A new gravity-wave parameterization for atmospheric circulation models with full respect to transience and conservation laws

Dissertation
zur
Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Rostock

vorgelegt von:
Matthäus Mai
geboren am 01.03.1989
in Rostock

eingereicht: 19.07.2023
verteidigt: 23.02.2024

Gutachter:
Prof. Dr. Erich Becker, Boulder, NorthWest Research Associates
Prof. Dr. Riwal Plougonven, Palaiseau, École Polytechnique

Universität Rostock
Traditio et Innovatio

LEIBNIZ INSTITUTE OF ATMOSPHERIC PHYSICS IAP

https://doi.org/10.18453/rosdok_id00004656

Abstract

Gravity waves are the main driver of the summer-to-winter-pole circulation in the mesosphere and lower thermosphere. In particular, the gravity wave driven residual circulation gives rise to the cold summer mesopause and the warm winter polar stratopause. However, most of the gravity waves are not resolved in low-resolution global circulation models used for long-term simulations like climate simulations. But, because of their impact on the mesosphere lower thermosphere region, they must be parameterized and conventional parameterizations are based on strong assumptions, namely the single-column and steady state approximation. In this thesis, we present a new framework for gravity wave parameterizations, that relaxes the steady state approximation - the Radiative Transfer gravity wave Parameterization (RTP). This transient parameterization provides the local energy density of a wave field, from which the wave-mean flow interactions are calculated. Results of idealised simulations show that the wave field behaves as theoretically expected in different background configurations. Online simulations in a low-resolution global circulation model in January conditions provide adequate temperature and horizontal wind profiles. This all takes place in consideration of the conservation of energy and momentum. With these features the RTP will improve the circulation in the mesosphere and lower thermosphere of low-resolution global circulation models.

Zusammenfassung

Schwerewellen sind die treibende Kraft der residuellen Zirkulation vom Sommerpol zum Winterpol in der Mesosphäre und unteren Thermosphäre. Insbesondere führt die schwerewellengetriebene residuelle Zirkulation zu der kalten Mesopause am Sommerpol und der warmen Stratopause am Winterpol. Jedoch sind die meisten Schwerewellen zu klein, um in grob aufgelösten Zirkulationsmodellen, welche für Langzeitmodellrechnungen wie Klimasimulationen genutzt werden, aufgelöst zu sein. Aber auf Grund ihrer Bedeutung für die Mesosphäre und untere Thermosphäre müssen sie parametrisiert werden. Konventionelle Parametrisierungen basieren auf starke Vereinfachungen die Säulennäherung und Gleichgewichtszuständen. In dieser Arbeit präsentieren wir eine neue Schwerewellenparametrisierung, die nicht auf Gleichgewichtszuständen basiert - die Radiative Transfer gravity wave Parameterization (RTP). Diese transiente Parametrisierung liefert die lokale Energiedichte eines Wellenfeldes, mit deren Hilfe die Wellegrundstromwechselwirkungen berechnet werden können. Ergebnisse von vereinfachten Simulationen in verschiedenen Grundstromkonfigurationen zeigen, dass das Verhalten des Wellenfeldes theoretischen Vorhersagen entspricht. Simulationen mit einem grobgliedrigen globalen Zirkulationsmodell unter Januarbedingungen liefern angemessene Temperatur- und horizontale Windprofile unter Berücksichtigung der Energie- und Impluserhaltung. Mit diesen Eigenschaften wird die RTP dazu beitragen die Zirkulation in der Mesosphäre und unteren Thermosphäre von grob aufgelösten Zirkulationsmodellen zu verbessern.

Contents

Nomenclature

Acronyms

CIRA86	COSPAR International Reference Atmosphere 86
COSPAR	COmittee on SPAce Research
ESA	Effective Spectral Area
GCM	General Circulation Model
GVA	Gaussian Variational Approach
HIAMCM	HIgh Altitude Mechanistic Circulation Model
IDEMIX	Internal wave Dissipation, Energy and MIXing
KMCM	Kühlungsborn Mechanistic Circulation Model
MERRA-2	Modern-Era Retrospective analysis for Research and Applications, Version 2
MLT	Mesosphere Lower Thermosphere
RTP	Radiative Transfer gravity wave Parameterization
SSW	Sudden Stratospheric Warming
WKB	Wentzel Kramers Brillouin approximation

Greek Symbols

α	Saturation parameter
χ^2	Error function
$\dot{\omega}_I$	Modulation of the intrinsic frequency
Γ	(Dry) adiabatic lapse rate
γ	Interference parameter
λ	Wave length
ω	Ground-based frequency
ω_I	Intrinsic frequency
Φ	Geopotential
ϕ	Meridional angle
ρ	Density
ρ_0	Surface density
ρ_r	Hydrostatic reference density
Θ	Potential temperature
φ	Azimuth angle

Latin Letters

$\dot{m}$	Refraction of the vertical wavenumber
$\dot{z}$	Vertical component of the velocity vector, vertical group velocity
$\mathbf{D}$	Tensor of the damping rates for ion drag
$\mathbf{S}_h$	Horizontal strain tensor

$\mathcal{A}$	Wave action density
$\mathcal{D}$	Diffusion parameter
$\mathcal{E}$	Spectral energy density
$\mathcal{E}_0$	Physical energy density
$\mathcal{H}$	Enthalpy
$\mathcal{N}$	Normalization factor
$\mathcal{S}$	Source function
$\vec{c}_g$	Ground-based group velocity vector
$\vec{k}$	Wavenumber vector
$\vec{v}$	Wind vector
$\vec{x}$	Position vector
$\vec{x}_h$	Horizontal component of the position vector
b	Buoyancy
c_p	Heat capacity at constant pressure
c_{ph}	Horizontal phase velocity
$E_{BG,\,kin}$	Kinetic energy of the background
E_{dep}	Energy deposition
E_{GW}	Energy of the gravity waves
E_{tot}	Total energy
E_{tr}	Gravity wave energy density tendency
F	Vertical flux of horizontal momentum
f	Coriolis frequency

F_f	Frontogenesis function
F_p	Vertical flux of pressure
g	Gravitational acceleration on earth
H	Scale height
K	Diffusion coefficient
k, l, m	Components of the wavenumber vector
m'	Modified vertical wavenumber
m^*	Characteristic vertical wavenumber
M_{BG}	Background momentum
M_{dep}	Momentum deposition / gravity wave drag
M_{GW}	Momentum of the gravity waves
M_{tot}	Total momentum
M_{tr}	Gravity wave momentum density tendency
N	Brunt-Väisälä frequency
p	Pressure
p_{00}	Surface pressure
Q_c	Convective heating
R	Specific gas constant
T	Temperature
T_r	Hydrostatic reference temperature
U, V, W	Components of the background wind vector
u, v, w	Components of the wind vector

Mathematical symbols and operators

δX	Variation of X
ΔX	Difference between X_1 and X_0
$\dot{X}$	Total time derivative
$\langle X \rangle$	Average over gravity wave scales
∇	Nabla operator, $\frac{\partial}{\partial x}\vec{e}_x + \frac{\partial}{\partial y}\vec{e}_y + \frac{\partial}{\partial z}\vec{e}_z$
∂_X	Partial derivative, $\frac{\partial}{\partial X}$
$\Re(X^*)$	Real part of X^*
$\tilde{X}$	$\frac{X^*}{\rho_r}$
X'	Perturbation (except for m')
X^*	Complex valued gravity wave property (except for m^*)
X_{GW}	Gravity wave related interaction with X

Chapter I

Introduction

The dynamics and thermal structure of the Mesosphere and Lower Thermosphere (MLT) are inevitably connected to the breaking of gravity waves. These gravity waves are mostly generated in the lower atmosphere by different mechanisms and propagate upward into the MLT region. During the propagation the gravity waves become dynamically unstable due to the refraction by the background wind (Doppler shift) and the exponentially decreasing background density. This instability triggers the transition to turbulence (e.g., Fritts and Alexander 2003). In the saturation theory of Lindzen (1981), the turbulence is represented by a vertical diffusion coefficient that keeps the gravity wave at the limit of static stability, thereby giving rise to dissipation of wave energy. Here, the absorption of a gravity wave at a critical level is a special form of this mechanism. The breakdown of gravity waves from dynamic or static instability gives rise to the subsequent turbulent dissipation of wave energy and wave-mean flow interactions. These interactions are known as energy and momentum deposition (e.g., Becker and McLandress 2009). The mean-flow effects are crucial to understand the dynamical and thermal structure of the MLT (e.g., Lindzen 1981; Smith 2012; Garcia and Solomon 1985; McLandress et al. 2006).

In low resolution General Circulation Models (GCM) used for long term simulations the relevant gravity waves can not be resolved. The typical effective horizontal resolution of these GCMs corresponds to horizontal wavelength of about 1000 km (gridpoint resolution of a few hundred kilometers) while the horizontal wavelengths of small- and medium-scale gravity waves are about 50 km to 500 km. To take the unresolved gravity waves into account we need to apply model equations where the gravity wave effects are represented by some parameterization.

Current routinely used gravity wave parameterizations are based on several strong assumptions that define the framework of these parameterizations. These assumptions are the steady-state and single-column approximations, as well as the representation of the gravity wave field by a superposition or a continuous spectrum of monochromatic wave solutions (e.g., Garcia et al. 2007; Hines 1997a; Hines 1997b).

The aim of this thesis is to develop a transient parameterization that includes the continuous generation of gravity waves in the troposphere, the vertical propagation into the middle atmosphere, and the dissipation of the gravity wave field due to saturation and due to molecular diffusion above the mesopause. While the steady-state approximation and the assumption of monochromatic wave solutions are relaxed, we retain the single-column approximation for the sake of technical feasibility. On the other hand, our aim is that the new parameterization includes gravity wave sources that are continuous functions of height and time. This is made possible by using the radiative transfer equation of for the gravity wave field as a framework (Hasselmann 1968; Müller and Olbers 1975; Müller and Natarov 2003; Olbers and Eden 2013; Quinn et al. 2020; Olbers et al. 2023). Furthermore, we will for this framework derive the precisely energy and momentum conserving set of equations for the mean flow and the gravity wave field.

In the following sections we review the importance of gravity waves for the MLT region and the problems of conventional parameterizations used in community models. An outline of this thesis is given at the end of this chapter.

1.1 Gravity waves in the mesosphere and lower thermosphere

The momentum deposition from gravity waves drives the summer-to-winter pole residual circulation in the MLT region which is known to give rise to the cold summer mesopause and the warm winter stratopause (Lindzen 1981; Holton 1983; Smith 2012; Limpasuvan et al. 2012). These gravity waves are generated in the troposphere and lower stratosphere either by flow over orography or non-orographic processes. Non-orographic gravity wave sources are convection, shear instability of the jet stream, fronts, and spontaneous emission due to imbalance of the tropospheric and stratospheric jets (Fritts and Alexander 2003; Plougonven and Zhang 2014).

The turbulent dissipation of gravity waves in the middle atmosphere gives rise to wave-mean flow interactions. The summer-to-winter pole residual circulation driven by the momentum deposition from gravity waves becomes that strong that it causes a reversal of the pole-to-pole temperature gradient (Lindzen 1981; Holton 1983). Instead of a negative summer-to-winter pole temperature gradient for the radiatively determined state there is a positive temperature gradient in the middle atmosphere (e.g., see Becker 2012, Fig. 1). The polar summer mesopause is colder in the new state and the polar winter stratopause is warmer. In particular, the summer polar mesopause is the coldest place of the terrestrial atmosphere with temperatures of about 130 K on average. The strong pole-to-pole temperature gradient demands a strong vertical zonal wind gradient to balance the thermal wind relations. This causes in the end a zonal wind reversal in regions with strong temperature gradients.

Gravity waves also influence the stratosphere. For example, gravity waves help to constrain the polar vortex and contribute to the Semi Annual Oscillation (Garcia and Sassi 1999) and the Semi-Biannual Oscillation (Baldwin et al. 2001).

1.2 State-of-the-art gravity-wave parameterizations

The gravity-wave parameterizations that are currently used in community GCMs rely on strong assumptions (steady-state and single-column) that are often in contradiction to the effective resolutions applied in these models. As mentioned earlier, a conventional parameterization employs monochromatic wave solutions that are obtained with the assumption of a temporally constant and horizontally uniform background state. The monochromatic wave solutions are derived from WKB theory to obtain the dependence of the vertical wavenumber on the background vertical wind profile (also known as Doppler shifting or vertical refraction) (Plumb 1977; Lindzen 1981). The tunable parameters of these WKB solutions are the horizontal wavelengths, ground-based horizontal phase speeds, and amplitudes of the gravity waves. The parameters are often called launch-level parameters, because the parameterizations assume that the gravity waves are launched at a certain height (typically in the troposphere for most GCMs), and that the gravity waves are absent below that level.

The steady-state and single-columns approximations no longer hold if the resolved flow of the GCMs varies horizontally and temporally at the scales of the parameterized gravity waves. Note in this context that gravity waves in the real atmosphere appear in wave packets that are localized in space and time, and that also the generation processes of gravity waves show a high degree of intermittency (e.g., Geller et al. 2013; Vadas et al. 2018; Suzuki et al. 2013)

The issues associated with the steady-state approximation can be avoided with a transient gravity wave scheme that allows for spatially and temporally continuous generation of gravity waves, as well as as for the propagation of gravity wave packets in a spatially and temporally varying background flow. The framework for such a description of gravity waves is given by the radiative transfer equation for the wave field (Hasselmann 1968; Müller and Olbers 1975; Olbers and Eden 2013; Muraschko et al. 2015; Bölöni et al. 2016). The first transient gravity-wave parameterization for a GCM was described for Bölöni et al. (2021) and Kim et al. (2021). In that parameterization, wave transience was realized by describing the gravity waves with the help of ray tracing. The gravity wave sources still had to be represented by launch-levels parameters, the amplitudes of the generated gravity waves are vertically integrated over the whole source region and then gravity wave packet is launched at the uppermost source level. Another attempt by Quinn et al. (2020) was based on a simplified form of the radiative transfer equation, showing that this concept can be considered as a feasible alternative to the ray-tracing approach. Note in this context that using the ray-tracing equations as a frame work for a gravity-wave scheme in a GCM requires a permanent mapping between the Lagrangian space of the gravity-wave rays and the flow field in physical space (Senf 2012; Vadas and Liu 2013).

1.3 Thesis overview

In Chapter II we recapitulate the anelastic primitive equations and the wave-mean flow interaction terms in steady state. Furthermore we introduce the dispersion and polarization relations of hydrostatic, mid-frequency gravity waves. Additional wave-mean flow interaction terms that are required in the transient case are derived at the end of Chapter II. Chapter III presents the theory of the Radiative Transfer gravity wave Parameterization (RTP). The radiative transfer equation, which accommodates the generation, propagation, and dissipation of a wave field, is the basis. We apply the Gaussian Variational Principle to derive two prognostic equations for the free parameters of the assumed gravity wave spectrum. Furthermore, we derive a turbulent vertical diffusion coefficient based on the saturation hypothesis of Lindzen (1981). In Chapter IV we perform one-dimensional, transient simulations where a single wave packet is generated and propagates through different idealised vertical profiles of the background horizontal winds and temperature until it dissipates. We discuss the consistency of these results and check them for energy and momentum conservation. Chapter V presents results when the new RTP is implemented in a GCM run at low resolution such that the single-column approximation can considered to be valid. These results are compared to results from the corresponding high-resolution GCM with resolved gravity waves. Some new insights provided by this comparison are discussed. A summary of the main results of this thesis and an overview about future tasks regarding applications of the RTP are given in Chapter VI. The latter include the extension of the RTP such as to include horizontal and downward propagation of gravity waves, the combination of orographic and non-orographic gravity waves in one parameterization and the generation of higher order gravity waves.

Chapter II

Gravity waves and interactions with the mean-flow

In this chapter we introduce the fundamental properties of gravity waves and the full transient wave-mean flow interaction terms. Section 2.1 summarizes the dispersion and the polarization relations of gravity waves that fulfill the mid-frequency approximation. This approximation is usually made in conventional Gravity wave schemes. We shall apply this approximation as well when specifying our new transient gravity wave scheme. In the Sections 2.2 and 2.3, we derive the complete set of wave-mean flow interaction terms required for momentum and energy conservation in the case of a transient gravity wave parameterization in the single-column approximation.

2.1 Governing equations of gravity waves

Linear theory describes monochromatic gravity waves as small departures from a stably stratified and horizontally uniform background atmosphere. The restoring force is the buoyancy, which results from the isentropic displacements of air parcels. Hydrostatic gravity wave solutions can be computed from the Eulerian fluid dynamical equations linearized about a horizontally uniform horizontal flow using linear perturbation theory. Here we use the Eulerian equation in the anelastic approximation according to Becker (2017) in order to eliminate sound waves. As a further approximation, we neglect the Coriolis force for the gravity waves. This means that the gravity waves can be described in the x, z-plane, where x is a horizontal coordinate in direction of horizontal propagation and z is the height above sea level. Note that this approximation excludes so-called inertia gravity waves. The resulting linear wave equations can be written as

$$\begin{aligned} d_t u^* &= -\partial_x \tilde{p} && (2.1)\\ d_t w^* &= -\partial_z \tilde{p} + g\frac{T^*}{T_r} && (2.2)\\ 0 &= \partial_x u^* + \left(\partial_z - \frac{1}{H}\right) w^* && (2.3)\\ d_t \frac{T^*}{T_r} &= -\frac{N^2}{g} w^* && (2.4) \end{aligned}$$

with the total time derivative defined as $d_t = \partial_t + U\partial_x$. Here, U is the horizontally uniform mean wind in the x direction. The gravity wave perturbations are denoted with stars and the pressure is scaled with the hydrostatic reference density $\tilde{p} = p^*/\rho_r(z)$. The temperature $T_r(z)$ is the hydrostatic reference temperature. Within the hydrostatic approximation the scale height H and the squared Brunt-Väisälä frequency are defined as

$$H(z) = \frac{RT_r}{g}, \qquad (2.5)$$

$$N^2(z) = \frac{g}{T_r}\left(\frac{g}{c_p} + \partial_z T_r\right), \qquad (2.6)$$

with the specific gas constant for dry air R, the gravitational acceleration g and the specific heat capacity of constant pressure for dry air c_p. In the following we drop the explicit notation of the slow altitude dependence of the scale height and the Brunt-Väisälä frequency. The reference density $\rho_r(z)$ can be written in terms of the reference temperature by applying the barometric formula

$$\rho_r(z) = \rho_0 \exp\left(-\int_0^z \frac{g}{RT_r}\,dz'\right) = \rho_0 \exp\left(-\int_0^z \frac{1}{H}\,dz'\right) \tag{2.7}$$

with the reference density of the air at surface level ρ_0. With the ansatz

$$X^* = X_0 \exp\left[\frac{1}{2}\int_0^z \frac{1}{H}\,dz' + \mathrm{i}(kx + mz - \omega t)\right] \tag{2.8}$$

for the gravity wave perturbations, the intrinsic frequency $\omega_I = \omega - kU$, and the Boussinesq approximation for the waves $\rho^*/\rho_r = -T^*/T_r$, we get from the set of equations (2.1)-(2.4)

$$\begin{aligned}
0 &= -\mathrm{i}\omega_I u^* + \mathrm{i}k\tilde{p} && (2.9)\\
0 &= -\mathrm{i}\omega_I w^* + \left(\mathrm{i}m + \frac{1}{2H}\right)\tilde{p} + g\frac{\rho^*}{\rho_r} && (2.10)\\
0 &= \mathrm{i}ku^* + \left(\mathrm{i}m - \frac{1}{2H}\right)w^* && (2.11)\\
0 &= -\mathrm{i}\omega_I \frac{\rho^*}{\rho_r} + \frac{N^2}{g}w^*, && (2.12)
\end{aligned}$$

Our sign convention is that the intrinsic frequency is always positive. Then, the vertical wavenumber m is positive (negative) for upward (downward) propagation of gravity wave phase lines in a framework moving with the mean flow (intrinsic framework). In the same way, the horizontal wavenumber k is positive (negative) for a wave that propagates in positive (negative) x-direction in the intrinsic frame work. Note that the vertical group velocity of a gravity wave is upward when the intrinsic phase propagation is downward. With this sign convention, the ground-based frequency ω is positive definite as long as the background wind in the direction of k is not faster than the phase speed ω/k.

The non-trivial solution of the determinant of the equation system (2.9)-(2.12) yields the dispersion relation for gravity waves that are not affected by the Coriolis force (also known as mid-frequency and high-frequency gravity waves):

$$\omega_I^2 = \frac{k^2 N^2}{k^2 + m^2 + \frac{1}{4H^2}}. \tag{2.13}$$

Here and in the following we drop the explicit notation of the slow altitude dependence of the scale height and the Brunt-Väisälä frequency. With the Boussinesq limit we assume for the vertical wavenumber that $|m| > 1/(2H)$. For wavenumbers equal or smaller the anelastic dispersion relation would have unrealistic imaginary branches (Fritts and Alexander 2003; Becker 2019; Vadas 2013). Waves with smaller absolute vertical wavenumbers are compressible and the anelastic equations can not be applied for them.

Assuming the hydrostatic approximation ($\partial_z \tilde{p} = -g\tilde{\rho}$), the set of equations (2.9)-(2.12) reduces to

$$\begin{aligned} 0 &= -\mathrm{i}\omega_I u^* + \mathrm{i}k\tilde{p} && (2.14)\\ 0 &= \mathrm{i}ku^* + \left(\mathrm{i}m - \frac{1}{2H}\right) w^* && (2.15)\\ 0 &= +\mathrm{i}\omega_I \frac{1}{g}\left(\mathrm{i}m + \frac{1}{2H}\right)\tilde{p} + \frac{N^2}{g} w^* && (2.16) \end{aligned}$$

and results in the dispersion relation for mid-frequency gravity waves:

$$\omega_I^2 = \frac{k^2 N^2}{m^2 + \frac{1}{4H^2}}. \tag{2.17}$$

The polarisation relations for mid-frequency gravity waves are

$$\begin{aligned} w^* &= \frac{\omega_I}{N}\frac{m - \frac{\mathrm{i}}{2H}}{m'} u^* && (2.18)\\ p^* &= -\rho_r \frac{N}{m'} u^* && (2.19)\\ \rho^* &= \frac{\rho_r}{g}\left[\frac{N}{m'}\frac{1}{2H} + \mathrm{i}N\frac{m}{m'}\right] u^* && (2.20) \end{aligned}$$

with $m' = -\sqrt{m^2 + \frac{1}{4H^2}}$. For the last polarization relation we assumed that $\partial_z(N/m') \to 0$ due to the slow dependence of the background state on z.

Ray-tracing equations

The ray equations describes the ray path $\vec{x}$ and the refraction of the wavenumber vector $\vec{k}$ along the ray (Marks and Eckermann 1995). For gravity waves with the dispersion relation $\omega = \omega(\vec{k}, \vec{x})$ the equations are

$$\dot{\vec{x}} = \frac{\partial \omega}{\partial \vec{k}} \tag{2.21}$$

$$\dot{\vec{k}} = -\frac{\partial \omega}{\partial \vec{x}}. \tag{2.22}$$

Here it is assumed that the background and the wave parameter are locally constant with time. This leads to a constant ground-based frequency along the ray ($\dot{\omega} = 0$). Note that $\dot{\vec{x}}$ is equal the ground-based group velocity $\vec{c}_g$.

The dispersion relation of the ground-based frequency of hydrostatic and mid-frequency gravity waves in single column in the x, z-plane is (see also Eq. 2.17)

$$\omega = \pm \frac{kN}{\sqrt{m^2 + \frac{1}{4H^2}}} + kU, \tag{2.23}$$

where we have the plus sign for positive horizontal wavenumbers k and the minus sign for negative ones. The vertical group velocity is

$$\dot{z} = \frac{\partial \omega}{\partial m} = -\frac{m}{m'^2}\omega_I. \tag{2.24}$$

For the case of zero vertical background wind the intrinsic and the ground-based vertical group velocities are identical. The refraction of the vertical wavenumber m is

$$\dot{m} = -\frac{\partial \omega}{\partial U}\frac{\partial U}{\partial z} - \frac{\partial \omega}{\partial N}\frac{\partial N}{\partial z} = \frac{m'\omega_I}{N}\partial_z U - \frac{\omega_I}{N}\partial_z N. \tag{2.25}$$

Here we neglect the vertical variability of the scale height H. This variability is rather small and neglected for the sake of comprehensibility of the formulation and the routinely model code. The modulation of the intrinsic frequency ω_I is

$$\dot{\omega}_I = \dot{\omega} - k\dot{U} = -\frac{m}{m'}\frac{\omega_I^2}{N}\partial_z U, \tag{2.26}$$

where we used the assumption of the locally time independent background ($\partial_t U = 0$).

2.2 Momentum deposition

In the following we compute the momentum deposition in the general case where the conventional steady-state assumption for parameterized gravity waves is relaxed. As will be shown, this leads to an additional term that, to the best of our knowledge, has not been considered in the literature so far.

For the full transient momentum deposition we need to consider the wave momentum, which is lost in the anelastic approximation of the horizontal momentum equation due to neglecting the density perturbation ρ'. To derive the wave momentum we start with the horizontal momentum equation in a inviscid compressible atmosphere in the x, z-plane and neglect Coriolis force

$$\partial_t(\rho u) + \partial_x\left(\rho u^2\right) + \partial_z(\rho u w) + \partial_x p = 0. \tag{2.27}$$

The horizontal wind and the density are separated mean flow variables and perturbations:

$$\begin{aligned} \rho &= \rho_r + \rho' && (2.28)\\ u &= U + u' && (2.29)\\ w &= w' && (2.30) \end{aligned}$$

For the vertical wind we assume that the mean value W is zero. Inserting the relations (2.28)-(2.30) into equation (2.27) and averaging over the gravity-wave scales (denoted by $\langle ... \rangle$) and retaining only second order terms related to the gravity wave perturbations leads to

$$\begin{aligned}&\partial_t\rho_r U + \partial_t\langle\rho' u'\rangle\\ &+\partial_x\rho_r UU + \partial_x\langle\rho_r u'u'\rangle + 2\partial_x\langle\rho' u' U\rangle\\ &+\partial_z\langle\rho_r u'w'\rangle + \partial_z\langle\rho' U w'\rangle + \partial_x p = 0\end{aligned} \tag{2.31}$$

where only the averages over the product of two perturbations are non zero. After applying the single column approximation we get

$$\partial_t\left(\rho_r U\right) = -\partial_t\langle\rho' u'\rangle - \partial_z\langle\rho_r u'w'\rangle - \partial_z\langle\rho' U w'\rangle. \tag{2.32}$$

On the right hand side we have the transient effect, which is the tendency of the wave momentum density $M_{GW} = \langle\rho' u'\rangle$, the well known gravity wave drag and another term, which is U times the vertical mass flux density owing to gravity waves. The tendency of the wave momentum density is related to an elastic term which arises from density fluctuations if the Coriolis force is considered (see eq. 187 in Achatz et al. 2017). This elastic term too can not be derived from the anelastic theory and is discussed in Achatz et al. (2017) and Wei et al. (2019).

The wave momentum is only nonzero when the scale height H is considered in the gravity wave equations. By looking at the polarisation relation (2.20) we can see, that without the scale height the density and the horizontal wind perturbations have a phase shift of $\pi/2$ and so the wave momentum $\langle\rho' u'\rangle = \langle\Re(\rho^*)\Re(u^*)\rangle$ is always zero.

Scale analyses reveals that we can neglect the mass flux density term rightaway. The estimations are for the scale height $1/(2H) \approx \mathcal{O}\left(10^{-4}\,\mathrm{m}^{-1}\right)$, the Brunt-Väisälä frequency $N \approx \mathcal{O}\left(10^{-2}\,\mathrm{s}^{-1}\right)$, the vertical scale $z \approx 2\pi/m$, the time scale $t \approx 2\pi/\omega_I$, the gravitational acceleration $g \approx \mathcal{O}\left(10^{1}\,\mathrm{ms}^{-2}\right)$, and the horizontal wind $U \approx \mathcal{O}\left(10^{1}\,\mathrm{ms}^{-1}\right)$. For gravity waves with $\omega_I \approx$

$\mathcal{O}\left(10^{-3}\,\mathrm{s}^{-1}\right)$ and $m \approx \mathcal{O}\left(10^{-2}\,\mathrm{m}^{-1}\right)$ we get

$$\frac{\partial_t\left(\rho' u'\right)}{\partial_z\left(\rho_r u' w'\right)} \approx \mathcal{O}\left(10^{-3}\right) \tag{2.33}$$

$$\frac{\partial_z\left(\rho' U w'\right)}{\partial_z\left(\rho_r u' w'\right)} \approx \mathcal{O}\left(10^{-4}\right). \tag{2.34}$$

The only dominant term is the gravity wave drag and the other two terms can be neglected. For gravity waves with a vertical wavenumber close to the inverse scale height ($m \approx \mathcal{O}\left(10^{-4}\,\mathrm{m}^{-1}\right)$) we get

$$\frac{\partial_t\left(\rho' u'\right)}{\partial_z\left(\rho_r u' w'\right)} \approx \mathcal{O}\left(10^{-1}\right) \tag{2.35}$$

$$\frac{\partial_z\left(\rho' U w'\right)}{\partial_z\left(\rho_r u' w'\right)} \approx \mathcal{O}\left(10^{-4}\right). \tag{2.36}$$

Here the wave momentum term is only one order smaller than the drag and should not be neglected rightaway. The density flux term otherwise is still orders of magnitude smaller and can be neglected. The resulting wave-mean flow interaction for the horizontal flow is

$$(\partial_t U)_{GW} = -\frac{1}{\rho_r}\partial_z\langle\rho_r u' w'\rangle - \frac{1}{\rho_r}\partial_t\langle\rho' u'\rangle. \tag{2.37}$$

For unambiguousness we define the gravity wave drag or classic momentum deposition as M_{dep} and the transient wave momentum density tendency as M_{tr}:

$$(\partial_t U)_{GW} = M_{dep} + M_{tr}. \tag{2.38}$$

The importance of the momentum density term is the largest compressible gravity waves with large vertical wavelengths. We restrict ourself to anelastic gravity waves where the influence of the momentum density term might be small, but we keep it to investigate its effects.

