Efficient modelling of the interaction of mesoscale gravity waves with unbalanced large-scale flows: Pseudomomentum-flux convergence versus direct approach

Junhong Wei

School of Atmospheric Sciences, and Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, Sun Yat-sen University, Guangzhou, China *

Gergely Bölöni

Institut für Atmosphäre und Umwelt, Goethe Universität Frankfurt am Main, Germany.

Ulrich Achatz †

Institut für Atmosphäre und Umwelt, Goethe Universität Frankfurt am Main, Germany.

* Former affiliation: Institut für Atmosphäre und Umwelt, Goethe Universität Frankfurt am Main, Germany.

†*Corresponding author address:* Institut für Atmosphäre und Umwelt, Goethe Universität Frankfurt am Main, Altenhöferallee 1, 60438 Frankfurt am Main, Germany.

E-mail: achatz@iau.uni-frankfurt.de

ABSTRACT

This paper compares two different approaches for the efficient modelling of sub-grid-scale inertia-gravity waves in a rotating compressible atmosphere. The first approach, denoted as pseudomomentum scheme, exploits the fact that in a Lagrangian-mean reference frame the response of a large-scale flow can only be due to forcing momentum. Present-day gravity-wave parameterizations follow this route. They do so, however, in a Eulerian-mean formulation. Transformation to that reference frame leads, under certain assumptions, to pseudomomentum-flux convergence by which the momentum is to be forced. It can be shown that this approach is justified if the large-scale flow is in geostrophic and hydrostatic balance. Otherwise, elastic and thermal effects might be lost. In the second approach, called direct scheme and not relying on such assumptions, the large-scale flow is forced both in the momentum equation, by anelastic momentum-flux convergence and an additional elastic term, and in the entropy equation, via entropy-flux convergence. A budget analysis based on one-dimensional wave packets suggests that the comparison between the above-mentioned two schemes should be sensitive to the following two parameters: 1) the intrinsic frequency, and 2) the wave packet scale. The smaller the intrinsic frequency is, the greater their differences are. More importantly, with high-resolution wave-resolving simulations as a reference, this study shows conclusive evidence that the direct scheme is more reliable than the pseudomomentum scheme, regardless whether one-dimensional or two-dimensional wave packets are considered. In addition, sensitivity experiments are performed to further investigate the relative importance of each term in the direct scheme, as well as the wave-mean-flow interactions during the wave propagation.

1. Introduction

As one of the most fundamental physical modes in meteorology, gravity waves (GWs) are ubiquitous buoyancy oscillations in the atmosphere. The sources of excited gravity waves include, among others, topographic forcing (Smith 1980; Menchaca and Durran 2017), convection (Alexander et al. 1995; Lane et al. 2001), the jets (Zhang 2004; Plougonven and Zhang 2014; Hien et al. 2018), frontal systems (Snyder et al. 1993; Griffiths and Reeder 1996), shear instability (Bühler et al. 1999; Bühler and McIntyre 1999). GW dynamics is also influenced by rotation, especially when the respective waves have relatively long horizontal wavelengths and short vertical wavelengths. These are called inertia-gravity waves (IGWs), and they are characterized by parcel oscillations that are elliptical instead of straight lines in the pure gravity wave case without earth rotation effect (Gill 1982; Nappo 2002). In addition to that, IGWs have relatively low intrinsic frequencies and small vertical group velocities. Many studies have documented the signals and the life cycles of IGWs associated with complex background flows, using observations (Wang and Geller 2003; Plougonven et al. 2003; Gong et al. 2012), numerical investigations of observed cases (Zhang et al. 2013, 2015), and idealized simulations (Wei and Zhang 2014, 2015; Wei et al. 2016).

GWs play a significant role in atmospheric dynamics on various spatial and temporal scales. For example, GWs can generate and modulate atmospheric turbulence (Shapiro 1980; Lane et al. 2004), and they can also initiate and organize moist convection (Zhang et al. 2001; Lane and Zhang 2011). As an important candidate, GWs can also potentially contribute significantly to building the atmospheric energy spectra (Callies et al. 2014; Bierdel et al. 2016; Sun et al. 2017). GWs are found to likely link the small-scale small-amplitude initial error to the rapid upscale error growth and thus limit mesoscale predictability (Zhang et al. 2007; Sun and Zhang 2016; Bierdel et al. 2018). The most important impact globally is due to the fact that GWs can travel over

large distances from their sources and transfer significant amounts of momentum and energy to high altitudes, which contributes to the forcing of the circulation and the variability of the middle atmosphere (Holton and Lindzen 1972; Houghton 1978; Lindzen 1981; Dunkerton 1997; Richter et al. 2010; Limpasuvan et al. 2012; Butchart 2014). The dynamics of the middle atmosphere can influence the tropospheric circulation by downward control (Haynes et al. 1991), and it can be very important for the forecasting of weather (Baldwin and Dunkerton 2001) and climate (Scaife et al. 2005, 2012).

Despite the increasing computational power, an important range of GW spatial scales remains unresolved in most atmospheric global circulation models (GCM) or in global numerical weather prediction (NWP) models (Alexander et al. 2010), and GW parameterizations are still applied in those models (Medvedev and Klaassen 1995; Hines 1997a,b; Lott and Miller 1997; Alexander and Dunkerton 1999; Warner and McIntyre 2001; Lott and Guez 2013). Most of the parameterization schemes are based on the Wentzel-Kramer-Brillouin (WKB) theory, but with some oversimplifications, including the assumption of a steady-state wave field and background flow, instantaneous GW propagation, and one-dimensional vertical propagation. Those above-mentioned oversimplifications are often made to ensure numerical stability and model efficiency, but they can lead to the neglect of important aspects of the interaction between GWs and mean flow (Bühler and McIntyre 1999, 2003, 2005; Dosser and Sutherland 2011; Bölöni et al. 2016). The spectral phase-space representation (Bühler and McIntyre 1999; Hertzog et al. 2002) turns out to be effective to avoid numerical instabilities due to caustics. The issue of caustics occurs when by wave mean-flow interactions a wave field loses its initial locally monochromatic character (Tabaei and Akylas 2007; Rieper et al. 2013a). Based on the spectral approach, a prognostic Lagrangian WKB GW ray tracing model has been implemented, using a multi-dimensional phase space spanned by physical space and the wavenumber space, and it has been successfully validated against wave-resolving

large-eddy simulation (LES) data in different idealized settings, including one-dimensional (1D) vertically propagating idealized wave packets with variable vertical wavenumbers in a non-rotating Boussinesq atmosphere (Muraschko et al. 2015), 1D wave packets propagating in a non-rotating compressible atmosphere (Bölöni et al. 2016), and two-dimensional (2D) wave packets of parameterized sub-mesoscale GWs interacting with resolved mesoscale GWs in a rotating Boussinesq atmosphere (Wilhelm et al. 2018). In addition, this approach has also been used to study the interaction between GWs and solar tides (Ribstein et al. 2015; Ribstein and Achatz 2016).

However, there is still no corresponding detailed validation of an IGW parameterization in a rotating compressible atmosphere against data from idealized wave resolving simulations. This gives one of the motivations for the current study. More importantly, there are actually two available approaches for IGW parameterization, here called pseudomomentum scheme and direct scheme, respectively. The pseudomomentum scheme exploits the fact that in a Lagrangian-mean reference frame the effect of GWs on the large-scale flow only appears in the momentum equation (Andrews and McIntyre 1978). Application of this theory to a Eulerian-mean reference frame leads to a pseudomomentum-flux convergence by which the large-scale momentum is to be forced (Andrews and McIntyre 1976, 1978). As will be shown below this is at least justified if the large-scale flow is in geostrophic and hydrostatic balance. The direct scheme does not rely on any balance assumption with regard to the large-scale flow, and the large-scale flow is forced by anelastic momentum-flux convergence in the momentum equation, an elastic term also in the momentum equation, and entropy-flux convergence in the entropy equation, as given by Grimshaw (1975) and Achatz et al. (2017). All present-day operational IGW parameterizations represent, one way or other, simplified versions of the pseudomomentum approach, where the vertical gradient of pseudomomentum-flux convergence forces the resolved flow, when wave dissipation occurs (Fritts and Alexander 2003; Kim et al. 2003), and neither elastic nor thermal effects are taken into

account. In contrast, to the best of our knowledge, there is still no numerical application of the direct scheme. This further motivates us to understand the differences between these two available schemes, to investigate which scheme performs better in validations against data from idealized wave resolving simulations, and to verify how trustworthy the pseudomomentum approach is when the large-scale flow is unbalanced.

This article is arranged as follows. The theory part related to the two available approaches for the IGW parameterization will be shown in section 2, followed in section 3 by a brief introduction to the numerical models, and in section 4 by a presentation of the various cases used. Section 5 will present the budget analysis of the wave-induced forcing terms from the two different schemes in different idealized wave packet profiles. Based on the prognostic Lagrangian WKB GW ray tracing model, both schemes will be verified against the data from wave resolving simulations in section 6. In section 7, sensitivity experiments are performed to further investigate the relative importance of each wave-induced forcing term, as well as the wave-mean-flow interactions during the wave propagation. Section 8 contains a summary.

2. Theory

For an explanation of the theoretical underpinnings of the two respective approaches we follow the presentation of Achatz et al. (2017) where, expanding on previous work by Grimshaw (1975), the theory is discussed mostly in non-dimensional form. We translate the essentials into dimensional form and choose, for easier tractability, a heuristic formulation. For all mathematical details the reader is referred back to Achatz et al. (2017).

Starting point are the compressible Euler equations on an f-plane, (e.g. Durran 1989), with Coriolis parameter f, without external sources or sinks,

$$\begin{aligned}
\frac{D\mathbf{u}}{Dt} + f\mathbf{e}_z \times \mathbf{u} &= -c_p\theta\nabla_h\pi && (1)\\
\frac{Dw}{Dt} &= -c_p\theta\frac{\partial\pi}{\partial z} - g && (2)\\
\frac{D\theta}{Dt} &= 0 && (3)\\
\frac{D\pi}{Dt} + \frac{R}{c_V}\pi\nabla\cdot\mathbf{v} &= 0 && (4)
\end{aligned}$$

where $\mathbf{u}$ and w are the horizontal and vertical components of the total wind $\mathbf{v}$, respectively. c_p and $c_V = c_p - R$ are the specific heat capacities at constant volume and pressure, respectively, with R the ideal gas constant of dry air. θ is potential temperature, π the Exner pressure, and g the gravitational acceleration.

Within this setting we consider a superposition of an exclusively altitude-dependent hydrostatic reference atmosphere at rest, a rather general synoptic-scale flow, and a locally monochromatic small-scale wave field,

$$\begin{aligned}
\mathbf{v} &= \mathbf{V} + \widetilde{\mathbf{V}} + \Re\left(\hat{\mathbf{v}}e^{i\phi}\right) + \ldots && (5)\\
\theta &= \bar{\theta} + \Theta + \widetilde{\Theta} + \Re\left(\hat{\theta}e^{i\phi}\right) + \ldots && (6)\\
\pi &= \bar{\pi} + \Pi + \widetilde{\Pi} + \Re\left(\hat{\pi}e^{i\phi}\right) + \ldots && (7)
\end{aligned}$$

Here $\bar{\theta}(z)$ and $\bar{\pi}(z)$ are potential temperature and Exner pressure of a horizontally symmetric and steady reference atmosphere. All other fields depend on all three spatial coordinates and on time. Using a $\alpha = 0,1$ we discriminate between a weakly stratified atmosphere, as in the troposphere, indicated by $\alpha = 1$, and a moderately strongly stratified case ($\alpha = 0$), characterizing

the stratosphere, so that

$$\bar{\theta} = \begin{cases} \bar{\theta}_0(z), & \text{if } \alpha = 0 \\ \bar{\theta}_0 + \bar{\theta}_1(z), & \text{if } \alpha = 1 \end{cases} \tag{8}$$

to be understood so that $\bar{\theta}_0$ is a constant in the second case, and $\bar{\theta}_1/\bar{\theta}_0 = \mathscr{O}(\varepsilon)$, where $\varepsilon \ll 1$ is the Rossby number. Capital letters indicate the leading-order and next-order (with tilde, so that $|\widetilde{\mathbf{V}}|/|\mathbf{V}| = \mathscr{O}(\varepsilon)$, e.g.) synoptic-scale flow, and hatted quantities the leading-order amplitudes of a locally monochromatic wave field, with local phase ϕ, yielding local wave number $\mathbf{k} = \nabla\phi$ and frequency $\omega = -\partial\phi/\partial t$. All additional contributions are next-order terms, in an expansion of all fields in terms of ε, and higher harmonics of the wave, induced by nonlinear interactions.

