16 and $\varphi = 15°$S for the eastward propagating GW member 9 in the "full" experiment.

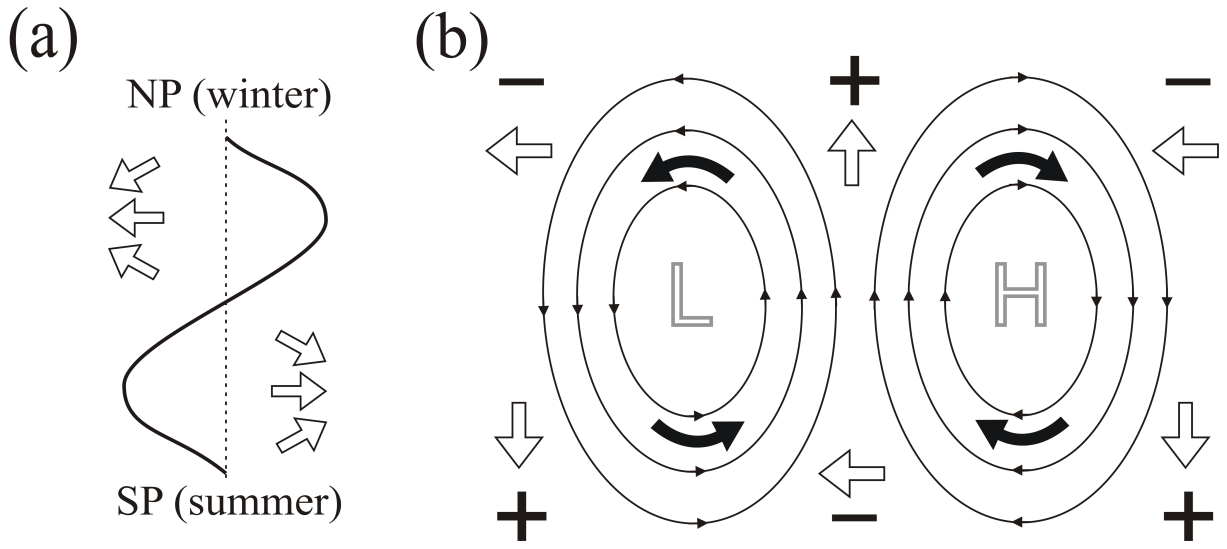

Figure 6. Schematic view on mean refraction of horizontal wave vector. Panel (a): $[\bar{u}]$ (solid line) induces refraction of $\boldsymbol{k}_h$ (open arrows) into the jet. Panel (b): Vorticity (bold arrows) and strain deformation (plus/minus signs) of a stationary planetary wave (solid streamlines) cause tendencies in k (horizontal open arrows) and l (vertical open arrows).