2.3 Energy deposition

To derive the wave-mean flow interaction for the thermodynamic equation we recapitulate the discussion of Becker (2017). The thermodynamic prognostic variable used in GCMs is usually the enthalpy per unit mass. This is also the case for the Kühlungsborn Mechanistic Circulation Model (KMCM) (see Chap. V). In the anelastic approximation the thermodynamic equation is

$$c_p d_t T = -wg\left(1 - \frac{\tilde{\rho}}{\rho_r}\right), \tag{2.39}$$

where $\tilde{\rho}$ is the deviation of the density from the hydrostatic reference density, that is $\rho = \rho_r(z) + \tilde{\rho}$. Equation (2.39) is extended by radiative and latent heating, turbulent and molecular diffusion of heat, turbulent and molecular frictional heating, as well as frictional heating to due ion drag. Here we neglect the radiative and latent heating since these terms are unimportant for the wave-mean flow interaction. Then, averaging the enthalpy equation over gravity-wave scales The average over the gravity wave scales (denoted by $\langle ... \rangle$). in the single column approximation leads to

$$\begin{aligned} c_p d_t T = & -wg\left(1 - \frac{\tilde{\rho}}{\rho_r}\right) \\ & + (K_z + \nu)(\partial_z \mathbf{v})^2 + K_h |\mathbf{S}_h|^2 + \mathbf{v} \cdot (\mathbf{D}\mathbf{v}) \\ & + (K_z + \nu)\langle(\partial_z \mathbf{v}')^2\rangle + K_h \langle|\mathbf{S}_h'|^2\rangle + \langle \mathbf{v}' \cdot (\mathbf{D}\mathbf{v}')\rangle - \frac{g}{T_r}\langle T'w' \rangle. \end{aligned} \tag{2.40}$$

Here $\mathbf{v}$ is the horizontal wind field, K_h and K_z are the horizontal and vertical turbulent diffusion coefficients, which are assumed to be mean-flow quantities, and ν is the molecular diffusion coefficient. The tensor $\mathbf{D}$ contains the damping rates for ion drag (Liu and Yeh 1969) and $\mathbf{S}_h$ is the horizontal strain tensor. The second row on the right hand side of equation (2.40) represents the frictional heating due to momentum diffusion and ion drag of the resolved flow. The last row includes the frictional heating due to the unresolved gravity waves and the negative buoyancy production of gravity wave kinetic energy. Due to the quadratic nature of the frictional heating in the horizontal velocity field, there are contributions from the resolved flow and unresolved gravity wave perturbations. The turbulent frictional heating represents the molecular frictional heating in the real atmosphere. It occurs as a result from gravity wave breaking and the so-induced forward kinetic energy cascade. In the thermosphere the molecular viscosity becomes important and give rise to direct damping of gravity waves (Vadas and Fritts 2005; Vadas 2007; Vadas and Liu 2009). The kinetic energy of the gravity waves is another prognostic variable and needs to be evaluated additionally. The prognostic equation for the gravity wave kinetic energy can be written as (Becker 2017):

$$\begin{aligned}\partial_t \frac{\langle \mathbf{v}'^2 \rangle}{2} = & - \langle \mathbf{v}' w' \rangle \partial_z \mathbf{v} - \frac{1}{\rho_r} \partial_z \langle p' w' \rangle + \frac{g}{T_r} \langle T' w' \rangle \\ & - (K_z + \nu) \langle (\partial_z \mathbf{v}')^2 \rangle - K_h \langle |\mathbf{S}'_h|^2 \rangle - \langle \mathbf{v}' \cdot (\mathbf{D} \mathbf{v}') \rangle .\end{aligned} \tag{2.41}$$

For a conventional steady state gravity wave parameterization the left hand side is assumed to be zero. This assumption must be relaxed for a transient gravity wave parameterization. In this general case we can eliminate the frictional heating and the buoyancy production using

$$\begin{aligned}(K_z + \nu) \langle (\partial_z \mathbf{v}')^2 \rangle + K_h \langle |\mathbf{S}'_h|^2 \rangle + \langle \mathbf{v}' \cdot (\mathbf{D} \mathbf{v}') \rangle - \frac{g}{T_r} \langle T' w' \rangle = & \\ - \frac{1}{\rho_r} \partial_z \langle p' w' \rangle - \langle \mathbf{v}' w' \rangle \partial_z \mathbf{v} - \partial_t \frac{\langle \mathbf{v}'^2 \rangle}{2} . &\end{aligned} \tag{2.42}$$

The first two terms on the right hand side are the well known energy deposition (Hines 1997a)

$$E_{dep} = - \frac{1}{\rho_r} \partial_z \langle p' w' \rangle - \langle \mathbf{v}' w' \rangle \partial_z \mathbf{v} \tag{2.43}$$

with the convergence of vertical pressure flux and the vertical momentum flux times the vertical shear of the horizontal winds. The last term on the right hand side is the transient energy density tendency

$$E_{tr} = -\partial_t \frac{\langle \mathbf{v}'^2 \rangle}{2}. \tag{2.44}$$

With these substitutions the thermodynamic equation (2.40) becomes

$$\begin{aligned} c_p d_t T = & -wg\left(1 - \frac{\tilde{\rho}}{\rho_r}\right) \\ & + (K_z + \nu)(\partial_z \mathbf{v})^2 + K_h |\mathbf{S}_h|^2 + \mathbf{v} \cdot (\mathbf{D}\mathbf{v}) \\ & + E_{dep} + E_{tr}. \end{aligned} \tag{2.45}$$

To highlight only the wave-mean flow interaction in the x, z-plane we can write:

$$(c_p d_t T)_{GW} = E_{dep} + E_{tr} = -\frac{1}{\rho_r}\partial_z \langle p'w' \rangle - \langle u'w' \rangle \partial_z U - \partial_t \frac{\langle u'u' \rangle}{2} \tag{2.46}$$

Chapter III

Radiative Transfer gravity wave Parameterization

In this chapter the theory of the Radiative Transfer gravity wave Parameterization (RTP) is presented. In the first section the radiative transfer equation is introduced, followed by the introduction of the Desaubies spectrum as a two-parameter representation of the wave spectrum in Section 3.2. To achieve prognostic equations for these parameters the Gaussian Variational Approach (GVA) is applied in Section 3.3. The calculation of the wave-mean flow interactions from the prognostic variables of the wave spectrum is part of Section 3.4. Section 3.7 specifies the different spectral areas of the wave field included in the RTP and GVA, as well as in the wave-mean flow interactions. Furthermore we present the diffusion scheme in Section 3.5 and the general form of the source functions used in this study in Section 3.6.

3.1 Radiative transfer equation of the gravity wave field

In the year 2003 Müller and Natarov proposed an internal gravity wave parameterization for the ocean, which was based on the radiative transfer equation for weakly interacting waves in a six dimensional phase space (Müller and Natarov 2003). However their approach has not been successful. Olbers and Eden (2013) simplified the radiative transfer equation by assuming a specific spectral form and integrating over all wavenumbers. Furthermore they assumed horizontal homogeneity, which leads to a partial differential equation for the energy of a vertical propagating oceanic wave field and its dissipation. This simplified model is called IDEMIX and was recently transferred to the atmosphere (Quinn et al. 2020).

The radiative transfer equation has the form (Müller and Olbers 1975; Hasselmann 1968)

$$\left(\partial_t + \dot{\vec{x}}\partial_{\vec{x}} + \dot{\vec{k}}\partial_{\vec{k}}\right)\mathcal{A}(\vec{k},\vec{x},t) = \mathcal{S}_{gen} + \mathcal{S}_{trans} + \mathcal{S}_{diss}, \tag{3.1}$$

with the wavenumber vector $\vec{k}$ and the position vector $\vec{x}$. The change of the wave action density spectrum $\mathcal{A}(\vec{k},\vec{x},t)$ along wave trajectories is described on the left hand side, where $\dot{\vec{x}}$ is the group velocity and $\dot{\vec{k}}$ the rate of refraction. On the right hand side are terms that give rise to non-conservative wave propagation. The generation of wave action is described by $\mathcal{S}_{gen}$, the transfer of wave action within the spectrum by $\mathcal{S}_{trans}$, and the dissipation by $\mathcal{S}_{diss}$. The wave action density spectrum would not change along wave group trajectories without these nonconservative terms (Müller and Natarov 2003).

The radiative transfer equation is based on three assumptions (Müller and Olbers 1975; Müller and Natarov 2003): The first one is the random phase approximation, which says that the phase of waves in the wave field is that much distorted that it is neither necessary nor possible to predict it. The second one is the well-known WKB approximation, where the wavelength of the waves is small compared with the scales of the background and where the wave field varies slowly in time to the time scales of the background. The last one is the weak interaction approximation. The waves are linear and propagate along their group trajectories. They are only slowly modified by dynamical processes.

The wave action density can be expressed in terms of the wave spectral energy density via the relation $\mathcal{A} = \mathcal{E}/\omega_I$. When using this as the parameter of the wave field and transforming the three dimensional wavenumber vector $\vec{k}$ to vertical wavenumber, wave frequency and azimuth coordinates (m, ω_I, φ), the radiative transfer equation is transformed into (Quinn et al. 2020; Olbers et al. 2019)

$$\left(\partial_t + \nabla \dot{\vec{x}}_h + \partial_z \dot{z} + \partial_{\omega_I} \dot{\omega}_I + \partial_m \dot{m} + \partial_\varphi \dot{\varphi}\right) \mathcal{E} = \frac{\dot{\omega}_I}{\omega_I} \mathcal{E} - m^2 \mathcal{D} \mathcal{E} + \mathcal{S}, \tag{3.2}$$

where $\vec{x}_h$ is the horizontal position vector. Using the single column approximation, equation (3.2) simplifies to

$$\left(\partial_t + \partial_z \dot{z} + \partial_{\omega_I} \dot{\omega}_I + \partial_m \dot{m}\right) \mathcal{E} = \frac{\dot{\omega}_I}{\omega_I} \mathcal{E} - m^2 \mathcal{D} \mathcal{E} + \mathcal{S}. \tag{3.3}$$

Equation (3.3) is the base of the RTP for a single wave field with a certain azimuth φ. For wave fields with different azimuths, apply equation (3.3) for each azimuth independently. The total spectral energy of the wave fields is

$$\mathcal{E}_{tot} = \frac{1}{n} \sum_{i}^{n} \mathcal{E}_i, \tag{3.4}$$

where the indices i and n denotes the azimuth and the number of azimuths respectively.

3.2 Wave field spectrum

In analogy to the IDEMIX model (Olbers and Eden 2013; Quinn et al. 2020) we assume that the spectral energy density can be decomposed in a physical and a spectral part

$$\mathcal{E}(z, t, m, \omega_I) = \mathcal{E}_0(z, t) \mathcal{E}(m, \omega_I), \tag{3.5}$$

where we omitted the azimuth dependency. The spectral part can be any arbitrary function of the vertical wavenumber m and intrinsic frequency ω_I. Following Fritts and VanZandt (1993), we assume in this thesis that the Desaubies spectrum applies for the spectral part (Desaubies 1976):

$$A(m) = A_0 \frac{\left(\frac{m}{m^*}\right)^s}{1 + \left(\frac{m}{m^*}\right)^{s+t}} \tag{3.6}$$

$$B(\omega_I) = B_0 \omega_I^{-p} \tag{3.7}$$

where A_0 is such that $\int_{-\infty}^{0} A(m)dm = 1$ and B_0 such that $\int_f^N B(\omega_I)d\omega_I = 1$. The characteristic vertical wavenumber m^* is the transition number between the spectral form for large and small wavenumbers (Fritts and VanZandt 1993). We set $(p, s, t) = (2, 1, 3)$ in accordance with observations (Guo et al. 2017). Smith et al. (1987) fitted the wavenumber spectrum to observational data and found that $m^* \sim \exp(-z/2H)$, but the data are too variable to support such a rule of thumb in general. Other studies also found that m^* is highly variable and does not follow a universal behaviour (VanZandt and Fritts 1989; Fritts and VanZandt 1993), whereas a universal m^{-3} power law is observed for larger vertical wavenumbers (Smith et al. 1987; Fritts and Alexander 2003). Hence, we define the characteristic vertical wavenumber m^* as an altitude and time dependent variable

$$m^* = m^*(z, t). \tag{3.8}$$

Note that the wavenumber m^* varies slowly in z and t with respect to the predominant vertical wave length and intrinsic period of the wave field. The result is an altitude and time dependent Desaubies spectrum and the resulting spectral energy density is

$$\mathcal{E}(z, t, m, \omega_I) = \mathcal{E}_0(z, t) A(m, m^*(z, t)) B(\omega_I) \tag{3.9}$$

3.3 Gaussian Variational Principle

We apply the GVA to our problem of the time derivatives of the two parameters of the gravity wave field These parameters were defined as $\mathcal{E}_0(z,t)$ and $m^*(z,t)$. To this end we first write the spectral energy density as:

$$\mathcal{E}(z,t,m,\omega_I) = \mathcal{E}_0(z,t)C(m,m^*(z,t),\omega_I) \tag{3.10}$$

with $C = A(m,m^*(z,t))B(\omega_I)$. For the temporal and spatial partial derivatives we get

$$\begin{aligned} \partial_t\mathcal{E} &= C\partial_t\mathcal{E}_0 + \partial_{m^*}C\mathcal{E}_0\partial_t m^* \qquad (3.11)\\ \partial_z\mathcal{E} &= C\partial_z\mathcal{E}_0 + \partial_{m^*}C\mathcal{E}_0\partial_z m^* \qquad (3.12)\end{aligned}$$

and the radiative transfer equation (Eq. 3.3) becomes

$$\begin{aligned} &C\partial_t\mathcal{E}_0 + \dot{z}C\partial_z\mathcal{E}_0 + \partial_{m^*}C\mathcal{E}_0\partial_t m^* + \dot{z}\partial_{m^*}C\mathcal{E}_0\partial_z m^* + \partial_z\dot{z}C\mathcal{E}_0\\ &+\partial_m\left(\dot{m}C\right)\mathcal{E}_0 + \partial_{\omega_I}\left(\dot{\omega}_I C\right)\mathcal{E}_0 = \frac{\dot{\omega}_I}{\omega_I}C\mathcal{E}_0 - m^2\mathcal{D}C\mathcal{E}_0 + \mathcal{S}. \end{aligned} \tag{3.13}$$

We define the error function χ as:[1]

$$\begin{aligned} \chi^2(z,t) = \iint C \Bigg[&\partial_t\mathcal{E}_0 + \dot{z}\partial_z\mathcal{E}_0 + \frac{\partial_{m^*}C}{C}\mathcal{E}_0\partial_t m^* + \dot{z}\frac{\partial_{m^*}C}{C}\mathcal{E}_0\partial_z m^* + \partial_z\dot{z}\mathcal{E}_0\\ &+ \frac{1}{C}\partial_m\left(\dot{m}C\right)\mathcal{E}_0 + \frac{1}{C}\partial_{\omega_I}\left(\dot{\omega}_I C\right)\mathcal{E}_0 - \frac{\dot{\omega}_I}{\omega_I}\mathcal{E}_0 + m^2\mathcal{D}\mathcal{E}_0 - \frac{\mathcal{S}}{C}\Bigg]^2\\ &dm\,d\omega_I. \end{aligned} \tag{3.14}$$

---

[1] The definition of the error function differs from the proposed definition in the chapter "The IDEMIX model: Parameterization of internal gravity waves for circulation models of ocean and atmosphere" (Olbers et al. 2019, eq. 3.22) to conserve the energy with respect to the definition of the norm of the Desaubies spectrum (see Sec. 3.2). A discussion of this issue is beyond the scope of this thesis.

As a result the variation of the error function is

$$\begin{aligned}
\delta\chi^2(z,t) = 2\int\!\!\int C \Bigg[& \partial_t\mathcal{E}_0 + \frac{\partial_{m^*}C}{C}\mathcal{E}_0\partial_t m^* \\
& + \dot{z}\partial_z\mathcal{E}_0 + \dot{z}\frac{\partial_{m^*}C}{C}\mathcal{E}_0\partial_z m^* + \partial_z\dot{z}\mathcal{E}_0 \\
& + \frac{1}{C}\partial_m\left(\dot{m}C\right)\mathcal{E}_0 + \frac{1}{C}\partial_{\omega_I}\left(\dot{\omega}_I C\right)\mathcal{E}_0 \\
& - \frac{\dot{\omega}_I}{\omega_I}\mathcal{E}_0 + m^2\mathcal{D}\mathcal{E}_0 - \frac{\mathcal{S}}{C}\Bigg]\,\delta\left(\partial_t\mathcal{E}_0\right)\,dm\,d\omega_I \\
+2\int\!\!\int \partial_{m^*}C\mathcal{E}_0 \Bigg[& \partial_t\mathcal{E}_0 + \frac{\partial_{m^*}C}{C}\mathcal{E}_0\partial_t m^* \\
& + \dot{z}\partial_z\mathcal{E}_0 + \dot{z}\frac{\partial_{m^*}C}{C}\mathcal{E}_0\partial_z m^* + \partial_z\dot{z}\mathcal{E}_0 \\
& + \frac{1}{C}\partial_m\left(\dot{m}C\right)\mathcal{E}_0 + \frac{1}{C}\partial_{\omega_I}\left(\dot{\omega}_I C\right)\mathcal{E}_0 \\
& - \frac{\dot{\omega}_I}{\omega_I}\mathcal{E}_0 + m^2\mathcal{D}\mathcal{E}_0 - \frac{\mathcal{S}}{C}\Bigg]\,\delta\left(\partial_t m^*\right)\,dm\,d\omega_I.
\end{aligned} \tag{3.15}$$

The error function reaches its minimum, if the functional derivatives with respect to the variation parameters become zero

$$\frac{\delta\chi^2(z,t)}{\delta\left(\partial_t\mathcal{E}_0\right)} = 0 \tag{3.16}$$

$$\frac{\delta\chi^2(z,t)}{\delta\left(\partial_t m^*\right)} = 0. \tag{3.17}$$

This leads to the equations

$$\begin{aligned}
\iint C \Bigg[& \partial_t\mathcal{E}_0 + \dot{z}\partial_z\mathcal{E}_0 + \frac{\partial_{m^*}C}{C}\mathcal{E}_0\partial_t m^* + \dot{z}\frac{\partial_{m^*}C}{C}\mathcal{E}_0\partial_z m^* + \partial_z\dot{z}\mathcal{E}_0 \\
& + \frac{1}{C}\partial_m\left(\dot{m}C\right)\mathcal{E}_0 + \frac{1}{C}\partial_{\omega_I}\left(\dot{\omega}_I C\right)\mathcal{E}_0 \\
& - \frac{\dot{\omega}_I}{\omega_I}\mathcal{E}_0 + m^2\mathcal{D}\mathcal{E}_0 - \frac{\mathcal{S}}{C}\Bigg]\,dm\,d\omega_I = 0
\end{aligned} \tag{3.18}$$

and

$$\begin{aligned}\iint \partial_{m^*}C \Bigg[&\partial_t\mathcal{E}_0 + \dot{z}\partial_z\mathcal{E}_0 + \frac{\partial_{m^*}C}{C}\mathcal{E}_0\partial_t m^* + \dot{z}\frac{\partial_{m^*}C}{C}\mathcal{E}_0\partial_z m^* + \partial_z\dot{z}\mathcal{E}_0 \\ &+ \frac{1}{C}\partial_m\left(\dot{m}C\right)\mathcal{E}_0 + \frac{1}{C}\partial_{\omega_I}\left(\dot{\omega}_I C\right)\mathcal{E}_0 \\ &- \frac{\dot{\omega}_I}{\omega_I}\mathcal{E}_0 + m^2\mathcal{D}\mathcal{E}_0 - \frac{\mathcal{S}}{C}\Bigg]\, dm\, d\omega_I = 0.\end{aligned} \tag{3.19}$$

Solving this set of coupled differential equations for $\partial_t\mathcal{E}_0$ and $\partial_t m^*$, using the ray equations (2.24) to (2.26) and applying the Desaubies spectrum (see Sec. 3.2) for the function C (see Eq. 3.10) leads to the final prognostic equations

$$\begin{aligned}\partial_t\mathcal{E}_0 = &+ a_0\partial_z\mathcal{E}_0 + a_1\partial_z m^*\mathcal{E}_0 - a_U\mathcal{E}_0\frac{\partial_z U}{N} \\ &+ a_N\mathcal{E}_0\frac{\partial_z N}{N} - a_{\mathcal{D}}\mathcal{D}\mathcal{E}_0 + a_{\mathcal{S}}\mathcal{S}\end{aligned} \tag{3.20}$$

$$\partial_t m^* = +b_0\frac{\partial_z\mathcal{E}_0}{\mathcal{E}_0} + b_1\partial_z m^* - b_U\frac{\partial_z U}{N} + b_N\frac{\partial_z N}{N} - b_{\mathcal{D}}\mathcal{D} + b_{\mathcal{S}}\frac{\mathcal{S}}{\mathcal{E}_0}. \tag{3.21}$$

The parameters a_i and b_i are functions of m^* and contain integrals over the wavenumber frequency spectrum. They have the general form

$$X = \frac{\iint f_1\, dm\, d\omega_I \iint f_2\, dm\, d\omega_I - \iint f_3\, dm\, d\omega_I \iint f_4\, dm\, d\omega_I}{\iint f_5\, dm\, d\omega_I \iint f_2\, dm\, d\omega_I + \iint f_6\, dm\, d\omega_I \iint f_4\, dm\, d\omega_I} \tag{3.22}$$

with $f_i = f_i(m, m^*, \omega_I)$. The specification of the integration intervals are subject of Section 3.7. The full derivation of the prognostic equations can be found in the Appendix A. The equations (3.20) and (3.21) are coupled partial differential equations with respect to height z and time t for the energy volume density $\mathcal{E}_0$ and the characteristic vertical wavenumber m^* of the height and time dependent wave spectrum. These equations form the basis of the Radiative Transfer Parameterization. Note that the transient vertical propagation of wave packets as well as the modulation of the vertical wavenumber spectrum by the height and time-dependent background state are described. This makes this parameterization fundamentally different from conventional gravity-wave parameterizations where the background is considered to be constant in time and only steady-state wave solutions are considered.

3.4 Fluxes, wave momentum density, and wave energy density

In the following we show how the wave-mean flow interaction terms can be calculated from the spectral energy density $\mathcal{E}(z, t, m, \omega_I, \phi)$. We know for the energy volume density for waves in x-direction, that

$$E_{GW} = \rho_r \langle u' u' \rangle. \tag{3.23}$$

This expression includes the kinetic and the potential energy density, which are equivalent for mid-frequency gravity waves that are not affected by the Coriolis force, as it is assumed for our parameterization. With $X' = \Re(X^*)$ and the identity for the ansatz (2.8)

$$\langle \Re(X^*) \Re(X^*) \rangle = \frac{1}{2} \langle X^{*\dagger} X^* \rangle \tag{3.24}$$

we can write

$$E_{GW} = \frac{1}{2} \rho_r \langle u^{*\dagger} u^* \rangle. \tag{3.25}$$

Since the energy density (3.25) is given by the integration over the spectral energy density[2],

$$E_{GW} = \iint \mathcal{E}(z, t, m, \omega_I) \, dm \, d\omega_I, \tag{3.26}$$

---

[2] For more than one azimuth, the result is

$$E_{GW} = \frac{1}{n} \sum_{i=1}^{n} \iint \mathcal{E}_i(z, t, m, \omega_I) \, dm \, d\omega_I$$

we can write

$$\frac{1}{2}\rho_r \langle u^{*\dagger} u^* \rangle = \iint \mathcal{E}(z,t,m,\omega_I)\, dm\, d\omega_I. \tag{3.27}$$

Using the polarization relations (2.18)-(2.20), we can write the momentum flux $F = \rho_r \langle u'v' \rangle$, the wave momentum density $M_{GW} = \langle \rho' u' \rangle$, the pressure flux $F_p = \langle p'w' \rangle$, and again for the sake of completeness the wave energy density $E_{GW} = \rho_r \langle u'u' \rangle$ in terms of the spectral energy density $\mathcal{E}$ as

$$F = \iint\limits_{ESA} \frac{\omega_I}{N}\frac{m}{m'}\mathcal{E}(z,t,m,\omega_I)\, dm\, d\omega_I \tag{3.28}$$

$$M_{GW} = \iint\limits_{ESA} \frac{1}{g}\frac{N}{m'}\frac{1}{2H}\mathcal{E}(z,t,m,\omega_I)\, dm\, d\omega_I \tag{3.29}$$

$$F_p = -\iint\limits_{ESA} \frac{\omega_I m}{m'^2}\mathcal{E}(z,t,m,\omega_I)\, dm\, d\omega_I \tag{3.30}$$

$$E_{GW} = \iint\limits_{ESA} \mathcal{E}(z,t,m,\omega_I)\, dm\, d\omega_I. \tag{3.31}$$

The full derivations of these expressions can be found in Appendix C. The integration intervals, denoted here as Effective Spectral Area (ESA), are subject to the next section. The resulting gravity-wave effects in the mean-flow equations for horizontal wind and enthalpy (compare also with Eqs. 2.38 and 2.46) are

$$(\partial_t U)_{GW} = -\frac{1}{\rho_r}\partial_z F - \frac{1}{\rho_r}\partial_t M_{GW} \tag{3.32}$$

$$(c_p d_t T_r)_{GW} = -\frac{1}{\rho_r}\partial_z F_p - \frac{1}{\rho_r} F \partial_z U - \frac{1}{\rho_r}\frac{1}{2}\partial_t E_{GW}. \tag{3.33}$$

These gravity-wave effects conserves on one hand the horizontal momentum, which is the sum of the vertical integrated horizontal momentum of the background and the gravity-wave momentum, and on the other hand the energy, which contains the sum of the vertical integrated enthalpy, kinetic background energy, and the wave energy.

3.5 Diffusion

The diffusion of gravity waves can be caused by different processes. In this thesis we restrict ourselves to the turbulent and molecular viscosity and heat conduction. The diffusion coefficient of the equations (3.20) and (3.21) is the sum of the diffusion coefficients of these two processes

$$\mathcal{D}(z,t) = \mathcal{D}_{sat}(z,t) + \mathcal{D}_{mol}(z,t), \tag{3.34}$$

where the index *sat* indicates that we assume that turbulent diffusion arise due to wave saturation according to Lindzen (1981).

Diffusion by saturation

In the framework of the nondissipative linear theory of gravity waves it is expected for the amplitude to grow exponentially with height due to decreasing density. Observations, however, reveal that this is not the case, that is, the wave amplitude becomes saturated (Lindzen 1981; Smith et al. 1987). Therefore we need a process to prevent the wave amplitude to keep growing exponentially. Lindzen described the process of wave saturation as a result of the onset of convective instability. That is, the amplitude of the wave is prevented to grow beyond the point where the stratification becomes statically unstable. For a dry atmosphere this point is defined via the adiabatic lapse rate Γ:

$$\Gamma = -\frac{g}{c_p} = \partial_z T. \tag{3.35}$$

The atmosphere is stable, if $\partial_z T$ is greater than the lapse rate. With the decomposition of the temperature T into a reference temperature and a temperature perturbation due waves ($T = \bar{T}(z) + T'(z,t)$), equation (3.35) becomes

$$-\frac{g}{c_p} = d_z\bar{T} + \partial_z T'. \tag{3.36}$$

Using the definition of the Brunt-Väisälä frequency (Eq. 2.6), the saturation condition (3.36) can be written as

$$0 = N^2 + g\frac{\partial_z T'}{\bar{T}}. \tag{3.37}$$

The wave buoyancy is $b' = gT'/\bar{T}$ such that (3.37) can also be written as

$$\partial_z b' + N^2 = 0. \tag{3.38}$$

Hence, the stratification is unstable if the local and instantaneous wave buoyancy fulfills

$$-\partial_z b' > N^2. \tag{3.39}$$

Lindzen calculated the point of saturation and the corresponding diffusion coefficient for every monochromatic wave individually. In our new transient parameterization scheme, we follow the approach that was suggested, for example, by Becker and McLandress (2009), and compute a single turbulent diffusion coefficient for the whole gravity-wave field. No matter what spectral part of the wave packet causes the atmosphere to become instable, the diffusion will act on the whole wave packet. Even when there are two wave packets with different azimuth directions at the same place. The left hand side of equation (3.39) combined with the identity (3.24)

$$\langle \partial_z b' \partial_z b' \rangle = \frac{1}{2} \langle \partial_z b^* \partial_z b^{*\dagger} \rangle. \tag{3.40}$$

From the ansatz (2.8) we get

$$\langle \partial_z b' \partial_z b' \rangle = \frac{1}{2} \left\langle \left(m^2 + \frac{1}{4H^2} \right) b^* b^{*\dagger} \right\rangle. \tag{3.41}$$

The spectral energy density is

$$\mathcal{E}(z, t, m, \omega_I, \varphi) = \frac{1}{2} \frac{\rho_r}{N^2} b^* b^{*\dagger} \tag{3.42}$$

where the averaging over the gravity wave scales is defined as

$$\left\langle X^* Y^{*\dagger} \right\rangle = \iiint X^* Y^{*\dagger} \, dm \, d\omega_I \, d\varphi. \tag{3.43}$$

With these relations we get from the saturation condition (3.39) applied to a single azimuth

$$\frac{1}{N^2 \rho_r} \iint \left(m^2 + \frac{1}{4H^2} \right) \mathcal{E}(z, t, m, \omega_I) \, dm \, d\omega_I > 1. \tag{3.44}$$

Note that gravity waves can be unstable below this threshold (e.g., Achatz 2005; Fritts and Alexander 2003). To account for the variety in instability mechanisms in a very simple way, we introduce a tunable parameter α such that the saturation condition becomes

$$\mathcal{C}\mathcal{E}_0 = \frac{\mathcal{E}_0}{\rho_r N^2} \iint \left(m^2 + \frac{1}{4H^2} \right) A(m) B(\omega_I) \, dm \, d\omega_I > \alpha^2. \tag{3.45}$$

With the saturation condition (3.45) we assume that all compartments of the wave packet are subject to constructive interference. The compartments are in phase and their amplitudes add up to the largest possible convective instability. This is, however, a rather unlikely scenario. Therefore, we introduce an additional azimuth-independent interference parameter γ with $\gamma \in [0, 1]$. An interference parameter of $\gamma = 0$ describes completely destructive interference, where the waves are that much out of phase that the total buoyancy amplitude is zero. The other limit of the parameter, $\gamma = 1$, corresponds to completely constructive interference. With this additional parameter, we can write the saturation condition as

$$\gamma \mathcal{C}\mathcal{E}_0 = \gamma \frac{\mathcal{E}_0}{\rho_r N^2} \iint \left(m^2 + \frac{1}{4H^2} \right) A(m) B(\omega_I) \, dm \, d\omega_I > \alpha^2. \tag{3.46}$$

In our new RTP, the diffusion coefficient is larger than zero if (3.46) is fulfilled:

$$\mathcal{D}_{sat} = \begin{cases} K_0 \left(\gamma \mathcal{C}\mathcal{E}_0 - \alpha^2 \right) & , \gamma \mathcal{C}\mathcal{E}_0 > \alpha^2 \\ 0 & , \text{else} \end{cases} \tag{3.47}$$

where the diffusion coefficient K_0 is another tunable parameter.