The synoptic-scale flow is assumed to vary on horizontal and vertical scales L_s and H_s, respectively, and on the time scale T_s. Following standard considerations one can show that for a rotating atmospheric layer with inertia f and characteristic temperature T_{00} these synoptic scales are represented by

$$L_s = \varepsilon^{\alpha/2}\sqrt{RT_{00}}/f \qquad H_s = \varepsilon^{5/2}\sqrt{RT_{00}}/f \qquad T_s = \varepsilon^{-1}/f \tag{9}$$

The various dynamical fields scale as

$$\mathbf{U} = \mathscr{O}\left(\varepsilon^{(2+\alpha)/2}\sqrt{RT_{00}}\right) \quad W = \mathscr{O}\left(\varepsilon^{7/2}\sqrt{RT_{00}}\right) \quad \Theta/\bar{\theta} = \mathscr{O}\left(\varepsilon^{1+\alpha}\right) \quad \Pi/\bar{\pi} = \mathscr{O}\left(\varepsilon^{1+\alpha}\right) \tag{10}$$

The wave field is assumed to vary on spatial and temporal scales L_w, H_w, and T_w shorter than those of the synoptic-scale flow

$$L_w = \varepsilon L_s \qquad H_w = \varepsilon H_s \qquad T_w = \varepsilon T_s \tag{11}$$

and its dynamical fields are those of a large-amplitude wave close to the margin of static instability, leading to

$$\hat{\mathbf{v}} = \mathscr{O}(\mathbf{V}) \quad \hat{\theta}/\bar{\theta} = \mathscr{O}\left(\Theta/\bar{\theta}\right) \quad \hat{\pi}/\bar{\pi} = \mathscr{O}(\varepsilon\,\Pi/\bar{\pi}) \tag{12}$$

With these scaling assumptions one obtains the following main results: (a) Frequency and wave number are connected by the dispersion relations of either geostrophic modes, not to be pursued here any further, or IGWs

$$\hat{\omega} = \omega - \mathbf{k}\cdot\mathbf{U} = \pm\sqrt{f^2 + N^2\frac{k^2+l^2}{m^2}} \tag{13}$$

where k, l, and m are the wavenumber components in x, y and z direction, respectively, and $N^2 = (g/\bar{\theta})d\bar{\theta}/dz$ the squared Brunt-Väisala frequency. Hence, with

$$\Omega(\mathbf{x},t,\mathbf{k}) = \mathbf{k}\cdot\mathbf{U}(\mathbf{x},t) \pm \sqrt{f^2 + N^2(z)\frac{k^2+l^2}{m^2}} \tag{14}$$

and its wavenumber gradient $\mathbf{c}_g = \nabla_{\mathbf{k}}\Omega$ the local group velocity, the prognostic equation for the local wavenumber is given by

$$\dot{\mathbf{k}} = \left(\frac{\partial}{\partial t} + \mathbf{c}_g\cdot\nabla\right)\mathbf{k} = -\nabla\Omega \tag{15}$$

(b) The IGW amplitudes satisfy the polarization relations, with $\mathbf{k}_h = k\mathbf{e}_x + l\mathbf{e}_y$ the horizontal wavenumber and $\hat{b} = g\hat{\theta}/\bar{\theta}_0$ the wave buoyancy amplitude,

$$\hat{\mathbf{u}} = \frac{\hat{\omega}\mathbf{k}_h - if\mathbf{e}_z\times\mathbf{k}_h}{(\hat{\omega}^2 - f^2)\,im}\hat{b} \tag{16}$$

$$\hat{w} = \frac{i\hat{\omega}}{N^2}\hat{b} \tag{17}$$

$$c_p\bar{\theta}_0\hat{\pi} = \frac{\hat{b}}{im} \tag{18}$$

(c) Higher harmonics are negligible to leading order. (d) The leading order amplitudes satisfy wave-action conservation

$$\frac{\partial\mathscr{A}}{\partial t} + \nabla\cdot(\mathbf{c}_g\mathscr{A}) = 0 \tag{19}$$

with $\mathscr{A} = E_w/\hat{\omega}$ the wave action, where

$$E_w = \frac{\bar{\rho}}{2}\left(\frac{|\hat{\mathbf{u}}|^2}{2} + \frac{\left|\hat{b}\right|^2}{2N^2}\right) \tag{20}$$

is wave energy, and $\bar{\rho}$ the leading-order reference-atmosphere density.

Of central interest here are the following results: (e) To leading order the synoptic-scale flow is horizontal

$$W = 0 \tag{21}$$

in geostrophic equlilbrium

$$\mathbf{U} = \mathbf{e}_z \times \nabla_h \Psi \tag{22}$$

with $\Psi = c_p \bar{\theta}_0 \Pi / f$ the stream function given by the leading-order synoptic-scale pressure fluctuations, and the leading order synoptic-scale flow is hydrostatic

$$g\frac{\Theta}{\bar{\theta}_0} = f\frac{\partial \Psi}{\partial z} + \begin{cases} g\left(\dfrac{\bar{\theta}_1}{\bar{\theta}_0}\right)^2 & \text{if } \alpha = 1 \\ -N^2 f \Psi / g & \text{if } \alpha = 0 \end{cases} \tag{23}$$

so that the thermal-wind relation

$$\nabla_h \frac{\Theta}{\bar{\theta}_0} = -\mathbf{e}_z \times \frac{f}{g}\frac{\partial \mathbf{U}}{\partial z} - (1-\alpha) N^2 \nabla_h \frac{\Psi}{\bar{\theta}_0} \tag{24}$$

holds. Here the first term on the right-hand side is the classic result for the weakly stratified troposphere, while the second term contributes under more strongly stratified conditions, e.g. in the stratosphere. (f) The wave impact on the synoptic-scale flow is exerted by wave-entropy flux-convergence in the entropy equation

$$\left(\frac{\partial}{\partial t} + \mathbf{U}\cdot\nabla\right)\Theta + N^2 \widetilde{W} = -\nabla_h \cdot \frac{1}{2}\Re\left(\hat{\mathbf{u}}\hat{\theta}^*\right) \tag{25}$$

where the asterisk indicates complex conjugation, and in the momentum equation

$$\begin{aligned} &\left(\frac{\partial}{\partial t} + \mathbf{U}\cdot\nabla\right)\mathbf{U} + f\mathbf{e}_z \times \widetilde{\mathbf{U}} \\ &= -c_p\left\{\bar{\theta}_0 \nabla_h \widetilde{\Pi} - \left[\alpha\bar{\theta}_1 + (1-\alpha)\Theta\right]\right\}\nabla_h \Pi - \frac{1}{\bar{\rho}}\nabla\cdot\frac{1}{2}\Re\left(\bar{\rho}\hat{\mathbf{v}}\hat{\mathbf{u}}^*\right) + \frac{f}{g}\mathbf{e}_z \times \frac{1}{2}\Re\left(\hat{\mathbf{u}}\hat{b}^*\right) \end{aligned} \tag{26}$$

by anelastic wave-momentum flux-convergence $-\frac{1}{\bar{\rho}}\nabla\cdot\frac{1}{2}\Re\left(\bar{\rho}\hat{\mathbf{v}}\hat{\mathbf{u}}^*\right)$ and an elastic term $\frac{f}{g}\mathbf{e}_z \times \frac{1}{2}\Re\left(\hat{\mathbf{u}}\hat{b}^*\right)$. As usual, only the next-order ageostrophic fields appear in the Coriolis and in the

pressure-gradient term while the leading-order geostrophic fields cancel each other. (g) If the leading-order synoptic-scale flow is in geostrophic and hydrostatic equilibrium, it is described completely by its streamfunction. A prognostic equation for the latter can be obtained from the entropy equation (25), the horizontal momentum equation (26), and the leading-order Exner-pressure equation

$$(1-\alpha)\left(\frac{\partial}{\partial t}+\mathbf{U}\cdot\nabla_h\right)\Pi+\frac{R}{c_v}\bar{\pi}^{(R-c_V)/R}\nabla\cdot\left(\bar{\pi}^{c_V/R}\widetilde{\mathbf{V}}\right)=0 \tag{27}$$

using geostrophy (22), the hydrostatic equilibrium (23), the leading-order horizontal character of the flow (21), and the IGW polarization relations. The result is a generalized quasigeostrophic potential-vorticity (PV) equation

$$\left(\frac{\partial}{\partial t}+\mathbf{U}\cdot\nabla_h\right)P=-\frac{\partial}{\partial x}\left[\frac{1}{\bar{\rho}}\nabla\cdot(\hat{\mathbf{c}}_g l\mathscr{A})\right]+\frac{\partial}{\partial y}\left[\frac{1}{\bar{\rho}}\nabla\cdot(\hat{\mathbf{c}}_g k\mathscr{A})\right] \tag{28}$$

for quasigeostrophic PV

$$P=\nabla_h^2\Psi+\frac{1}{\bar{\rho}}\frac{\partial}{\partial z}\left(\bar{\rho}\frac{f^2}{N^2}\frac{\partial\Psi}{\partial z}\right) \tag{29}$$

Here $\hat{\mathbf{c}}_g=\mathbf{c}_g-\mathbf{U}$ is the intrinsic group velocity. One sees that quasigeostrophic PV is driven by the vertical curl of the convergence of IGW pseudomomentum flux $\hat{\mathbf{c}}_g\mathbf{k}_h\mathscr{A}$, and this determines completely the time development of the leading-order synoptic-scale flow.

Of considerable importance is the following observation, following a similar argument given in a simplified Boussinesq context by Bühler (2009): *If* (a) in the entropy equation (25) the entropy fluxes are neglected and (b) the anelastic momentum-flux convergence and the elastic term in the horizontal-momentum equation (26) are replaced by pseudomomentum-flux convergence, i.e.

$$-\nabla_h\cdot\frac{1}{2}\Re\left(\hat{\mathbf{u}}\hat{\theta}^*\right)\;\;\rightarrow\;\;0 \tag{30}$$

$$-\frac{1}{\bar{\rho}}\nabla\cdot\Re\left(\frac{\bar{\rho}}{2}\hat{\mathbf{v}}\hat{\mathbf{u}}^*\right)+\frac{f}{2g}\mathbf{e}_z\times\Re\left(\hat{\mathbf{u}}\hat{b}^*\right)\;\;\rightarrow\;\;-\frac{1}{\bar{\rho}}\nabla\cdot(\hat{\mathbf{c}}_g\mathbf{k}\mathscr{A}) \tag{31}$$

then one obtains the same PV equation (28), *provided the synoptic-scale flow is horizontal, geostrophic, and hydrostatic*. Hence the dynamics will obey (28) and the synoptic-scale flow

will develop in a realistic manner. The corresponding calculations are nearly the same, only a bit simpler, as those leading from the original equations (25) and (26) to (28), described by Achatz et al. (2017), and are therefore not detailed here any further. Therefore a recipe for an IGW parameterization can be to implement it exclusively into the horizontal-momentum equation, via pseudomomentum-flux convergence (Andrews and McIntyre 1976). In a single-column approximation, where all horizontal flux gradients are neglected, this is what present-day operational IGW parameterizations do (e.g. Alexander and Dunkerton 1999; Warner and McIntyre 2001; Scinocca 2002, 2003; Orr et al. 2010), and this pseudomomentum approach guarantees that a geostrophically and hydrostatically balanced flow is affected correctly by parameterized IGWs. However, the basis for this approach is even more fundamental. In their seminal work Andrews and McIntyre (1978) show that in a Lagrangian-mean reference frame the effect of arbitrary-amplitude GWs on a large-scale flow indeed only occurs in the momentum equation, by an appropriate forcing. Nonetheless an issue remains how to transfer this result to the Eulerian-mean reference frame that atmospheric models use, both for weather and climate applications. Andrews and McIntyre (1976), in the derivation of their generalized Eliassen-Palm-flux convergence for the zonal-mean zonal-wind equation, use a scaling where weak wave amplitudes are assumed, and where hence approximate geostrophy and hydrostaticity of the mean flow allow the derivation of an approximate thermal-wind balance that can be used to avoid GW entropy-flux convergence and its application in a mean-flow entropy equation. Due to the weak-amplitude assumption advection of zonal momentum by the residual mean flow turns out to be negligible as well, so that one is left with a direct acceleration of the zonal-mean zonal wind by a the convergence of pseudomomentum (or generalized Eliassen-Palm) flux. The reliance of the pseudomomentum-flux result of Achatz et al. (2017) on large-scale balance appears to be directly related to this, however without the need for weak wave amplitudes. The question may be asked what happens if the required balance conditions

are not fulfilled anymore. Although the theory of Achatz et al. (2017) exploits synoptic scaling, and hence derives geostrophic and hydrostatic balance, one can see (25) and (26) as intermediate results not explicitly relying on the large-scale flow being balanced, while the derivation of (28) makes heavy use of large-scale flow balance. It might therefore be more realistic to use in a more direct approach the unmodified entropy-flux convergence, anelastic momentum-flux convergence and elastic term. This could be the case, e.g. if large-scale flow at different scales are of interest, or also ageostrophic flow such as the residual circulation. A central goal of the present work is a comparison of the two approaches in validations against wave resolving simulations, and it will be shown that the direct approach is considerably more reliable.

In a first comparison between the pseudomomentum and the direct approach one can easily point out that the pseudomomentum approach does not include any elastic or heating effect by IGWs while, mathematically speaking, each momentum-flux component

$$MF_{ij} = \bar{\rho}\frac{1}{2}\Re\left(\hat{v}_i^*\hat{v}_j\right) \tag{32}$$

with i being x or y, and j being x, y, or z, can be related to its corresponding pseudomomentum-flux component

$$PMF_{ij} = k_i\mathscr{A}\hat{c}_{g,j} \tag{33}$$

by

$$MF_{ij} = \gamma_{ij}PMF_{ij} \tag{34}$$

with

$$\gamma_{xx} = 1 + \frac{(l/k)^2+1}{(\hat{\omega}/f)^2-1} = 1 + \frac{f^2}{N^2}\frac{m^2}{k^2} \tag{35}$$

$$\gamma_{xy} = \gamma_{yx} = 1 \tag{36}$$

$$\gamma_{xz} = \gamma_{yz} = 1 + \frac{(f/\hat{\omega})^2}{1-(f/\hat{\omega})^2} = 1 + \frac{f^2}{N^2}\frac{m^2}{k^2+l^2} \tag{37}$$

$$\gamma_{yy} = 1 + \frac{(k/l)^2+1}{(\hat{\omega}/f)^2-1} = 1 + \frac{f^2}{N^2}\frac{m^2}{l^2} \tag{38}$$

One sees that the absolute value of the momentum flux is always greater than or equal to the absolute value of the pseudomomentum flux. The difference becomes notable in the low-frequency limit where $\hat{\omega} = \mathcal{O}(f)$, if

$$\left| \frac{m}{\sqrt{k^2+l^2}} \right| \geq \frac{N}{f} \tag{39}$$

i.e. the ratio between horizontal and vertical wavelength must be of the order or larger than N/f. As the whole goes with the square of this ratio, corresponding factors get rapidly quite large. E.g., for typical values of $N/f = 10^2$, wavelength ratios of 10^2 or $2 \cdot 10^2$ lead to a factor 2 or 5, respectively. Also of interest is that in some cases certain contributions to the pseudo-momentum flux vanish, i.e. $PMF_{xx} = 0$ if $k = 0$ and $PMF_{yy} = 0$ if $l = 0$, while the corresponding momentum-flux is non-zero. As we will show below this can have consequences for the structure of the mean-flow response. Finally we note that all $\gamma_{ij} = 1$ in the non-rotational case. Also the elastic terms and the GW entropy fluxes vanish then. Hence, the difference between the approaches exists only outside the tropics, and it is negligible for mid-frequency and high-frequency GWs as well.