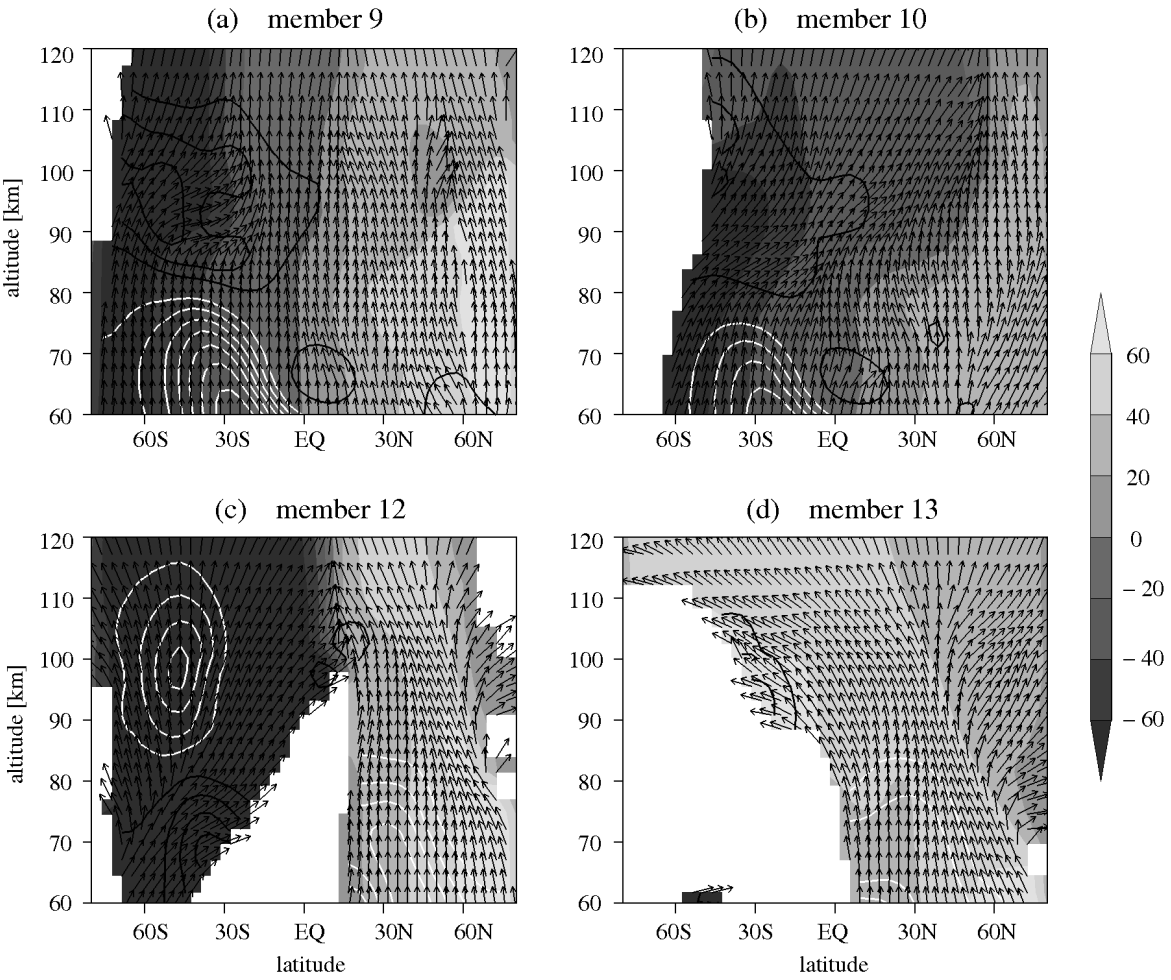

Figure 7. Meridional projection of the temporally and zonally averaged group velocities ($[\bar{c}_{g\varphi}]$, $100\,[\bar{c}_{gz}]$) (arrows) for several GW ensemble members 9 (a), 10 (b), 12 (c) and 13 (d). The horizontal BG wind in wave direction $[\bar{u}_h]$, is plotted in contours with an interval of 10 m/s, positive (black) and negative (white). The initial meridional position of the GW field at the lower boundary, i.e. at $\hat{z}_B = 20$ km, where the rays have been initialized, is shown using shadings with an interval of 20°.

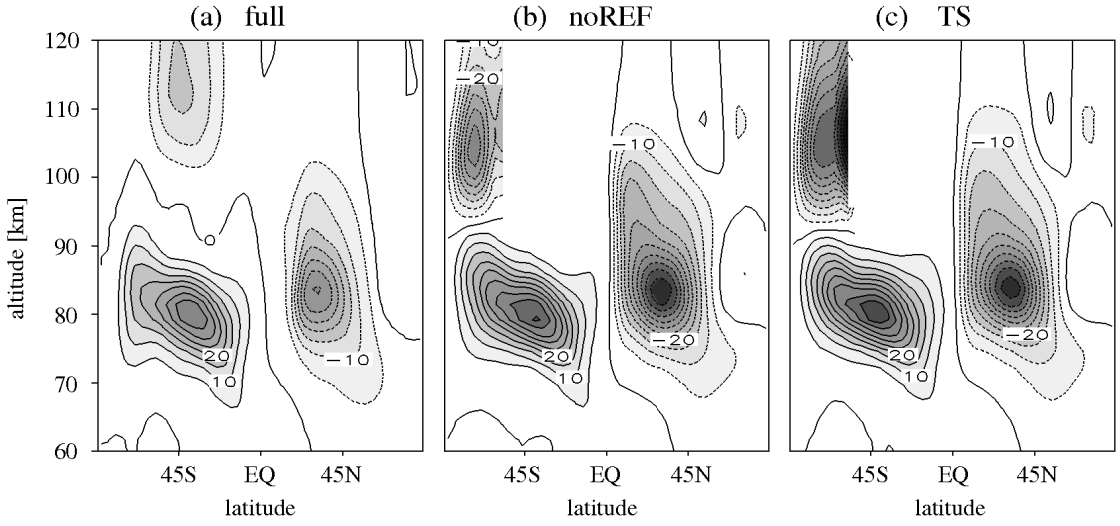

Figure 8. Temporally and zonally average zonal GW force f_λ in the "full" (a), "noREF" (b) and "TS" experiment (c). Contour interval is 5 $\mathrm{ms}^{-1}\,\mathrm{day}^{-1}$ and negative value with dashed lines.