If there are wave packets with different azimuths, the contributions to equation (3.46) from these azimuths add up. This leads to the situation that there can be saturation even if the saturation condition is not fulfilled for any of the individual azimuths (see also Hines 1997a; Becker and McLandress 2009). For several discrete azimuths equation (3.46) becomes

$$\frac{\gamma}{n}\sum_{i=1}^{n}\frac{\mathcal{E}_{0,i}}{\rho_r N^2}\iint\left(m^2+\frac{1}{4H^2}\right)A_i(m)B_i(\omega_I)\,dm\,d\omega_I > \alpha^2 \tag{3.48}$$

where the azimuths are enumerated by the index i and $1/n$ normalizes $1/n\sum_{i=1}^{n}1$ to one. Note that each azimuth has its own spectrum. The turbulent vertical diffusion coefficient that results from the contributions from azimuths considered in the parameterizations can then be written as

$$\mathcal{D}_{sat}=\begin{cases} K_0\left(\frac{\gamma}{n}\sum_{i=1}^{n}\mathcal{C}_i\mathcal{E}_{0,i}-\alpha^2\right) & ,\ \frac{\gamma}{n}\sum_{i=1}^{n}\mathcal{C}_i\mathcal{E}_{0,i}>\alpha^2 \\ 0 & ,\ \text{else}\end{cases}\ . \tag{3.49}$$

Molecular viscosity

Molecular viscosity acts on the small-end of the forward energy cascade where the energy is eventually dissipated. In gravity wave schemes this process is often parameterized by some turbulent diffusion (see also Lindzen 1981; Becker and McLandress 2009), as it is also done in the gravity wave scheme proposed in this thesis. With increasing altitude in the thermosphere, however, gravity waves are increasingly damped directly by molecular viscosity (see also Vadas 2007). The reason is that the kinematic molecular viscosity grows exponentially with altitude because of the decreasing density and the associated increase of the free path of molecules. To account for this process we add the kinematic molecular viscosity to the turbulent viscosity (see Eq. 3.34).

A simple expression for the kinematic molecular is

$$\mathcal{D}_{mol} = \frac{\rho_0}{\rho} K_{mol}. \tag{3.50}$$

Here, ρ_0 and K_{mol} are the density and kinematic molecular viscosity, respectively, at sea level.

3.6 Source function

The source function of the wave field can be any function that is continuous in space and time, as well as in the spectral domain:

$$\mathcal{S} = \mathcal{S}(z, t, m, \omega_I, \varphi). \tag{3.51}$$

Note that unlike to conventional gravity wave scheme, there are no launch level parameter for a certain altitude and time , and there is not a set of monochromatic waves which are launched. We separate the source function into a physical, a spectral, and an azimuthal part

$$\mathcal{S}(z, t, m, \omega_I, \varphi) = \mathcal{S}(z, t)\mathcal{S}(m, \omega_I)\mathcal{S}(\varphi) \tag{3.52}$$

We define the spectral part in this thesis to be a Desaubies spectrum

$$\mathcal{S}(m, \omega_I) = A(m, m_S^*)B(\omega_I). \tag{3.53}$$

The parameter m_S^* is the characteristic vertical wavenumber related to the assumed gravity wave source. This number can depend on time and altitude in the single-column picture. Note that the wavenumber m_S^* is different from the characteristic vertical wavenumber $m^*(z, t)$.

The azimuthal part $\mathcal{S}(\varphi)$ can be any function of the azimuths (e.g. the orientation of the front or the direction of the horizontal wind approaching a mountain).

There can be a superposition of different source functions for a single azimuth

$$\mathcal{S}(z,t,m,\omega_I) = \sum_j \mathcal{S}_j(z,t,m,\omega_I), \tag{3.54}$$

where the different sources are enumerated by the index j.

3.7 Spectral areas

Until this point there was no specification of the integration intervals and areas spanned in the m, ω_I-plane. As will be shown below, there are two spectral areas necessary for the RTP. The first one is the total spectral area, which is used to specify the prognostic equations. The second one is the effective spectral area, that is used to compute the wave fluxes and, hence, the resulting wave-mean flow interactions.

There are certain restrictions to the integration intervals because of the assumption of mid-frequency waves and due to numerical limitations for the (characteristic) vertical wavenumber. To make sure that there are only mid-frequency waves inside of the spectrum, we set the lower and upper limits for the intrinsic frequency ω_I to

$$\omega_B = 5f_0 \tag{3.55}$$

and

$$\omega_T = \frac{N}{5}, \tag{3.56}$$

respectively, where $f_0 = 7.3 \cdot 10^{-5}\mathrm{s}^{-1}$ is a fixed Coriolis frequency, and N is the local Brunt-Väisälä frequency. The vertical wavenumber m can range from $-\infty$ to 0 for upward propagating waves. However, a wavenumber of $m = 0$ would result in an infinite vertical wavelength which is not suitable. We associate the upper limit for the vertical wavenumber (that is, the smallest absolute vertical wavenumber) with half of the maximum altitude of the model Z using

$$m_0 = -\frac{2\pi}{Z/2}\frac{1}{\mathrm{m}}. \tag{3.57}$$

An infinite negative vertical wavenumber would cause the integrals of $a_{\mathcal{D}}$ and $b_{\mathcal{D}}$ (see Eqs. 3.20 and 3.21) to become infinite. Numerical test calculations showed that a sufficiently large lower limit is (see also Quinn et al. 2020)

$$m_1 = -\frac{2\pi}{100}\frac{1}{\mathrm{m}}. \tag{3.58}$$

These limits m_0 and m_1 are also the limits for the characteristic vertical wavenumber m^*.

The upper limit m_0 deviates from the Boussinesq limit, which would require $|m| > 1/(2H)$ (see Sec. 2.1), when the model height Z is larger than $8\pi H$. Since the largest faction of wave energy is centered around the characteristic vertical wavenumber m^*, and since m^* is usually within the Boussinesq limit, we neglect the Boussinesq limit as a constraint for the entire spectrum. In other words, we assume that a violation of the Boussinesq limit by only a small part of the wave spectrum is of minor importance as long as the characteristic vertical wavenumber m^* fullfils $|m^*| > 1/(2H)$. A verification for this approach is presented in Section 4.5, where we inspect extreme cases for which even m^* exceeds the Boussinesq limit.

Total spectral area

The total spectral area as defined by the limits (3.55)-(3.58) is used for the computation of the prognostic equations (3.20) and (3.21) and the saturation condition (3.45). These integration intervals are the same at every altitude and independent of time. That is, all integrals occurring in (3.20) and (3.21) are evaluated according to

$$X = \int_{\omega_B}^{\omega_T} \int_{m_1}^{m_0} f(m, m^*, \omega_I)\, dm\, d\omega_I, \tag{3.59}$$

where f is an arbitrary function. When applying the GVA (Gaussian Variational Approach) to the radiative transfer equations, we assume that only $\mathcal{E}_0(z,t)$ and $m^*(z,t))$ can vary in order to find the minimum of the error function. Hence, we implicitly assume that the integration intervals are fixed in this formalism. As a consequence, it is irrelevant if all waves generated at the source are preserved at any altitude and time or if they were "removed" from the spectrum at a lower altitude and earlier time because of a critical level or other effects that remain to be specified.

Effective spectral area

The Effective Spectral Area (ESA) is a reduction of the total spectral area. Similar to ray-tracing, where the path of the wave through the changing background is traced, we trace the ESA and determine where the ESA is reduced due to critical layers, reflection, and effects of slow changes of $H(z)$ and $N(z)$. This is schematically visualised in Figure 3.1 for an idealized and temporally fixed vertical profile of the horizontal wind $U(z)$. There are three layer (1, 2, and 3). While there is a positive horizontal wind (middle column in Fig. 3.1) at layer 2, the horizontal wind is zero at layer 1 and 3. The wind is assumed to be constant in time. The schematic shows the modulation and reduction of the ESA of two wave packets (left and right column). The wave packet in the left column is assumed to propagate against the horizontal wind, while the wave packet in the right column is assumed to propagate with the wind. At layer 1 there is an initial ESA which is already smaller than the total spectral area due to processes at lower altitudes. The wave packet in the left column is Doppler shifted to larger intrinsic frequencies and smaller absolute vertical wavenumbers (larger vertical wavelengths) at layer 2. As a result, every wave parcel of the wave packet that exceeds the limits (3.56) and/or (3.57) is cropped from the ESA. The remaining wave

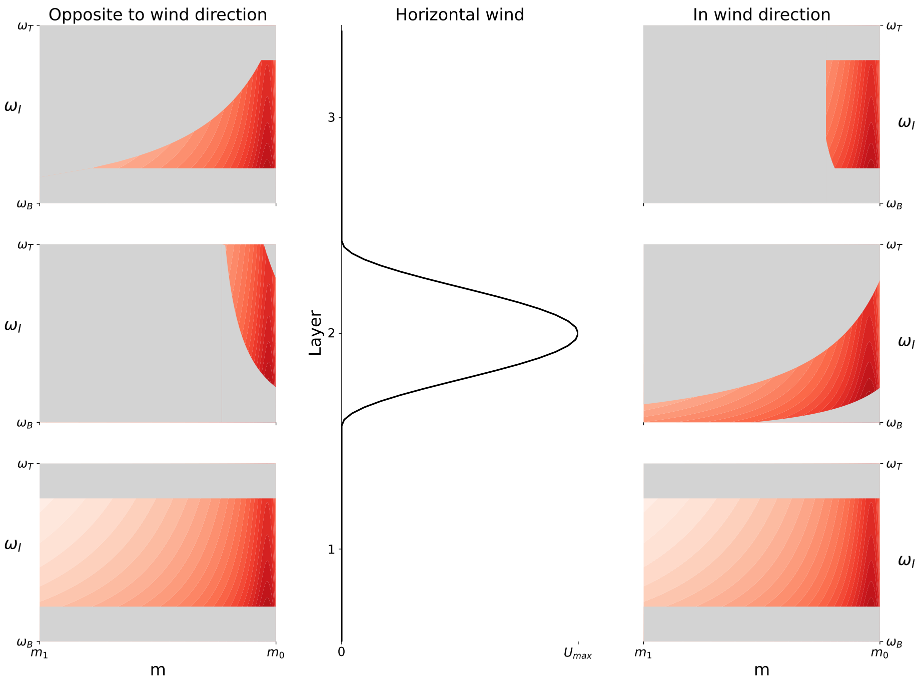

Figure 3.1: Erosion of the effective spectral area due to Doppler shifting. In red colors: the effective spectral area with intensities according to $A(m)B(\omega_I)$; In gray: the total spectral area; Left panel: Reflection. Wave number frequency pairs are omitted due to exceeding the upper frequency limit ω_T; Middle panel: Sketch of the wind jet; Right panel: Critical layer. Wave number frequency pairs are omitted due to exceeding the lower frequency limit ω_B and/or the wavenumber limit m_1; Bottom: spectral areas below the wind jet at layer 1; Middle: spectral areas at the peak of the wind jet at layer 2; Top: spectral areas above the wind jet at layer 3.

spectrum is shifted back to smaller intrinsic frequencies and larger absolute vertical wavenumbers at layer 3. A comparison of the ESA at layer 3 with the ESA at layer 1 shows which parts of the wave spectrum at level 1 have been subject to reflection. The wave packet in the right column is Doppler shifted to smaller intrinsic frequencies and larger absolute vertical wavenumbers (smaller vertical wavelength) at layer 2. As a result, every wave parcel that exceeds the limits (3.55) and/or (3.58) is cropped. The reduced wave packet is shifted back to larger intrinsic frequencies and smaller absolute vertical wavenumbers (longer vertical wavelengths) at layer 3. From the comparison of the ESAs at layers 3 and 1 we can identify which parts of the wave packet have encountered critical layers. For the new gravity wave scheme proposed in this thesis we consider also effects on the ESA that result from the effects of slow changes of $H(z)$ and $N(z)$ with height on the intrinsic frequency and vertical wavenumber.

In the following we describe the mathematical approach for the reduction of the ESA. We take advantage of the WKB assumption that the vertical and temporal variations of the background are slow with respect to the scales of the waves. This allows us to diagnose the ESA at any time from the surface layer to the model top[3]

To calculate the refraction of the vertical wavenumber m and the modulation of the intrinsic frequency ω_I of a spectral element inside the wave packet we assume that the horizontal wavenumber k is constant along its path:

$$k = \pm\sqrt{\frac{\omega_I^2\left(m^2 + \frac{1}{4H^2}\right)}{N^2}} = const \tag{3.60}$$

This assumption follows from the single-column approximation. Note that the horizontal wavenumber k is positive (negative) for a wave that propagates in positive (negative) x-direction in the intrinsic frame work. Our second assumption is that the ground-based frequency ω is constant along the path of the wave:

$$\omega = \omega_I - kU = const \tag{3.61}$$

This assumption applies when the temporal changes of the background flow is very slow compared to the intrinsic period of the wave. Note that we only invoke this assumption to determine the ESA. The new gravity wave scheme takes the transience of the background flow fully into account otherwise.

---

[3] A more advanced approach is to follow the path of the wave packet in time and determine the reduction of the ESA from previous times.

With the conditions (3.60) and (3.61) we get for the changes of the intrinsic frequency and the vertical wavenumber of a given wave parcel:

$$\omega_{I,a+1} = \omega_{I,a}\left(1 \mp \frac{U_{a+1}-U_a}{N_a}\sqrt{m_a^2+\frac{1}{4H_a^2}}\right) \tag{3.62}$$

$$m_{a+1} = -\sqrt{\frac{N_{a+1}^2}{N_a^2}\frac{m_a^2+\frac{1}{4H_a^2}}{\left(1 \mp \frac{U_{a+1}-U_a}{N_a}\sqrt{m_a^2+\frac{1}{4H_a^2}}\right)^2} - \frac{1}{4H_{a+1}^2}}. \tag{3.63}$$

The minus (plus) signs correspond to a positive (negative) horizontal wavenumber k, and a and $a+1$ are two adjacent vertical layers.

At any layer with a nonzero source $\mathcal{S}_z(z)$ (from the separation $\mathcal{S}(z,t) = \mathcal{S}_z(z)\mathcal{S}_t(t)$) we define the ESA to span a rectangle area (see Fig. 3.1, layer 1) with the limits

$$\omega_B \le \omega_b \le \omega_I \le \omega_t \le \omega_T \tag{3.64}$$

$$m_1 \le m_{11} \le m \le m_{00} \le m_0, \tag{3.65}$$

where ω_b and ω_t are the lower and upper intrinsic frequency limits and m_{11} and m_{00} the lower and upper vertical wavenumber limits of the generated wave field. For layers between the surface layer and the lowest layer with a nonzero source $\mathcal{S}_z(z)$ the ESA is zero. At any layer with $\mathcal{S}_z(z) = 0$ we determine the ESA with the equations (3.62) and (3.63) from the ESA of the layer below.[4]

Note that the integrals over the ESA at layer $a + 1$

$$I = \iint\limits_{ESA} f(m, m^*, \omega_I)\, dm\, d\omega_I \tag{3.66}$$

have integration limits which depend on the limits from the previous layer

---

[4] When we demand $\mathcal{S}(z,t)$ to be nonzero instead of only $\mathcal{S}_z(z)$ then we can have the case where $\mathcal{S}_t(t) = 0$ and $\mathcal{S}_z(z) \neq 0$ at a certain time. This would immediately leads to a vanishing ESA at any layer. For the condition that $\mathcal{S}(z,t)$ has to be nonzero we would need to follow the path of the wave field in time.

(e.g., $\omega_{I,a+1} = \omega_{I,a+1}(\omega_{I,a}, m_a)$). In Appendix B we describe our approximate numerical method to solve these integrals using finite elements.

Chapter IV

Idealised simulations

In this chapter we perform different simulations where a single wave packet is generated and propagates through different idealised vertical profiles of horizontal winds and temperature until it dissipates. Parameters of the wave packet, which are the same for all simulations, idealised wind and temperature profiles, and the calculation of energy and momentum budgets are presented in Section 4.1. The propagation and interaction of a wave packet in a homogeneous background is shown in Section 4.2. The influence of different wind shears on the propagation of a wave packet is subject to Section 4.3 and the influence of a temperature gradient to Section 4.4. The final Section (4.5) is about two wave packets with opposite horizontal phase velocities propagating through different wind and temperature profiles from the COSPAR International Reference Atmosphere (CIRA-86).

4.1 Simulation environment

All simulations in this chapter are for one single column. The vertical and temporal resolutions are

$$\Delta z = 1000\,\mathrm{m} \tag{4.1}$$

$$\Delta t = 60\,\mathrm{s}. \tag{4.2}$$

The altitude range is from $z_{min} = 0$ to $z_{max} = 110\,\mathrm{km}$ with 111 equidistant layers. The duration of the simulations is $t_{max} = 2\,\mathrm{days}$. The parameters ρ_0, c_p, g, and R are:

$$\begin{aligned} \rho_0 &= 1.2\,\frac{\mathrm{kg}}{\mathrm{m}^3} && (4.3)\\ c_p &= 1005\,\frac{\mathrm{J}}{\mathrm{kgK}} && (4.4)\\ g &= 9.81\,\frac{\mathrm{m}}{\mathrm{s}^2} && (4.5)\\ R &= 287\,\frac{\mathrm{J}}{\mathrm{kgK}}. && (4.6)\\ & && (4.7) \end{aligned}$$

The coefficients for the molecular and turbulent diffusion mechanisms are

$$K_{mol} = 1.33 \cdot 10^{-5}\,\frac{\mathrm{m}^2}{\mathrm{s}} \tag{4.8}$$

$$K_0 = 50\,\frac{\mathrm{m}^2}{\mathrm{s}} \tag{4.9}$$

$$\alpha^2 = 0.66 \tag{4.10}$$

$$\gamma = 1. \tag{4.11}$$

Source function

The source function describes the excitation of a wave packet in the troposphere. We write

$$\mathcal{S}(z,t,m,\omega_I) = \mathcal{S}_z(z)\mathcal{S}_t(t)A(m,m^*(z,t))B(\omega_I) \tag{4.12}$$

with

$$\mathcal{S}_z(z) = \begin{cases} \frac{S_0}{2}\left(1+\cos\left(\frac{\pi(z-z_0)}{\Delta z_S}\right)\right) & ,|z-z_0| < \Delta z_S \\ 0 & ,\text{else} \end{cases} \tag{4.13}$$

$$\mathcal{S}_t(t) = \begin{cases} \sin\left(\frac{\pi t}{\Delta t_S}\right) & ,t < 2\Delta t_S \\ 0 & ,\text{else.} \end{cases} \tag{4.14}$$

The characteristic vertical wavenumber of the source is set to

$$m^*_{\mathcal{S}} = -\frac{2\pi}{2000}\,\frac{1}{\mathrm{m}}, \tag{4.15}$$

which corresponds to the characteristic vertical wavenumbers that are typical for the troposphere (VanZandt and Fritts 1989). The other parameters for the source function are

$$S_0 = \frac{5\cdot 10^{-8}}{|m^*_{\mathcal{S}}|}\,\frac{\mathrm{kg}}{\mathrm{s}^3} \tag{4.16}$$

$$z_0 = 10000\,\mathrm{m} \tag{4.17}$$

$$\Delta z_S = 5000\,\mathrm{m} \tag{4.18}$$

$$\Delta t_S = 3600\,\mathrm{s}. \tag{4.19}$$

The effective spectral area in the source region is limited by

$$5f_0 \le \omega_I \le \frac{N}{5} \tag{4.20}$$

$$-\frac{2\pi}{100}\,\frac{1}{\mathrm{m}} \le m \le -\frac{2\pi}{7000}\,\frac{1}{\mathrm{m}}. \tag{4.21}$$

Horizontal wind profiles

In this thesis we present simulations with idealised jets or step functions and more realistic wind profiles interpolated from CIRA86 data (Fleming et al. 1990). The idealised profiles are defined as

$$U_{step}(z) = U_0 + U_{max}\left(1 + \exp\left(-\frac{(z-z_0)}{\Delta z_U}\right)\right) \tag{4.22}$$

$$U_{jet}(z) = \begin{cases} U_0 + \frac{U_{max}}{2}\left(1 + \cos\left(\frac{\pi(z-z_0)}{\Delta z_U}\right)\right) & , |z-z_0| < \Delta z_U \\ U_0 & , \text{else} \end{cases} . \tag{4.23}$$

The first function U_{step}, a continuous approximation of the Heaviside function, is defined by the main wind velocity U_0, the maximum deviation of this velocity U_{max}, the position z_0 of $(U_0 + U_{max})/2$ and the width Δz_U. This function is used to describe a rapid but continuous change of the horizontal wind U with altitude. The function U_{jet} describes a wind jet, with the parameter U_0 for the main velocity, U_{max} for the largest deviation from U_0, z_0 for the altitude of the peak velocity $U_0 + U_{max}$, and $2\Delta z_U$ for the full depth of the jet. A positive U corresponds to an eastward horizontal wind. We also use realistic wind profiles. These are taken from Fleming et al. (1990) and are approximated by polynomials for the purpose of our simulations. Such profiles will be introduced in the context of the respective simulations.

Temperature profiles

Our simulations use an idealised step function for the temperature profile, or more realistic temperature profiles interpolated from CIRA86 data (Fleming et al. 1990). The idealised profile is defined as

$$T_{step}(z) = T_0 + T_{max}\left(1 + \exp\left(-\frac{(z-z_0)}{\Delta z_T}\right)\right) . \tag{4.24}$$

The function T_{step}, a continuous approximation of the Heaviside function, is defined by the main temperature T_0, the maximum deviation of this temperature T_{max}, the position z_0 of $(T_0 + T_{max})/2$ and the width Δz_T. This function is used to describe a rapid but continuous change of the temperature T with altitude. We also use realistic temperature profiles. These are taken from Fleming et al. (1990) and are approximated by polynomials for the purpose of our simulations. Such profiles will be introduced in the context of the respective simulations.

Energy and momentum budgets

To validate the conservation of energy and momentum we evaluate the energy and momentum budgets. The energy budget consists of three compartments, the enthalpy $\mathcal{H} = c_p \rho_r T$ of the background flow, the background kinetic energy $E_{BG,\,kin} = \rho_r U^2$, and the wave energy E_{GW}. The temporal evolution of these energies relative to the initial state can be written as

$$\Delta\mathcal{H}(t) = \mathcal{N}_E \left(\int\limits_0^{z_{max}} c_p \rho_r T(z,t)\, dz - \int\limits_0^{z_{max}} c_p \rho_r T(z,0)\, dz \right) \tag{4.25}$$

$$\Delta E_{GW}(t) = \mathcal{N}_E \left(\int\limits_0^{z_{max}} E_{GW}(z,t)\, dz - \int\limits_0^{z_{max}} E_{GW}(z,0)\, dz \right) \tag{4.26}$$

$$\Delta E_{BG,\,kin}(t) = \mathcal{N}_E \left(\int\limits_0^{z_{max}} \rho_r U^2(z,t)\, dz - \int\limits_0^{z_{max}} \rho_r U^2(z,0)\, dz \right), \tag{4.27}$$

where $\mathcal{N}_E$ is a normalization factor that is defined as

$$\begin{aligned} \mathcal{N}_E = \max\Big(& \int E_{GW}(z,t_1)\, dz - \int E_{GW}(z,0)\, dz, \\ & \ldots, \\ & \int E_{GW}(z,t_n)\, dz - \int E_{GW}(z,0)\, dz \Big) \end{aligned} \tag{4.28}$$

for each particular simulation. Here, $t_1, ldots, t_n$ denote the individual time steps of a simulation. The relative error of the total energy is then given by

$$\Delta E_{tot}(t) = \Delta\mathcal{H}(t) + \Delta E_{GW}(t) + \Delta E_{BG,\,kin}(t). \tag{4.29}$$

The momentum budget consists of two compartments, namely the wave momentum M_{GW} and the momentum of the background $M_{BG} = \rho_r U$. The temporal evolution of wave momentum and background momentum relative to the initial conditions can be written as

$$\Delta M_{GW}(t) = \mathcal{N}_M \left(\int\limits_0^{z_m max} M_{GW}(z,t)\, dz - \int\limits_0^{z_{max}} M_{GW}(z,0)\, dz \right) \tag{4.30}$$

$$\Delta M_{BG}(t) = \mathcal{N}_M \left(\int\limits_0^{z_{max}} \rho_r U(z,t)\, dz - \int\limits_0^{z_{max}} \rho_r U(z,0)\, dz \right), \tag{4.31}$$

where $\mathcal{N}_M$ is a normalization factor that is defined as

$$\begin{aligned} \mathcal{N}_M = \max \Bigg(& \int M_{GW}(z,t_1)\, dz - \int M_{GW}(z,0)\, dz, \\ & \ldots, \\ & \int M_{GW}(z,t_n)\, dz - \int M_{GW}(z,0)\, dz \Bigg). \end{aligned} \tag{4.32}$$

Again, the individual time steps of the simulations are $t_1, \ldots, t_n$. The change of total momentum is

$$\Delta M_{tot}(t) = \Delta M_{GW}(t) + \Delta M_{BG}(t). \tag{4.33}$$

That is, the momentum is conserved for $\Delta M_{tot}(t_i) = 0$ for $i = 1, \ldots, n$.

4.2 Wave packets in an isothermal atmosphere at rest

First we investigate the generation, propagation, and breaking of a wave packet in a highly idealized background atmosphere with

$$U(z,0) = 0 \tag{4.34}$$

$$T(z,0) = 240\,\mathrm{K}. \tag{4.35}$$

The background can only change with time due to the wave-mean flow interactions (2.38) and (2.46). For this case we shall assess the conventional and the correct energy and momentum budgets.

The temporal evolutions of the energy density E_{GW}, the characteristic vertical wavenumber m^*, and the diffusion coefficient $\mathcal{D}$ can be found in Figure 4.1, the temporal evolutions of the gravity wave drag M_{dep}, the wave momentum density tendency M_{tr}, and the horizontal wind U in Figure 4.2, and the temporal evolutions of the energy deposition E_{dep}, the energy density tendency E_{tr} and the background temperature T in Figure 4.3. In the following we describe the wave-mean flow interaction during the generation of the wave packet, the propagation and dissipation of the wave packet, and the energy and momentum budgets during the life-time of the wave packet.

Generation of the wave packet

During the first two hours, the time of wave generation, wave energy is taken from the enthalpy via the energy density tendency E_{tr}, which leads immediately to a cooling of the source region (Fig. 4.3 e and f). After two hours the heating by the energy density tendency and the cooling by the energy deposition roughly balance each other and the cold temperature in the source region prevails (Fig. 4.3 d, e, and f). The colder source region is

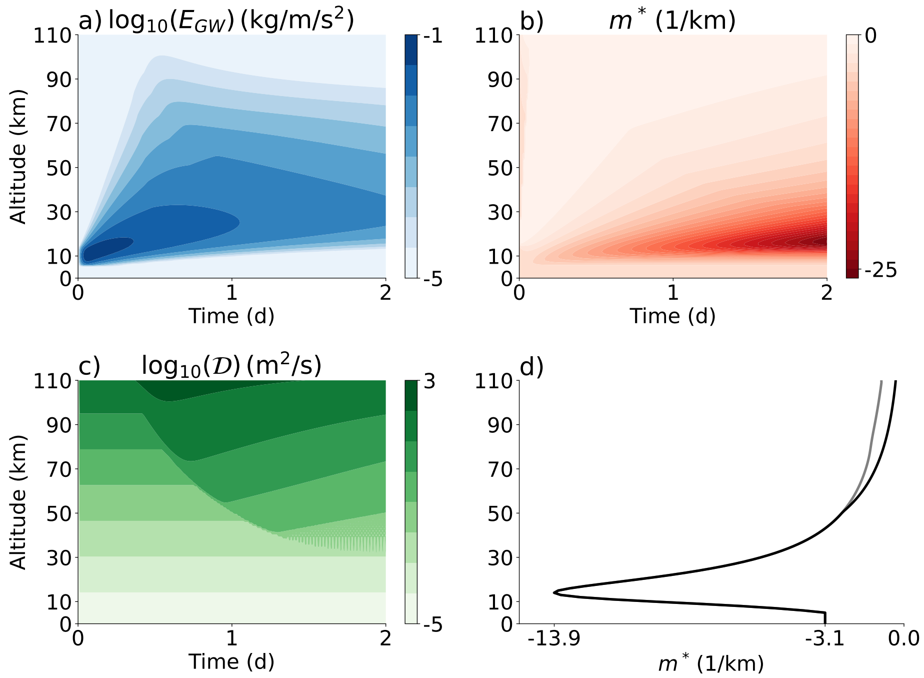

Figure 4.1: Time series of the logarithm of energy density (a), the characteristic vertical wavenumber (b), and the logarithm of the diffusion coefficient (c) for the life cycle of a wave packet in the case of a isothermal background at rest. Panel (d) shows the characteristic vertical wavenumbers with (black line) and without (gray line) diffusion at the time $t = 1\,\mathrm{d}$.

mainly the result of the negative energy deposition in the early stage of the simulation. This negative energy deposition is due to the vertical transport of potential energy by the wave packet away from the source region. The energy density tendency also contributes to the initial cooling in the source region.

There is also a transformation of horizontal background momentum to wave momentum during the generation of the waves via the momentum density tendency M_{tr}, but the amount of transferred momentum is negligibly small compared with the amount of momentum transported vertically and deposited at higher altitudes by the waves (Fig. 4.2 d and e). There is hardly any contribution of the momentum density tendency to the induced negative

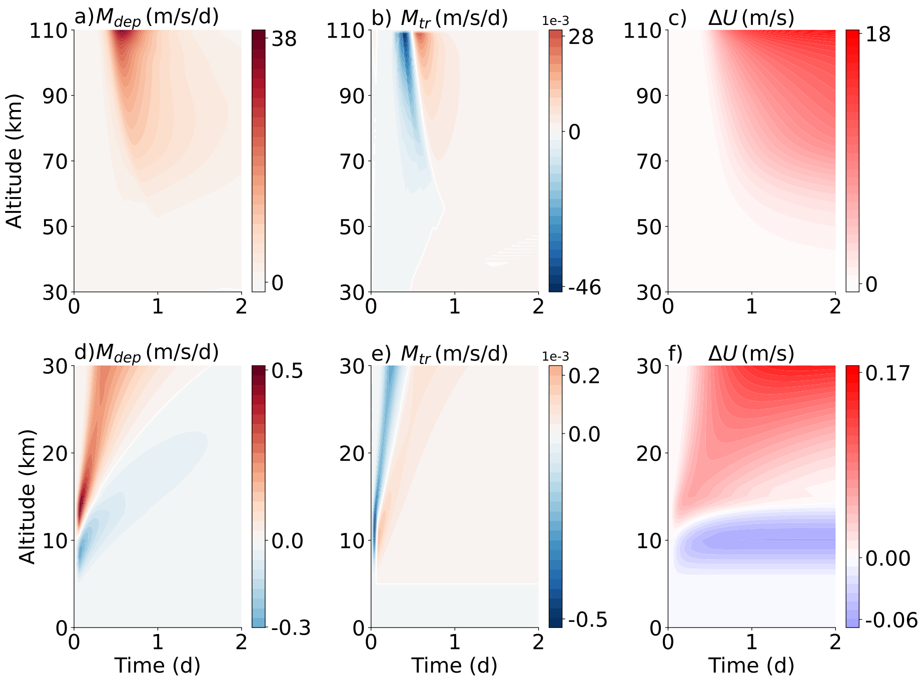

Figure 4.2: The time series of the gravity wave drag M_{dep} (a, d), the wave momentum density tendency M_{tr} (b, e), and the horizontal wind difference $\Delta U = U(t) - U(0)$ (c, f) during the propagation through a homogeneous background show the contributions of the momentum deposition and the wave momentum density tendency to the local wind variations. Top row: Time series for the upper model region between 30 km and 110 km; Bottom row: Time series for the lower model region between 0 km and 30 km. The two regions are separated for a better visualization. Note that the values for the momentum density tendency are scaled with the factor 1×10^3.

winds and therefore negative momentum in the source region (Fig. 4.2 f). The wind response is mainly caused by the vertical transport of positive horizontal momentum.