3. Description of the numerical models

In this section, the numerical code used for the validation cases is described, used either in a wave-resolving high-resolution mode with subgrid-scale (SGS) turbulence parameterization, to

provide reference data, or in a low-resolution mode with WKB module switched on, for validations of the two parameterization approaches.

a. *PincFloit-LES*

The Pseudo-incompressible Flow solver with implicit turbulence modeling (PincFloit) solves on an f-plane the pseudo-incompressible equations of Durran (1989) in flux form (Klein 2009; Rieper et al. 2013b)

$$\frac{\partial}{\partial t}(\rho\mathbf{v}) + \nabla\cdot(\rho\mathbf{v}\mathbf{v}) + f\mathbf{e}_z\times\rho\mathbf{u} = -c_p\bar{P}\nabla\pi' - \rho g\frac{\theta'}{\bar{\theta}}\mathbf{e}_z + \nabla\cdot\Sigma \tag{40}$$

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = \nabla\cdot\left(\frac{\nu_t}{Pr}\nabla\rho\right) \tag{41}$$

$$\nabla\cdot(\bar{P}\mathbf{v}) = \frac{\rho Q}{c_p\bar{\pi}} \tag{42}$$

$$\rho\theta = \bar{\rho}\bar{\theta} = \bar{P} \tag{43}$$

where ρ and $\bar{\rho}$ are total density and reference-atmosphere density, respectively, and $\theta' = \theta - \bar{\theta}$ the deviation of potential temperature from its reference-atmosphere value. Σ is the viscous stress tensor with elements

$$\Sigma_{ij} = \eta\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3}\delta_{ij}\nabla\cdot\mathbf{v}\right) \tag{44}$$

where η is the dynamic shear viscosity coefficient. ν_t is the turbulent viscosity (see below), Pr the turbulent Prandtl number, here taken to be 1/2, while Q is the total heating, here zero in all wave resolving simulations. The elliptic problem for the pressure is solved using the HYPRE package (Falgout et al. 2006)

In the original numerical implementation of Rieper et al. (2013b) the code uses an implicit scheme for the parameterization of SGS turbulence. More recently Remmler et al. (2015) have shown in benchmark validations against direct numerical simulations of GW breaking that although this approach works quite well, it is surpassed in performance in some cases by the dynamic

Smagorinsky method (Germano et al. 1991). The present implementation therefore calculates the resolved fluxes by an upstream-centred scheme for conservation laws (MUSCL) following van Leer et al. (1991) with a modern total variation diminishing (TVD) limiter (Toro 1999; Kemm 2010), and parameterizes turbulence effects by the dynamic Smagorinsky approach. Following corresponding rules (Germano et al. 1991; Remmler et al. 2015) a turbulent Smagorinsky coefficient C_S^2 is determined, then averaged for stability reasons over a local 5-point smoothing window along all the available spatial directions, and finally a turbulent kinematic viscosity is determined via

$$\nu_t = C_S^2 \Delta^2 S \tag{45}$$

where

$$\Delta = \begin{cases} (\Delta x \Delta y \Delta z)^{1/3}, & \text{3D case} \\ (\Delta x \Delta z)^{1/2}, & \text{2D case.} \end{cases} \tag{46}$$

and

$$S = \sqrt{\sum_{i,j} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right)^2} \tag{47}$$

The turbulent viscosity is then used in the dynamic shear viscosity coefficient via

$$\eta = \eta_m + \rho \nu_t \tag{48}$$

where η_m is the molecular dynamic shear viscosity.

b. PincFloit/MS-GWaM

The numerical implementation of the transient interaction processes between IGW and large-scale flow is achieved by coupling the Multi-Scale Gravity-Wave Model (MS-GWaM), a Lagrangian phase-space WKB ray tracer, to the PincFloit code in a coarse-resolution setting. MS-GWaM is based on the methods described by Muraschko et al. (2015), Bölöni et al. (2016), and

Wilhelm et al. (2018). It predicts the spectral distribution of wave action by a phase-space wave-action density $\mathscr{N}$ so that wave-action density and energy density are

$$\mathscr{A}(\mathbf{x},t) = \int d^3k \mathscr{N}(\mathbf{x},\mathbf{k},t) \tag{49}$$

$$E_w(\mathbf{x},t) = \int d^3k (\hat{\omega}\mathscr{N})(\mathbf{x},\mathbf{k},t) \tag{50}$$

Its development is predicted by

$$\frac{\partial \mathscr{N}}{\partial t} + \mathbf{c}_g \cdot \nabla \mathscr{N} + \dot{\mathbf{k}} \cdot \nabla_{\mathbf{k}} \mathscr{N} = 0 \tag{51}$$

with $\dot{\mathbf{k}}$ given by (15). In the Lagrangian approach wavenumber and phase-space wave-action density are followed along rays parallel to the local $\mathbf{c}_g$ and $\dot{\mathbf{k}}$. These respond to changes in the resolved horizontal flow $\mathbf{u}$ and stratification N^2, where from here on we denote the total resolved flow by lower-case letters. If coupled to MS-GWaM, PincFloit is used without molecular and turbulent viscosities, since IGWs should lead to the leading SGS effects.

In the *direct approach* the IGW impact on the resolved flow is mediated by anelastic momentum-flux convergence, the elastic term, and by entropy-flux convergence. Since in pseudo-incompressible dynamics the non-diffusive continuity equation, i.e. (41) with $\nu_t = 0$, is equivalent, together with the divergence constraint (42) and the equation of state (43), to the entropy equation

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla\right) \theta = \frac{Q}{c_p \bar{\pi}} \tag{52}$$

the IGW impact on the resolved flow in PincFloit is obtained by replacing in (40) – (42)

$$\nabla \cdot \Sigma \rightarrow -\frac{\rho}{\bar{\rho}} \nabla \cdot (\bar{\rho} \langle \mathbf{v}_{GW} \mathbf{u}_{GW} \rangle) + \rho \frac{f}{\bar{\theta}} \mathbf{e}_z \times \langle \mathbf{u}_{GW} \theta_{GW} \rangle \tag{53}$$

$$\nabla \cdot \left(\frac{\nu_t}{Pr} \nabla \rho\right) \rightarrow 0 \tag{54}$$

$$\frac{\rho Q}{c_p \bar{\pi}} \rightarrow -\rho \nabla \cdot \langle \mathbf{u}_{GW} \theta_{GW} \rangle \tag{55}$$

Here the subscript *GW* indicates IGW fields and the angle brackets stand for a local average over typical IGW wavelengths. Following the considerations above, the momentum fluxes are obtained

from

$$\bar{\rho}\left\langle v_{GW,j}v_{GW,i}\right\rangle = \int d^3k\,\gamma_{ij}\,c_{g,j}k_i\mathscr{N} \tag{56}$$

From the polarization relations one also obtains

$$\left\langle \mathbf{u}_{GW}\,\theta_{GW}\right\rangle = -\mathbf{e}_z \times \frac{\bar{\theta}fN^2}{g\bar{\rho}}\int d^3k\,\mathbf{k}\frac{\mathscr{N}m}{\hat{\omega}\left|\mathbf{k}\right|^2} \tag{57}$$

In the *pseudomomentum approach* only the momentum equation is forced by the IGWs, via pseudomomentum-flux convergence leading to the replacements

$$\nabla\cdot\Sigma \;\rightarrow\; -\frac{1}{\bar{\rho}}\nabla\cdot\int d^3k\,\mathbf{c}_g\mathbf{k}\mathscr{N} \tag{58}$$

$$\nabla\cdot\left(\frac{\nu_t}{Pr}\nabla\rho\right) \;\rightarrow\; 0 \tag{59}$$

$$\frac{\rho Q}{c_p\bar{\pi}} \;\rightarrow\; 0 \tag{60}$$

in (40) – (42).

At every Runge-Kutta step, information is exchanged between IGWs and large-scale flow dynamics. MS-GWaM determines the wave-induced fluxes based on either direct or pseudomomentum scheme, and updates the resolved large-scale flow in PincFloit. After integrating the PincFloit model, the new resolved large-scale flow information is given to MS-GWaM, which solves the ray tracing equations and yields updated parameterized wave-induced fluxes, thus closing the circle. Within the last step we also use a saturation scheme for mimicking turbulent wave breakdown. Whenever the local wave amplitudes are large enough to make the wave field statically unstable, a turbulent viscosity is switched on that reduces the wave amplitudes within one time step to the saturation level. More details on the methods used by MS-GWaM can be found in Muraschko et al. (2015), Bölöni et al. (2016), and Wilhelm et al. (2018).

4. Wave-packet validation cases

In this section, we introduce the various wave-packet cases used for the validation of the two approaches at stake. In all of these an isothermal ($T = 240$K) atmosphere at rest is initially superposed to a locally monochromatic Gaussian IGW packet. Therein

$$\begin{pmatrix} \mathbf{v}_{GW} \\ \theta_{GW} \end{pmatrix} = \Re \left[\begin{pmatrix} \tilde{\mathbf{V}} \\ \frac{\bar{\theta}}{g}\tilde{B} \end{pmatrix} (x,z)\, e^{i\mathbf{k}\cdot\mathbf{x}} \right] \tag{61}$$

where

$$\tilde{B} = a_0 \frac{N^2}{m} \exp[-\frac{(y-y_0)^2}{2\sigma_y^2} - \frac{(z-z_0)^2}{2\sigma_z^2}] \tag{62}$$

and $\tilde{\mathbf{V}}$ can be obtained from $\tilde{B}$ using the polarization relations (16) and (17). $m < 0$ is the initial vertical wave number and the horizontal wave-number components are $(k,l) = k_h(\sin\alpha, \cos\alpha)$ with $k_h > 0$ and $0 \leq \alpha \leq \pi/2$. The factor a_0 indicates the wave-packet amplitude with regard to the static-instability threshold, reached when $a_0 = 1$, where the isentropes in the wave-packet begin to overturn. The positive dispersion-relation branch is taken, so that $m < 0$ corresponds to a local group velocity pointing upwards. The wave amplitudes are modulated in the horizontal in y-direction to be reminiscent of a zonally symmetric amplitude distribution inducing a zonal wind, as is the paradigmatic picture in parameterized IGW mean-flow interactions. Four cases are considered: (1) A horizontally symmetric IGW field with $1/\sigma_y = 0$ and $1/\sigma_x = 0$ (1DWP), (2) and (3) IGW fields modulated in y-direction, with $\sigma_y < \infty$ and $1/\sigma_x = 0$, where the angle α between the horizontal wavenumber and the horizontal direction of modulation of the wavepacket is either 0 (2DWP00) or $\pi/2$ (2DWP90). In both of these cases the vertical group velocity is so small that a simulation of a wavepacket from initialization to the point where it becomes statically unstable would be computationally very demanding. Since the IGW group velocity is about linearly increasing with the Coriolis parameter we have therefore also decided to follow in a (4)

further case (2DWP90-HAMP-HF) the approach of Lelong and Dunkerton (1998a,b) of considering an atmosphere with a Coriolis parameter an order of magnitude larger. In order to keep the same ratio $(k_h/m)N/f$ as above, the horizontal wavelength in that case is one order of magnitude smaller. The launch amplitude has there also been chosen stronger than in all other cases, in order to further reduce the required computational time. The details of the settings are given in Table 1. Particularly, in each wave-resolving PincFloit-LES simulation, there are at least 16 grid points per wavelength of the initial IGW along each available horizontal direction, and 10 grid points per vertical wavelength of the initial IGW. In each PincFloit/MS-GWaM simulation with IGW parameterization, the resolution along each available direction is close to the wavelength of the initial parameterized IGW along the corresponding direction, and the large-scale flow induced by the parameterized IGWs can still be resolved in the coarse-resolution PincFloit/MS-GWaM simulation. The resolution in the wave resolving simulations, using PincFloit-LES is too coarse to really justify these as LES. It is, however, sufficient for our present purposes, as we are not addressing turbulent wave breaking itself.

5. Incipient budget analysis

In the following we first discuss the various cases in terms of the initial contributions

$$MFC_{ij} = -\frac{1}{\bar{\rho}}\frac{\partial}{\partial x_j}\int d^3k\,\gamma_{ij}\,c_{g,j}k_i\mathcal{N} \tag{63}$$

$$PMFC_{ij} = -\frac{1}{\bar{\rho}}\frac{\partial}{\partial x_j}\int d^3k\,c_{g,j}k_i\mathcal{N} \tag{64}$$

to the momentum-flux convergence or pseudomomentum-flux convergence, respectively, in terms of the elastic terms

$$ET_i = \left(\frac{f}{\bar{\theta}}\mathbf{e}_z \times \langle \mathbf{u}_{GW}\,\theta_{GW}\rangle\right)_i \tag{65}$$

and in terms of the wave-induced heating

$$-c_p\bar{\pi}\nabla\cdot\langle\mathbf{u}_{GW}\theta_{GW}\rangle \tag{66}$$

In the case of 1DWP all horizontal derivatives in the various flux convergences vanish. Hence, in the pseudomomentum scheme the vertical gradient of the pseudomomentum flux in the x momentum equation (referred to as $PMFCxz$) is the only nonzero forcing term. In the direct scheme, the vertical gradient of the momentum flux and the elastic term in the x momentum equation (referred to as $MFCxz$ and ETx, respectively) are the only two nonzero forcing terms. The comparison of the above-mentioned three forcing terms (i.e., $PMFCxz$ in the pseudomomentum scheme, $MFCxz$ and ETx in the direct scheme) is given in Fig. 1a. In the case of 1DWP, the signal of $MFCxz$ is apparently stronger than that of $PMFCxz$. Compared with $MFCxz$, ETx is relatively small but noticeable.

For further illustration, Figure 2 shows the dependence of the above-mentioned three forcing terms on zonal wavelength (upper panel) and vertical wave packet scale σ_z (lower panel). In each sensitivity test, all the other parameters are kept the same as those in 1DWP. For waves with relatively short horizontal wavelengths (hence intrinsic frequencies well above the inertia frequency), the signal of $PMFCxz$ may be still close to that of $MFCxz$, and ETx is negligible. As the horizontal wavelength increases, the signal of $PMFCxz$ becomes weaker, while $MFCxz$ is getting stronger instead. A large separation is seen between $PMFCxz$ and $MFCxz$ for waves with relatively long horizontal wavelengths, and ETx becomes more noticeable but still secondary compared with $MFCxz$. In the sensitivity test to the changes of σ_z only, the pattern of $PMFCxz$ is the same as $MFCxz$, but the signal of $MFCxz$ is approximately three times as strong as that of $PMFCxz$. The signals of both positive branch (i.e., upper branch) and negative branch (i.e., lower branch) of $MFCxz$ are getting weaker as σ_z increases. However, the negative branch is apparently

much more sensitive to the change of σ_z, compared with the positive branch. This asymmetric behavior is caused by the vertically decreasing density. The center value of ETx is not sensitive to the change of σ_z, and σ_z only changes the vertical width of ETx. Generally speaking, in the current study, ETx appears to be secondary compared with $MFCxz$. However, with larger σ_z, ETx gets more important. Hence, it is not only the IGW aspect ratio between vertical and horizontal wavelength that determines, as described above, the relative magnitude of the dynamical terms in the direct and pseudomomentum approach, but also the IGW amplitude distribution, as here exemplified using the wave-packet width.