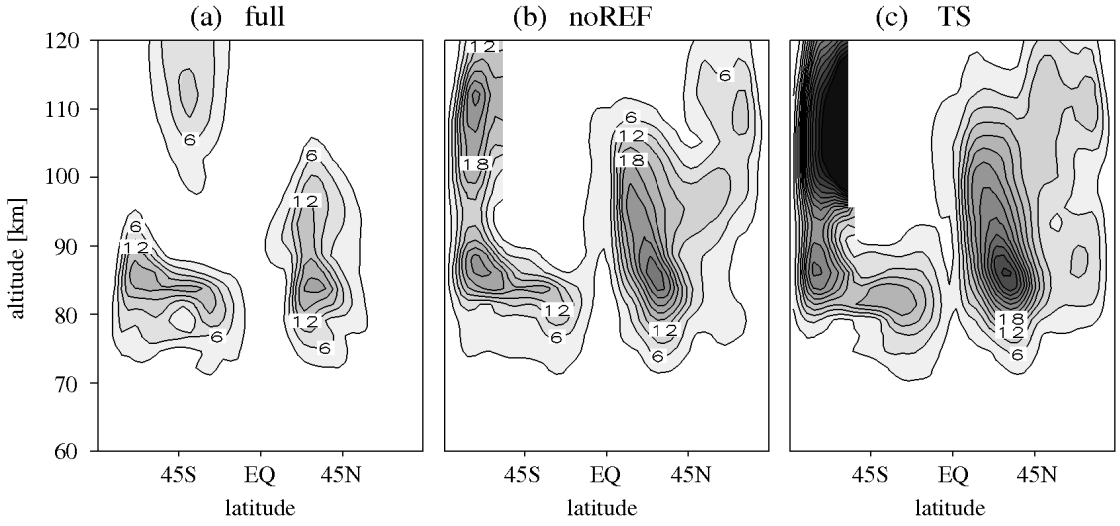

Figure 9. Diurnal amplitudes of the zonal GW force f_λ in $\mathrm{ms}^{-1}\,\mathrm{day}^{-1}$ for the "full" (a), "noREF" (b) and "TS" experiment (c). Contour interval is 3 $\mathrm{ms}^{-1}\,\mathrm{day}^{-1}$.

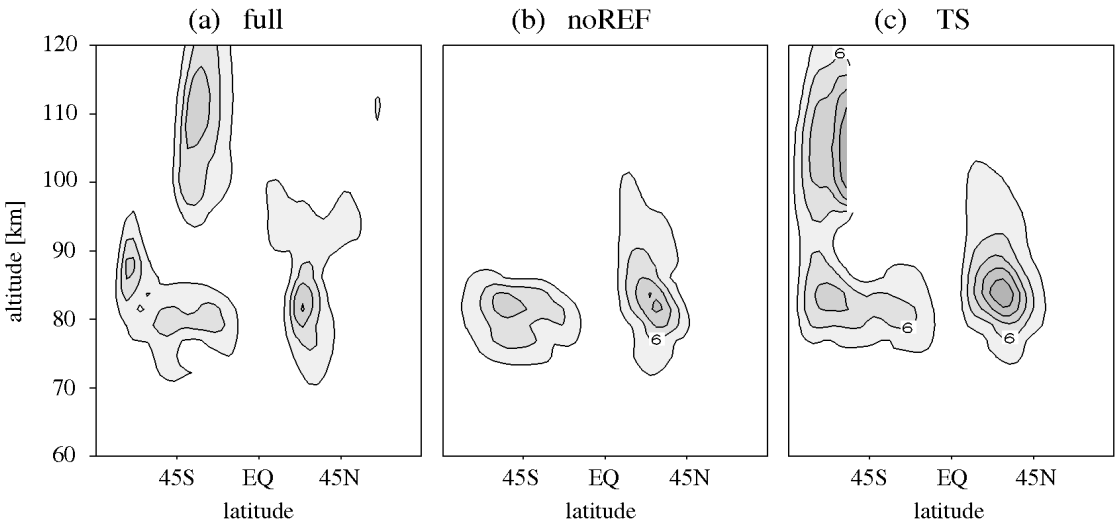

Figure 10. Same as fig.9, but for the amplitude of the meridional GW force on the diurnal tide.