Propagation and dissipation of the wave packet

Specifications of the source spectrum allows for the generation of a broad spectrum of gravity waves with different vertical group velocities. This causes the wave packet to spread in the vertical with time, where the fast waves reach the model top within half a day while the slower waves barely propagate within the time frame of Figure 4.1 a. Using the characteristic vertical wavenumber m^* (Fig. 4.1 b), we can estimate the predominant intrinsic vertical group velocity. A wavenumber m^* close to zero corresponds to waves with large vertical group velocities while a large absolute value of m^* corresponds to small vertical group velocities.

The diffusion coefficient $\mathcal{D}$ consists of the time-independent kinematic molecular viscosity coefficient $\mathcal{D}_{mol}$, which increases exponentially with height, and the saturation diffusion coefficient $\mathcal{D}_{sat}$ (Fig. 4.1 c), which depends on time and height. There are two reasons why the saturation condition (3.46) at a certain altitude and time is fulfilled and a nonzero saturation diffusion coefficient occurs. Either the energy density is large enough at high altitudes (e.g., see the first day of the simulation in Figure 4.1 a and c) or the absolute value of the wavenumber m^* is large (Fig. 4.1 b and c after the first day of the simulation). The diffusion acts on waves with small vertical wave lengths (large absolute value of the vertical wavenumber m) stronger than on waves with larger vertical wave length (see Eq. 3.3). This leads to a shifting of the wavenumber m^* closer to zero compared with a simulation without any diffusion (Fig 4.1 d).

While the wave packet propagates through the initially isothermal background at rest, the waves deposit positive horizontal momentum (taken from the background at the source level) into the background (the wind variation caused by the wave momentum tendency M_{tr} plays only a minor role) and causes a positive acceleration of the wind (Fig. 4.2 a, b, and c). The wave-induced wind refracts the following waves to smaller vertical wave lengths. As a result, these waves become saturated which intensifies the momentum depositions and accelerates the wind even further and so forth.

The refraction of the waves by the wave-induced wind and the diffusion act in opposite direction on the wavenumber m^*. The wave-induced wind

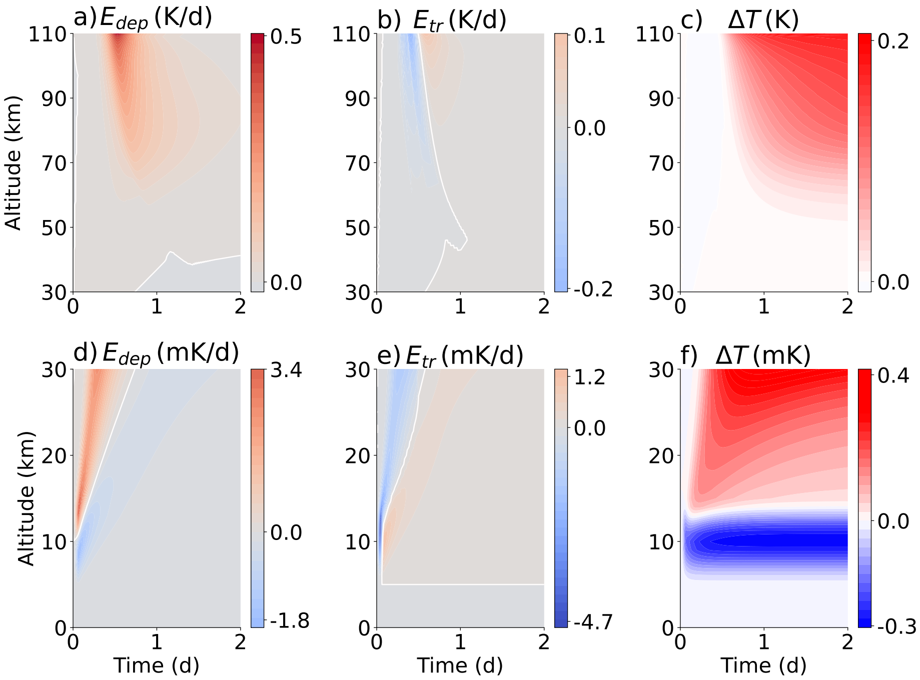

Figure 4.3: The time series of the energy deposition E_{dep} (a, d), the energy density tendency E_{tr} (b, e), and the temperature difference $\Delta T = T(t) - T(0)$ (c, f) during the propagation through a homogeneous background show the contributions of the energy deposition and the energy density tendency to the local temperature variations. Top row: Time series for the upper model region between 30 km and 110 km; Bottom row: Time series for the lower model region between 0 km and 30 km. The two regions are separated for a better visualization. The value-to-color codes of (a) and (b), and for (d) and (e) are the same.

refracts the wavenumber m^* to larger absolute values (Fig. 4.1 d, gray line). The diffusion acts stronger on these waves and damps them, which pushes m^* again closer to zero (Fig. 4.1 d, black line).

As explained in Section 2.3, the wave-mean flow interactions of gravity waves consist not only of the acceleration or deceleration of the mean flow that result from momentum deposition and the transient term M_{tr} introduced in this theses (see Eq. 2.38), but also of heating or cooling of the mean temperature (enthalpy) field that result from energy deposition and the transient term E_{tr} introduced in this study (see Eq. 2.46). Additionally, modelers usually apply the vertical diffusion coefficient from gravity wave schemes to the mean flow. These mean-flow diffusion tendencies are additional gravity-wave related effects on the mean flow. The upper panels of Figure 4.3 show that the heating of the temperature field in higher altitudes is a result from the energy deposition and the transient term, where the energy deposition is the main contributor.

Energy and momentum budgets

To estimate the importance of the new transient wave-mean flow interaction terms introduced in this thesis to the conventional approach, we performed an additional simulation where the tendencies of the momentum and energy densities were neglected in the mean-flow equations. The energy and momentum budgets from both simulations are shown in Figure 4.4.

From the dashed black lines in Figure 4.4 we conclude that neither energy nor momentum are conserved, whether just the conventional wave-mean flow interactions are considered (Fig. 4.4 c and d) or whether also the newly introduced transient terms (Fig. 4.4 a and b, where the energy loss is roughly 1/1000th of the wave energy, which is not visible in a) are included in the computations.

When only the conventional wave-mean flow interaction terms are taken into account, we get an increase of total energy and total momentum with increasing wave energy and wave momentum, respectively. The wave energy and wave momentum come out of nowhere during the generation of the wave packet.

For the new, transient wave-mean flow interactions we see that the increase of wave energy and the decrease of mean-flow enthalpy are of the same amount during the generation of the wave packet and total energy is conserved. Hence, during the generation of the waves, enthalpy is converted to

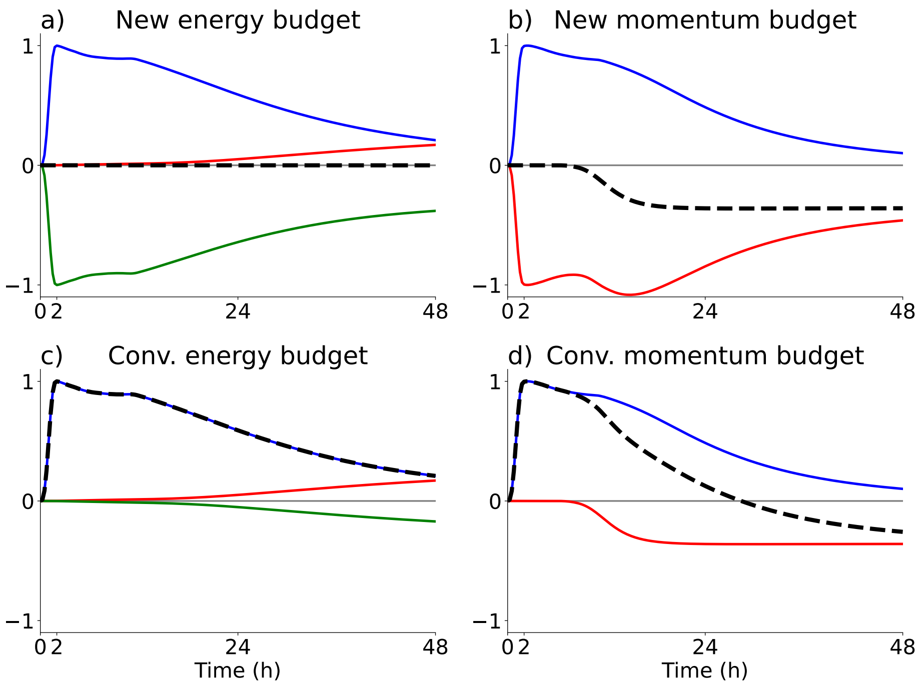

Figure 4.4: Temporal evolution of vertically integrated energies (a, c) and momenta (b, d) for the idealized example of a wave packet propagating through a homogeneous background (with zero initial mean wind). The initial state energies and momenta are subtracted and the resulting deviations from the initial states are normalized with the maximum changes of the wave energy and the wave momentum. Top row: Budgets with the new transient wave-mean flow interactions. Bottom row: Budgets with the conventional wave-mean flow interactions. a) and c): Change of wave energy (blue line), change of background kinetic energy (red line), change of enthalpy (green line), and change of total energy (dashed black line); b) and d): change of wave momentum (blue line), change of background momentum (red line), and change of total momentum (dashed black line).

wave energy. Similar effects are visible by looking at the momentum budgets in the new, transient case. There, the background momentum is converted to wave momentum (Fig. 4.4 b).

However, after around six hours, the momentum (and energy) conservation is somewhat violated when the new transient interactions are included. The reason is that the diffusion is not strong enough to damp the whole wave packet until it reaches the upper layer of the model. This leads to vertical transport of momentum (and energy) out of the system at the upper model layer. While the loss of momentum is on the order of the wave momentum (Fig. 4.4 b), the corresponding loss of energy is roughly 1/1000th of the maximum wave energy (which is not visible in Fig. 4.4 a).

To prevent the loss of energy and momentum at the upper layer the whole wave packet needs to be strongly damped below the upper layer. In the real atmosphere the reason for the gravity waves to break in the MLT region are strong wind variations and shear (e.g., Becker and Vadas 2018, Fig. 13). For this purpose we applied a strong horizontal wind shear in the upper model domain for an additional critical layer absorption. The wind $U_{step}(z)$ has the parameter $U_0 = 0$, $U_{max} = 60\,\mathrm{m/s}$, $z_U = 90\,\mathrm{km}$, and $\Delta z_U = 20\,\mathrm{km}$ (see Eq. 4.22). Hence, with this critical layer absorbtion the energy and momentum are conserved over the whole simulation period for the new, transient wave-mean flow interaction terms (Fig. 4.5). The remaining small numerical error (not visible in Fig. 4.5) can be further reduced by smaller time steps.

From the energy budget we can see that enthalpy is converted to wave energy during the generation of the wave packet (Fig. 4.5). When the waves are breaking, however, the wave is converted to background kinetic energy and enthalpy. There is a net conversion of enthalpy to background kinetic energy for this transient wave event. Note that the net energy conversions in steady state for gravity waves dissipating in the mesosphere are different (e.g., Becker and McLandress 2009). Such differences emphasize the importance of a transient gravity wave scheme in order to improve the simulation of gravity wave events in models (Muraschko et al. 2015; Bölöni et al. 2016; Kim et al. 2021).

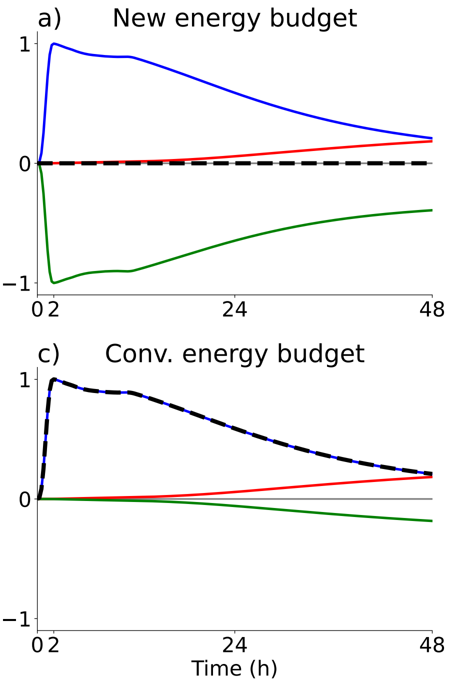

Figure 4.5: Same as Fig. 4.4, but for a simulation with additional critical layer absorption of gravity waves in the upper model domain that prevent spurious loss of wave energy and momentum at the upper model layer (see text for further details).

4.3 Wave packets in a wind shear

Now we investigate the behaviour of a wave packet propagating into two different wind jets and the resulting wave-mean flow interactions. The first wind jet points in the same direction as the horizontal phase speeds of the wave packet. For this case the wind jet is defined to be positive. The second wind jet points in the opposite direction of the horizontal propagation direction of the wave packet. This wind jet is defined to be negative.

Positive wind jet

When a monochromatic wave with positive horizontal wavenumber propagates into a positive wind jet, the wave encounters a critical layer when the jet amplitude is large enough. This is the case when the ground-based horizontal phase speed is equal the local horizontal wind:

$$c_{ph} - U(z) = 0 \tag{4.36}$$

with

$$c_{ph} = \frac{\omega_I}{k} + U_0, \tag{4.37}$$

where U_0 and ω_I are the horizontal wind and the intrinsic frequency at some launch level, respectively. The more the local wind becomes comparable to the ground-based phase velocity, the more the vertical wave length tends to zero and the vertical group velocity slows down. When (4.36) is fulfilled, the vertical group velocity becomes zero, the vertical wavenumber diverges, and the wave stops to propagate vertically any further. Note that since the vertical gradients of the gravity-wave horizontal wind and temperature perturbations increase dramatically close to a critical level, any viscosity will dissipate a monochromatic wave in the idealized linear case. Accordingly, Lindzen (1981) assumed that a critical level leads to the total attenuation of a monochromatic wave. We apply this assumption to the spectral elements of our transient gravity wave scheme that encounter a critical level (see Sec. 3.7).

To validate our transient gravity wave scheme with regard to the behaviour of a wave packet encountering a critical layer, we apply an isothermal initial background with a temperature of $T(z, 0) = 240\,\mathrm{K}$ and a positive wind jet of $U(z, 0) = U_{jet}(z)$ with $U_0 = 0$, $U_{max} = 20\,\frac{\mathrm{m}}{\mathrm{s}}$, $z_0 = 30\,\mathrm{km}$, and $\Delta z_U = 10\,\mathrm{km}$ (see Eq. 4.23). Furthermore we apply an additional critical layer absorption in the upper model layers to conserve the energy and momentum (see Sec. 4.2). This critical layer regime has the parameters $U(z, 0) = U_{step}(z)$ with $U_0 = 0$, $U_{max} = 60\,\mathrm{m/s}$, $z_U = 90\,\mathrm{km}$, and $\Delta z_U = 20\,\mathrm{km}$ (see Eq. 4.22).

When the wave packet propagates into the wind jet, no wave energy reaches the altitudes above the jet (Fig. 4.6 a and c). The whole wave packet is attenuated due to critical layers. The characteristic vertical wavenumber m^* tends to large absolute values below the jet maximum (Fig. 4.6 b and d). Not all spectral elements of the broad spectrum reach a critical layer at the same time and altitude.

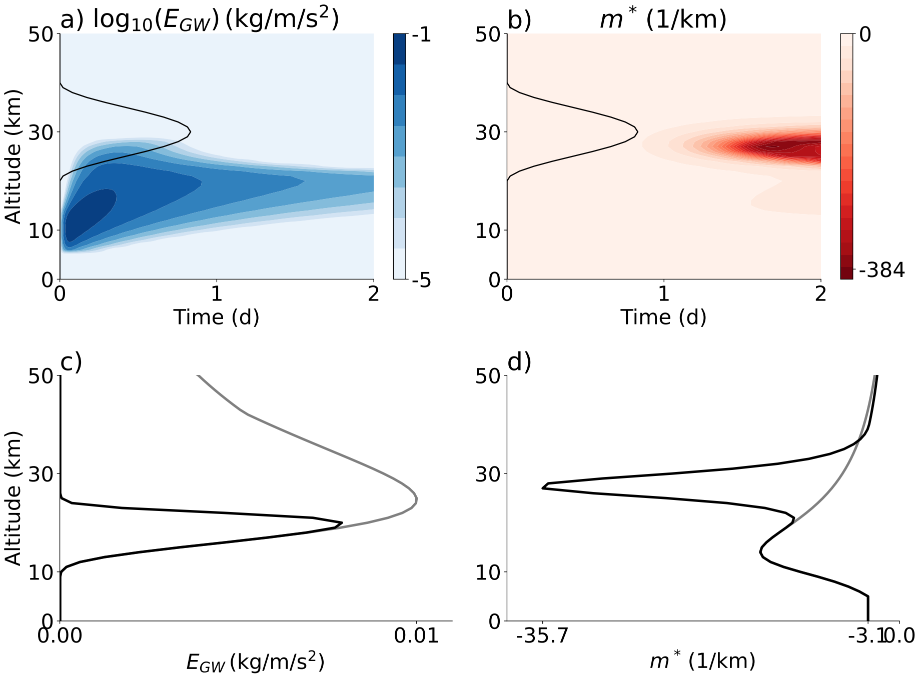

Figure 4.6: Time series of the logarithm of energy density (a) and the characteristic vertical wavenumber (b) for the life cycle of a wave packet in the case of a positive wind jet (indicated by black lines). Energy densities (c) and characteristic vertical wavenumbers (d) in the case of a positive wind jet (black line) and in the case of an isothermal atmosphere at rest (gray line) at the time $t = 1\,\mathrm{d}$. The wave packet is absorbed due to critical layers.

As described in Section 3.7, the attenuation of spectral elements due to critical layers is accounted for in our transient gravity wave scheme via the removal of these elements if their vertical wavenumber is less than the lower wavenumber limit m_1 (Eq. 3.58). More specifically, the characteristic vertical wavenumber m^* and the effective spectral area are shifted by the wind (a sketch for the effective spectral area can be found on the right hand side of Fig. 3.1). All waves that with vertical wavenumbers less than the limit m_1 are omitted. The upward momentum and energy fluxes from these spectral elements are deposited in into the mean flow. In fact, there is a flux of spectral wave elements (hereafter: spectral flux) beyond the lower wavenumber limit m_1 from both saturation and attenuation of waves due to critical layers. The results confirm that the new RTP can well simulate the behaviour of spectral elements encountering critical layer.

The wave-mean flow interactions and the changes of the horizontal wind and temperature can be found in the Figures 4.7 and 4.8, respectively. The energy and momentum budgets can be found in Figure 4.9.

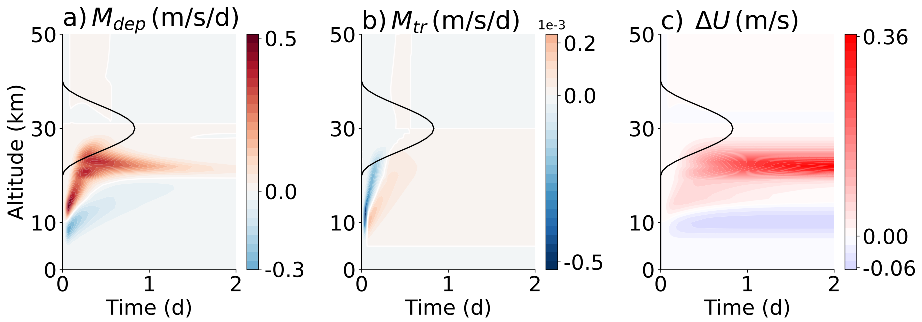

Figure 4.7: Time series of the gravity wave drag M_{dep} (a), the wave momentum density tendency M_{tr} (b), and the horizontal wind difference $\Delta U = U(t) - U(0)$ (c) for the life cycle of a wave packet in the case of a positive wind jet (indicated by black lines). Note that the values for the wave momentum density tendency are scaled with the factor 1×10^3. The gravity wave drag is the main contributor of the mean wind response.

The attenuation of spectral elements by critical layers and saturation beneath the jet gives rise to acceleration and heating of the background (Figs. 4.7 c and 4.8 c). The kinetic energy for the accelerated jet is extracted from the wave energy, which in turn is generated earlier in the simulation from the conversion of mean-flow enthalpy as a result of the transient source function (Fig. 4.9). So there is in total a conversion from enthalpy into

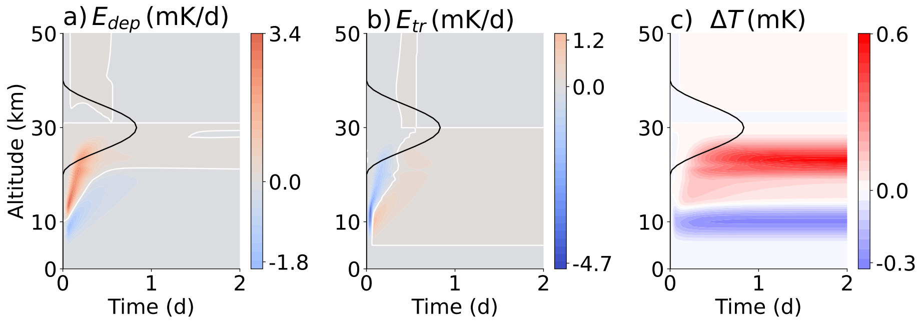

Figure 4.8: Time series of the energy deposition E_{dep} (a), the energy density tendency E_{tr} (b), and the temperature difference $\Delta T = T(t) - T(0)$ (c) for the life cycle of a wave packet in the case of a positive wind jet (indicated by black lines). The value-to-color codes of (a) and (b) are the same. The main contributor to the mean temperature response in the jet region is the energy deposition.

wave energy and then into mean-flow kinetic energy. These energy conversion are somewhat reminiscent of the energy conversions during an idealized

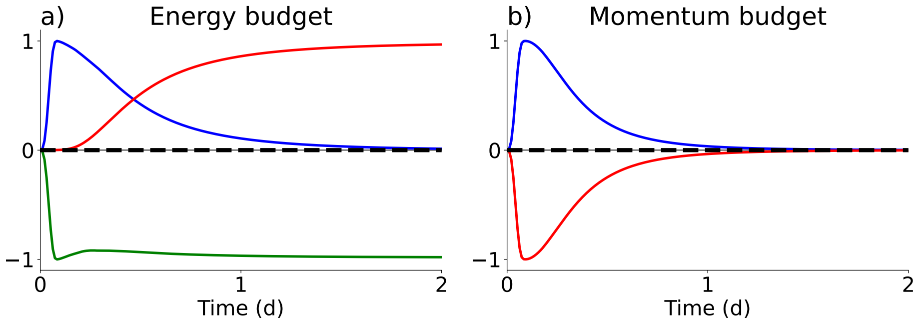

Figure 4.9: Temporal evolution of vertically integrated energies (a) and momenta (b) for the idealized example of a wave packet propagating through a positive wind jet (in an isothermal atmosphere). The initial state energies and momenta are subtracted and the resulting deviations from the initial states are normalized with the maximum changes of the wave energy and the wave momentum. a) Change of wave energy (blue line), change of background kinetic energy (red line), change of enthalpy (green line), and change of total energy (dashed black line); b) change of wave momentum (blue line), change of background momentum (red line), and change of total momentum (dashed black line).

baroclinic life cycle (Hoskins and Simmons 1975). Note, however, that no explicit available potential energy of the mean flow is involved in our case. The initial generation of wave energy (potential and kinetic energy of the waves are generated simultaneously) is induced by our prescribed sources function. Also note that this energetics for gravity wave generation is distinct from the energetics proposed in Becker et al. (2022a) for the spontaneous emission. In that case, wave energy is extracted from the vertical shear of the quasi-geostrophic mean flow.

Negative wind jet

Gravity waves can be reflected when they encounter a wind jet that is opposite to their horizontal propagation direction. If the wind jet is strong enough, the WKB theory predicts for a monochromatic wave that the squared vertical wavenumber can become zero and change sign, which leads to the reflection of the wave (see e.g., Sutherland 2018). For the hydrostatic dispersion relation as used in RTP, the intrinsic frequency ω_I can become rather larger than the Brunt-Väisälä frequency N than that the vertical wavenumber m becomes zero (see Eqs. 3.62 and 3.63). From equation (3.62) we can estimate the wind amplitude, where the intrinsic frequency is equal the Brunt-Väisälä frequency

$$|U_{a+1}| = \frac{N}{|-\sqrt{m_a^2 + \frac{1}{4H^2}}|}\left(\frac{N}{\omega_{I,a}} + 1\right), \tag{4.38}$$

when we set $\omega_{I,a+1} = N$ and $U_a = 0$. $\omega_{I,a}$ and m_a are the intrinsic frequency and vertical wavenumber in the source region. We assume that spectral elements of the wave packet with intrinsic frequencies larger than the Brunt-Väisälä frequency are subject to reflection. However, the RTP spectrum has an upper frequency limit of $\omega_T < N$ (waves close to N are not mid-frequency waves).

Similar to our method to account for critical layers we assume in the case of wave reflection that every spectral element which exceeds the upper frequency limit is going to be reflected (see also the left hand side of Fig. 3.1). These waves are removed from the effective spectral area, giving rise to energy and momentum deposition. This wave-mean flow interaction from wave reflection is part of the new parameterization, because downward propagating gravity waves are not allowed. In a more comprehensive scheme, the omitted waves should be subject to a downward propagating branch, but such an extension is beyond the scope of this thesis.

To investigate the behaviour of our parameterization for a wave packet encountering a jet opposite to the horizontal direction of wave propagation we apply an initial background with a constant temperature of $T(z,0) = 240\,\mathrm{K}$ and a negative wind jet of $U(z,0) = U_{peak}$ with $U_0 = 0$, $U_{max} = -20\,\frac{\mathrm{m}}{\mathrm{s}}$, $z_0 = 30\,\mathrm{km}$, and $\Delta z_U = 10\,\mathrm{km}$ (see Eq. 4.23). An additional the critical layer absorption is applied in the upper model layers to conserve energy and momentum (see Sec. 4.2). This critical layer regime has the parameters $U(z,0) = U_{step}(z)$ with $U_0 = 0$, $U_{max} = 60\,\mathrm{m/s}$, $z_U = 90\,\mathrm{km}$, and $\Delta z_U = 20\,\mathrm{km}$ (see Eq. 4.22).

The response of the wave packet to the negative jet can be found in Figure 4.10. When the wave packet propagates into the jet, the wave energy is reduced and when the wave packet leaves the jet, the wave packet regains energy (Fig. 4.10 a and c). Simultaneously the characteristic vertical wavenumber tends to small absolute values in the vicinity of the jet maximum and tends back to larger absolute values while leaving the jet (Fig. 4.10 b and d). A smaller absolute value of the vertical wavenumber corresponds to a larger vertical group velocity $\dot{z}$ (see Eq. 2.24). The response of the wavepacket to the negative jet is, that the vertical group velocity increases with increase of the jet and the wavepacket is stretched. This leads to the reduction of energy density until the jet maximum. Above the jet maximum the vertical group velocity decreases and the wave packet is compressed. This leads to an increased energy density. However, because wave reflection has caused that spectral elements were removed from the effective spectral area, the

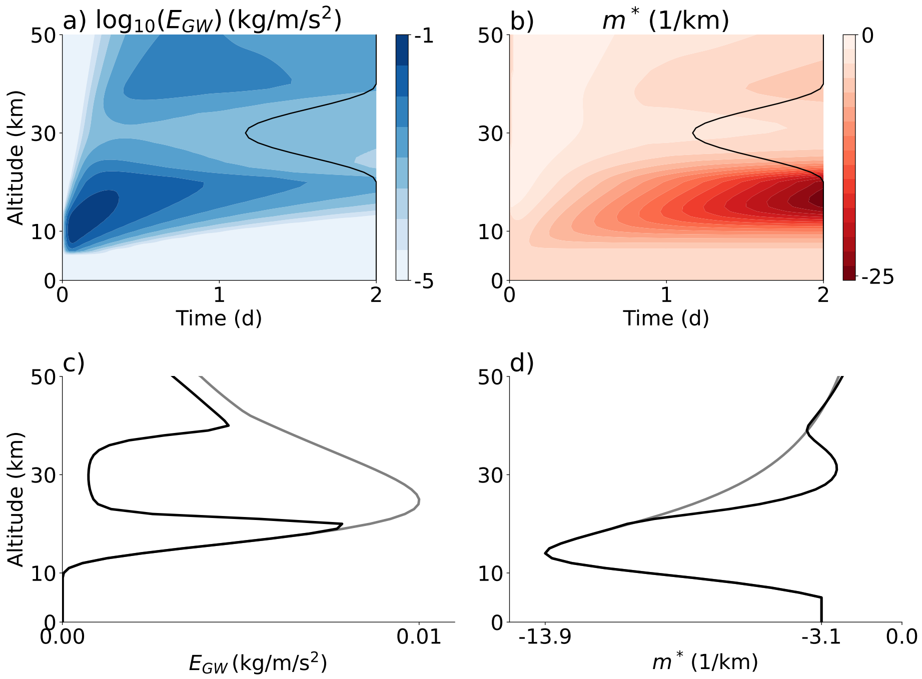

Figure 4.10: Time series of the logarithm of energy density (a) and the characteristic vertical wavenumber (b) for the life cycle of a wave packet in the case of a negative wind jet (indicated by black lines). Energy densities (c) and characteristic vertical wavenumbers (d) in the case of a negative wind jet (black line) and in the case of an isothermal atmosphere at rest (gray line) at the time $t = 1\,\mathrm{d}$. The wave packet is refracted and partly reflected.

wave energy density above the jet is smaller than that below the jet. This is further illustrated in Figure 4.10 c which shows the loss of wave energy compared with the case of a wave packet in an isothermal background at rest (see also Fig. 4.1).

The conventional and new, transient wave-mean flow interactions and the response of the mean wind are shown in Figure 4.11 and wave-mean flow interactions and the mean temperature response are shown in Figure 4.12. The wave packet deposits momentum, while it propagates into the jet (Fig. 4.11 a). This leads to a deceleration of the jet by the wave packet (Fig. 4.11 c). Above the jet maximum the opposite happens. The

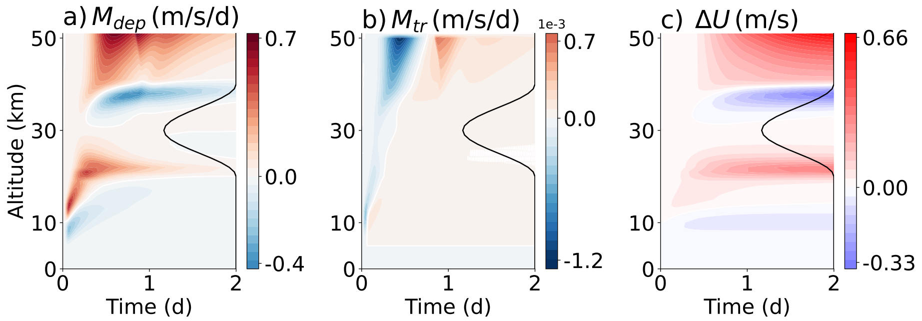

Figure 4.11: Time series of the gravity wave drag M_{dep} (a), the wave momentum density tendency M_{tr} (b), and the horizontal wind difference $\Delta U = U(t) - U(0)$ (c) for the life cycle of a wave packet in the case of a negative wind jet (indicated by black lines). Note that the values for the momentum density tendency are scaled with the factor 1×10^3.

wave packet takes momentum from the mean wind and the jet is accelerated by the wave packet. The momentum density tendency does not play a role here (Fig. 4.11 b).