In the case of 2DWP00, the nonzero IGW-induced forcing terms in the direct approach only include MFC_{yy}, MFC_{yz}, and ET_y in the y-momentum equation, and there is no forcing in the x-momentum equation and thermodynamic equation. Together with the wave-energy density initial vertical cross sections of all non-zero terms are shown in Figure 3. There is a dipole pattern at the altitude of wave packet center for MFC_{yy}, causing a horizontally asymmetric forcing effect. Instead, the dipole structure in MFC_{yz} results in a vertically asymmetric forcing effect. ET_y has consistent negative signals with minimum at the wave packet center. It is worth mentioning that the forcing from MFC_{yy} is comparable with that from MFC_{yz}, and it is very important in this case. Similar to 1DWP, ET_y appears to be secondary but not negligible in this case. To facilitate a comparison between pseudomomentum and direct approach, the corresponding γ_{ij} are shown in the upper left corner of the subplots of the contributors to the momentum-flux convergence, indicating how much stronger the respective momentum flux is than the corresponding pseudomomentum flux. Both γ_{yy} and γ_{yz} are as large as 3.2, which indicates much weaker signals of both $PMFC_{yy}$ and $PMFC_{yz}$ in the pseudomomentum approach, compared to MFC_{yy} and MFC_{yz} in the direct approach. This again suggests that the forcing from MFC_{yy}, which is neglected in operational IGW schemes, is comparable with that from MFC_{yz} in this case.

Case 2DWP90 is interesting for several reasons. Figure 4 shows the initial distribution of all non-zero IGW-induced forcing terms in the direct approach. These include MFC_{yy} in the y-momentum equation, MFC_{xz} and ET_x in the x-momentum equation, and the IGW-induced heating term in the thermodynamic equation. In contrast to these, $PMFC_{xz}$ is the only non-zero term in the pseudo-momentum approach, i.e. only x-momentum is forced. Both MFC_{yy} and ET_x are due to rotational effects. Among these, only MFC_{yy} and the wave-induced heating term are horizontally asymmetric, while in the pseudomomentum approach no horizontal asymmetry is forced initially.

6. Numerical simulations of different idealized wave packet profiles prior to turbulent dissipation

In this section, the direct and pseudomomentum approach will be validated against wave resolving simulations with unbalanced flows. The initial condition settings of the respective numerical experiments are given by the three idealized wave packet profiles described in the previous section (i.e., 1DWP, 2DWP00, and 2DWP90). In each wave-resolving PincFloit-LES simulation, there are 16 or $\sim$17 grid points per wavelength of the initial IGW along each available horizontal direction, and 10 grid points per vertical wavelength of the initial IGW. As we have convinced ourselves from 2D simulations with increased resolutions (not shown), these resolutions are sufficient for the laminar phase of the occurring IGW-mean-flow interactions, while the later turbulent IGW breaking would require a finer resolution. In each PincFloit/MS-GWaM simulation, the resolution along each available direction is close to the wavelength of the initial IGW in the corresponding direction, and the large-scale flow induced by the parameterized IGWs can still be resolved in the coarse-resolution PincFloit/MS-GWaM simulation. In the PincFloit-LES simulations, the wave packets of the resolved IGWs are initialized in two-dimensional domains for 1DWP and 2DWP00, and three-dimensional domain settings are employed for 2DWP90. Instead, in the PincFloit/MS-

GWaM simulations, 1DWP uses a single vertical column, and two-dimensional domain settings are employed for all two-dimensional wave packet cases. The details of the model designs are given in Table 1 and Table 2.

Figure 5 compares the large-scale x-velocity component and the large-scale y-velocity component between the PincFloit-LES simulation, and the PincFloit/MS-GWaM simulations using either direct or pseudomomentum approach. Similar to the procedure in Bölöni et al. (2016) and Wilhelm et al. (2018), the large-scale wind component in the PincFloit-LES simulation is obtained by a running mean over two wavelengths of the initial IGW. The selected time for this comparison in Fig. 5 is at t=1500 min, which is approximately one and a half inertial periods. Since the wave energy of the 1DWP case does not depend on the horizontal direction, the induced large-scale zonal velocity component and meridional velocity component do not change along the horizontal direction, and the large-scale vertical velocity component is zero everywhere in the domain. As depicted in Fig. 5, the direct scheme is able to reproduce the signals in the data from the wave resolving simulation with very small errors. The pseudomomentum scheme, on the contrary, gives a much weaker large-scale horizontal wind compared with the wave resolving data, and thus large model errors could be expected.

Figure 6 shows corresponding results, again at $t = 1500$min, but now for case 2DWP00. For the two-dimensional wave packets in the current study, in addition to the induced large-scale horizontal wind on both directions, the large-scale vertical velocity component is nonzero. Again, the direct scheme successfully captures the signals in the data from the wave resolving simulation. Even though the pattern from the results based on the pseudomomentum scheme is still similar to the wave resolving simulation for all the three wind components, its signals are much weaker. Note that the fact that the pseudomomentum scheme generates much weaker large-scale wind components is consistent with the budget analysis shown in the previous section.

Fig. 7 shows corresponding results for the 2DWP90 case. It again demonstrates that the direct scheme is more reliable than the pseudomomentum scheme, with the data from the wave resolving simulation as a reference. Two further things are worth noting here: (1) Arguably the best agreement between pseudomomentum approach and wave resolving simulation is visible in the x-wind component. This could be expected since this is the horizontal wind component that can also contain balanced contributions, given by the y-derivative of the balanced part of the pressure. The other side of the medal is, however, that the unbalanced response in the meridional circulation is met well at best in structure but definitely not in strength. (2) All three large-scale wind components produced by the pseudomomentum scheme for this particular case are highly horizontally symmetrical, which is quite different from the results from the direct scheme and the wave resolving data. This suggests that the horizontally asymmetrical forcing is apparently negligible in the pseudomomentum scheme in this case. As suggested in Fig. 4 from the previous section, this is again consistent with the budget analysis in the pseudomomentum scheme for the 2DWP90 case. Indeed here $l = 0$ and hence $PMF_{yy} = 0$, while MF_{yy} is nonzero and explains the asymmetry of the response.

7. Sensitivity experiments

Various sensitivity tests have been done. In the first of these we have convinced ourselves that our findings even hold for IGW amplitudes very close to overturning instability. To this purpose we have considered a high-amplitude two-dimensional wave packet case (2DWP90-HAMP). Here, the initial setting is the same as that in 2DWP90, except that the initial wave amplitude at the wave packet center reaches the limit of convective instability. Therefore, a stronger wave-mean-flow interaction is expected. To be compared to Fig. 7 for 2DWP90, Figure 8 shows the comparisons of the large-scale mean wind among three model codes at the same time for 2DWP90-HAMP.

Similar large-scale wind patterns are found in cases 2DWP90 (Fig. 7) and 2DWP90-HAMP (Fig. 8), except that the signals in 2DWP90-HAMP are approximately four times as strong as those in 2DWP90. With a stronger wave-mean-flow interaction, this case continues to support the statement that the direct approach is generally more reliable than the pseudomomentum approach.

Next, expanding on the budget analysis in section 5, we have investigated the relative importance of the elastic effect and the heating term for the results from the direct approach. Fig. 9 shows results from three sensitivity experiments, where either the elastic term has been set to zero, or the heating term, or both. These are to be compared to the results from the direct scheme for the 2DWP90-HAMP case in Fig. 8. These sensitivity experiments are denoted by 2DWP90-HAMP-NoET, 2DWP90-HAMP-NoHeat, and 2DWP90-HAMP-NoET-NoHeat, respectively. With PincFloit-LES as a reference, all of the above-mentioned sensitivity experiments demonstrate a much stronger negative large-scale zonal mean wind at the domain center, especially when the heating effect is switched off. This implies that it is important to take into account the heating term in order to get the correct large-scale flow response. The elastic term, however, is, at least in this study, noticeable but still of secondary importance.

In all of the cases considered so far, the vertical group velocity is so small that a simulation of a wavepacket from initialization to the point where it becomes statically unstable by propagation in increasingly attenuated altitudes, hence growing in amplitude, would be computationally very demanding. Since the IGW group velocity is about linearly increasing with the Coriolis parameter we have therefore also decided to follow in a further case (2DWP90-HAMP-HF) the approach of Lelong and Dunkerton (1998a,b) of considering an atmosphere with a Coriolis parameter an order of magnitude larger. In order to keep the same ratio $(k_h/m)N/f$ as above, the horizontal wavelength in that case is one order of magnitude smaller. The comparison among the three model codes for the 2DWP90-HAMP-HF case at a late stage of the simulation is shown in Fig.

10. Although the large-scale x-wind, with its balanced components, is relatively reasonable in the simulations with the pseudomomentum scheme, the direct scheme is the only one able to reproduce the response of the $y-z$ circulation in its full strength. It is also able to match the horizontal asymmetry of the response, while the pseudomomentum approach misses this feature.

Finally, Fig. 11 illustrates the importance of the wave-mean-flow interaction, as well as the role of the saturation scheme, by comparing for the same case 2DWP90-HAMP-HF the large-scale $y-z$ circulation and the wave energy among two sensitivity experiments and their corresponding control simulation, all using PincFloit/MS-GWaM with the direct approach, but at a much later stage, where we do not have PincFloit-LES data available. Here, in the first sensitivity experiment the coupling process between the large-scale mean wind and the gravity waves has been turned off by assuming that the large-scale mean wind has no impact on changing the wavenumber of the gravity waves via the ray tracing equation. In the second sensitivity experiment the saturation scheme has been turned off by assuming that the wave action density associated with each ray is always conserved during its vertical propagation. Differences are easily identified between each above-mentioned sensitivity experiment and its corresponding control simulation. The findings of Bölöni et al. (2016) are confirmed that the direct wave-mean-flow interaction dominates the dynamics to leading order, while turbulent wave breaking leads to next-order modifications.

8. Concluding remarks and discussion

This article investigates and compares two different approaches for the efficient modelling of subgrid-scale inertia-gravity waves (IGW) in a rotating compressible atmosphere. The first approach, denoted as pseudomomentum scheme, is closely related to the fundamental result of Andrews and McIntyre (1978) that in a Lagrangian-mean reference frame GW effects on the large-scale flow only occur in the momentum equation. It can be shown that this holds also in the

Eulerian-mean reference frame used by atmospheric models, at least if the large-scale flow is in hydrostatic and geostrophic balance. The GW forcing can then be expressed by a pseudomomentum-flux convergence in the large-scale momentum equation. Present-day gravity-wave parameterizations follow this route, exclusively applying a corresponding forcing, and leaving out elastic and thermal effects that may arise in the absence of large-scale-flow balance. In the second approach, called direct scheme, the large-scale flow is forced both in the momentum equation, by anelastic momentum-flux convergence, and in the entropy equation, via entropy-flux convergence. In addition, also an elastic term is taken into account in the momentum equation. This second approach does not rely on any balance assumption with regard to the large-scale flow, and it differs from the pseudomomentum approach wherever rotational effects matter, i.e. outside of the tropics and for low-frequency IGWs.

It has been the purpose of the current study to understand the difference between the pseudomomentum and the direct approach. To this end wave-resolving simulations have been done, to provide reference data against which to validate coarse-resolution model simulations with either of the two approaches providing the basis for a prognostic gravity-wave parameterization. All simulations have been done with the code PincFloit, solving the pseudo-incompressible equations by a finite volume method. At high resolutions PincFloit has been used together with a dynamical scheme for the parameterization of the largest turbulent eddies (PincFloit-LES). In the coarse-resolution simulations it has been coupled to the prognostic gravity wave model MS-GWaM, solving the coupled WKB equations using a Lagrangian approach. This model has been used either in pseudomomentum or direct mode.

The dynamics of 1D IGW packets, with an amplitude only depending on time and the vertical direction, already shows that the comparative performance of the two approaches is sensitive to the following two parameters: 1) the intrinsic frequency, and 2) the wave packet scale. The smaller

549 the intrinsic frequency is, the larger the discrepancy between pseudomomentum-flux convergence
550 and momentum-flux convergence is, hence the greater the differences between two schemes are.
551 Pseudomomentum-flux convergence is then much weaker than momentum-flux convergence. The
552 elastic term is also non-zero in that case, but it generally appears to be of comparatively smaller
553 importance, and only begins exhibiting some effect if the wave-packet scale is rather large.

554 Further 2D case studies, with IGW envelope depending on time and y and z, meant to be reminis-
555 cent of a zonally symmetric IGW distribution in interaction with a zonally symmetric large-scale
556 flow, confirm the expectations from the 1D analyses. They show that only the direct approach is
557 able to capture the response in the unbalanced components of the large-scale flow. This response
558 is strongly underestimated by the pseudomomentum approach. It will be interesting to investigate,
559 e.g., what consequences this has for the correct description of the IGW impact on the residual
560 circulation. Even in the balanced response, however, some structural asymmetries can only be
561 captured by the direct approach, while the more symmetric pseudomomentum-flux convergence
562 is not able to do so. In these 2D cases it is especially the horizontal entropy-flux convergence that
563 seems to matter, while the elastic term does not seem to be that important. It should be stressed
564 that these differences are only relevant for low-frequency IGWs, and outside the tropics, while for
565 high-frequency GWs and in the tropics pseudomomentum and direct approach are approximately
566 equivalent. Finally, in the present settings we re-confirm the findings of Bölöni et al. (2016) that
567 the direct wave-mean-flow interaction takes a leading role, as compared to turbulence effects.

568 In summary, our results show that the direct approach is more reliable than its pseudomomentum
569 counterpart. In some cases, the pseudomomentum approach yields a balanced response that is of
570 the right strength, however not quite with the right structure. Even more conspicuous, however,
571 appears the difference in the reproduction of the potentially most unbalanced part of the response.
572 The meridional-circulation response from the pseudomementum scheme is generally too weak.

As the direct scheme is not much more expensive than the pseudomomemntum scheme, it seems advisable to rather use it for GW parameterizations in atmospheric models, at least outside the tropics. Nonetheless, our analyses have not yet considered global applications where irreversible GW effects due to wave breaking will weigh in more strongly. For these the comparison between the two approaches, beyond the scope of the present study, is still to be done. Our results give considerable motivation for this next step.

Acknowledgments. We benefited from the insightful comments from two anonymous reviewers on an earlier version of the manuscript. The authors also thank the German Research Foundation (DFG) for partial support through the research unit *Multiscale Dynamics of Gravity Waves (MS-GWaves)* and through Grants AC 71/8-2, AC 71/9-2, AC 71/10-1, AC 71/10-2, AC 71/11-2, AC 71/12-2, BO 5071/2-2, and BO 5071/1-2. Calculations for this research were conducted on the LOEWE-CSC high performance computer of the Goethe Universität Frankfurt.

References

Achatz, U., B. Ribstein, F. Senf, and R. Klein, 2017: The interaction between synoptic-scale balanced flow and a finite-amplitude mesoscale wave field throughout all atmospheric layers: weak and moderately strong stratification. *Quart. J. Roy. Meteor. Soc.*, **143**, 342–361, doi:10.1002/qj.2926.

Alexander, M. J., and T. J. Dunkerton, 1999: A spectral parameterization of mean-flow forcing due to breaking gravity waves. *J. Atmos. Sci.*, **56**, 4167–4182.