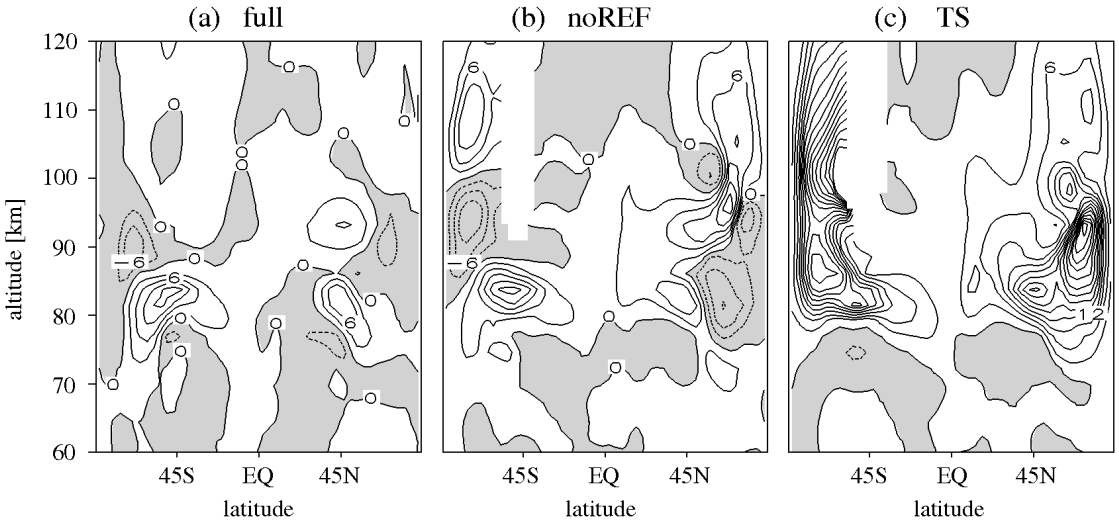

Figure 11. Real part of the equivalent Rayleigh friction coefficient of the zonal tidal wind for the "full" (a), "noREF" (b) and "TS" experiment (c) in $10^{-6}s^{-1}$ with an interval of $3 \times 10^{-6}s^{-1}$. Negative values are shaded.

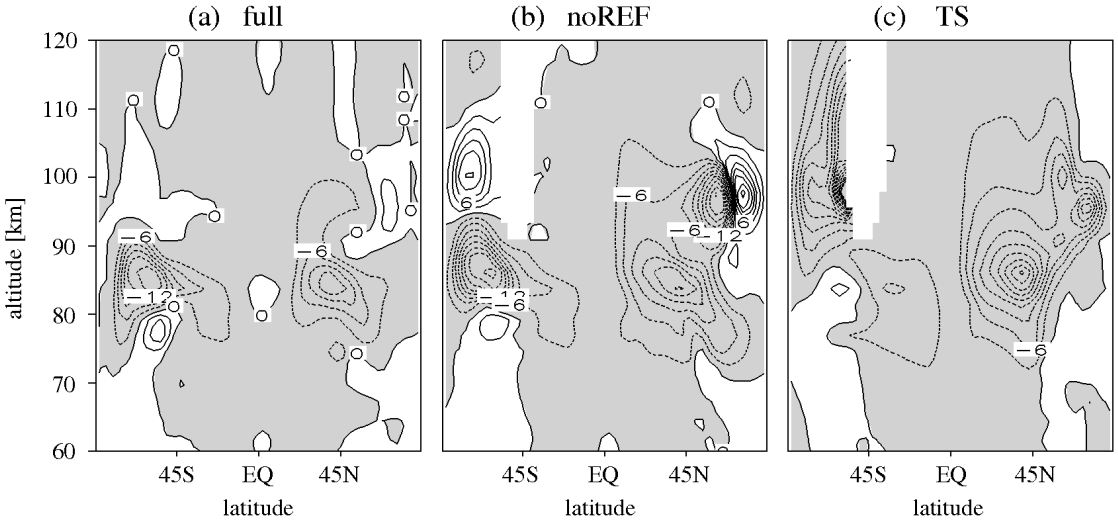

Figure 12. Same as fig.11, but for the imaginary part.