Similar to the momentum deposition there is a energy deposition, when the wave packet propagates into the jet (Fig. 4.11 a) and the mean temperature increases (Fig. 4.11 c). When the wave packet leaves the jet, where it regains its energy, it also takes potential energy from the background. This leads to an decreased mean temperature in that region. The transient energy deposition does not contribute to the cooling and heating in the jet region (Fig. 4.12 b).

From the energy budget (Fig. 4.13) we can see, that in the end there is a conversion of background kinetic energy to enthalpy. This conversion takes place in the region where the jet is decelerated. This is consistent with the fact that a reduced amplitude of the mean wind means a reduced mean-flow kinetic energy. In the same region the temperature increases and so does the enthalpy. Higher up where the jet is accelerated, the wind and temperature changed are opposite. However, due to the exponential decrease of density the amount of converted energy is smaller such that background

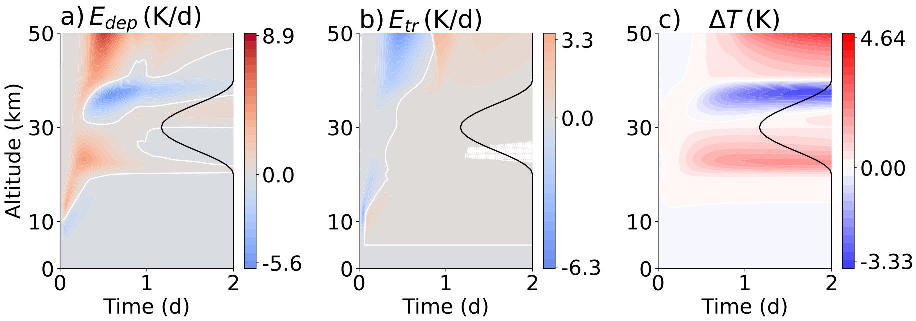

Figure 4.12: Time series of the energy deposition E_{dep} (a), the energy density tendency E_{tr} (b), and the temperature difference $\Delta T = T(t) - T(0)$ (c) for the life cycle of a wave packet in the case of a negative wind jet (indicated by black lines). The value-to-color codes of (a) and (b) are the same.

kinetic energy is converted to enthalpy when integrating over the vertical model domain.

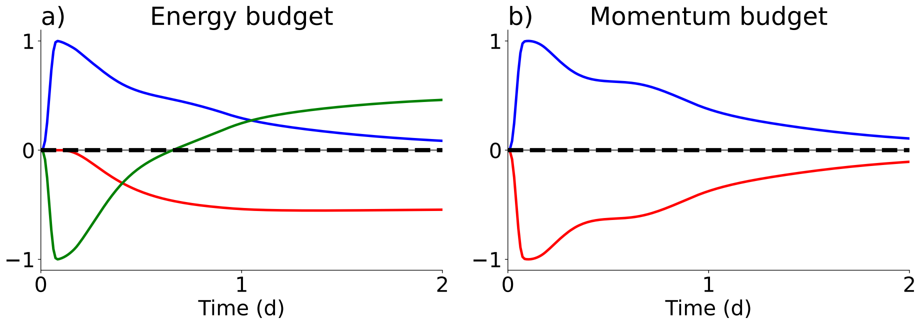

Figure 4.13: Temporal evolution of vertically integrated energies (a) and momenta (b) for the idealized example of a wave packet propagating through a negative wind jet (in an isothermal atmosphere). The initial state energies and momenta are subtracted and the resulting deviations from the initial states are normalized with the maximum changes of the wave energy and the wave momentum. a) Change of wave energy (blue line), change of background kinetic energy (red line), change of enthalpy (green line), and change of total energy (dashed black line); b) change of wave momentum (blue line), change of background momentum (red line), and change of total momentum (dashed black line).

4.4 Wave packets in a temperature gradient

The next initial background condition is a homogeneous wind of $U(z,0) = 0$ with the critical layer regime and a vertical temperature gradient. The temperature profile is a step function $T(z,0) = T_{step}(z)$ (see Eq. 4.24) with $T_0 = 240\,\mathrm{K}$, $T_{max} = 60\,\mathrm{K}$, $z_0 = 30\,\mathrm{km}$, and $\Delta z_T = 20\,\mathrm{km}$. A vertical temperature gradient also leads to a vertically varying scale height H and Brunt-Väisälä frequency N (see Eqs. 2.5 and 2.6).

The results for the energy density and the characteristic vertical wavenumber are shown in Figure 4.14. There is no visible difference for the time series of the energy density or the characteristic vertical wavenumber between the cases of a temperature gradient and a homogeneous background (Figs. 4.14 and 4.1). Only by looking at the differences of the energy densities and the characteristic vertical wavenumbers of the two cases at the time $t = 1\,\mathrm{d}$ (black and gray line in Fig. 4.14 c and d) we can discern small differences between these two simulations. An increasing Brunt-Väisälä frequency leads to a deceleration of the vertical propagation of the wave packet (increase of $|m^*|$, decrease of $\dot{z}$), resulting in a larger energy density below 30 km when compared to the case with isothermal initial condition.

The Figures 4.15 and 4.16 show the the temporal evolution of the wave-mean flow interactions terms and the resulting mean wind and temperature changes. When compared to the former case with an isothermal initial atmosphere at rest (Sec. 4.2, Figs. 4.2 and 4.3), the wave-mean flow interactions in layers of larger buoyancy frequencies (dashed lines in Figs. 4.15 and 4.16) results in stronger mean-flow responses beneath the maximum of N and weaker responses above the maximum.

The change of the Brunt-Väisälä frequency N with altitude together with the single column approximation (constant horizontal wavenumber k) and the dispersion relation (Eq. 2.17) implies a vertical refraction of the wave packet that is independent from the background wind. More specifically,

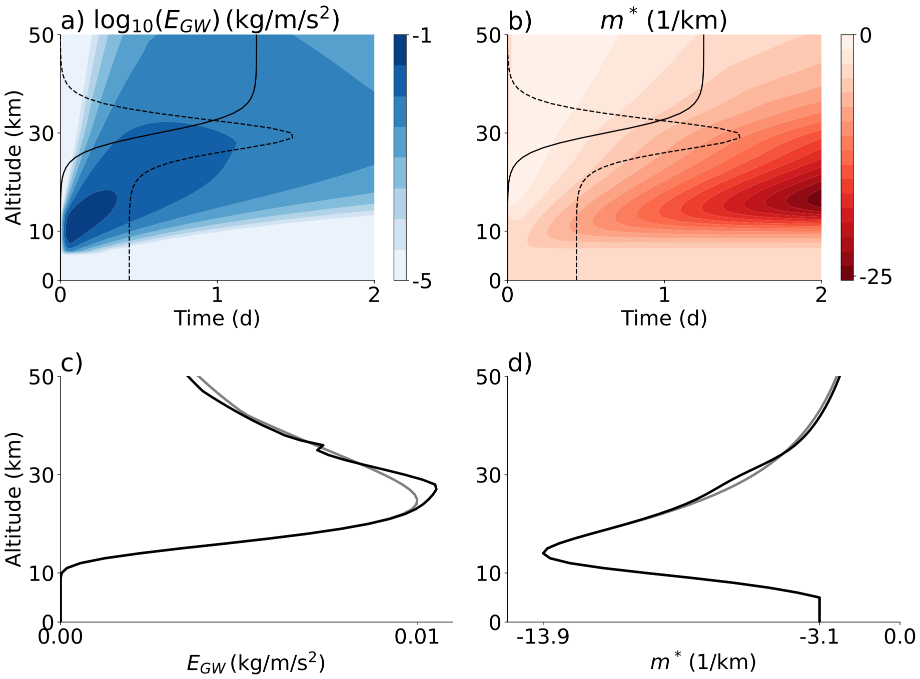

Figure 4.14: Time series of the logarithm of energy density (a) and the characteristic vertical wavenumber (b) for the life cycle of a wave packet for the case of a positive initial vertical temperature gradient (indicated by solid black lines) and zero initial mean flow. The dashed black lines indicate the Brunt-Väisälä frequency. Energy densities (c) and characteristic vertical wavenumbers (d) for the case of a positive initial vertical temperature gradient and zero initial mean flow (black line) and for the case of an isothermal atmosphere at rest (gray line) at the time $t = 1\,\mathrm{d}$.

according to the dispersion relation (Eq. 3.63), an increase of N^2 with altitude results in a corresponding increase of $(m^2 + \frac{1}{4H^2})\omega_I^2$. As long as any wave-induced changes of the mean wind are negligible, ω_I does not change. Hence, m^2 has to assume larger values, which is visible as an increase of the absolute characteristic vertical wavenumber m^* in Figure 4.14 b and d. The whole spectrum is shifted to smaller wave lengths and a spectral flux over the lower wavenumber limit m_1 arises. Similar to the positive wind jet case (see Sec. 4.3, Figs. 4.7 and 4.8) this causes the deposition of energy and momentum in the altitude regime where $\partial_z N > 0$. At higher

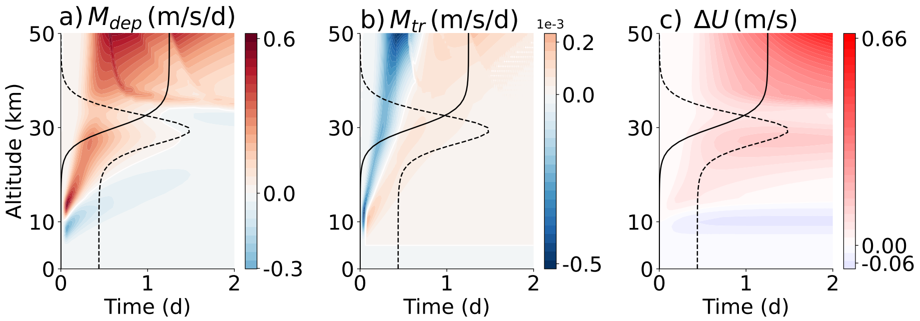

Figure 4.15: Time series of the gravity wave drag M_{dep} (a), the wave momentum density tendency M_{tr} (b), and the horizontal wind difference $\Delta U = U(t) - U(0)$ (c) for the life cycle of a wave packet in the case of a positive temperature gradient (indicated by solid black lines). The dashed black lines indicate the Brunt-Väisälä frequency. Note that the values for the momentum density tendency are scaled with the factor 1×10^3.

altitudes where N decreases with z, the characteristic vertical wavenumber and, hence, the whole spectrum, is shifted toward smaller absolute vertical wavenumbers (larger wave lengths). As a result, the spectral flux ceases and the wave-mean flow interactions are weaker than below the maximum of N.

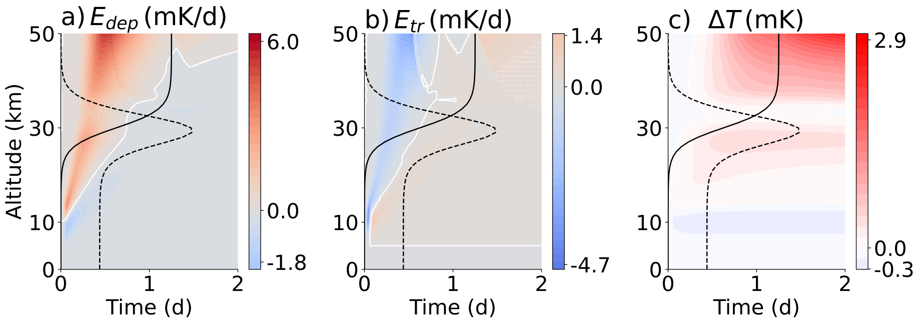

Figure 4.16: Time series of the energy deposition E_{dep} (a), the energy density tendency E_{tr} (b), and the temperature difference $\Delta T = T(t) - T(0)$ (c) for the life cycle of a wave packet in the case of a positive temperature gradient (indicated by solid black lines). The dashed black lines indicate the Brunt-Väisälä frequency. The value-to-color codes of (a) and (b) are the same.

If the N^2 decreases further with altitude, then there is a point where the product $\left(m^2 + \frac{1}{4H^2}\right)\omega_I^2$ can not decrease anymore unless the vertical wavenumber m^2 becomes negative. This is the point where the wave is reflected. We simulated such a case by using an initial temperature profile that has negative vertical gradient in the stratosphere (see Fig. 4.17, we used $T_0 = 300\,\mathrm{K}$ and $T_{max} = -60\,\mathrm{K}$ with respect to Eq. 4.4)). The results from this simulation show that the wave-mean flow interaction caused by our method to account for wave reflection (see Sec. 4.3) are comparatively small. In summary, the influence of temperature variations on the wave packet is of minor importance compared with the influence of a horizontal wind shear.

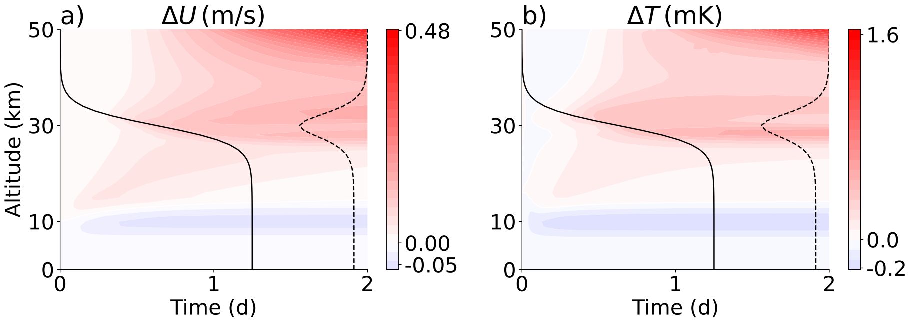

Figure 4.17: Time series of the horizontal wind difference $\Delta U = U(t) - U(0)$ (a) and temperature difference $\Delta T = T(t) - T(0)$ (b) for the life cycle of a wave packet in the case of a negative temperature gradient (indicated by solid black lines). The dashed black lines indicate the Brunt-Väisälä frequency.

4.5 Wave packets in realistic temperature and wind profiles

In this section we present and discuss simulation results for the propagation of two wave packets with horizontal intrinsic phase velocities in opposite directions. One wave packet has positive intrinsic horizontal phase velocities and the other one negative (hereafter: eastward and westward waves). These wave packets propagate through but do not interact with January mean profiles of zonal wind and temperature at 60°S and 60°N taken from CIRA86 data (see Sec. 4.1). The diffusion coefficient $\mathcal{D}_{sat}$ is calculated from the superposition of both wave packets and is the same for both (see Eq. 3.49). The interference parameter is set to $\gamma = 0.2$. Note that for this current formulation of the RTP with two horizontal azimuths, the relation between the intrinsic phase velocity and the horizontal wind needs to be considered in the Equations (3.20) and (3.21), as well as for the calculation of the effective spectral area for each azimuth (Sec. 3.7).

The simulations presented in the following are offline simulations, that is, wave-mean flow interactions are not taken into account, and the background wind and temperature profiles are fixed. We test the RTP for two background wind and temperature profiles that corresponds to the zonal-mean state at 60°S and at 60°N, respectively, during January. We compare the results with the expected behaviour according to figure 5 c in Becker (2012).

In the summer hemisphere we expect from Becker (2012) that the westward gravity waves are filtered in the lower stratosphere and that the eastward gravity waves propagate to the upper mesosphere where they dissipate and deposit momentum. The resulting wave-mean flow interactions of the RTP simulation in 60°S January zonal-mean background wind and temperature profiles are shown in Figure 4.18. The westward wave packet encounters critical layers already in the westward stratospheric jet and provides a negligible amount of deposited energy and momentum. The eastward wave packet propagates through the westward stratospheric jet into the MLT region. There the waves become saturated above 75 km altitude and reach their critical layers around an altitude of 90 km. They provide a temporarily momentum deposition of around 300 m/s/d and energy deposition of around 5 K/d. These results are in agreement with the expected behaviour.

In the winter hemisphere we expect from Becker (2012) that the eastward gravity waves are filtered in the lower stratosphere and that the west-

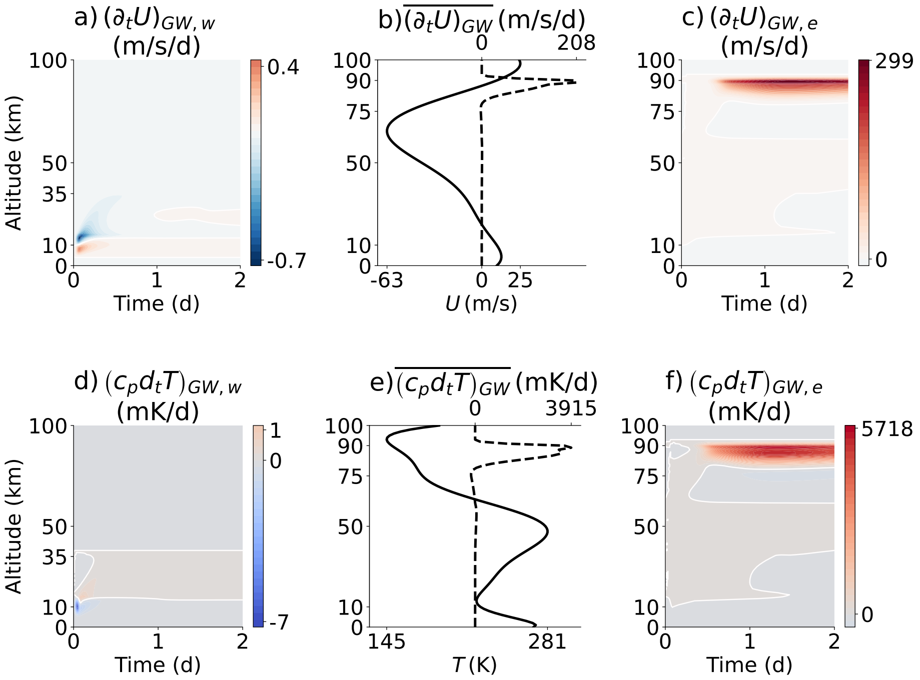

Figure 4.18: Time series of the wave-mean flow interaction $(\partial_t U)_{GW}$ (gravity wave drag plus wave momentum density tendency (see Eq. 2.38) of the westward (a) and the eastward (c) wave packet, and time series of wave-mean flow interaction $(c_p d_t T)_{GW}$ (energy deposition plus wave energy density tendency, see Eq. 2.46) of the westward (d) and the eastward (f) wave packet for a background wind and temperature profile that corresponds to the zonal-mean state at 60°S during January. The middle column shows the zonal wind and the time averaged wave-mean flow interaction $(\partial_t U)_{GW}$ (b) and the temperature and the time averaged wave-mean flow interaction $(c_p d_t T)_{GW}$ (e). The time averages are defined as $\overline{X} = \frac{1}{t_{max}} \int\limits_0^{t_{max}} X(t)\,dt$, where $t_{max} = 2\,\text{days}$.

ward gravity waves propagate to the upper mesosphere where they dissipate and deposit momentum. The resulting wave-mean flow interactions of the RTP simulation in 60°N January zonal-mean background wind and temperature profiles are shown in Figure 4.19. We can see that only the westward waves contribute to the wave-mean flow interactions in the MLT region. The eastward wave packet encounters critical layers in the troposphere and

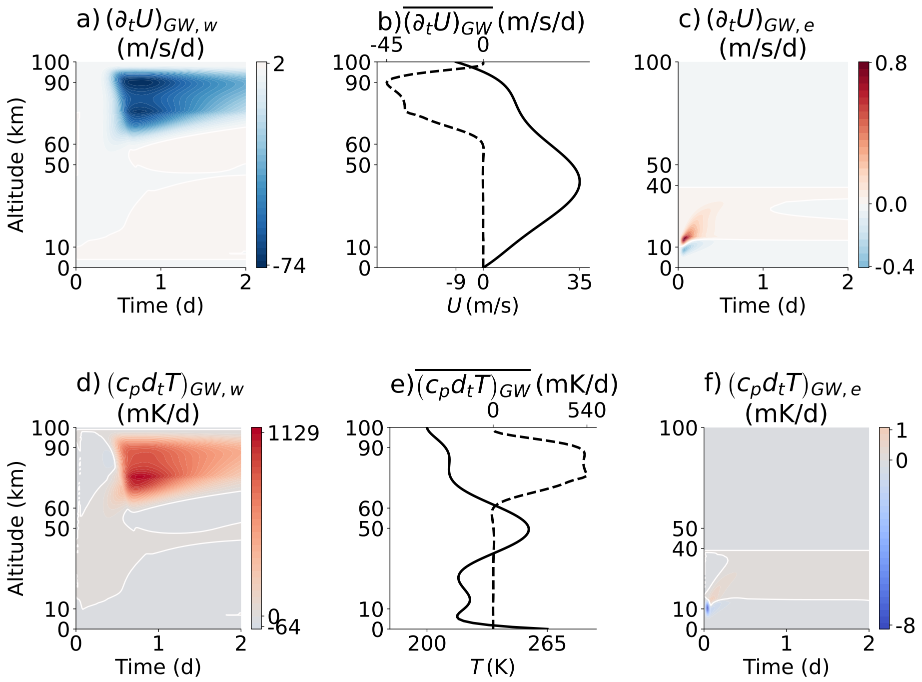

Figure 4.19: Time series of the wave-mean flow interaction $(\partial_t U)_{GW}$ (gravity wave drag plus wave momentum density tendency (see Eq. 2.38) of the westward (a) and the eastward (c) wave packet, and time series of wave-mean flow interaction $(c_p d_t T)_{GW}$ (energy deposition plus wave energy density tendency, see Eq. 2.46) of the westward (d) and the eastward (f) wave packet for a background wind and temperature profile that corresponds to the zonal-mean state at 60°N during January. The middle column shows the zonal wind and the time averaged wave-mean flow interaction $(\partial_t U)_{GW}$ (b) and the temperature and the time averaged wave-mean flow interaction $(c_p d_t T)_{GW}$ (e). The time averages are defined as $\overline{X} = \frac{1}{t_{max}} \int_0^{t_{max}} X(t)\, dt$, where $t_{max} = 2\,\text{days}$.

stratosphere and provides only a negligible amount of deposited energy and momentum per unit mass. The westward wave packet propagates through the eastward jet into the MLT. These waves become saturated above 60 km altitude and reach their critical layers around an altitude of about 90 km. The deeper distribution of energy and momentum deposition compared with 60°S is accompanied by smaller magnitudes of the energy and momentum deposition per unit mass, reaching maximum values of up to around 1 K/d and −75 m/s/d, respectively. Such a difference between winter and summer conditions is as expected (e.g., Lindzen 1981; Becker and Schmitz 2003; Becker 2012) and the new RTP reproduces this behaviour.

Note, however, that we simulated only two wave packets without feedback on the background flow. Furthermore there is no realistic generation mechanisms of gravity waves (orographic or non-orographic) and the simulation is too short to provide an average picture that could be compared to the climatologies for summer and winter as simulated by a GCM (either with parameterized or resolved gravity waves). Nevertheless, these offline results show that the RTP is able to provide reasonable altitudes and magnitudes of energy and momentum deposition. Results that illustrate the performance of the RTP when implemented in a GCM will be presented in the next chapter.

On the characteristic vertical wavenumber and the Boussinesq limit

In Section 3.7 we defined the upper limit m_0 of the vertical wavenumber spectrum. This limit may violate the Boussinesq approximation for the gravity waves as was assumed in Section 2.1. We argued that a violation of the Boussinesq approximation by only a small part of the wave spectrum is of minor importance as long as the characteristic vertical wavenumber m^* fullfils $|m^*| > 1/(2H)$. Here we test this condition for the previous two offline simulations by looking for regions where the ratio $|m^*|/(\frac{1}{2H})$ is smaller than one. The results are shown in the Figures 4.20 and 4.21. The ratio $|m^*|/(\frac{1}{2H})$ for the two simulations is nearly everywhere larger than two. There is a small region for all wave packets within the first hours of the simulations and altitudes above 60 km where the absolute characteristic vertical wavenumber m^* is smaller than the inverse scale height. However, since these altitudes are characterized by nearly zero energy density early in both simulations, the characteristic vertical wavenumber contains no relevant

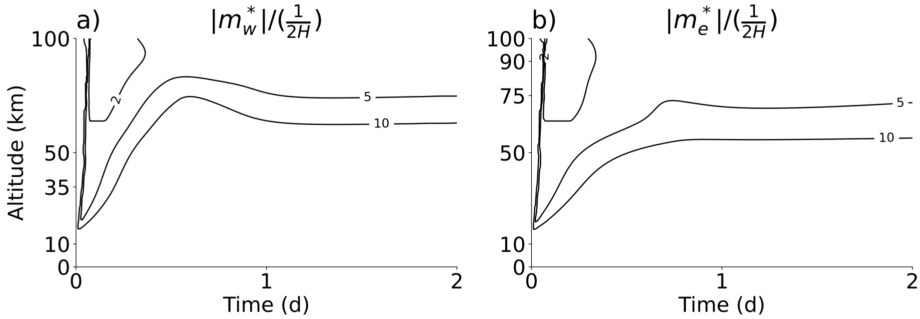

Figure 4.20: Time series of the ratio $|m^*|/(\frac{1}{2H})$ for the westward (a) and eastward (b) wave packet for a background wind and temperature profile that corresponds to the zonal-mean state at 60°S during January. The contours show values of 2, 5, and 10.

information. When energy and momentum deposition become significant at later times, m^* is at least five times larger than $1/(2H)$. Note that in regimes where the gravity waves dissipate from saturation (damping by turbulent vertical diffusion), the characteristic vertical wavenumber tends to be closer to $1/(2H)$ (e.g., see Fig 4.20 b above 75 km). In summary, the assumption that $|m^*| > 1/(2H)$ for tropospheric gravity waves applies reasonably well in our test simulations with the new RTP.

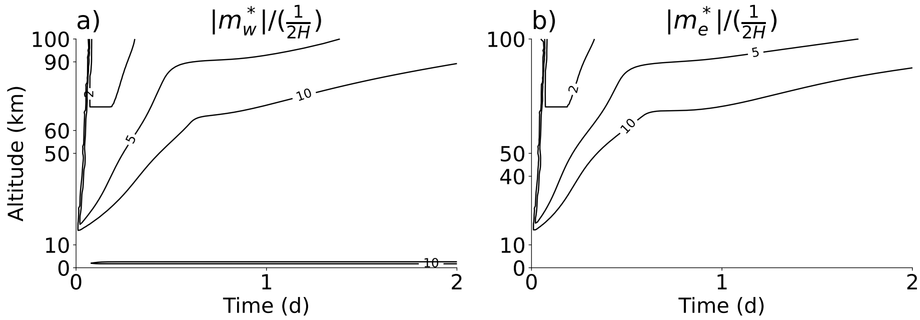

Figure 4.21: Time series of the ratio $|m^*|/(\frac{1}{2H})$ for the westward (a) and eastward (b) wave packet for a background wind and temperature profile that corresponds to the zonal-mean state at 60°N during January. The contours show values of 2, 5, and 10.

Chapter V

Simulation results using a general circulation model

In this chapter we perform a simulation with a GCM for 90 days with a fully implemented RTP. The used GCM is the Kühlungsborn Mechanistic Circulation Model (KMCM). The enthalpy equation of the model is completed by the conventional plus the transient energy deposition. The transient momentum deposition term is neglected due to its minor role in the wave-mean flow interaction (see Sec. 4.2). Therefore, only the conventional momentum deposition (gravity wave drag) from the RTP is added to the horizontal momentum equation of the GCM. The initial condition is an atmosphere without effects of parameterized gravity waves in the mid of December. After a spin-up time of a few days, effects of the parameterized gravity waves become visible. In the following we analyze model results from the first of January on.

A short description of KMCM and the simulation environment of the RTP can be found in Section 5.1. In Section 5.2, we will show and discuss monthly means of zonally averaged fields and wave-mean flow interactions. We will compare this model output with corresponding results from a gravity wave resolving simulation. Section 5.3 deals with the behaviour of the parameterized gravity wave drag during a Sudden Stratospheric Warming (SSW) that develops spontaneously in the February of the simulation. These results are compared with the dynamics during a SSW simulated with the gravity wave resolving model. The last Section (5.4) addresses the additional highly transient and localized heating of the background by the energy density tendency due to gravity wave breaking triggered by thermal tides in the MLT region.

The aim of this chapter is to show that the RTP works reasonably well when implemented in a GCM. Furthermore, we shall illustrate instances where the transience of the new scheme is essential to describe certain dynamics features in agreement with the gravity wave resolving model results.

5.1 Model descriptions

The KMCM is a free running GCM with a standard spectral dynamical core. It has a terrain-following vertical coordinate and a staggered vertical grid (Simmons and Burridge 1981). For our simulations with the new RTP, the triangular spectral truncation is at a horizontal wavenumber 32, which corresponds to a horizontal grid spacing of around 450 km and a shortest resolved horizontal wavelength of $\lambda_k \sim 1350\,\text{km}$. We apply 80 full vertical layers from about $100\,m$ above the surface up to $p = 1.3 \times 10^{-6}\,\text{hPa}$ ($z \sim 175\,\text{km}$) with a vertical spacing of $\Delta z \sim 1\,\text{km}$ in the troposphere and $\Delta z \sim 2\,\text{km}$ in the MLT region. The horizontal grid contains 48 Gaussian latitudes and 96 equidistant longitudes. Relevant components of an atmospheric climate model are represented in a simplified but explicit manner: radiative transfer, water vapor transport, large-scale condensation, moist convection, full surface energy budget with a slap ocean, macroturbulent and molecular horizontal and vertical diffusion, and ion drag. Further detail of the current version of the KMCM can be found in Becker (2017) and references therein.

The RTP has been implemented in the KMCM such that gravity wave computations are performed at every model grid point. We apply four azimuthal directions with $\varphi_1 = 0$ for eastward, $\varphi_2 = \pi$ for westward, $\varphi_3 = \pi/2$ for northward, and $\varphi_4 = 3\pi/2$ for southward intrinsic horizontal propagation of the parameterized gravity waves. At each point of the KMCM horizontal grid the prognostic variables $\mathcal{E}_0(z,t,\varphi_i)$ and $m^*(z,t,\varphi_i)$, and the wave-mean flow interactions are calculated by summing up the contribution from the different azimuths:

$$(\partial_t U)_{GW} = -\frac{1}{4}\partial_z \left(\cos(\varphi_1)F(\varphi_1) + \cos(\varphi_2)F(\varphi_2)\right) \tag{5.1}$$

$$(\partial_t V)_{GW} = -\frac{1}{4}\partial_z \left(\sin(\varphi_3)F(\varphi_3) + \sin(\varphi_4)F(\varphi_4)\right) \tag{5.2}$$

$$\begin{aligned}(c_p d_t T)_{GW} = & -\frac{1}{4}\sum_{i=1}^{4}\partial_z F_p(\varphi_i) \\ & -\frac{1}{4}\left(\cos(\varphi_1)F(\varphi_1) + \cos(\varphi_2)F(\varphi_2)\right)\partial_z U \\ & -\frac{1}{4}\left(\sin(\varphi_3)F(\varphi_3) + \sin(\varphi_4)F(\varphi_4)\right)\partial_z V \\ & -\frac{1}{8}\sum_{i=1}^{4}\partial_t E_{GW}(\varphi_i),\end{aligned} \tag{5.3}$$

where the factor 1/4 normalizes the sum over the four azimuths.