Alexander, M. J., J. R. Holton, and D. R. Durran, 1995: The gravity wave response above deep convection in a squall line simulation. *J. Atmos. Sci.*, **52**, 2212–2226, doi:https://doi.org/10.1175/1520-0469(1995)052⟨2212:TGWRAD⟩2.0.CO;2.

595 Alexander, M. J., and Coauthors, 2010: Review article, recent developments in gravity-wave ef-
596 fects in climate models and the global distribution of gravity-wave momentum flux from obser-
597 vations and models. *Quart. J. Roy. Meteor. Soc.*, **136**, 1103–1124, doi:10.1002/qj.637.

598 Andrews, D. G., and M. E. McIntyre, 1976: Planetary waves in horizontal and vertical shear: The
599 generalized eliassen-palm relation and the mean zonal acceleration. *J. Atmos. Sci.*, **33**, 2031–
600 2048, doi:https://doi.org/10.1175/1520-0469(1976)033⟨2031:PWIHAV⟩2.0.CO;2.

601 Andrews, D. G., and M. E. McIntyre, 1978: An exact theory of nonlinear waves on a Lagrangian-
602 mean flow. *J. Fluid Mech.*, **89**, 609– 646.

603 Baldwin, M. P., and T. J. Dunkerton, 2001: Stratospheric harbingers of anomalous weather
604 regimes. *Science*, **294**, 581–584.

605 Bierdel, L., T. Selz, and G. C. Craig, 2018: Theoretical aspects of upscale error growth on the
606 mesoscales: Idealised numerical simulations. *Quart. J. Roy. Meteor. Soc.*, **144**, 682–694, doi:
607 10.1002/qj.3236.

608 Bierdel, L., C. Snyder, S. Park, and W. C. Skamarock, 2016: Accuracy of rotational and divergent
609 kinetic energy spectra diagnosed from flight-track winds. *J. Atmos. Sci.*, **73**, 3273–3286, doi:
610 10.1175/JAS-D-16-0040.1.

611 Bölöni, G., B. Ribstein, J. Muraschko, C. Sgoff, J. Wei, and U. Achatz, 2016: The interaction
612 between atmospheric gravity waves and large-scale flows: An efficient description beyond the
613 nonacceleration paradigm. *J. Atmos. Sci.*, **73**, 4833–4852, doi:10.1175/JAS-D-16-0069.1.

614 Bühler, O., 2009: *Waves and Mean Flows*, chap. Forcing of mean vortical mode, 182 – 183.
615 Cambridge Univ. Press.

Bühler, O., and M. E. McIntyre, 1999: On shear-generated gravity waves that reach the mesosphere. part ii: Wave propagation. *J. Atmos. Sci.*, **56**, 3764–3773.

Bühler, O., and M. E. McIntyre, 2003: Remote recoil: a new wave-mean interaction effect. *J. Fluid Mech.*, **492**, 207–230.

Bühler, O., and M. E. McIntyre, 2005: Wave capture and wave-vortex duality. *J. Fluid Mech.*, **534**, 67–95, doi:10.1017/S0022112005004374.

Bühler, O., M. E. McIntyre, and J. F. Scinocca, 1999: On shear-generated gravity waves that reach the mesosphere. part i: Wave generation. *J. Atmos. Sci.*, **56**, 3749–3763.

Butchart, N., 2014: The brewer-dobson circulation. *Rev. Geophys.*, **52**, 157–184, doi:10.1002/2013RG000448.

Callies, J., R. Ferrari, and O. Bühler, 2014: Transition from geostrophic turbulence to inertia-gravity waves in the atmospheric energy spectrum. *Proc. Natl. Acad. Sci. USA*, **111**, 17 033–17 038, doi:https://doi.org/10.1073/pnas.1410772111.

Dosser, H. V., and B. R. Sutherland, 2011: Anelastic internal wave packet evolution and stability. *J. Atmos. Sci.*, **68**, 2844–2859, doi:10.1175/JAS-D-11-097.1.

Dunkerton, T. J., 1997: The role of gravity waves in the quasi-biennial oscillation. *J. Geophys. Res.*, **102**, 26 053–26 076.

Durran, D. R., 1989: Improving the anelastic approximation. *J. Atmos. Sci.*, **46**, 1453–1461.

Falgout, R. D., J. E. Jones, and U. M. Yang, 2006: *Numerical Solution of Partial Differential Equations on Parallel Computers*, chap. The design and implementation of hypre, a library of parallel high performance preconditioners, 267 – 294. Springer.

637 Fritts, D., and M. J. Alexander, 2003: Gravity wave dynamics and effects in the middle atmo-
638 sphere. *Rev. Geophys.*, **41(1)**, 1003, doi:10.1029/2001RG000106.

639 Germano, M., U. Piomelli, P. Moin, and W. H. Cabot, 1991: A dynamic subgrid-scale eddy vis-
640 cosity model. *Phys. Fluids A*, **3**, 1760–1765.

641 Gill, A. E., 1982: *Atmosphere-Ocean Dynamics*. 1st ed., Academic Press, 662 pp.

642 Gong, J., D. L. Wu, and S. D. Eckermann, 2012: Gravity wave variances and propagation derived
643 from airs radiances. *Atmos. Chem. Phys.*, **12**, 1701–1720, doi:10.5194/acp-12-1701-2012.

644 Griffiths, M., and M. J. Reeder, 1996: Stratospheric inertia-gravity waves generated in a numerical
645 model of frontogenesis. i: Model solutions. *Quart. J. Roy. Meteor. Soc.*, **122**, 1153–1174.

646 Grimshaw, R., 1975: Nonlinear internal gravity waves in a rotating fluid. *J. Fluid Mech.*, **71**,
647 497–512.

648 Haynes, P. H., C. J. Marks, M. E. McIntyre, T. G. Shepherd, and K. P. Shine, 1991: On the
649 downward control of extratropical diabatic circulations by eddy-induced mean zonal forces. *J.*
650 *Atmos. Sci.*, **48**, 651–678.

651 Hertzog, A., C. Souprayen, and A. Hauchecorne, 2002: Eikonal simulations for the formation
652 and the maintenance of atmospheric gravity wave spectra. *J. Geophys. Res.*, **107**, 4145, doi:
653 10.1029/2001JD000815.

654 Hien, S., J. Rolland, S. Borchert, L. Schoon, C. Zülicke, and U. Achatz, 2018: Spontaneous
655 inertia-gravity wave emission in the differentially heated rotating annulus experiment. *J. Fluid.*
656 *Mech.*, **838**, 5–41, doi:10.1017/jfm.2017.883.

657 Hines, C. O., 1997a: Doppler-spread parameterization of gravity-wave momentum deposition in
658 the middle atmosphere. part 1: Basic formulation. *J. Atmos. Sol. Terr. Phys.*, **59**, 371–386.

659 Hines, C. O., 1997b: Doppler-spread parameterization of gravity-wave momentum deposition in
660 the middle atmosphere. part 2: Broad and quasi monochromatic spectra, and implementation. *J.*
661 *Atmos. Sol. Terr. Phys.*, **59**, 387–400.

662 Holton, J. R., and R. S. Lindzen, 1972: An updated theory for the quasi-biennial cycle of the
663 tropical stratosphere. *J. Atmos. Sci.*, **29**, 1076–1080.

664 Houghton, J. T., 1978: The stratosphere and mesosphere. *Quart. J. Roy. Meteor. Soc.*, **104**, 1–29,
665 doi:10.1002/qj.49710443902.

666 Kemm, F., 2010: A comparative study of tvd-limiters, well-known limiters and an introduction of
667 new ones. *Int. J. Numer. Methods Fluids*, doi:10.1002/fld.2357, URL http://dx.doi.org/10.1002/
668 fld.2357.

669 Kim, Y.-J., S. D. Eckermann, and H.-Y. Chun, 2003: An overview of the past, present and fu-
670 ture gravity-wave drag parametrization for numerical climate and weather prediction models.
671 *Atmos.–Ocean*, **41**, 65–98, doi:10.3137/ao.410105.

672 Klein, R., 2009: Asymptotics, structure, and integration of sound-proof atmospheric flow equa-
673 tions. *Theor. Comput. Fluid Dyn.*, **23**, 161–195.

674 Lane, T. P., J. D. Doyle, R. Plougonven, M. A. Shapiro, and R. D. Sharman, 2004: Observations
675 and numerical simulations of inertia-gravity waves and shearing instabilities in the vicinity of a
676 jet stream. *J. Atmos. Sci.*, **61**, 2692–2706.

677 Lane, T. P., M. J. Reeder, and T. L. Clark, 2001: Numerical modeling of gravity wave gener-
678 ation by deep tropical convection. *J. Atmos. Sci.*, **58**, 1249–1274, doi:https://doi.org/10.1175/
679 1520-0469(2001)058⟨1249:NMOGWG⟩2.0.CO;2.

Lane, T. P., and F. Zhang, 2011: Coupling between gravity waves and tropical convection at mesoscales. *J. Atmos. Sci.*, **68**, 2582–2598, doi:10.1175/2011JAS3577.1.

Lelong, M., and T. Dunkerton, 1998a: Inertia-gravity wave breaking in three dimensions. Part I: Convectively stable waves. *J. Atmos. Sci.*, **55**, 2473–2488.

Lelong, M. P., and T. J. Dunkerton, 1998b: Inertia-gravity wave breaking in three dimensions. Part II: Convectively unstable waves. *J. Atmos. Sci.*, **55**, 2489–2501.

Limpasuvan, V., J. H. Richter, Y. J. Orsolini, F. Stordal, and O. K. Kvissel, 2012: The roles of planetary and gravity waves during a major stratospheric sudden warming as characterized in waccm. *J. Atmos. Sol. Terr. Phys.*, **78–79**, 84–98, doi:10.1016/j.jastp.2011.03.004.

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

Lott, F., and L. Guez, 2013: A stochastic parameterization of the gravity waves due to convection and its impact on the equatorial stratosphere. *J. Geophys. Res.*, **118**, 8897–8909, doi:10.1002/jgrd.50705.

Lott, F., and M. J. Miller, 1997: A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Quart. J. Roy. Meteor. Soc.*, **123**, 101–127.

Medvedev, A. S., and G. P. Klaassen, 1995: Vertical evolution of gravity wave spectra and the parameterization of associated gravity wave drag. *J. Geophys. Res.*, **100**, 25 841–25 853.

Menchaca, M. Q., and D. R. Durran, 2017: Mountain waves, downslope winds, and low-level blocking forced by a midlatitude cyclone encountering an isolated ridge. *J. Atmos. Sci.*, **74**, 617–639, doi:10.1175/JAS-D-16-0092.1.

Muraschko, J., M. D. Fruman, U. Achatz, S. Hickel, and Y. Toledo, 2015: On the application of wentzel-kramer-brillouin theory for the simulation of the weakly nonlinear dynamics of gravity waves. *Quart. J. Roy. Meteor. Soc.*, **141**, 676–697, doi:10.1002/qj.2381.

Nappo, C. J., 2002: *An Introduction to Atmospheric Gravity Waves*. 1st ed., Academic Press, 279 pp.

Orr, A., P. Bechtold, J. Scinocca, M. Ern, and M. Janiskova, 2010: Improved middle atmosphere climate and forecasts in the ecmwf model through a nonorographic gravity wave drag parameterization. *J. Climate*, **23**, 5905–5926.

Plougonven, R., H. Teitelbaum, and V. Zeitlin, 2003: Inertia gravity wave generation by the tropospheric midlatitude jet as given by the fronts and atlantic storm-track experiment radio soundings. *J. Geophys. Res.*, **108**, 4686, doi:10.1029/2003JD003535.

Plougonven, R., and F. Zhang, 2014: Internal gravity waves from atmospheric jets and fronts. *Rev. Geophys.*, **52**, 33–76, doi:10.1002/2012RG000419.

Remmler, S., S. Hickel, M. D. Fruman, and U. Achatz, 2015: Validation of large-eddy simulation methods for gravity wave breaking. *J. Atmos. Sci.*, **72**, 3537–3562, doi:https://doi.org/10.1175/JAS-D-14-0321.1.

Ribstein, B., and U. Achatz, 2016: The interaction between gravity waves and solar tides in a linear tidal model with a 4-d ray-tracing gravity-wave parameterization. *J. Geophys. Res.*, **121**, 8936–8950, doi:10.1002/2016JA022478.

Ribstein, B., U. Achatz, and F. Senf, 2015: The interaction between gravity waves and solar tides: Results from 4-d ray tracing coupled to a linear tidal model. *J. Geophys. Res.*, **120**, 6795–6817, doi:10.1002/2015JA021349.

723 Richter, J. H., F. Sassi, and R. R. Garcia, 2010: Towards a physically based gravity wave
724 source parameterization in a general circulation model. *J. Atmos. Sci.*, **67**, 136–156, doi:
725 10.1175/2009JAS3112.1.

726 Rieper, F., U. Achatz, and R. Klein, 2013a: Range of validity of an extended wkb theory for
727 atmospheric gravity waves: one-dimensional and two-dimensional case. *J. Fluid Mech.*, **729**,
728 330–363, doi:10.1017/jfm.2013.307.

729 Rieper, F., S. Hickel, and U. Achatz, 2013b: A conservative integration of the pseudo-
730 incompressible equations with implicit turbulence parameterization. *Mon. Wea. Rev.*, **141**, 861–
731 886, doi:10.1175/MWR-D-12-00026.1.

732 Scaife, A. A., J. R. Knight, G. K. Vallis, and C. K. Folland, 2005: A stratospheric influence on
733 the winter nao and north atlantic surface climate. *Geophys. Res. Let.*, **32**, (18) L18 715,, doi:
734 10.1029/2005GL023226.

735 Scaife, A. A., and Coauthors, 2012: Climate change projections and stratosphere-troposphere
736 interaction. *Clim. Dyn.*, **38**, 2089–2097, doi:10.1007/s00382-011-1080-7.

737 Scinocca, J. F., 2002: The effect of back-reflection in the parameterization of non-orographic
738 gravity-wave drag. *J. Meteorol. Soc. Japan*, **80**, 939–962.

739 Scinocca, J. F., 2003: An accurate spectral nonorographic gravity wave drag parameterization for
740 general circulation models. *J. Atmos. Sci.*, **60**, 667–682.

741 Shapiro, M. A., 1980: Turbulent mixing within tropopause folds as a mechanism for the exchange
742 of chemical constituents between the stratosphere and troposphere. *J. Atmos. Sci.*, **37**, 994–1004.

743 Smith, R. B., 1980: Linear theory of stratified hydrostatic flow past an isolated mountain. *Tellus*,
744 **32**, 348–364, doi:10.1111/j.2153-3490.1980.tb00962.x.

Snyder, C., W. C. Skamarock, and R. Rotunno, 1993: Frontal dynamics near and following frontal collapse. *J. Atmos. Sci.*, **50**, 3194–3211.