The spectral limits of the gravity wave spectrum are set to

$$m_1 = -\frac{2\pi}{100}\frac{1}{\mathrm{m}} \tag{5.4}$$

$$m_0 = -\frac{2\pi}{12000}\frac{1}{\mathrm{m}} \tag{5.5}$$

for the vertical wavenumbers, and

$$\omega_B = \begin{cases} 5f, & f > f_0 \\ 5f_0, & \text{else} \end{cases} \tag{5.6}$$

$$\omega_T = \frac{N}{5}. \tag{5.7}$$

for the intrinsic frequency. Here f is the Coriolis parameter and the buoyancy frequency N is taken from KMCM. The frequency spectrum is limited to a minimum value of $7.3 \times^{1} 0{-}5\,\mathrm{s}^{-1}$ to prevent the generation of gravity waves with nearly zero vertical group velocities in the tropics. The parameters for the diffusion induced by saturation (see Sec. 3.5) are set to $K_0 = 50\,\mathrm{m}^2/\mathrm{s}$, $\alpha^2 = 0.66$, and $\gamma = 0.2$. The molecular viscosity coefficient K_{mol} is taken from the KMCM and is computed using the method described in Becker and Vadas (2020, their Appendix A).

As source functions we only take into account the generation of non-orographic waves in the troposphere. The non-orographic generation mechanisms are fronts, jets, convection, and a constant background generation. Orographic gravity waves are currently not considered in our RTP. The effective spectral area in the source region has the limits

$$m_1 \leq m \leq -\frac{2\pi}{8000}\frac{1}{\mathrm{m}} \tag{5.8}$$

$$\omega_B \leq \omega_I \leq \omega_T \tag{5.9}$$

which are applied in a horizontally uniform fashion for the sake of simplicity. From Section 3.6 we know that we can decompose the source function $\mathcal{S}(z, t, m, \omega_I, \varphi)$ into a physical part $\mathcal{S}(z, t)$, a spectral part $\mathcal{S}(m, \omega_I)$, and an azimuthal part $\mathcal{S}(\varphi)$. The spectral part is also horizontally uniform with a constant characteristic vertical source wavenumber of

$$m_S^* = -\frac{2\pi}{4000}\frac{1}{\mathrm{m}}. \tag{5.10}$$

The physical and azimuthal parts depend on the generation mechanisms and therefore also on the horizontal (azimuthal) direction. The explicit physical and azimuthal source functions can be found in Appendix D. Here we only describe the types of source functions. They all have in common, that they are only nonzero in the troposphere between $p = 950\,\mathrm{hPa}$ and $p = 120\,\mathrm{hPa}$ ($z < 15\,\mathrm{km}$) and are zero at higher altitudes. The source function $\mathcal{S}_{const}$

for the constant background generation is time-independent, but latitude-dependent with maxima around 60°N and 60°S. To parameterize the generation of gravity waves by fronts we apply a combination of the methods used by Charron and Manzini (2002) and Mirzaei et al. (2014). The source function $\mathcal{S}_{front}$ is nonzero, when the frontogenesis function F_f (see App. B3, Eq. D6) exceeds a threshold of $1 \times 10^{-1}\,\mathrm{K}/(100\,\mathrm{km\,h})$. This source function depends on the horizontal direction of the front and on the amplitude of the frontogenesis function. For gravity waves generated by jets and moist convection we apply the corresponding methods proposed by Mirzaei et al. (2014). Gravity waves are generated in jet exit regions when the wind parameter u_c (see App. D, Eq. D12) is larger than the threshold of $3.5\,\mathrm{m/s}$. The amplitude of the source function $\mathcal{S}_{jet}$ depends on the amplitude of the wind parameter u_c and the horizontal directions of the geostrophic flow. Finally, when the heating by moist convection Q_c exceeds a threshold of $2 \times 10^{-1}\,\mathrm{K/h}$, gravity waves are generated uniformly in all azimuthal directions with the amplitude of the source function $\mathcal{S}_{conv}$ depending on the moist convective heating Q_c. The complete source function from these processes is a sum of the source functions of the individual generation mechanisms:

$$\begin{aligned}\mathcal{S}(z,t,m,\omega_I,\varphi) = \mathcal{S}(m,\omega_I)\,(&\mathcal{S}_{jet}(z,t,\varphi) + \mathcal{S}_{front}(z,t,\varphi)\\ &+ \mathcal{S}_{conv}(z,t,\varphi) + \mathcal{S}_{const}(z,t,\varphi))\end{aligned} \tag{5.11}$$

5.2 Monthly averaged zonal means

Model output from the KMCM has been averaged over the whole January of our simulation and over all longitudes, resulting into monthly zonal means of the zonal wind and temperature, as well as of the momentum deposition (Eq. (3.32)) and energy deposition (Eq. (3.33)) from the parameterized gravity waves. These model results will be compared to corresponding results obtained with the HIgh Altitude Mechanistic general Circulation Model (HI-AMCM) (Becker and Vadas 2020; Becker et al. 2022a; Becker et al. 2022b).

The HIAMCM is a further development of the KMCM and extends far into the thermosphere. The most important property of the HIAMCM in our context is that gravity waves are not parameterized but simulated explicitly, which is subject to the effective resolution of the model. According to Becker et al. (2022a), the HIAMCM reliably resolves gravity waves down to horizontal wavelengths of about 200 km. Here we use output from the HIAMCM nudged to MERRA-2 reanalysis for the December 2016 to February 2017 period. Our 30 day average from the HIAMCM is from 22 December 2016 to 21 January 2017. This period was characterized by a relatively strong polar vortex.

The resolved gravity wave drag from the HIAMCM is computed as the convergence of the vertical momentum flux from the T256 model output minus the corresponding momentum flux convergence when spectrally truncating the model output at T32. This difference between the two corresponds to the drag from gravity waves having horizontal wavelengths between 156 km (the smallest resolved waves in T256) and 1350 km (the smallest resolved waves in T32). We can estimate the largest horizontal wavelength from the KMCM simulation from the wavenumber and frequency limits of the RTP. The largest horizontal wavelength is around 1000 km, which is comparable to the largest wavelength we get from the HIAMCM simulation. We do not have the energy deposition from the gravity wave resolved simulation, so a comparison is not possible.

The zonal mean winds and the zonal gravity wave drag are shown in Figure 5.1. The mean temperatures and the thermal wave-mean flow interactions (energy deposition plus energy density tendency) can be found in Figure 5.2. We can see that there is a good agreement in the summer hemisphere (southern hemisphere) of the parameterized and resolved gravity wave drag, as well as their effects on the zonal-mean zonal wind and temperature. The region of significant gravity wave drag is in the latitudinal region between 90°S and 30°S around the mesopause ($p \sim 0.001\,\mathrm{hPa}$) (Fig. 5.1). The maximum of the parameterized gravity wave drag is larger,

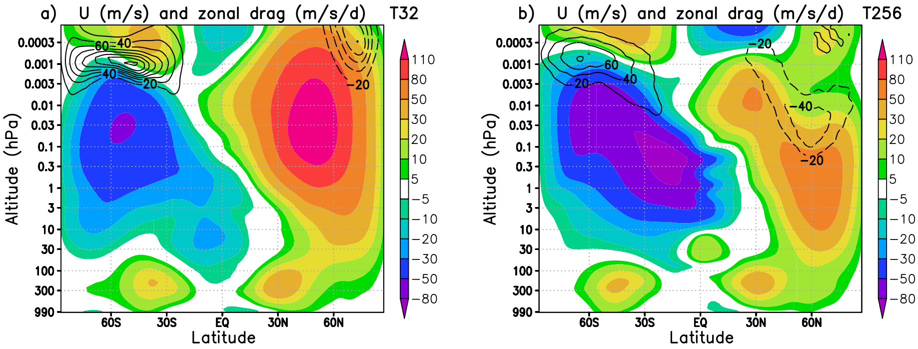

Figure 5.1: Zonally and monthly averaged zonal wind (colors) and zonal gravity-wave drag (contours for $\pm 20, 40, 60, 80, 100\,\mathrm{m/s/d}$) for the KMCM with T32 spectral resolution and the new RTP (a), and the HIAMCM with T256 spectral resolution and explicit simulation of gravity waves (b). See text for further details.

but can be adjusted with the tunable parameters in the RTP. The parameterized gravity wave drag causes the wind reversal above the mesopause like in the high-resolution simulation. The summer polar mesopause region is characterized by temperatures below 140 K (Fig. 5.2). This substantial deviation from the radiatively determined state is mainly caused by the gravity wave driven summer-to-winter pole residual circulation and the resulting adiabatic cooling beneath the summer polar mesopause. The energy deposition around the mesopause has amplitudes of a few K/d and acts as a direct heating against the adiabatic cooling. The fact that energy deposition (which largely corresponds to turbulent frictional heating) needs to be taken into account in models to simulate the cold summer mesopause along with realistic gravity wave drag and gravity wave amplitudes was first emphasized by Lübken (1997).

In the winter hemisphere the situation is different. Our T32 simulation with the new RTP exhibits a strong polar vortex that extends into the mesopause region (Fig. 5.1 a). This pattern is not very realistic and differs from the HIAMCM results, where the polar vortex weakens with increasing altitude from the lower mesosphere on such that an additional wind maximum develops in the subtropical upper mesosphere at about 30°N and

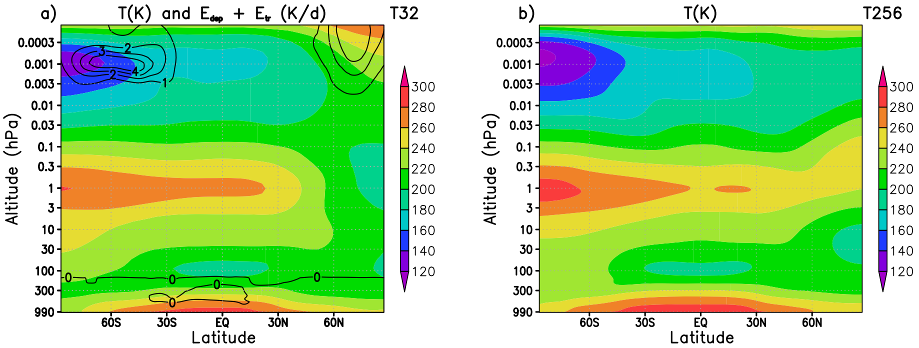

Figure 5.2: Zonally and monthly averaged temperature (colors) for the KMCM with T32 spectral resolution and the new RTP (a), and the HIAMCM with T256 spectral resolution and explicit simulation of gravity waves (b). The contours in (a) shows the energy deposition plus the energy density tendency from the RTP (contours for $1, 2, 3, 4, 5\,\mathrm{K/d}$). See text for further details.

0.005 hPa. There is poor agreement between the parameterized and the resolved gravity wave drag in the winter middle atmosphere. While in the gravity wave resolving simulation the westward gravity wave drag is strongest in the lower mesosphere at 60°N and extends to higher altitudes towards lower latitudes, it is located close to the pole above the mesopause in the T32 simulations with parameterized gravity waves. Also the temperature patterns in the winter middle atmosphere differ significantly (Fig. 5.2). The warm winter stratopause of the gravity wave resolving simulation can not be reproduced with the current configuration of the new RTP.

These differences in the winter hemisphere are a consequence of missing generation mechanisms for the parameterized gravity waves. More specifically, there are no orographic gravity waves in the KMCM simulation. On the other hand, it is very well known that orographic gravity waves have a high impact on the circulation of the entire winter middle atmosphere (e.g., McLandress et al. 2013). A simulation with parameterized orographic gravity waves that were generated over the Rocky Mountains, the Himalayas, and eastern Siberia by McLandress and McFarlane (1993) showed that these orographic waves break in the stratosphere/lower mesosphere and give rise to a much weaker zonal flow in the winter middle atmosphere. Orographic gravity waves further contribute to separate the stratospheric polar night jet from the tropospheric jet in the subtropics, to reduce the magnitude of the tropospheric jet, and to enhance the westward wind shear in the subtropical upper troposphere (Alexander et al. 2010). All these effects affect the propagation and instability of the non-orographic gravity waves (particularly those from jets and fronts) such that these gravity waves deposit their momentum and energy at lower altitudes in the mesosphere. The overall result is a much weaker polar vortex and with a different pattern than what would be the case without orographic gravity waves. We simulated such an hypothetical state without orographic gravity waves with the current version of the T32 KMCM. Indeed our GCM with the new RTP is equipped with a parameterization of just non-orographic gravity waves but misses the effects of orographic gravity waves (Fig. 5.1 a).

5.3 Gravity waves and sudden stratospheric warmings

A Sudden Stratospheric Warming (SSW) is a temporary increase of the stratospheric temperature by typically several 10 K. A SSW occurs in three out of four northern hemisphere winters, but barely in the southern hemisphere winters. The first SSW was observed in 1952 over Berlin (Scherhag 1952), and in September 2002 the first SSW was observed in the southern hemisphere (Krüger et al. 2005). SSWs occur when the polar vortex is disrupted by planetary waves such that the vortex breaks down and the zonal-mean eastward winds before the warming turn into weakly westward winds at middle and high latitudes for about several days or longer. The warming of the polar stratosphere is caused by the residual descent of air masses. This downwelling is caused by the intensified residual circulation in the upper stratosphere which is driven by the large planetary Rossby wave amplitudes that also cause the disruption of the vortex. A stratospheric warming is usually accompanied by a cooling of the mesosphere (Labitzke 1972). With changing winds in the stratosphere, the westward propagating gravity waves no longer break in the mesosphere but a lower altitudes. Instead, eastward gravity waves propagate into the mesosphere during the wind reversal, in analogy to the summer mesosphere. This leads to an anomalous equatorward residual flow, which is balanced by a reduced residual downwelling (or even upwelling) over the polar cap and therefore results in the mesospheric cooling (Holton 1983). An overview of the state-of-the-art knowledge about SSWs can be found in the review of Baldwin et al. (2021).

We inspected the simulations with the KMCM (gravity waves parameterized with the RTP) and the HIAMCM (resolved gravity waves) from Section 5.2 with regard to the dynamics of simulated SSWs. A SSW took place in each simulation. By defining the wind reversal at $p = 1\,\mathrm{hPa}$ as the starting point of the SSW and setting the model time to day zero, we can qualitatively compare the zonal gravity wave drag during the SSWs. The time series of the zonal-mean zonal mean winds and the zonal gravity wave drags are shown in Figure 5.3. In both simulations the wind reversal from eastward to westward winds causes a reversal of the zonal gravity wave drag from westward to eastward. At the end of each SSW a new polar vortex is established and the direction of the gravity wave drag changes back to westward. During the wind reversal, the eastward gravity wave drag is in both cases located at the upper part of the westward wind region. Because the polar night jet before and after the SSW is much stronger and extends to higher altitudes in the KMCM simulation, also the westward gravity wave drag during these periods is located at higher altitudes when compare to the HIAMCM. By looking at the gravity wave drag in Figure 5.3 a we can see that it takes roughly two days after the decline of the westward gravity wave drag for the eastward gravity wave drag to evolve, and even about 10 more days to reach its maximum. This is presumably because after the sudden wind reversal in the stratosphere and mesosphere, the eastward gravity waves

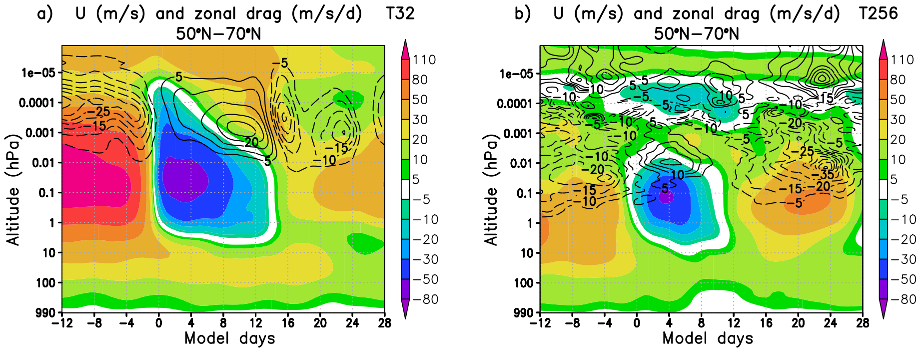

Figure 5.3: Temporal evolution of the zonal-mean zonal wind and the eastward gravity wave drag, averaged from 50°N-70°N during a sudden stratospheric warming as simulated by the free-running KMCM with the new RTP (a) and the gravity wave-resolving HIAMCM nudged to MERRA-2 reanalysis during January and February 2017 (b). The zonal wind is shown by colors and the gravity wave drag is shown by black contours (for $\pm 5, 10, 15, 20, \ldots$ m/s/d).

need some time to propagate into the upper mesosphere. This is different from the recovery phase of the polar vortex in Figure 5.3 a, which is characterized by a smoother transition from westward to eastward flow such that the gravity wave drag in the upper mesosphere changes more synchronously with (and opposite to) the wind direction in the stratopause region and lower mesosphere. The SSW simulated by the HIAMCM shows different features regarding the wind reversal, the location of the maximum westward and eastward gravity wave drag from (presumably) primary gravity waves, and the gravity wave drag in the lower thermosphere which is eastward and presumably results from secondary gravity waves (e.g., Becker et al. 2022b). Nevertheless, the overall behavior of the KMCM below the mesospause is dynamically consistent and comparable to the HIAMCM.

In summary, the RTP reproduces the expected behaviour of gravity waves during a SSW in a dynamically consistent fashion. Differences to the HIAMCM simulation are mainly due to the fact that the self-induced SSW in the free-running KMCM develops from an unrealistically strong polar vortex (see the above discussion about missing orographic gravity waves in our current KMCM version). Furthermore, the SSW simulated in the nudged HIAMCM is an event with a vortex displacement. Such events usually show dynamical features that are different from the more regular SSW events where the vortex is split into two fragments. A closer inspection of the KMCM data revealed that the KMCM with the RTP simulated such a split-vortex event.

5.4 Gravity waves and thermal tides

The impact of gravity waves on the amplitudes and phases of thermal tides and the impact of thermal tides on the refraction and breaking of gravity waves have been subject to many studies during the last decades. Thermal tides are forced mainly by the daily cycle of the absorption of solar insolation by water vapor and clouds in the troposphere and by ozone in the stratosphere. Another forcing process is the daily cycle of cumulus convection in the tropics (Grieger et al. 2004; Achatz et al. 2008). When the thermal tides propagate into the MLT region, they cause regular and significant variations of winds and temperature. This is illustrated in Figure 5.4 c in terms of the zonal wind variations in the northern winter MLT from the KMCM simulation with the new RTP. The question of how thermal tides affect the gravity wave drag in the zonal mean has been discussed controversially in the literature. One result is that in the summer hemisphere, the zonal gravity wave drag from a conventional parameterization (single column and steady-state approximations) is shifted to higher altitudes and, hence, exerts a larger forcing per unit mass (Miyahara and Wu 1989). Another study found quite the opposite result, namely that the drag of resolved waves in the summer hemisphere is shifted to lower altitudes and smaller forcing per unit mass (Becker 2017). The modulation of thermal tides regarding their amplitudes and phases by gravity waves has been the topic of many model studies (e.g., Senf and Achatz 2011; Ortland and Alexander 2006), but a consensus on how the waves modulate the tides has not been reached yet. The question appears to be further complicated by the fact that the modulation of the tides by gravity waves is highly sensitive to the details of the gravity wave spectrum (Ortland and Alexander 2006). Furthermore, the comparison of the simulated tidal modulation using conventional gravity wave parameterizations with the simulated tidal modulation using three-dimensional, transient ray-tracing reveals significantly different results regarding the gravity wave effects on the amplitudes and phases of the tides (Senf and Achatz 2011).

We inspect the interaction of the parameterized gravity waves with the thermal tides in the KMCM simulation by plotting the conventional energy deposition and the new energy density tendency over the temperature deviation from the daily mean temperature (Fig. 5.4 a and b), and the zonal and meridional gravity wave drag over the zonal and meridional winds, respectively (Fig. 5.4 c and d). More specifically, at the location 57°N and 180°E on day 44 of the simulation we can infer a clear semi-diurnal tidal pattern in the zonal and meridional winds. These patterns are typical for the current KMCM simulation up to the SSW discussed in the previous section. From Figure 5.4 a we can see that the conventional energy deposition is in phase with the positive temperature deviation. There is no such pattern visible for the energy density tendency in panel b. By comparing Figure 5.4 a with Figure 5.4 c we see that the energy deposition is in phase with the variation of the zonal wave drag in higher altitudes. The westward gravity wave drag coincides with the reversal of the zonal wind from eastward to westwards winds above about 0.001 hPa, thereby reinforcing the zonal wind tendency associated with the tide. A similar pattern can be inferred for the meridional gravity wave drag and the meridional wind.

From Section 4.4 we know that temperature variations have a minor impact on the propagation and dissipation of the parameterized gravity waves. Therefore we conclude that the energy deposition does not arise from the tidal structure in the temperature but from the tidal structure in the zonal and meridional winds. Indeed, the energy deposition is roughly in phase with the westward gravity wave drag as mentioned above. The westward gravity wave drag arises from the breaking of westward gravity waves due to saturation (and critical layers) which is induced by the tidal wind changing from strongly eastward to weakly westward. This so-induced westward gravity wave drag propagates downward in time with the phase of the tide (like for gravity waves, downward phase propagation of the tides is equivalent to

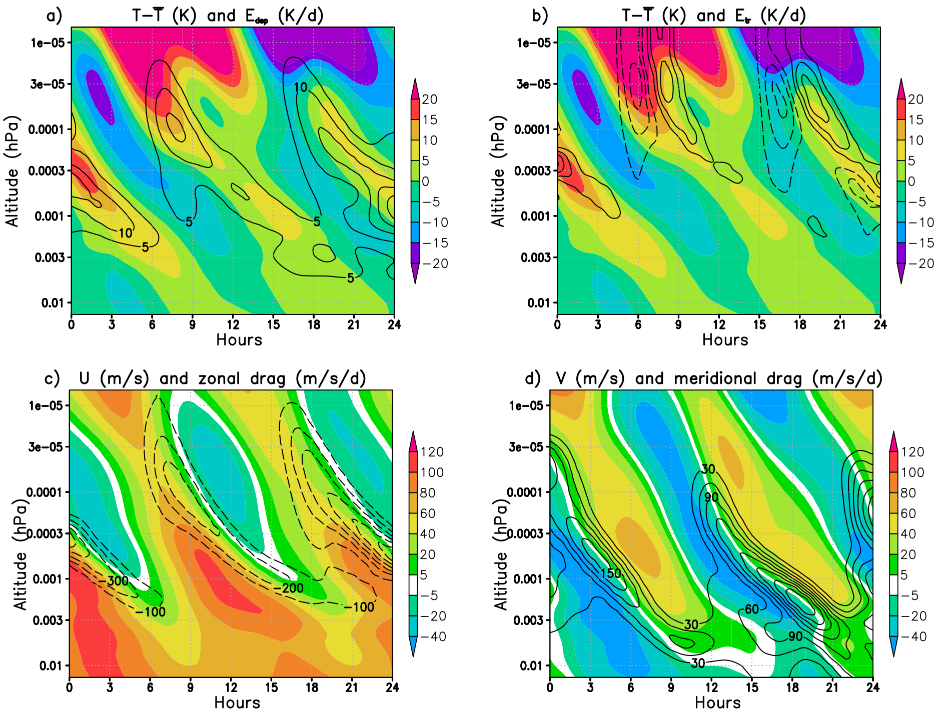

Figure 5.4: Temporal evolution of the daily temperature variations in colours and the energy deposition in contours (a), the daily temperature variations in colours and the energy density tendency in contours (b), the zonal background wind in colours and zonal gravity wave drag in contours (c), and the meridional background wind in colours and meridional gravity wave drag in contours (d) during a single day as simulated by the free-running KMCM with the new RTP. The intervals of the black contours for the different panels are $5, 10, 15, 20, 25\,\mathrm{K/d}$ for (a), $\pm 1, 2, 3\,\mathrm{K/d}$ for (b), $-100, -200, -300, \ldots\,\mathrm{m/s/day}$ for (c), and $30, 60, 90, \ldots\,\mathrm{m/s/day}$ for (d). The daily average is defined as $\overline{T} = \frac{1}{24\,\mathrm{h}} \int_{0}^{24\,\mathrm{h}} T(t)\,dt$.

upward group propagation). A similar result can be found for the northward gravity wave drag and the northward wind phase of the tide. A gravity wave drag from eastward and southward propagating gravity waves cannot be found in Figure 5.4. Such gravity waves are filtered at lower altitudes.

In summary, the results of the interaction of the parameterized gravity waves with tides agrees with the results in Becker (2017). The tidal wind variations trigger the gravity wave drag; therefore, the zonal-mean gravity wave drag has to be located at lower altitudes than would be the case without the tidal wind variations being considered in the background flow for the gravity waves. Note that we neglect the time-dependency of the background flow in the RTP which would lead to a modulation of the ground-based gravity-wave frequencies (Senf and Achatz 2011). A further investigation of the feed back of the gravity wave drag on the tides is beyond the scope of this thesis and subject to future investigations.

Chapter VI

Conclusions

In this thesis we developed a new, transient gravity wave parameterization for the application in general circulation models of the atmosphere. This parameterization is based on the radiative transfer equation for the gravity-wave field. As a major and common simplification we applied the single-column approximation. To solve the radiative transfer equation we assumed the Desaubies spectrum with two free parameters, and we then applied the Gaussian Variational Principle to derive two prognostic equations for these parameters, one prognostic equation is for the integrated energy density of the wave field in physical space, and a second prognostic equation for the characteristic vertical wavenumber of the Desaubies spectrum. We introduced the concept of the effective spectral area. This area includes all spectral elements that are part of the wave field at a certain altitude and time. All other spectral elements that are part of the total spectral area have been attenuated at lower altitudes and earlier in time, and have thereby caused wave-mean flow interaction. In this thesis we derived the complete set of wave-mean flow interaction terms that apply in the transient case. These terms include the usual momentum and energy deposition, as well as additional forcing and heating terms that are related to the transient changes of momentum and energy of the wave field. The new, complete set of transient wave-mean flow interaction terms ensures that the whole system of mean flow and wave field is energy and momentum conserving. By testing the new, transient Radiative Transfer Parameterization (RTP) in idealised simulations we were able to show that the system is indeed energy and momentum conserving.

We introduced and validated new methods to account for reflection and critical layers within the RTP. Both methods result in reductions of the effective spectral area as a result of reflection and critical layers. While the wave-mean flow interaction caused by critical layers is physically consistent, the reduction of the spectral area due to reflection accounts for the fact that our formulation of the RTP excludes the downward propagation of waves. We furthermore showed that vertical temperature gradients play only a minor role for the refraction that leads to dissipation, critical layer filtering, or reflection of gravity waves as compared to wind shears.

The RTP was implemented into a low-resolution version of the KMCM. The comparison of the monthly averaged zonal means of zonal wind and temperature during January with the results from a high-resolution, gravity-wave resolving HIAMCM simulation reveals good agreement in the MLT region of the summer hemisphere (cold mesopause and wind reversal), but large differences in the winter MLT region. The latter result mainly from the missing parameterization of orographic gravity waves when using the RTP in its present form. On the other hand, the simulation of a self-induced SSW using the KMCM with the RTP showed that the reversal from westward to eastward gravity wave drag in the winter mesosphere and the the overall temporal evolution of the SSW are reasonably well reproduced. Furthermore, we inspected the instantaneous variations of the local wave-mean flow interactions that are triggered by the semi-diurnal tide in the winter upper mesosphere at middle latitudes. Results showed that the resulting zonal-mean gravity wave drag is located at lower altitudes than it would be the case without the tidal wind variations being considered in the background flow for the gravity waves.

Overall, our idealized simulations and GCM applications show that the RTP is a powerful gravity wave parameterization in single column approximation. It accommodates different generation mechanisms in a single framework, takes the time needed for the vertical propagation of wave packets into account, and simulates the gravity wave drag in the MLT region that drives the summer-to-winter-pole meridional circulation.

The RTP can be further improved to include gravity waves from orographic sources, as well as secondary gravity waves. Particularly orographic gravity waves are essential to simulate a realistic circulation in the winter hemisphere, as was discussed in the previous chapter. Furthermore, secondary gravity waves are likely important in the winter upper mesopause region (e.g., Becker and Vadas 2018; Harvey et al. 2022). Such further developments of the RTP are necessary for routinely using the new scheme in GCMs. In addition, the single-column approximation should be dropped to take also horizontal propagation into account (Senf and Achatz 2011; Sato et al. 2012).

Source functions for the generation of orographic gravity waves from field variables resolved by a GCM are well known (e.g., McFarlane 1987). What might pose a difficulty when defining an orographic gravity-wave source function in the context of the new RTP is the narrow vertical wavenumber spectrum of orographic gravity waves. The vertical wavenumber of any monochromatic orographic gravity wave is proportional to the ratio of the Brunt-Väisälä frequency and the horizontal wind that flows over the orography, that is, $m \sim N/U$. We can set the characteristic vertical wavenumber to $m^* \sim N/U$ when applying the Desaubies spectrum to orographic gravity waves and start with a narrow effective spectral area when applying an orographic source function in the RTP at the lowest atmospheric model level. The question arises whether the so generated gravity wave field merges well enough with the broad vertical wavenumber spectrum of non-orographic gravity waves generated higher up in the free troposphere, or whether these two wave fields are too distinct to be well represented in a single framework.

If the latter is the case, then we would need additional prognostic equations for orographic gravity waves, but we could still use a common saturation condition for all gravity waves. Regarding possible source functions of higher-order gravity waves, we expect to face the same difficulties as for orographic gravity waves, namely the question whether the vertical wavenumber spectrum of the primary gravity waves merges well enough with the wavenumber spectrum of the secondary gravity waves.

The lateral propagation of gravity waves is potentially important for the spatial distribution of the wave field (Sato et al. 2012). Furthermore, offline simulations using three-dimensional ray-tracing models showed that the three-dimensional propagation has a strong impact on the momentum fluxes (Senf and Achatz 2011). Thus, a relaxation of the single-column approximation is advisable. An intermediate step from the single-column approximation to the three-dimensional propagation of wave packets is the two-dimensional propagation of wave packets, e.g., lateral in the direction of the horizontal wave vector of the gravity wave field for a specific azimuth and vertical propagation. This would require the refraction of the horizontal wavenumber (similar to the refraction of the vertical wavenumber in this thesis), but not a change of the horizontal propagation direction. Therefore we would need to relax the assumption of a constant horizontal wavenumber that was made in this thesis implicit with applying the single-column-approximation. Such a scheme with the partial relaxation of the single-column approximation and partially allowing for horizontal refraction would result in a gravity wave parameterization with three prognostic equations per azimuth, namely for the integrated energy density, the characteristic vertical wavenumber, and a characteristic horizontal wavenumber if we define a horizontal wavenumber spectrum with a characteristic horizontal wavenumber similar to the vertical wavenumber spectrum. A varying characteristic horizontal wavenumber would also require a new ansatz for the ESA, where a constant horizontal wavenumber in the Lagrangian sense was assumed in this thesis. The two-dimensional propagation would allow the model to horizontally exchange gravity wave energy (and momentum and potential energy flux as well). However, even with these complications we would still not account for the meridional propagation of gravity waves that contribute to the zonal drag at higher latitudes as was suggested by Sato et al. (2012). For a fully three-dimensional and transient gravity wave scheme based on the RTE, the complete horizontal refraction would need to be accounted for. Such a method may be developed in the future based on the experiences with a two-dimensional RTP discussed above.

We used the KMCM to develop and test the new parameterization. The KMCM has an easily comprehensible model code in comparison to comprehensive community models. Nevertheless, the new RTP should also work successfully when implemented into a comprehensive community model, like the UA-ICON climate model. As discussed above, this would require to add orographic gravity waves. The RTP is expected to accommodate also secondary gravity waves after some further development work. This all offers a new pathway to substantially improved parameterizations of gravity waves in middle atmosphere community climate models as compared to conventional methods.

Bibliography

Achatz, U. (2005). “On the role of optimal perturbations in the instability of monochromatic gravity waves”. In: *Physics of Fluids* 17.9, p. 094107.

Achatz, U., Grieger, N., and Schmidt, H. (2008). “Mechanisms controlling the diurnal solar tide: Analysis using a GCM and a linear model”. In: *Journal of Geophysical Research: Space Physics* 113.A8.

Achatz, U. et al. (2017). “The interaction between synoptic-scale balanced flow and a finite-amplitude mesoscale wave field throughout all atmospheric layers: weak and moderately strong stratification”. In: *Quarterly Journal of the Royal Meteorological Society* 143.702, pp. 342–361.

Alexander, M. J., May, P. T., and Beres, J. H. (2004). “Gravity waves generated by convection in the Darwin area during the Darwin Area Wave Experiment”. In: *Journal of Geophysical Research: Atmospheres* 109.D20.

Alexander, M. J. 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”. In: *Quarterly Journal of the Royal Meteorological Society* 136.650, pp. 1103–1124.

Amiramjadi, M. et al. (2020). “The Spatiotemporal Variability of Nonorographic Gravity Wave Energy and Relation to Its Source Functions”. In: *Monthly Weather Review* 148.

Baldwin, M. P. et al. (2001). “The quasi-biennial oscillation”. In: *Reviews of Geophysics* 39.2, pp. 179–229.

Baldwin, M. P. et al. (2021). “Sudden Stratospheric Warmings”. In: *Reviews of Geophysics* 59.1.

Becker, E. (2012). “Dynamical Control of the Middle Atmosphere”. In: *Space Sci Rev* 168, pp. 283–314.

— (2017). “Mean-Flow Effects of Thermal Tides in the Mesosphere and Lower Thermosphere”. In: *Journal of the Atmospheric Sciences* 74.6, pp. 2043–2063.

— (2019). *Anelastic equations as a tool to understand gravity waves and mean-flow effects.* Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS); Leibniz-Institut für Atmosphärenphysik (IAP). (Last accessed: 28 Dec 2022). DOI: `10.5446/40516`.

Becker, E. and McLandress, C. W. (2009). “Consistent Scale Interaction of Gravity Waves in the Doppler Spread Parameterization”. In: *Journal of The Atmospheric Sciences - J ATMOS SCI* 66, pp. 1434–1449.

Becker, E. and Schmitz, G. (2003). “Climatological Effects of Orography and Land–Sea Heating Contrasts on the Gravity Wave–Driven Circulation of the Mesosphere”. In: *Journal of the Atmospheric Sciences* 60.1, pp. 103–118.

Becker, E. and Vadas, S. L. (2018). “Secondary Gravity Waves in the Winter Mesosphere: Results From a High-Resolution Global Circulation Model”. In: *Journal of Geophysical Research: Atmospheres* 123.5, pp. 2605–2627.

— (2020). “Explicit Global Simulation of Gravity Waves in the Thermosphere”. In: *Journal of Geophysical Research: Space Physics* 125.10.

Becker, E. et al. (2022a). “A High-Resolution Whole-Atmosphere Model With Resolved Gravity Waves and Specified Large-Scale Dynamics in the Troposphere and Stratosphere”. In: *Journal of Geophysical Research: Atmospheres* 127.2, e2021JD035018.

Becker, E. et al. (2022b). “Multi-step vertical coupling during the January 2017 sudden stratospheric warming”. In: *Journal of Geophysical Research: Space Physics*. submitted.

Bölöni, G. et al. (2016). “The Interaction between Atmospheric Gravity Waves and Large-Scale Flows: An Efficient Description beyond the Nonacceleration Paradigm". In: *Journal of the Atmospheric Sciences* 73.12, pp. 4833–4852.

Bölöni, G. et al. (2021). "Toward Transient Subgrid-Scale Gravity Wave Representation in Atmospheric Models. Part I: Propagation Model Including Nondissipative Wave–Mean-Flow Interactions". In: *Journal of the Atmospheric Sciences* 78.4, pp. 1317–1338.

Charron, M. and Manzini, E. (2002). "Gravity Waves from Fronts: Parameterization and Middle Atmosphere Response in a General Circulation Model". In: *Journal of the Atmospheric Sciences* 59.5, pp. 923–941.

Desaubies, Y. J. F. (1976). "analytical representaion of internal wave spectra". In: *J. Phys. Oceanogr.* 6, pp. 976–981.

Fleming, E. L. et al. (1990). "Zonal mean temperature, pressure, zonal wind and geopotential height as functions of latitude". In: *Advances in Space Research* 10.12, pp. 11–59.

Fritts, D. C. and Alexander, M. J. (2003). "Gravity wave dynamics and effects in the middle atmosphere". In: *Reviews of Geophysics* 41.1.

Fritts, D. C. and VanZandt, T. E. (1993). "Spectral Estimates od Gravity Wave Energy and Momentum Fluxes, I, energy dissipation acceleration, and constraints". In: *J. Atmos. Sci.* 50.22, pp. 3685–3694.

Garcia, R. R. and Sassi, F. (1999). "Modulation of the mesospheric semiannual oscillation by the quasibiennial oscillation". In: *Earth Planet Sp.* 51, pp. 563–569.

Garcia, R. R. and Solomon, S. C. (1985). "The effect of breaking gravity waves on the dynamics and chemical composition of the mesosphere and lower thermosphere". In: *Journal of Geophysical Research* 90, pp. 3850–3868.

Garcia, R. R. et al. (2007). "Simulation of secular trends in the middle atmosphere, 1950–2003". In: *Journal of Geophysical Research: Atmospheres* 112.D9.

Geller, M. A. et al. (2013). “A Comparison between Gravity Wave Momentum Fluxes in Observations and Climate Models”. In: *Journal of Climate* 26.17, pp. 6383–6405.

Grieger, N., Schmitz, G., and Achatz, U. (2004). “The dependence of the nonmigrating diurnal tide in the mesosphere and lower thermosphere on stationary planetary waves”. In: *Journal of Atmospheric and Solar-Terrestrial Physics* 66.6, pp. 733–754.

Guo, Y., Liu, A. Z., and Gardner, C. S. (2017). “First Na lidar measurements of turbulence heat flux, thermal diffusivity, and energy dissipation rate in the mesopause region”. In: *Geophysical Research Letters* 44, pp. 5782–5790.

Harvey, V. L. et al. (2022). “Evaluation of Polar Winter Mesopause Wind in WACCMX+DART”. In: *Journal of Geophysical Research: Atmospheres* 127.15, e2022JD037063.

Hasselmann, K. (1968). “Weak-Interaction Theory Of Ocean Waves”. In: *Basic Developments in Fluid Dynamics*. Ed. by Maurice holt. Academic Press, pp. 117–182.

Hines, C. O. (1997a). “Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere. Part 1: Basic formulation”. In: *Journal of Atmospheric and Solar-Terrestrial Physics* 59.4, pp. 371–386.

— (1997b). “Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere. Part 2: Broad and quasi monochromatic sepctra, and implementation”. In: *Journal of Atmospheric and Solar-Terrestrial Physics* 59.4, pp. 387–400.

Holton, J. R. (1983). “The Influence of Gravity Wave Breaking on the General Circulation of the Middle Atmosphere”. In: *Journal of Atmospheric Sciences* 40.10, pp. 2497–2507.

Hoskins, B. J. (1982). “The Mathematical Theory of Frontogenesis”. In: *Annual Review of Fluid Mechanics* 14.1, pp. 131–151.

Hoskins, B. J. and Simmons, A. J. (1975). “A multi-layer spectral model and the semi-implicit method”. In: *Quarterly Journal of the Royal Meteorological Society* 101.429, pp. 637–655.

Kim, Y.-H. et al. (2021). “Toward Transient Subgrid-Scale Gravity Wave Representation in Atmospheric Models. Part II: Wave Intermittency Simulated with convective Sources”. In: *Journal of the Atmospheric Sciences* 78.4, pp. 1339–1357.

Koch, S. E. and Dorian, P. (1988). “A mesoscale gravity wave event observed during CCOPE. III - Wave environment and probable source mechanisms”. In: *Monthly Weather Review* 116, pp. 2570–2592.

Krüger, K., Naujokat, B., and Labitzke, K. (2005). “The Unusual Midwinter Warming in the Southern Hemisphere Stratosphere 2002: A Comparison to Northern Hemisphere Phenomena”. In: *Journal of The Atmospheric Sciences* 62, pp. 603–613.

Labitzke, K. (1972). “Temperature Changes in the Mesosphere and Stratosphere Connected with Circulation Changes in Winter”. In: *Journal of Atmospheric Sciences* 29.4, pp. 756–766.

Limpasuvan, V. et al. (Apr. 2012). “The roles of planetary and gravity waves during a major stratospheric sudden warming as characterized in WACCM”. In: *Journal of Atmospheric and Solar-Terrestrial Physics* 78-79.

Lindzen, R. S. (1981). “Turbulence and stress owing to gravity wave and tidal breakdown”. In: *Journal of Geophysical Research: Oceans* 86.C10, pp. 9707–9714.

Liu, C. H. and Yeh, K. C. (1969). “Effect of ion drag on propagation of acoustic-gravity waves in the atmospheric F region”. In: *Journal of Geophysical Research (1896-1977)* 74.9, pp. 2248–2255.

Lübken, F.-J. (1997). “Seasonal variation of turbulent energy dissipation rates at high latitudes as determined by insitu measurements of neutral density fluctuations”. In: 102, pp. 13, 441–13, 456.

Marks, C. J. and Eckermann, S. D. (1995). “A Three-Dimensional Nonhydrostatic Ray-Tracing Model for Gravity Waves: Formulation and Preliminary Results for the Middle Atmosphere”. In: *Journal of Atmospheric Sciences* 52.11, pp. 1959–1984.

McFarlane, N.A. (1987). “The Effect of Orographically Excited Gravity Wave Drag on the General Circulation of the Lower Stratosphere and Troposphere”. In: *Journal of Atmospheric Sciences* 44.14, pp. 1775–1800.

McLandress, C. et al. (2006). “Large-scale dynamics of the mesosphere and lower thermosphere: An analysis using the extended Canadian Middle Atmosphere Model”. In: *Journal of Geophysical Research: Atmospheres* 111.D17.

McLandress, C. et al. (2013). “Dynamical Control of the Mesosphere by Orographic and Nonorographic Gravity Wave Drag during the Extended Northern Winters of 2006 and 2009”. In: *Journal of the Atmospheric Sciences* 70.7, pp. 2152–2169.

McLandress, C. W. and McFarlane, N. A. (1993). “Interactions between Orographic Gravity Wave Drag and Forced Stationary Planetary Waves in the Winter Northern Hemisphere Middle Atmosphere”. In: *Journal of the Atmospheric Sciences* 50, pp. 1966–1990.

Miller, J. E. (1948). “ON THE CONCEPT OF FRONTOGENESIS”. In: *Journal of Atmospheric Sciences* 5.4, pp. 169–171.

Mirzaei, M. et al. (2014). “Structure, Energy, and Parameterization of Inertia–Gravity Waves in Dry and Moist Simulations of a Baroclinic Wave Life Cycle”. In: *Journal of the Atmospheric Sciences* 71, pp. 2390–2414.

Miyahara, S. and Wu, D. (1989). “Effects of solar tides on the zonal mean circulation in the lower thermosphere: solstice condition”. In: *Journal of Atmospheric and Terrestrial Physics* 51.7, pp. 635–647.

Müller, P. and Natarov, A. (2003). “The Internal Wave Action Model IWAM”. In: *Near-Boundary Processes and Their Parameterization: Proc. ‘Aha Huliko’a Winter Workshop, Honolulu, HI, University of Hawaii*, pp. 95–105.

Müller, P. and Olbers, D. (1975). “On the dynamics of internal waves in the deep ocean”. In: *J. Geophys. Res.* 80, pp. 3848–3860.

Muraschko, J. et al. (2015). “On the application of Wentzel–Kramer–Brillouin theory for the simulation of the weakly nonlinear dynamics of gravity waves”. In: *Quarterly Journal of the Royal Meteorological Society* 141.688, pp. 676–697.

Olbers, D. and Eden, C. (2013). “A Global Model for the Diapycnal Diffusivity Induced by Internal Gravity Waves”. In: *Journal of Physical Oceanography* 43.8, pp. 1759–1779.

Olbers, D. et al. (2019). “The IDEMIX model: Parameterization of internal gravity waves for circulation models of ocean and atmosphere”. In: Eden, C. and Iske, A. *Energy Transfers in Atmosphere and Ocean*, pp. 87–125.

Olbers, D. et al. (2023). “A Model of Energy and Spectral Shape for the Internal Gravity Wave Field in the Deep Sea: The Parametric IDEMIX Model”. In: *Journal of Physical Oceanography* 53.5, pp. 1337–1354.

Ortland, D. A. and Alexander, M. J. (2006). “Gravity wave influence on the global structure of the diurnal tide in the mesosphere and lower thermosphere”. In: *Journal of Geophysical Research: Space Physics* 111.A10.

Plougonven, R. and Zhang, F. (2014). “Internal gravity waves from atmospheric jets and fronts”. In: *Reviews of Geophysics* 52.1, pp. 33–76.

Plumb, R. A. (1977). “The Interaction of Two Internal Waves with the Mean Flow: Implications for the Theory of the Quasi-Biennial Oscillation”. In: *Journal of Atmospheric Sciences* 34.12, pp. 1847–1858.

Quinn, B., Eden, C., and Olbers, D. (2020). “Application of the IDEMIX concept for Internal Gravity Waves in the Atmosphere”. In: *Journal of the Atmospheric Sciences*, pp. 1–59.

Sato, K. et al. (2012). “Gravity Wave Characteristics in the Southern Hemisphere Revealed by a High-Resolution Middle-Atmosphere General Circulation Model”. In: *Journal of the Atmospheric Sciences* 69.4, pp. 1378–1396.

Scherhag, R. (1952). “Die explosionsartige Stratosphärenerwärmung des Spätwinter 1951/52”. In: *Ber. Deut. Wetterdienst* 38, pp. 51–63.

Senf, F. (2012). “On the interaction between thermal tides and gravity waves in the middle atmosphere”. doctoralthesis. Universitätsbibliothek Johann Christian Senckenberg, pp. XVIII, 114.

Senf, F. and Achatz, U. (2011). “On the impact of middle-atmosphere thermal tides on the propagation and dissipation of gravity waves”. In: *Journal of Geophysical Research: Atmospheres* 116.D24.

Simmons, A. J. and Burridge, D. M. (1981). “An Energy and Angular-Momentum Conserving Vertical Finite-Difference Scheme and Hybrid Vertical Coordinates”. In: *Monthly Weather Review* 109.4, pp. 758–766.

Smith, A. K. (2012). “Interactions Between the Lower, Middle and Upper Atmosphere”. In: *Space Sci Rev* 168, pp. 1–21.

Smith, S. A., Fritts, D. C., and Vanzandt, T. E. (1987). “Evidence for a Saturated Spectrum of Atmospheric Gravity Waves”. In: *Journal of Atmospheric Sciences* 44.10, pp. 1404–1410.

Sutherland, B. R. (2018). *Internal Gravity Waves.* Camebridge University Press. ISBN: 978-1-108-45708-8.

Suzuki, S. et al. (2013). “Typhoon-induced concentric airglow structures in the mesopause region”. In: *Geophysical Research Letters* 40.22, pp. 5983–5987.

Vadas, S. L. (2007). “Horizontal and vertical propagation and dissipation of gravity waves in the thermosphere from lower atmospheric and thermospheric sources”. In: *Journal of Geophysical Research: Space Physics* 112.A6.

— (2013). “Compressible f-plane solutions to body forces, heatings, and coolings, and application to the primary and secondary gravity waves generated by a deep convective plume”. In: *Journal of Geophysical Research: Space Physics* 118.5, pp. 2377–2397.

Vadas, S. L. and Fritts, D. C. (2005). “Thermospheric responses to gravity waves: Influences of increasing viscosity and thermal diffusivity”. In: *Journal of Geophysical Research: Atmospheres* 110.D15.

Vadas, S. L. and Liu, H.-L. (2009). “Generation of large-scale gravity waves and neutral winds in the thermosphere from the dissipation of convectively generated gravity waves”. In: *Journal of Geophysical Research: Space Physics* 114.A10.

— (2013). “Numerical modeling of the large-scale neutral and plasma responses to the body forces created by the dissipation of gravity waves from 6 h of deep convection in Brazil”. In: *Journal of Geophysical Research: Space Physics* 118.5, pp. 2593–2617.

Vadas, S. L. et al. (2018). “The Excitation of Secondary Gravity Waves From Local Body Forces: Theory and Observation”. In: *Journal of Geophysical Research: Atmospheres* 123.17, pp. 9296–9325.

VanZandt, T. E. and Fritts, D. C. (1989). “A Theory of Enhanced Saturation of the Gravity Wave Spectrum Due to Increases in Atmospheric Stability”. In: *PAGEOPH* 130.10, pp. 399–420.

Wei, J., Bölöni, G., and Achatz, U. (2019). “Efficient Modeling of the Interaction of Mesoscale Gravity Waves with Unbalanced Large-Scale Flows: Pseudomomentum-Flux Convergence versus Direct Approach”. In: *Journal of the Atmospheric Sciences* 76.9, pp. 2715–2738.

Zülicke, C. and Peters, D. (2008). “Parameterization of Strong Stratospheric Inertia–Gravity Waves Forced by Poleward-Breaking Rossby Waves”. In: 136, pp. 98–119.

Appendix A

Gaussian Variational Principle

In the following the steps between the equations

$$\iint C \left[\partial_t \mathcal{E}_0 + \dot{z} \partial_z \mathcal{E}_0 + \frac{\partial_{m^*} C}{C} \mathcal{E}_0 \partial_t m^* + \dot{z} \frac{\partial_{m^*} C}{C} \mathcal{E}_0 \partial_z m^* + \partial_z \dot{z} \mathcal{E}_0 \right.$$
$$\left. + \frac{1}{C} \partial_m (\dot{m} C) \mathcal{E}_0 + \frac{1}{C} \partial_{\omega_I} (\dot{\omega}_I C) \mathcal{E}_0 - \frac{\dot{\omega}_I}{\omega_I} \mathcal{E}_0 + m^2 \mathcal{D} \mathcal{E}_0 - \frac{\mathcal{S}}{C} \right] dm \, d\omega_I = 0 \tag{A1}$$

$$\iint \partial_{m^*} C \left[\partial_t \mathcal{E}_0 + \dot{z} \partial_z \mathcal{E}_0 + \frac{\partial_{m^*} C}{C} \mathcal{E}_0 \partial_t m^* + \dot{z} \frac{\partial_{m^*} C}{C} \mathcal{E}_0 \partial_z m^* + \partial_z \dot{z} \mathcal{E}_0 \right.$$
$$\left. + \frac{1}{C} \partial_m (\dot{m} C) \mathcal{E}_0 + \frac{1}{C} \partial_{\omega_I} (\dot{\omega}_I C) \mathcal{E}_0 - \frac{\dot{\omega}_I}{\omega_I} \mathcal{E}_0 + m^2 \mathcal{D} \mathcal{E}_0 - \frac{\mathcal{S}}{C} \right] dm \, d\omega_I = 0 \tag{A2}$$

and the final prognostic equations

$$\partial_t \mathcal{E}_0 = a_0 \partial_z \mathcal{E}_0 + a_1 \partial_z m^* \mathcal{E}_0 - a_U \mathcal{E}_0 \frac{\partial_z U}{N} + a_N \mathcal{E}_0 \frac{\partial_z N}{N} - a_{\mathcal{D}} \mathcal{D} \mathcal{E}_0 + a_{\mathcal{S}} \mathcal{S} \tag{A3}$$

$$\partial_t m^* = b_0 \frac{\partial_z \mathcal{E}_0}{\mathcal{E}_0} + b_1 \partial_z m^* - b_U \frac{\partial_z U}{N} + b_N \frac{\partial_z N}{N} - b_{\mathcal{D}} \mathcal{D} + b_{\mathcal{S}} \frac{\mathcal{S}}{\mathcal{E}_0} \tag{A4}$$

are shown. We start with defining some integrals

$$\begin{aligned} I_{\mathcal{E}_0 1} &= \int\int C\, dm\, d\omega_I && \text{(A5)} \\ I_{\mathcal{E}_0 2} &= \int\int \partial_{m^*} C\, dm\, d\omega_I && \text{(A6)} \\ I_{m^* 1} &= \int\int \partial_{m^*} C\, dm\, d\omega_I && \text{(A7)} \\ I_{m^* 2} &= \int\int \frac{\partial_{m^*} C \partial_{m^*} C}{C}\, dm\, d\omega_I && \text{(A8)} \end{aligned}$$

and solve (A1) and (A2) for $\partial_t \mathcal{E}_0$ and $\partial_t m^*$ respectively

$$\begin{aligned} \partial_t \mathcal{E}_0 =& \frac{1}{I_{\mathcal{E}_0 1} I_{m^* 2} - I_{\mathcal{E}_0 2} I_{m^* 1}} \times \\ \times & \Bigg[- \int\int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \dot{z}\, dm\, d\omega_I \cdot \partial_z \mathcal{E}_0 \\ & - \int\int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \dot{z} \frac{\partial_{m^*} C}{C}\, dm\, d\omega_I \cdot \partial_z m^* \mathcal{E}_0 \\ & - \int\int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \partial_z \dot{z}\, dm\, d\omega_I \cdot \mathcal{E}_0 \\ & - \int\int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \frac{\partial_m \left(\dot{m} A(m, m^*) \right)}{A(m, m^*)}\, dm\, d\omega_I \cdot \mathcal{E}_0 && \text{(A9)} \\ & - \int\int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \frac{\partial_{\omega_I} \left(\dot{\omega}_I B(\omega_I) \right)}{B(\omega_I)}\, dm\, d\omega_I \cdot \mathcal{E}_0 \\ & + \int\int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \frac{\dot{\omega}_I}{\omega_I}\, dm\, d\omega_I \cdot \mathcal{E}_0 \\ & - \int\int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) m^2 \mathcal{D}\, dm\, d\omega_I \cdot \mathcal{E}_0 \\ & + \int\int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \frac{1}{C} \mathcal{S}\, dm\, d\omega_I \Bigg] \end{aligned}$$

$$
\begin{aligned}
\partial_t m^* =& \frac{1}{I_{\mathcal{E}_0 1} I_{m^* 2} - I_{\mathcal{E}_0 2} I_{m^* 1}} \times \\
&\times \Bigg[- \int\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \dot{z}\, dm\, d\omega_I \cdot \frac{\partial_z \mathcal{E}_0}{\mathcal{E}_0} \\
&\quad - \int\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \dot{z} \frac{\partial_{m^*} C}{C}\, dm\, d\omega_I \cdot \partial_z m^* \\
&\quad - \int\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \partial_z \dot{z}\, dm\, d\omega_I \\
&\quad - \int\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \frac{\partial_m \left(\dot{m} A(m, m^*)\right)}{A(m, m^*)}\, dm\, d\omega_I \\
&\quad - \int\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \frac{\partial_{\omega_I} \left(\dot{\omega}_I B(\omega_I)\right)}{B(\omega_I)}\, dm\, d\omega_I \\
&\quad + \int\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \frac{\dot{\omega}_I}{\omega_I}\, dm\, d\omega_I \\
&\quad - \int\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) m^2 \mathcal{D}\, dm\, d\omega_I \\
&\quad + \int\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \frac{1}{C} \mathcal{S}\, dm\, d\omega_I \frac{1}{\mathcal{E}_0} \Bigg] .
\end{aligned}
\tag{A10}
$$

We can rewrite the group velocity and and refraction and frequency modulation terms with the ray equations from Section 2.1. The vertical group velocity is

$$\dot{z} = -\frac{m}{m'^2} \omega_I \tag{A11}$$

with $m' = -\sqrt{m^2 + 1/(4H^2)}$. Because of the scale height $H(z)$ the vertical group velocity is altitude dependent, but this term negligible small.

The refraction term for the frequency $\sim \partial_{\omega_I} \left(\dot{\omega}_I B(\omega_I)\right)$ vanishes. With the frequency spectrum from Section 3.2

$$B(\omega_I) = B_0 \omega_I^{-2} \tag{A12}$$

and the frequency modulation

$$\dot{\omega}_I = -\frac{m}{m'} \frac{\omega_I^2}{N} \partial_z U \tag{A13}$$

we can see that

$$\partial_{\omega_I}\left(\dot{\omega}_I B(\omega_I)\right) = 0. \tag{A14}$$

The wavenumber refraction term $\sim \partial_m\left(\dot{m}A(m,m^*)\right)$ becomes

$$\begin{aligned}\partial_m\left(\dot{m}A(m,m^*)\right) =& -\frac{\frac{1}{m}\left(1-3\left(\frac{m}{m^*}\right)^4\right)}{1+\left(\frac{m}{m^*}\right)^4}A(m,m^*)\frac{\omega_I}{N}\partial_z N\\ &+\frac{\frac{1}{4H^2}\left(1-3\left(\frac{m}{m^*}\right)^4\right)+2m^2\left(1-\left(\frac{m}{m^*}\right)^4\right)}{m'm\left(1+\left(\frac{m}{m^*}\right)^4\right)}\times\\ &\times A(m,m^*)\frac{\omega_I}{N}\partial_z U\end{aligned} \tag{A15}$$

by applying the partial derivative of the Desaubies spectrum from Section 3.2

$$\partial_m A(m,m^*) = \frac{1}{m}\frac{1-3\left(\frac{m}{m^*}\right)^4}{1+\left(\frac{m}{m^*}\right)^4}A(m,m^*) \tag{A16}$$

and the ray equation

$$\dot{m} = \frac{\omega_I}{N}\left(m'\partial_z U - \partial_z N\right). \tag{A17}$$

Furthermore we get

$$\partial_{m^*}A(m,m^*) = \frac{\frac{1}{m^*}\left(3\left(m/m^*\right)^4-1\right)}{1+\left(m/m^*\right)^4}A(m,m^*)+\frac{\partial_{m^*}A_0(m^*)}{A_0(m^*)}A(m,m^*). \tag{A18}$$

The resulting prognostic equations (A9) and (A10) are

$$
\begin{aligned}
\partial_t \mathcal{E}_0 =& \frac{1}{I_{\mathcal{E}_0 1} I_{m^* 2} - I_{\mathcal{E}_0 2} I_{m^* 1}} \times \\
& \times \Bigg[+ \int \int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \frac{m}{m'^2} \omega_I \, dm \, d\omega_I \cdot \partial_z \mathcal{E}_0 \\
& \quad + \int \int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \frac{m}{m'^2} \omega_I \frac{\partial_{m^*} C}{C} \, dm \, d\omega_I \cdot \partial_z m^* \mathcal{E}_0 \\
& \quad + \int \int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \frac{\frac{1}{m} \left(1 - 3 \left(\frac{m}{m^*} \right)^4 \right)}{1 + \left(\frac{m}{m^*} \right)^4} \frac{\omega_I}{N} \, dm \, d\omega_I \cdot \mathcal{E}_0 \partial_z N \\
& \quad - \int \int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \times \\
& \qquad \times \frac{\frac{1}{4H^2} \left(1 - 3 \left(\frac{m}{m^*} \right)^4 \right) + 2m^2 \left(1 - \left(\frac{m}{m^*} \right)^4 \right)}{m' m \left(1 + \left(\frac{m}{m^*} \right)^4 \right)} \frac{\omega_I}{N} \, dm \, d\omega_I \cdot \mathcal{E}_0 \partial_z U \\
& \quad - \int \int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \frac{m}{m'} \frac{\omega_I}{N} \, dm \, d\omega_I \cdot \mathcal{E}_0 \partial_z U \\
& \quad - \int \int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) m^2 \mathcal{D} \, dm \, d\omega_I \cdot \mathcal{E}_0 \\
& \quad + \int \int \left(C I_{m^* 2} - \partial_{m^*} C I_{\mathcal{E}_0 2} \right) \frac{1}{C} \mathcal{S} \, dm \, d\omega_I \Bigg]
\end{aligned}
\tag{A19}
$$

and

$$
\begin{aligned}
\partial_t m^* = & \frac{1}{I_{\mathcal{E}_0 1} I_{m^* 2} - I_{\mathcal{E}_0 2} I_{m^* 1}} \times \\
\times & \left[+ \int\!\!\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \frac{m}{m'^2} \omega_I \, dm \, d\omega_I \cdot \frac{\partial_z \mathcal{E}_0}{\mathcal{E}_0} \right. \\
& + \int\!\!\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \frac{m}{m'^2} \omega_I \frac{\partial_{m^*} C}{C} \, dm \, d\omega_I \cdot \partial_z m^* \\
& + \int\!\!\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \frac{\frac{1}{m}\left(1 - 3\left(\frac{m}{m^*}\right)^4\right)}{1 + \left(\frac{m}{m^*}\right)^4} \frac{\omega_I}{N} \, dm \, d\omega_I \cdot \partial_z N \\
& - \int\!\!\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \times \\
& \quad \times \frac{\frac{1}{4H^2}\left(1 - 3\left(\frac{m}{m^*}\right)^4\right) + 2m^2\left(1 - \left(\frac{m}{m^*}\right)^4\right)}{m' m \left(1 + \left(\frac{m}{m^*}\right)^4\right)} \frac{\omega_I}{N} \, dm \, d\omega_I \cdot \partial_z U \\
& - \int\!\!\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \frac{m}{m'} \frac{\omega_I}{N} \, dm \, d\omega_I \cdot \partial_z U \\
& - \int\!\!\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) m^2 \mathcal{D} \, dm \, d\omega_I \\
& \left. + \int\!\!\int \left(-C I_{m^* 1} + \partial_{m^*} C I_{\mathcal{E}_0 1}\right) \frac{1}{C} \mathcal{S} \, dm \, d\omega_I \frac{1}{\mathcal{E}_0} \right] .
\end{aligned}
\tag{A20}
$$