Sun, Y. Q., R. Rotunno, and F. Zhang, 2017: Contributions of moist convection and internal gravity waves to building the atmospheric -5/3 kinetic energy spectra. *J. Atmos. Sci.*, **74**, 185–201, doi: 10.1175/JAS-D-16-0097.1.

Sun, Y. Q., and F. Zhang, 2016: Intrinsic versus practical limits of atmospheric predictability and the significance of the butterfly effect. *J. Atmos. Sci.*, **73**, 1419–1438, doi:10.1175/JAS-D-15-0142.1.

Tabaei, A., and T. R. Akylas, 2007: Resonant long-short wave interactions in an unbounded rotating stratified fluid. *Stud. Appl. Math.*, **119**, 271–296.

Toro, E. F., 1999: *Riemann solvers and numerical methods for fluid dynamics, A practical introduction, 2nd ed.* Berlin: Springer.

van Leer, B., W.-T. Lee, and P. Roe, 1991: Characteristic time-stepping or local preconditioning of the euler equations. *AIAA paper*, **91-1552, 1991**.

Wang, L., and M. A. Geller, 2003: Morphology of gravity-wave energy as observed from 4 years (1998-2001) of high vertical resolution u.s. radiosonde data. *J. Geophys. Res.*, **108**, 4489, doi: 10.1029/2002JD002786.

Warner, C. D., and M. E. McIntyre, 2001: An ultrasimple spectral parameterization for nonorographic gravity waves. *J. Atmos. Sci.*, **58**, 1837–1857.

Wei, J., and F. Zhang, 2014: Mesoscale gravity waves in moist baroclinic jet-front systems. *J. Atmos. Sci.*, **71**, 929–952, doi:10.1175/JAS-D-13-0171.1.

Wei, J., and F. Zhang, 2015: Tracking gravity waves in moist baroclinic jet-front systems. *J. Adv. Model. Earth Syst.*, **7**, 67–91, doi:10.1002/2014MS000395.

Wei, J., F. Zhang, and J. H. Richter, 2016: An analysis of gravity wave spectral characteristics in moist baroclinic jet-front systems. *J. Atmos. Sci.*, **73**, 3133–3155, doi:10.1175/JAS-D-15-0316.1.

Wilhelm, J., T. Akylas, G. Bölöni, J. Wei, B. Ribstein, R. Klein, and U. Achatz, 2018: Interactions between meso- and sub-mesoscale gravity waves and their efficient representation in mesoscale-resolving models. *J. Atmos. Sci.*, **75**, 2257–2280, doi:10.1175/JAS-D-17-0289.1.

Zhang, F., 2004: Generation of mesoscale gravity waves in upper-tropospheric jet-front systems. *J. Atmos. Sci.*, **61**, 440–457, doi:https://doi.org/10.1175/1520-0469(2004)061⟨0440:GOMGWI⟩2.0.CO;2.

Zhang, F., N. Bei, R. Rotunno, C. Snyder, and C. C. Epifanio, 2007: Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics. *J. Atmos. Sci.*, **64**, 3579–3594, doi:10.1175/JAS4028.1.

Zhang, F., S. E. Koch, C. A. Davis, and M. L. Kaplan, 2001: Wavelet analysis and the governing dynamics of a large-amplitude mesoscale gravity-wave event along the east coast of the united states. *Quart. J. Roy. Meteor. Soc.*, **127**, 2209–2245, doi:10.1002/qj.49712757702.

Zhang, F., J. Wei, M. Zhang, K. P. Bowman, L. L. Pan, E. Atlas, and S. C. Wofsy, 2015: Aircraft measurements of gravity waves in the upper troposphere and lower stratosphere during the start08 field experiment. *Atmos. Chem. Phys.*, **15**, 7667–7684, doi:10.5194/acp-15-7667-2015.

786 Zhang, F., M. Zhang, J. Wei, and S. Wang, 2013: Month-long simulations of gravity waves over
787 north america and north atlantic in comparison with satellite observations. *Acta Meteor. Sinica*,
788 **27**, 446–454, doi:10.1007/s13351-013-0301-x.

LIST OF TABLES

TABLE 1. Initial model settings used in the idealized reference cases, including a one-dimensional wave packet case (1DWP), and two two-dimensional wave-packet cases with the angle between the initial horizontal wavenumber vector and the plane of modulation of the wave-packet amplitude being either 0 degree (2DWP00) or 90 degrees (2DWP90), respectively. The sensitivity experiment (2DWP90-HAMP-HF) is also included. In this table, NA indicates that the corresponding setting is not available.

Model Design	**1DWP**	**2DWP00**	**2DWP90**	**2DWP90-HAMP-HF**
initial wavelength λ_x (km) in x-direction	300	∞	300	30
initial wavelength λ_y (km) in y-direction	∞	300	∞	∞
initial vertical wavelength λ_z (km)	1	1	1	1
branch of intrinsic frequency (unitless)	+1	+1	+1	+1
initial wave packet scale in y direction σ_y (km)	∞	1500	1500	150
initial wave packet scale in z direction σ_z (km)	5	5	5	5
wave amplitude factor a_0 (unitless)	0.5	0.5	0.5	1.0
Coriolis parameter f (s^{-1})	10^{-4}	10^{-4}	10^{-4}	10^{-3}
buoyancy frequency N (s^{-1})	~ 0.02	~ 0.02	~ 0.02	~ 0.02
domain size in x direction L_x (km)	300	NA	300	30
domain size in y direction L_y (km)	NA	9000	9000	900
domain size in z direction L_z (km)	100	100	100	80
location y_0 of the wave packet center in y direction (km)	NA	4500	4500	450
location z_0 of the wave packet center in z direction (km)	30	30	30	30
Number of grid points in x direction in PincFloit-LES	16	1	16	32
Number of grid points in y direction in PincFloit-LES	1	512	512	128
Number of grid points in z direction in PincFloit-LES	1000	1000	1000	2400
resolution $\triangle x$ in x direction in PincFloit-LES (km)	18.75	NA	18.75	$\sim$0.94
resolution $\triangle y$ in y direction in PincFloit-LES (km)	NA	$\sim$17.58	$\sim$17.58	$\sim$7.03
resolution $\triangle z$ in z direction in PincFloit-LES (km)	0.1	0.1	0.1	$\sim$0.03
Number of grid points in x direction in PincFloit/MS-GWaM	1	1	1	1
Number of grid points in y direction in PincFloit/MS-GWaM	1	32	32	32
Number of grid points in z direction in PincFloit/MS-GWaM	100	100	100	100
resolution $\triangle x$ in x direction in PincFloit/MS-GWaM (km)	300	NA	300	30
resolution $\triangle y$ in y direction in PincFloit/MS-GWaM (km)	NA	281.25	281.25	28.125
resolution $\triangle z$ in z direction in PincFloit/MS-GWaM (km)	1	1	1	0.8

TABLE 2. Synopsis of the relevant general model parameters in both PincFloit/MS-GWaM and PincFloit-LES. In all the simulations, periodic boundaries are assumed in each available horizontal direction, while rigid boundaries are assumed in vertical direction. The time step Δt is determined through the CFL criterion ($CFL = 0.5$), while a 1-s upper threshold for the time step is used (i.e., $\Delta t_{max} = 1s$). The smoothing parameter indicates the total number of grid cells used for a local smoothing along all the available spatial directions. In PincFloit/MS-GWaM, the initial total number of ray volumes n_{ray} corresponds to the product of the number of grid cells in the 5-σ segment or box, and the corresponding number per grid cell and spatial direction (i.e., $\tilde{n}_{z,ray}$ in 1D wave packet case; $\tilde{n}_{z,ray}$ and $\tilde{n}_{y,ray}$ in 2D wave packet case). In this table, NA indicates that the corresponding parameter is not available.

Configuration Parameter	**PincFloit/MS-GWaM**	**PincFloit-LES**
Boundary condition in x direction	Periodic boundary	Periodic boundary
Boundary condition in y direction	Periodic boundary	Periodic boundary
Boundary condition in z direction	Rigid boundary	Rigid boundary
Resolution	Please check Table 1.	Please check Table 1.
Time step Δt	dynamic, $CFL = 0.5$, $\Delta t_{max} = 1s$	dynamic, $CFL = 0.5$, $\Delta t_{max} = 1s$
Wave number interval Δm_0	$\Delta m_0 = 10^{-5} m^{-1}$	NA
$\tilde{n}_{x,ray}$ (1D wave packet case)	1	NA
$\tilde{n}_{y,ray}$ (2D wave packet case)	25	NA
$\tilde{n}_{z,ray}$ (both 1D and 2D wave packet cases)	25	NA
Initial total number of ray volumes n_{ray} (1D wave packet case)	$n_{ray} = \frac{5\sigma_z}{\Delta z}\tilde{n}_{z,ray}$	NA
Initial total number of ray volumes n_{ray} (2D wave packet case)	$n_{ray} = \frac{5\sigma_z}{\Delta z}\tilde{n}_{z,ray} \times \frac{5\sigma_y}{\Delta y}\tilde{n}_{y,ray}$	NA
Smoothing parameter	Wave-induced flux: $n_{s,flux} = 3$	Dynamic Smagorinsky: $n_{s,dyn} = 5$

TABLE 3. Initial model settings used in the sensitivity experiment.

Sensitivity Experiment	**Model Design**
2DWP90-HAMP	Same as 2DWP90, except that the initial wave amplitude factor a_0 is 1.0.
2DWP90-HAMP-NoET	Same as 2DWP90-HAMP in PincFloit/MS-GWaM with the use of direct approach, except that the elastic term is ignored.
2DWP90-HAMP-NoHeat	Same as 2DWP90-HAMP in PincFloit/MS-GWaM with the use of direct approach, except that the wave-induced heating term is ignored.
2DWP90-HAMP-NoET-NoHeat	Same as 2DWP90-HAMP in PincFloit/MS-GWaM with the use of direct approach, except that both elastic term and wave-induced heating term are ignored.
2DWP90-HAMP-HF	Same as 2DWP90-HAMP, except for the below four major changes: (1) Coriolis parameter is increased from 10^{-4} s^{-1} to 10^{-3} s^{-1}, (2) Initial zonal wavelength λ_x is decreased from 300 km to 30 km, (3) Initial wave packet scale in y direction σ_y is decreased from 1500 km to 150 km, (4) A much finer resolution along the x and z directions in the wave resolving simulation. Please refer to Table 1 for more details on its model setting. The main purpose of this experiment is to have model data with more inertial periods by choosing a much higher value for the Coriolis parameter, in order to have a much longer vertical propagation of the unstable wave packet.
2DWP90-HAMP-HF-DeCoupled	Same as 2DWP90-HAMP-HF in PincFloit/MS-GWaM with the use of direct approach, except that the induced mean wind has no effect on the gravity waves.
2DWP90-HAMP-HF-NoSat	Same as 2DWP90-HAMP-HF in PincFloit/MS-GWaM with the use of direct approach, except that the saturation scheme is switched off.

LIST OF FIGURES

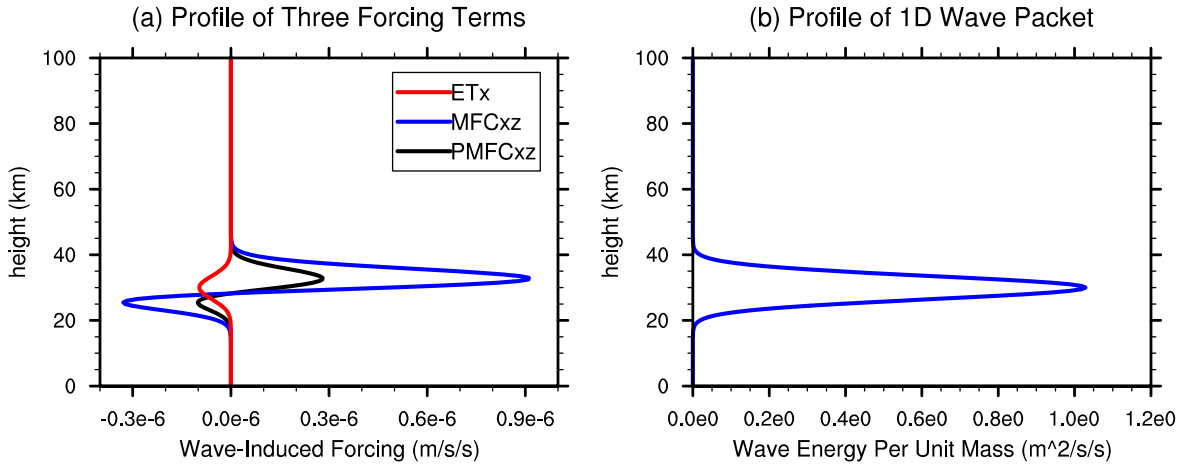
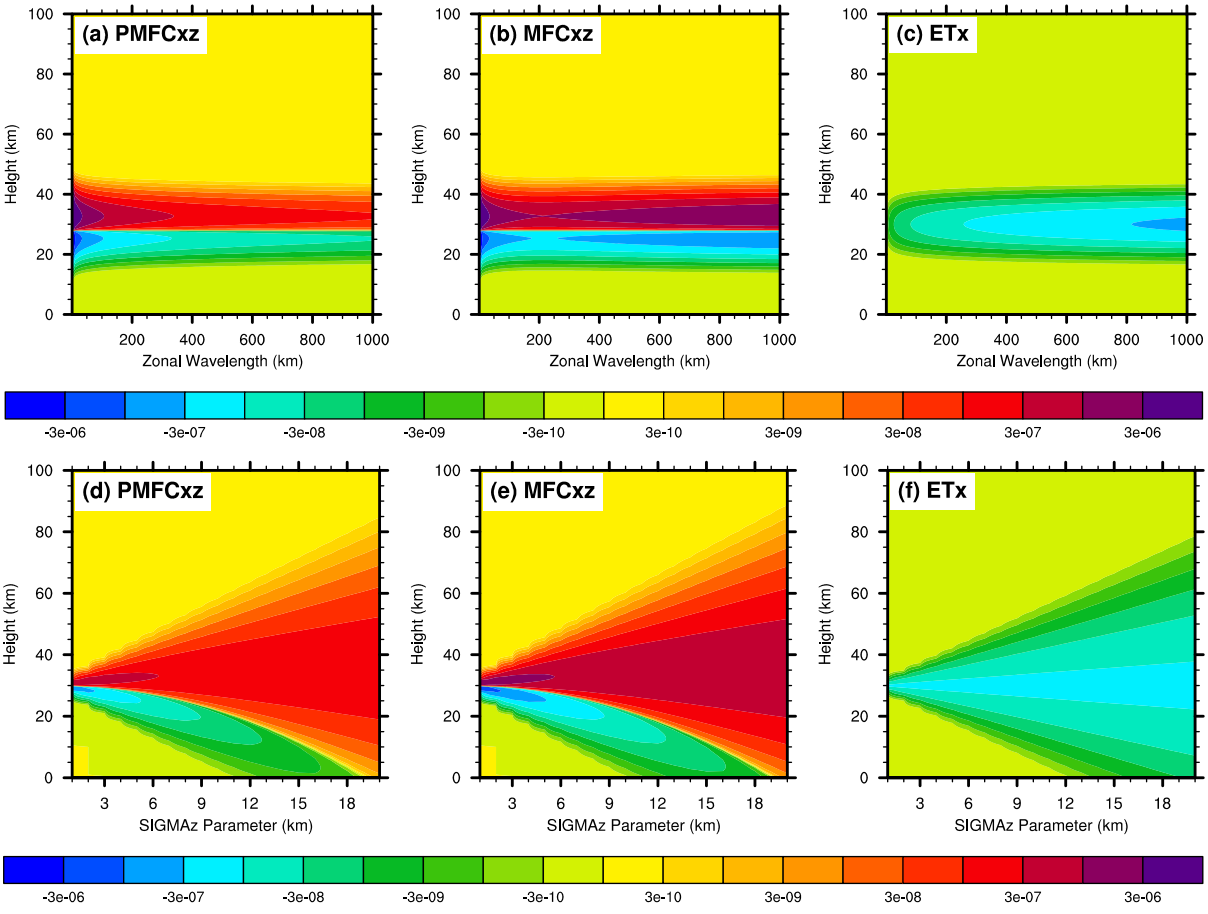


FIG. 1. The initial vertical profiles in the 1DWP case (i.e., a prescribed one-dimensional wave packet as a reference in this study): (a) three wave-induced forcing terms (unit: m/s^2), and (b) wave energy per unit mass (unit: m^2/s^2). Those three forcing terms include $PMFC_{xz}$ (black line) in the pseudomomentum scheme, and MFC_{xz} (blue line) and ET_x (red line) in the direct scheme. Please also refer to Table 1 for the details of the experiment design in 1DWP.



892 FIG. 2. The comparison among three wave-induced forcing terms (unit: m/s^2) in a wide range of one-
893 dimensional wave packet cases: (a,d) $PMFC_{xz}$ in the pseudomomentum scheme, and (b,e) MFC_{xz} and (c,f) ET_x
894 in the direct scheme. The upper panel demonstrates the sensitivity to the change of zonal wavelengths, and the
895 lower panel demonstrates the sensitivity to the change of the parameter σ_z (i.e., vertical wave packet scale). In
896 each sensitivity study, all the other parameters are kept the same as those in the initial setting of 1DWP. Please
897 also refer to Table 1 for the details of the experiment design in 1DWP.

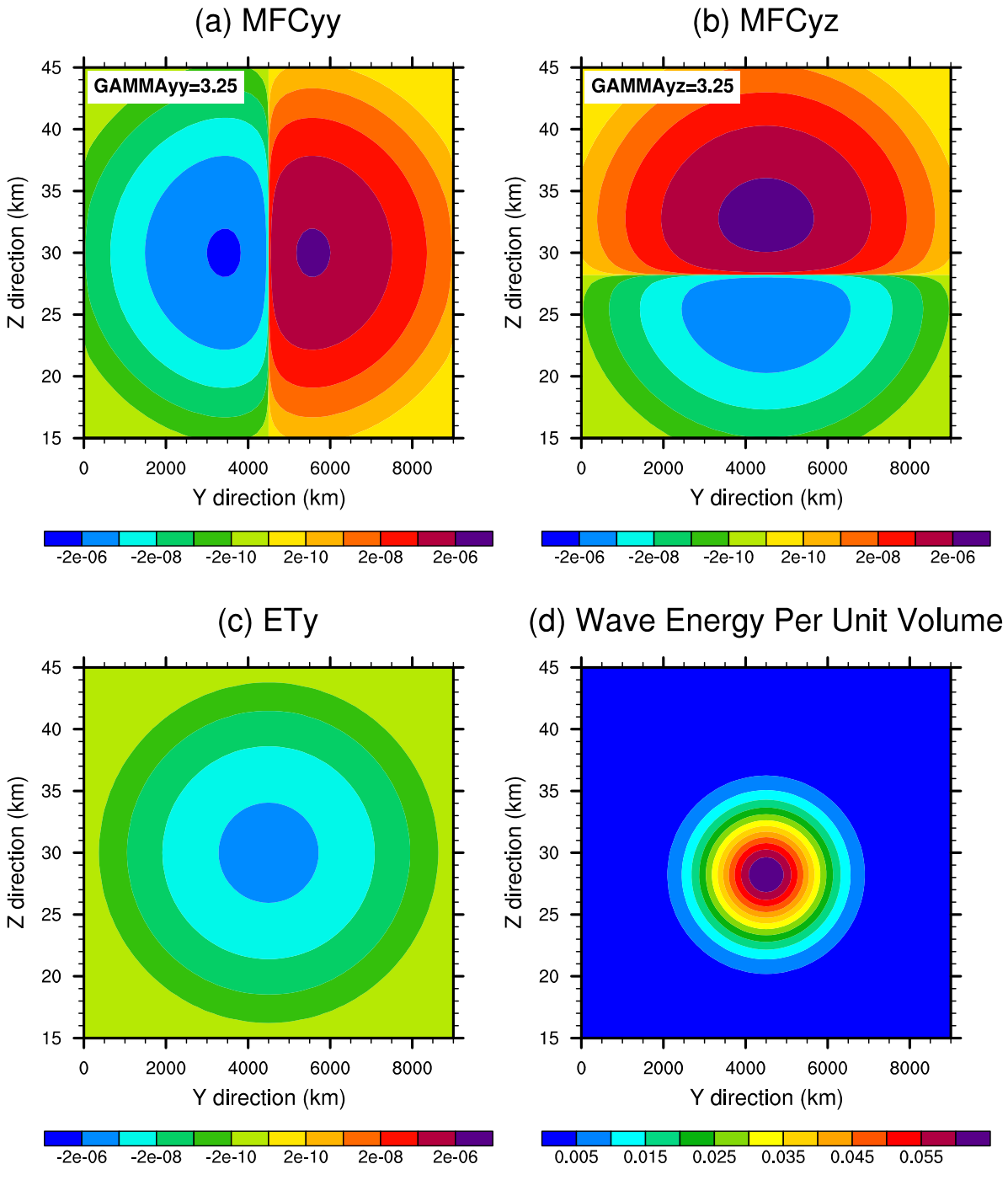


FIG. 3. Initial vertical cross sections of the contributors to the direct approach in the 2DWP00 case (with horizontal wavenumber in the plane of modulation of the wave-packet amplitude), i.e. (a) MFC_{yy} (unit: m/s^2), (b) MFC_{yz} (unit: m/s^2), and (c) ET_y (unit: m/s^2), together with (d) the wave energy per unit volume E_w (unit: Pa). In the case of the momentum-flux convergence contributors, γ_{ij} is shown in the top left corner, indicating how much stronger the momentum flux is as compared to the pseudomomentum flux.

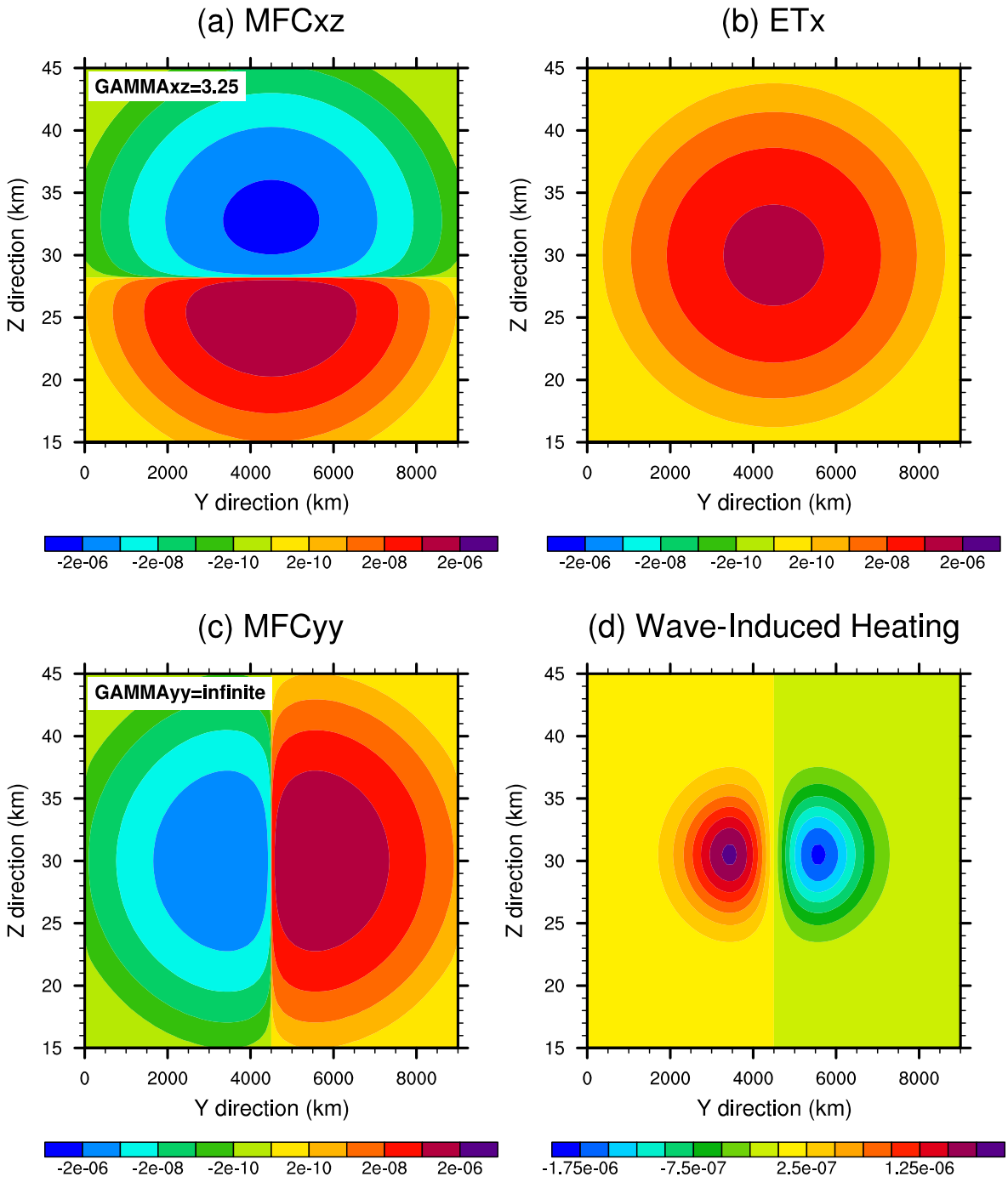


FIG. 4. Initial vertical cross sections of the contributors to the direct approach in the 2DWP90 case (with horizontal wavenumber orthogonal to the plane of modulation of the wave-packet amplitude), i.e. (a) MFC_{xz} (unit: m/s^2), (b) ET_x (unit: m/s^2), (c) MFC_{yy} (unit: m/s^2), and (d) the wave-induced heating (unit: K/s). In the case of the momentum-flux convergence contributors, γ_{ij} is shown in the top left corner, indicating how much stronger the momentum flux is as compared to the pseudomomentum flux.

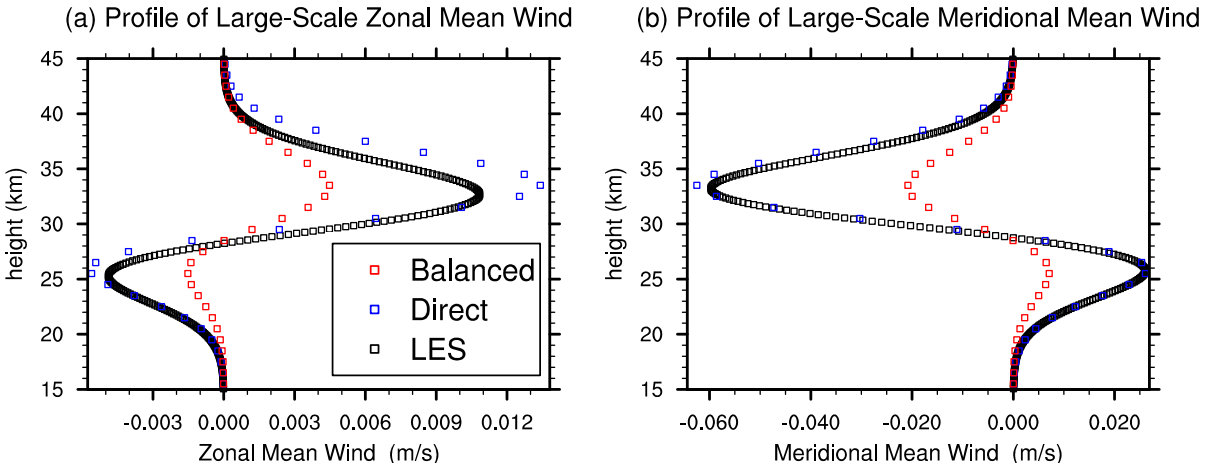
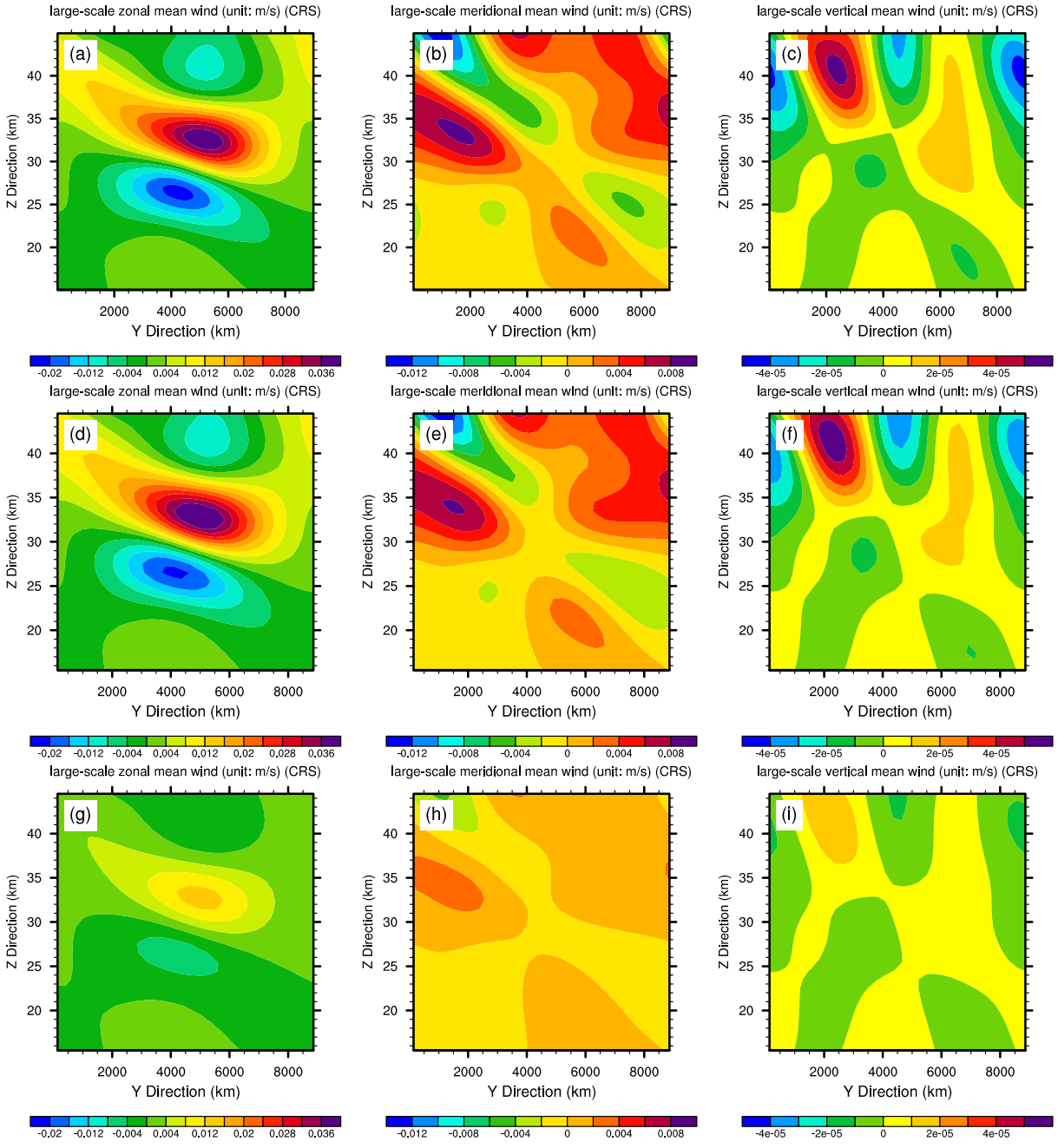