We can rewrite (A19) and (A20) as

$$
\begin{aligned}
\partial_t \mathcal{E}_0 = \frac{1}{A_1 B_2 - A_2 B_1} & \left[+ \left(A_3 B_2 - B_3 A_2\right) \partial_z \mathcal{E}_0 + \left(A_4 B_2 - B_4 A_2\right) \partial_z m^* \mathcal{E}_0 \right. \\
& - \left(A_5 B_2 - B_5 A_2\right) \mathcal{E}_0 \frac{\partial_z U}{N} - \left(A_6 B_2 - B_6 A_2\right) \mathcal{E}_0 \frac{\partial_z U}{N} \\
& + \left(A_7 B_2 - B_7 A_2\right) \mathcal{E}_0 \frac{\partial_z N}{N} \\
& \left. - \left(A_8 B_2 - B_8 A_2\right) \mathcal{D} \mathcal{E}_0 + \left(A_9 B_2 - B_9 A_2\right) \mathcal{S} \right]
\end{aligned}
\tag{A21}
$$

and

$$\begin{aligned}\partial_t m^* = \frac{1}{A_1B_2 - A_2B_1} \Bigg[& + \left(-A_3B_1 + B_3A_1\right) \frac{\partial_z \mathcal{E}_0}{\mathcal{E}_0} + \left(-A_4B_1 + B_4A_1\right) \partial_z m^* \\ & - \left(-A_5B_1 + B_5A_1\right) \frac{\partial_z U}{N} - \left(-A_6B_1 + B_6A_1\right) \frac{\partial_z U}{N} \\ & + \left(-A_7B_1 + B_7A_1\right) \frac{\partial_z N}{N} \\ & - \left(-A_8B_1 + B_8A_1\right) \mathcal{D} + \left(-A_9B_1 + B_9A_1\right) \frac{\mathcal{S}}{\mathcal{E}_0} \Bigg] \end{aligned} \tag{A22}$$

with respect to the following definitions

$$A_1 = I_{\omega_I 1} \int A(m, m^*)\, dm \tag{A23}$$

$$A_2 = I_{\omega_I 1} \int \partial_{m^*} A(m, m^*)\, dm \tag{A24}$$

$$A_3 = I_{\omega_I 2} \int A(m, m^*) \frac{m}{m'^2}\, dm \tag{A25}$$

$$A_4 = I_{\omega_I 2} \int \partial_{m^*} A(m, m^*) \frac{m}{m'^2}\, dm \tag{A26}$$

$$A_5 = I_{\omega_I 2} \int A(m, m^*) \frac{m}{m'}\, dm \tag{A27}$$

$$A_6 = I_{\omega_I 2} \int A(m, m^*) \frac{\frac{1}{4H^2}\left(1-3\left(\frac{m}{m^*}\right)^4\right) + 2m^2\left(1-\left(\frac{m}{m^*}\right)^4\right)}{m' m\left(1+\left(\frac{m}{m^*}\right)^4\right)}\, dm \tag{A28}$$

$$A_7 = I_{\omega_I 2} \int A(m, m^*) \frac{\frac{1}{m}\left(1 - 3\left(\frac{m}{m^*}\right)^4\right)}{1 + \left(\frac{m}{m^*}\right)^4}\, dm \tag{A29}$$

$$A_8 = I_{\omega_I 1} \int A(m, m^*) m^2\, dm \tag{A30}$$

$$A_9 = I_{\omega_I 1} \int A(m, m^*_{\mathcal{S}})\, dm \tag{A31}$$

$$
\begin{aligned}
B_1 &= I_{\omega_I 1} \int \partial_{m^*} A(m, m^*)\, dm && \text{(A32)} \\
B_2 &= I_{\omega_I 1} \int \frac{(\partial_{m^*} A(m, m^*))^2}{A(m, m^*)}\, dm && \text{(A33)} \\
B_3 &= I_{\omega_I 2} \int \partial_{m^*} A(m, m^*) \frac{m}{m'^2}\, dm && \text{(A34)} \\
B_4 &= I_{\omega_I 2} \int \frac{(\partial_{m^*} A(m, m^*))^2}{A(m, m^*)} \frac{m}{m'^2}\, dm && \text{(A35)} \\
B_5 &= I_{\omega_I 2} \int \partial_{m^*} A(m, m^*) \frac{m}{m'}\, dm && \text{(A36)} \\
B_6 &= I_{\omega_I 2} \textstyle\int \partial_{m^*} A(m, m^*) \frac{\frac{1}{4H^2}\left(1-3\left(\frac{m}{m^*}\right)^4\right)+2m^2\left(1-\left(\frac{m}{m^*}\right)^4\right)}{m'm\left(1+\left(\frac{m}{m^*}\right)^4\right)}\, dm && \text{(A37)} \\
B_7 &= I_{\omega_I 2} \int \partial_{m^*} A(m, m^*) \frac{\frac{1}{m}\left(1-3\left(\frac{m}{m^*}\right)^4\right)}{1+\left(\frac{m}{m^*}\right)^4}\, dm && \text{(A38)} \\
B_8 &= I_{\omega_I 1} \int \partial_{m^*} A(m, m^*) m^2\, dm && \text{(A39)} \\
B_9 &= I_{\omega_I 1} \int \frac{\partial_{m^*} A(m, m^*)}{A(m, m^*)} A(m, m^*_{\mathcal{S}})\, dm && \text{(A40)} \\
& && \text{(A41)}
\end{aligned}
$$

and

$$
\begin{aligned}
I_{\omega_I 1} &= \int B(\omega_I)\, d\omega_I && \text{(A42)} \\
I_{\omega_I 2} &= \int B(\omega_I) \omega_I\, d\omega_I. && \text{(A43)}
\end{aligned}
$$

With the substitutions

$$
\begin{aligned}
a_i &= \frac{A_i B_2 - B_i A_2}{A_1 B_2 - A_2 B_1} && \text{(A44)} \\
b_i &= \frac{-A_i B_1 + B_i A_1}{A_1 B_2 - A_2 B_1} && \text{(A45)}
\end{aligned}
$$

we get

$$
\begin{aligned}
\partial_t \mathcal{E}_0 = &+ a_3 \partial_z \mathcal{E}_0 + a_4 \partial_z m^* \mathcal{E}_0 - a_5 \mathcal{E}_0 \frac{\partial_z U}{N} - a_6 \mathcal{E}_0 \frac{\partial_z U}{N} \\
&+ a_7 \mathcal{E}_0 \frac{\partial_z N}{N} - a_8 \mathcal{D} \mathcal{E}_0 + a_9 \mathcal{S}
\end{aligned}
\qquad \text{(A46)}
$$

and

$$\begin{aligned}\partial_t m^* = &+ b_3 \frac{\partial_z \mathcal{E}_0}{\mathcal{E}_0} + b_4 \partial_z m^* - b_5 \frac{\partial_z U}{N} \\ &- b_6 \frac{\partial_z U}{N} + b_7 \frac{\partial_z N}{N} - b_8 \mathcal{D} + b_9 \frac{\mathcal{S}}{\mathcal{E}_0}.\end{aligned} \tag{A47}$$

After renaming the coefficients and combining a_5 and a_6 ($a_U = a_5 + a_6$), and b_5 and b_6 ($b_U = b_5 + b_6$) we arrived at the final expressions mentioned in Section 3.3

$$\begin{aligned}\partial_t \mathcal{E}_0 = &+ a_0 \partial_z \mathcal{E}_0 + a_1 \partial_z m^* \mathcal{E}_0 - a_U \mathcal{E}_0 \frac{\partial_z U}{N} \\ &+ a_N \mathcal{E}_0 \frac{\partial_z N}{N} - a_{\mathcal{D}} \mathcal{D} \mathcal{E}_0 + a_{\mathcal{S}} \mathcal{S}\end{aligned} \tag{A48}$$

and

$$\partial_t m^* = +b_0 \frac{\partial_z \mathcal{E}_0}{\mathcal{E}_0} + b_1 \partial_z m^* - b_U \frac{\partial_z U}{N} + b_N \frac{\partial_z N}{N} - b_{\mathcal{D}} \mathcal{D} + b_{\mathcal{S}} \frac{\mathcal{S}}{\mathcal{E}_0}. \tag{A49}$$

Appendix B

Discretization of the effective spectral area

To achieve the numerical solution we need to take several steps. The first step is to filter out every vertical wavenumber from the last source layer i which would exceed the integration limits (3.57) and (3.58) of the wavenumber integration (see Sec. 3.7) at the layer j. The remaining wavenumbers are within

$$m_1 \leq m_j \leq m_0. \tag{B1}$$

Therefore we need to calculate the difference of U between the layers i and j

$$\Delta U_j = U_j - U_i. \tag{B2}$$

Then we use equation (3.63)

$$m_{j,i} = -\sqrt{\frac{N_i^2}{N_j^2} \frac{m_j^2 + \frac{1}{4H_j^2}}{1 \pm \frac{\Delta U_j}{N_j} \sqrt{m_j^2 + \frac{1}{4H_j^2}}} - \frac{1}{4H_i^2}} \tag{B3}$$

to get (together with $m_j = m_0$) the wavenumbers $m_{j,i,0}$ which will exactly reach the limit m_0 at layer j and for $m_j = m_1$ the wavenumber $m_{j,i,1}$. The wavenumbers $m_{j,i,0}$ and $m_{j,i,1}$ can become positive or complex. Then they are set to the values of the previous layer $j-1$ (e.g., $m_{j,i,1} = m_{j-1,i,1}$). For the limit $m_j = m_1$ there is another case possible, where the frequencies ω_b and ω_t are the same for wavenumbers These cut wavenumber m_c is calculated with

$$m_c(i) = \begin{cases} -\sqrt{\left(\frac{\omega_t}{\omega_b} - 1\right)^2 \frac{N_i^2}{(\Delta U_j)^2} - \frac{1}{4H_i^2}}, & \Delta U_j \quad < 0 \\ -\sqrt{\left(\frac{\omega_b}{\omega_t} - 1\right)^2 \frac{N_i^2}{(\Delta U_j)^2} - \frac{1}{4H_i^2}}, & \Delta U_j \quad > 0 \end{cases} . \tag{B4}$$

If this cut wavenumber fulfils $m_c(i) > m_{j,i,1}$, then

$$m_{j,i,1} = m_c(i). \tag{B5}$$

Now we need to make sure to not bring back wavenumbers, which are already filtered out in the previous layer $j-1$

$$m_{j,i,0} = \begin{cases} m_{j,i-1,0}, & m_{j,i,0} \quad > m_{j,i-1,0} \\ m_{j,i,0}, & m_{j,i,0} \quad < m_{j,i-1,0} \end{cases} \tag{B6}$$

and

$$m_{j,i,1} = \begin{cases} m_{j,i-1,1}, & m_{j,i,1} \quad < m_{j,i-1,1} \\ m_{j,i,1}, & m_{j,i,1} \quad > m_{j,i-1,1} \end{cases} . \tag{B7}$$

A spectrum in the limits $(m_{j,i,1}, m_{j,i,0})$ starting at the layer i will not exceed the general boundaries for the vertical wavenumber at the layer j. If the source spectrum $[m_{11}, m_{00}]$ (e.g.,see Eq. 4.21) is inside of $[m_{j,i,1}, m_{j,i,0}]$ the we get

$$m_{j,i,1} = m_{11} \tag{B8}$$

$$m_{j,i,0} = m_{00} \tag{B9}$$

A sketch of the different wavenumber boundaries $[m_{j,i,1}, m_{j,i,0}]$ for different layers j form a source layer i is shown in Figure B1.

The second step is to discretize the wavenumber spectrum as well as the

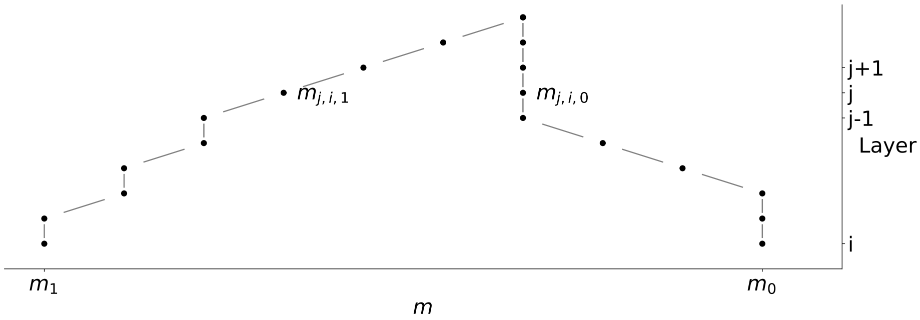

Figure B1: Sketch of the background dependency of the vertical wavenumber limits. Only waves in the set of $[m_{j,i,1}, m_{j,i,0}]$ can reach the layer j from the source layer i without exceeding the general boundaries at any layer between i and j.

frequency spectrum at layer i. The continuous wavenumber spectrum $m = [m_{j,i,1}, m_{j,i,0}]$ of layer j will be discretized in dependence of the characteristic vertical wavenumber m_j^*. For $k = [0, n-1] \in \mathbb{N}$:

$$m_{j,k} = \begin{cases} m_{j,i,0}, & k \ = 0 \\ -\sqrt{m^*(j)^2 \tan\left[-\frac{1}{n-2}\left(\tan^{-1}\left(\frac{m_{j,i,0}^2}{m_j^{*2}}\right) - \tan^{-1}\left(\frac{m_{j,i,1}^2}{m_j^{*2}}\right)\right) + \tan^{-1}\left(\frac{m_{j,k-1}^2}{m_j^{*2}}\right)\right]}, & k \ > 1 \end{cases} \tag{B10}$$

with n nodes. This method leads to areas of the same size between two nodes (see Fig. B2). We start from the uppermost source layer i with a unique discretized wavenumber spectrum for each layer j. With that we make sure, that we retain all nodes up to layer j and that we represent the region of high energy density inside the spectrum around m_j^* properly.

For the frequency limits we need the upper limit ω_T and the lower limit ω_B for spectrum at layer j. The frequency $\omega_I(i)$ depends on the wavenumber and frequency from the previous layer (see Eq. 3.62), so we have an upper and lower frequency limit for each $m_{j,k}$

$$\omega_{t,j,k} = \omega_t \tag{B11}$$

$$\omega_{b,j,k} = \omega_b. \tag{B12}$$

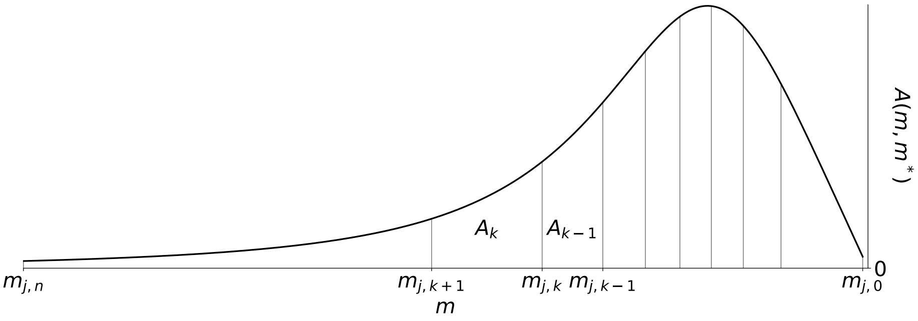

Figure B2: Sketch of the discretized vertical wavenumber spectrum in dependence of m_j^*

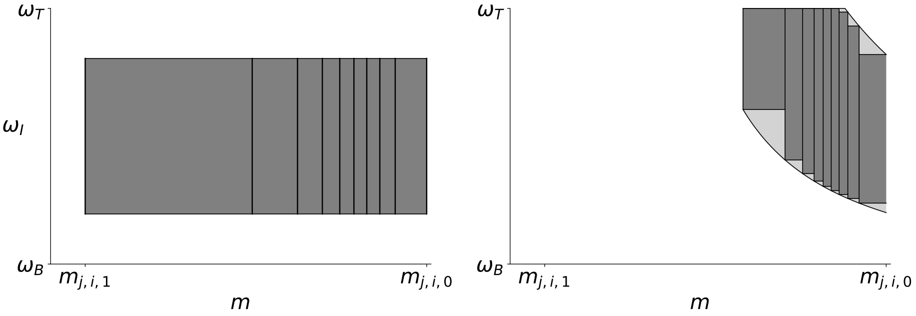

Figure B3: Left: the discretized spectral area at layer i; right: the discretize spectral area at layer j, in light gray the actual spectral area

The result is a spectrum, which is discretized into $n-1$ rectangles (see Fig. B3).

Third:

We calculate the change of $m_{j,k}$, $\omega_{b,j,k}$, and $\omega_{t,j,k}$ from layer j to layer $j+1$: We use the equations (3.62) and (3.63)

$$\omega_{b/t,j+1,k} = \omega_{b/t,j,k}\left(1 \mp \frac{U_{j+1} - U_j}{N_j}\sqrt{m_{j,k}^2 + \frac{1}{4H_j^2}}\right) \tag{B13}$$

$$m_{j+1,k} = -\sqrt{\frac{N_{j+1}^2}{N_j^2}\frac{m_{j,k}^2 + \frac{1}{4H_j^2}}{\left(1 \mp \frac{U_{j+1}-U_j}{N_j}\sqrt{m_{j,k}^2 + \frac{1}{4H_j^2}}\right)^2} - \frac{1}{4H_{j+1}^2}} \tag{B14}$$

with the additional conditions

$$\begin{aligned} \omega_B &\leq \omega_{b,j,k} \leq \omega_T \qquad \text{(B15)} \\ \omega_B &\leq \omega_{t,j,k} \leq \omega_T \qquad \text{(B16)} \end{aligned}$$

to filter out frequencies which are reflected ($\omega_I > \omega_T$) or are sttenuated due to a critical layer ($\omega_I < \omega_B$).

Final step:

We solve the double integrals (they have an analytic solution)

$$I_k = \int_{\omega_b}^{\omega_t} \int_{m_k}^{m_{k-1}} f(m, m^*, \omega_I)\, dm\, d\omega_I \qquad \text{(B17)}$$

with

$$\omega_b = \begin{cases} \omega_b(k-1), & \omega_b(k-1) \quad > \omega_b(k) \\ \omega_b(k), & \text{else} \end{cases} \qquad \text{(B18)}$$

and

$$\omega_t = \begin{cases} \omega_t(k-1), & \omega_t(k-1) \quad < \omega_t(k) \\ \omega_t(k), & \text{else} \end{cases} . \qquad \text{(B19)}$$

These conditions make sure that we stay within the spectral limits (see right panel of Fig. B3). The solution for the integral over the remaining spectrum is

$$I = \sum_{k=1}^{n-1} I_k. \qquad \text{(B20)}$$

Appendix C

Derivation of the fluxes, and the wave momentum density

In the following, we present the derivations of the vertical momentum flux $\rho_r\langle u'w'\rangle$, the wave momentum density $\langle \rho' u'\rangle$, and the vertical pressure flux $\rho_r\langle p'w'\rangle$:

$$
\begin{aligned}
\rho_r\langle u'w'\rangle &= \rho_r\langle \Re(u^*)\Re(w^*)\rangle \\
&\qquad (\langle \Re(X_1^*)\Re(X_2^*)\rangle = \frac{1}{4}\langle X_1^* X_2^{*\dagger} + X_1^{*\dagger} X_2^*\rangle) \\
&= \frac{1}{4}\rho_r\langle u^* w^{*\dagger} + u^{*\dagger} w^*\rangle \\
&\qquad \text{(eq. 2.18)} \\
&= \frac{1}{2}\rho_r\langle \frac{\omega_I}{N}\frac{m}{m'} u^{*\dagger} u^*\rangle \\
&\qquad (\frac{1}{2}\rho_r\langle f(m,\omega_I) u^* u^{*\dagger}\rangle) = \iint f(m,\omega_I)\mathcal{E}(z,t,m,\omega_I)\,dm\,d\omega_I \\
&= \iint \frac{\omega_I}{N}\frac{m}{m'}\mathcal{E}(z,t,m,\omega_I)\,dm\,d\omega_I
\end{aligned}
\tag{C1}
$$

$$
\begin{aligned}
\langle \rho' u' \rangle &= \langle \Re(\rho^*)\Re(u^*)\rangle \\
&\qquad (\langle \Re(X_1^*)\Re(X_2^*)\rangle = \frac{1}{4}\langle X_1^* X_2^{*\dagger} + X_1^{*\dagger} X_2^*\rangle) \\
&= \frac{1}{4}\langle \rho^* u^{*\dagger} + \rho^{*\dagger} u^*\rangle \\
&\qquad (\text{eq. } 2.20) \\
&= \frac{1}{2}\frac{\rho_r}{g}\langle \frac{N}{m'}\frac{1}{2H} u^{*\dagger} u^*\rangle \\
&\qquad (\frac{1}{2}\rho_r\langle f(m,\omega_I) u^* u^{*\dagger}\rangle) = \iint f(m,\omega_I)\mathcal{E}(z,t,m,\omega_I)\,dm\,d\omega_I \\
&= \frac{1}{g}\iint \frac{N}{m'}\frac{1}{2H}\mathcal{E}(z,t,m,\omega_I)\,dm\,d\omega_I
\end{aligned}
\tag{C2}
$$

$$
\begin{aligned}
\langle p' w' \rangle &= \langle \Re(p^*)\Re(w^*)\rangle \\
&\qquad (\langle \Re(X_1^*)\Re(X_2^*)\rangle = \frac{1}{4}\langle X_1^* X_2^{*\dagger} + X_1^{*\dagger} X_2^*\rangle) \\
&= \frac{1}{4}\langle p^* w^{*\dagger} + p^{*\dagger} w^*\rangle \\
&\qquad (\text{eqs. } 2.18 \text{ and } 2.19) \\
&= -\frac{1}{2}\rho_r\langle \frac{\omega_I m}{m'^2} u^{*\dagger} u^*\rangle \\
&\qquad (\frac{1}{2}\rho_r\langle f(m,\omega_I) u^* u^{*\dagger}\rangle) = \iint f(m,\omega_I)\mathcal{E}(z,t,m,\omega_I)\,dm\,d\omega_I \\
&= -\iint \frac{\omega_I m}{m'^2}\mathcal{E}(z,t,m,\omega_I)\,dm\,d\omega_I
\end{aligned}
\tag{C3}
$$

Appendix D

Source functions

In the following, the source functions for the generation of non-orographic gravity waves applied in the GCM simulations presented in Chapter V are introduced. The following field variables (and if needed their derivatives) are assumed to be known (and in practice are taken from the GCM): the temperature T, the horizontal winds U and V, the pressure p and the surface pressure p_{00}, the Coriolis frequency f, the geopotential Φ, and the moist-convective heating Q_c.

Constant background generation

The constant background generation source function describes a weak generation of gravity waves of unspecified origin

$$\mathcal{S}_{const}(z,t,\varphi) = \begin{cases} S_{const} S_0 \sin(\phi)\phi, & 120\,\mathrm{hPa} < p < 950\,\mathrm{hPa} \\ 0, & \text{else} \end{cases} \tag{D1}$$

with the parameters

$$S_0 = 4 \times 10^{-1} \,\frac{\mathrm{kg}}{\mathrm{ms}^2} \tag{D2}$$

$$S_{const} = 1 \times \frac{1}{86400}\mathrm{s}^{-1}. \tag{D3}$$

The energy parameter S_0 defines how much gravity wave energy is generated during the time $1/S_{const}$ where S_{const} is a generation rate. With the current setting, a gravity wave energy density of $4 \times 10^{-1}\,\mathrm{kg/(ms^2)}$ is generated within one day.

Generation by convection

To construct a source function for the generation of gravity waves by moist convection we apply the method of Mirzaei et al. (2014). The usual approach is that latent heating acts as a forcing function (Alexander et al. 2004). The source function for the generation of gravity waves by convection is

$$\mathcal{S}_{conv}(z,t,\varphi) = \begin{cases} \rho S_{conv} Q_c^2 \left(\frac{1}{\partial_z \Theta_0} \frac{L_h}{L_z}\right)^2, & Q_c > Q_{th} \\ 0, & \text{else} \end{cases} \tag{D4}$$

at altitudes corresponding to $120\,\mathrm{hPa} < p < 950\,\mathrm{hPa}$ with the threshold set to $Q_{th} = 5 \times 10^{-5}\,\mathrm{K/s}$ or in other units $Q_{th} = 2 \times 10^{-1}\,\mathrm{K/h}$. We use $L_h/L_z = 100$ as suggested by Mirzaei et al. (2014) and Amiramjadi et al. (2020) for the aspect ratio of the convective zone. The vertical potential temperature gradient is set to $\partial_z \Theta_0 = 3.1 \times 10^{-3}\,\mathrm{K/m}$. The generation rate is assumed to be

$$S_{conv} = 4 \times 10^{-2} \frac{1}{86400}\mathrm{s}^{-1}. \tag{D5}$$

Note that convective cells in the real atmosphere are on a horizontal scale of a few kilometers which even high resolution models can not resolve. The convective heating Q_c is a parameter averaged on the spatial and temporal scale of the GCM. The parameter describes the convective activity in a certain region.

Generation by fronts

To parameterize the generation of gravity waves by fronts we apply a source function that is a combination of the generation functions from Charron and Manzini (2002) and Mirzaei et al. (2014). Isothermal compression by wind deformation and convergence lead to frontogenesis, where large horizontal temperature gradients generate gravity waves. Since low resolution models usually do not resolve fronts, the frontogenesis function (Miller 1948; Hoskins 1982) is applied. This function indicates when fronts would occur in the case of sufficiently high resolution (Charron and Manzini 2002). The frontogenesis function is (e.g., Mirzaei et al. 2014)

$$\begin{aligned} F_f = & - \frac{\left((\partial_x\Theta)^2 + (\partial_y\Theta)^2\right)(\partial_x U + \partial_y V)}{2\sqrt{(\partial_x\Theta)^2 + (\partial_y\Theta)^2}} \\ & - \frac{\left((\partial_x\Theta)^2 - (\partial_y\Theta)^2\right)(\partial_x U - \partial_y V)}{2\sqrt{(\partial_x\Theta)^2 + (\partial_y\Theta)^2}} \\ & - \frac{\partial_x\Theta\partial_y\Theta\,(\partial_y U + \partial_x V)}{\sqrt{(\partial_x\Theta)^2 + (\partial_y\Theta)^2}} \end{aligned} \tag{D6}$$

with the potential temperature

$$\Theta = T\left(\frac{p_{00}}{p}\right)^{R/c_p}, \tag{D7}$$

where p_{00} is the mean surface pressure at sea level. Gravity waves are generated in the cross-front directions φ_{front} and $\varphi_{front}+\pi$ (Charron and Manzini

2002). The azimuth angle is

$$\varphi_{front} = \tan^{-1}\left(\cos(\phi)\frac{\partial_y\Theta}{\partial_x\Theta}\right). \tag{D8}$$

The source functions for waves in zonal directions are

$$\mathcal{S}_{front}(z,t,\varphi_{1,2}) = \begin{cases} S_{front}\rho F_f^2\cos^2(\varphi_{front})\left(\frac{gL_z}{\Theta_0 f^2}\right)^2, & F_f > F_{th} \\ 0, & \text{else} \end{cases} \tag{D9}$$

and for waves in meridional directions

$$\mathcal{S}_{front}(z,t,\varphi_{3,4}) = \begin{cases} S_{front}\rho F_f^2\sin^2(\varphi_{front})\left(\frac{gL_z}{\Theta_0 f^2}\right)^2, & F_f > F_{th} \\ 0, & \text{else} \end{cases} \tag{D10}$$

with the threshold set to $F_{th} = 2.7\times10^{-10}\,\mathrm{K/(ms)}$ or in other units $F_{th} = 1\times10^{-1}\,\mathrm{K/(100\,km\,h)}$. The azimuths of the gravity waves are indicated by φ_i with φ_1 for eastward, φ_2 for westward, φ_3 for northward, and φ_4 for southward direction. The generation rate is set to

$$S_{front} = 2\times10^{-1}\frac{1}{86400}\mathrm{s}^{-1} \tag{D11}$$

Furthermore, we use the reference potential temperature $\Theta_0 = 300\,\mathrm{K}$ and the vertical scale of the fronts $L_z = 2000\,\mathrm{m}$ from Mirzaei et al. (2014).

Generation by jets

To construct a source function for jet generated gravity waves we apply the methods of Zülicke and Peters (2008) and Mirzaei et al. (2014). Important for jet-generated gravity waves are the exit regions of jet streaks (Koch and Dorian 1988). The Lagrangian wind speed deceleration

$$fu_c = f\frac{U_{ag}V_g - U_gV_{ag}}{\sqrt{U_g^2+V_g^2}}, \tag{D12}$$

with the geostrophic winds

$$U_g = -\frac{\partial_y \Phi}{f} \tag{D13}$$

$$V_g = \frac{\partial_x \Phi}{f}, \tag{D14}$$

where Φ is the geopotential, and the ageostrophic winds

$$U_{ag} = U - U_g \tag{D15}$$

$$V_{ag} = V - V_g, \tag{D16}$$

are used to diagnose flow imbalances in jet exit regions (Zülicke and Peters 2008). The energy of the generated gravity waves is proportional to the squared cross-stream ageostrophic wind u_c if a certain wind threshold u_{th} is exceeded (Mirzaei et al. 2014). The source function for jet generated gravity is

$$\mathcal{S}_{jet}(z,t,\varphi) = \begin{cases} S_{jet}\rho u_c^2 \frac{\cos^2(\varphi)U^2+\sin^2(\varphi)V^2}{U^2+V^2}, & u_c > u_{th} \\ 0, & \text{else} \end{cases} \tag{D17}$$

at altitudes corresponding to $120\,\text{hPa} < p < 950\,\text{hPa}$ and with the threshold set to $u_{th} = 3.5\,\text{m/s}$. The additional free parameter (beside the threshold) is the generation rate which is set to

$$S_{jet} = 16 \times 10^{-2} \frac{1}{86400} \text{s}^{-1}. \tag{D18}$$

Acknowledgment

As all of the following explicitly and implicitly mentioned persons know, I am a man of few words. Nevertheless I would like to express my thanks to the following persons and groups: First of all my doctoral adviser Prof. Dr. rer. nat. habil. Erich Becker for giving me the opportunity to work on the field of gravity waves, a field I did not know anything about before this thesis, and his endless trust in and support of my approaches during the last years against all odds. Furthermore the department of (theory and) modelling at the IAP Kühlungsborn and here especially Dr. Serhat Can for countless fruitful discussions not only about atmospheric physics but also topics far away from that. The former head of the IAP Prof. Dr. Franz-Josef Lübken and his successor Prof. Dr. Claudia Stolle for the administrative support over the last years. Prof Dr. Carsten Eden and the DFG project TRR-181 without which this thesis would not exist. And last but not least Prof. Dr. Ulrich Achatz and Prof. Dr. Dirk Olbers for a lot of valuable discussions and comments. Without their prework over the last decades on gravity waves and their parameterization the TRR-181 subproject W1 and subsequent this thesis would also not exist.