FIG. 5. Comparison of two horizontal mean wind components (unit: m/s) among three model codes at t=1500 *min* (approximately one and a half inertial periods) in the 1DWP case. Those three codes include a high-resolution wave-resolving simulation (black boxes), and two coarse-resolution simulations with gravity wave parameterization based on the direct or pseudomomentum approach (blue boxes and red boxes, respectively). The two mean wind components include (a) zonal wind component and (b) meridional wind component. The mean vertical wind is zero in all simulations. Please refer to Table 1 and Table 2 for the details of the simulations.



915 FIG. 6. The comparison of all three mean wind components (unit: m/s) among simulations using the same
916 codes as in Fig. 5, again at t=1500 min (approximately one and a half inertial periods) but now for the 2DWP00
917 case. Results from the high-resolution wave-resolving simulation are shown in the upper row, with those from
918 coarse-resolution PincFloit/MS-GWaM simulations using the direct scheme (middle row) and the pseudomo-
919 mentum scheme (bottom row) farther below. The three mean wind components include the horizontal $x-$ (left
920 column) and $y-$ (middle column) wind components, and vertical wind component (right column). Please refer
921 to Table 1 and Table 2 for the details of the simulations.

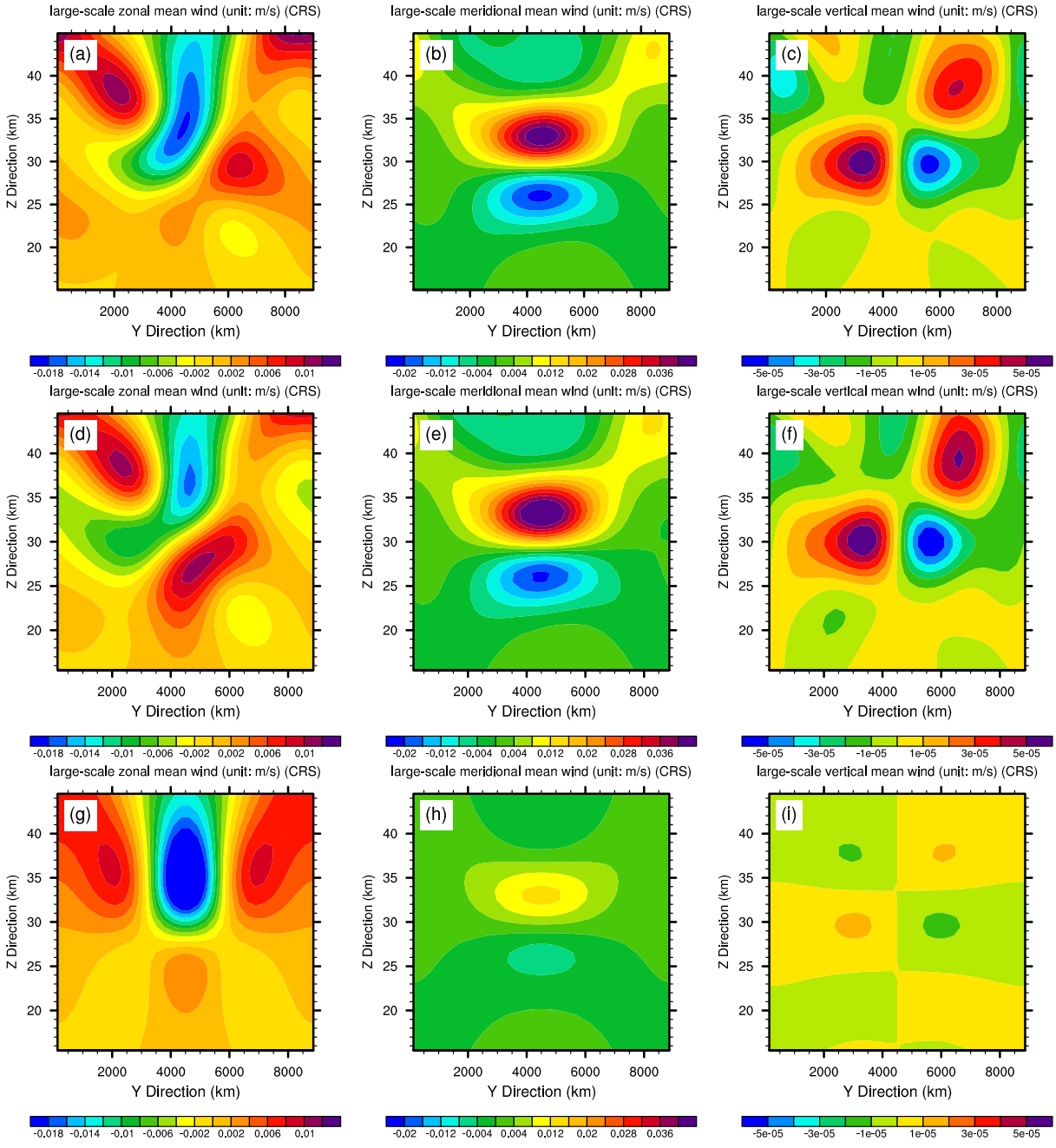
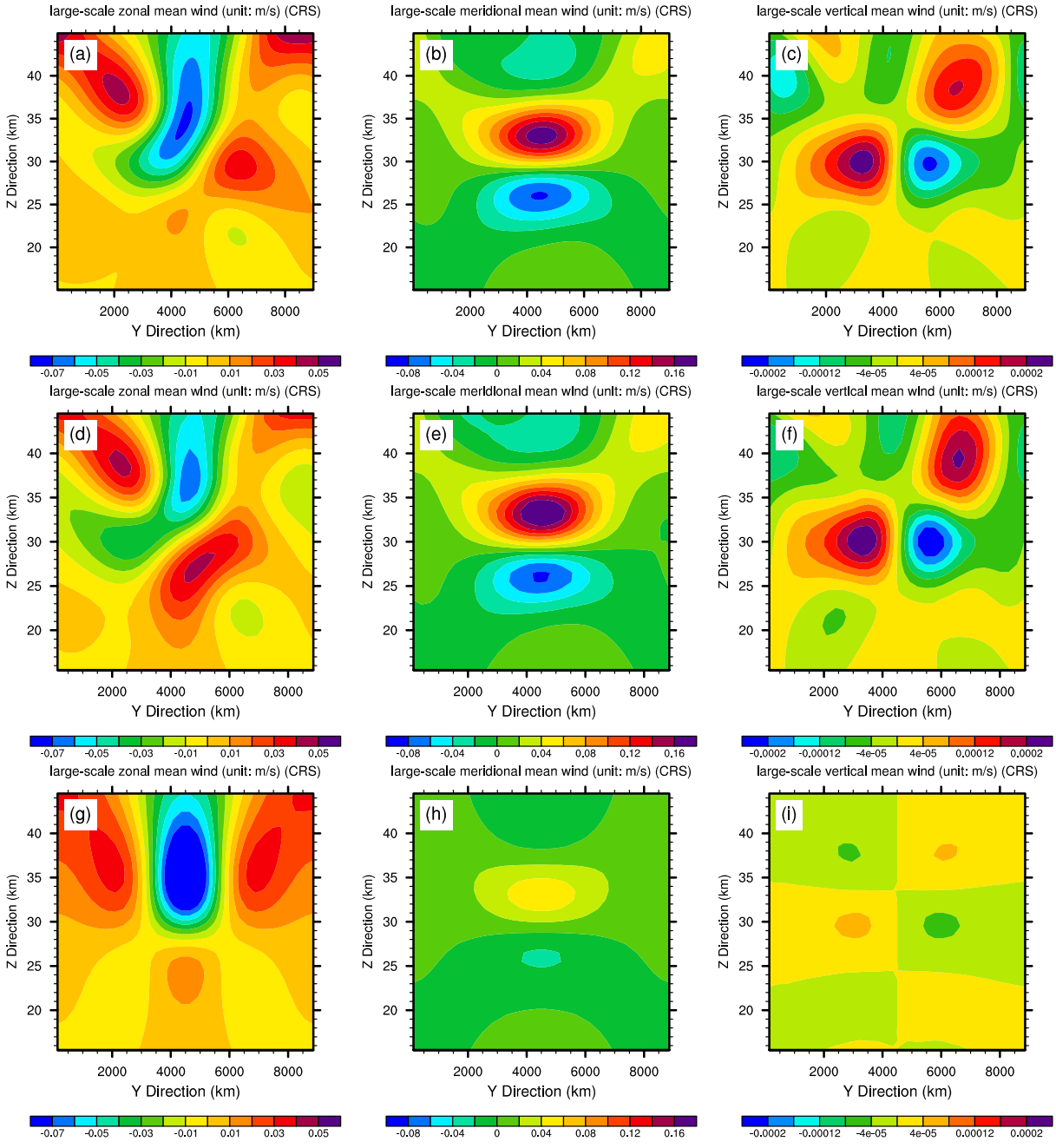
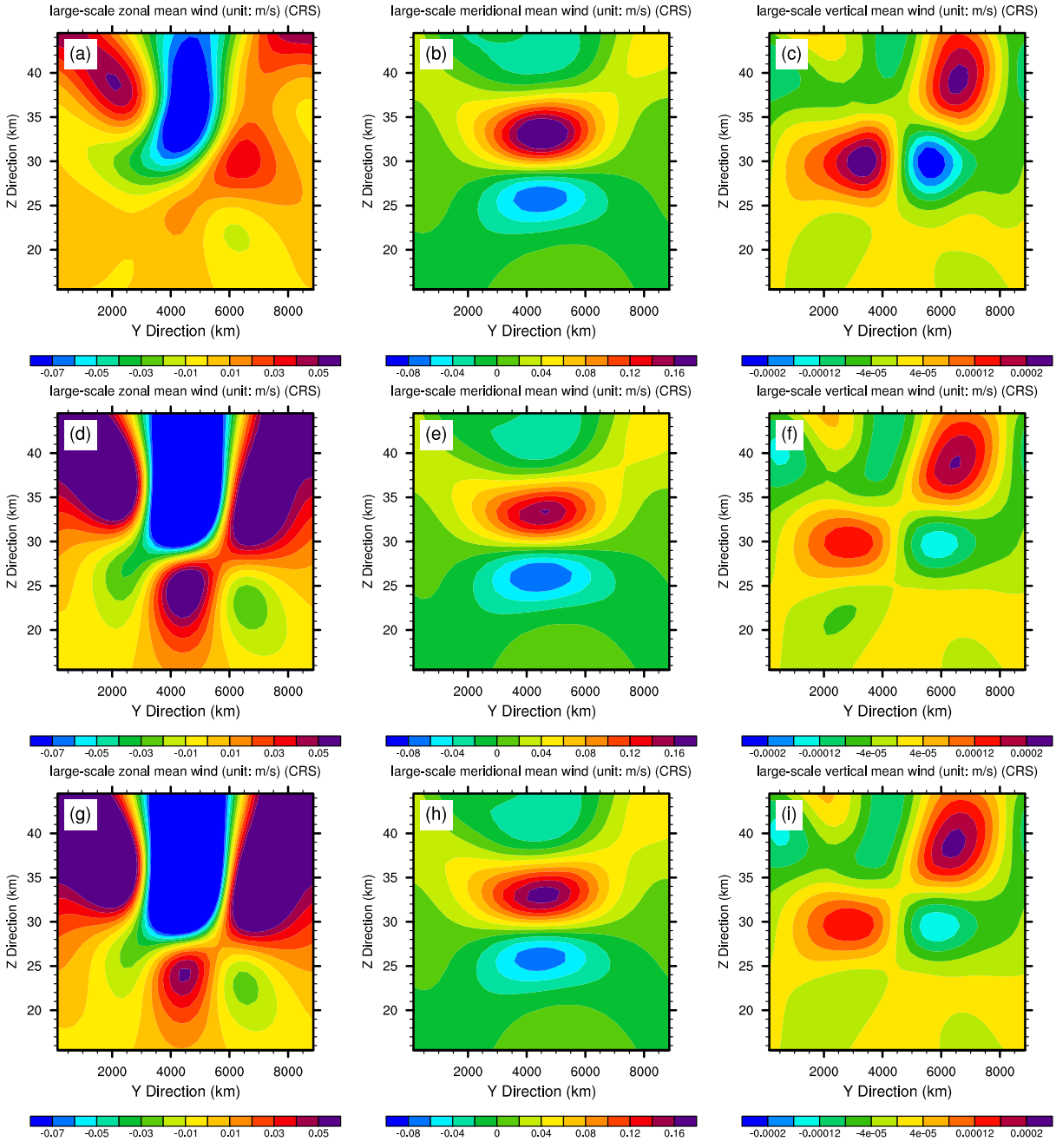


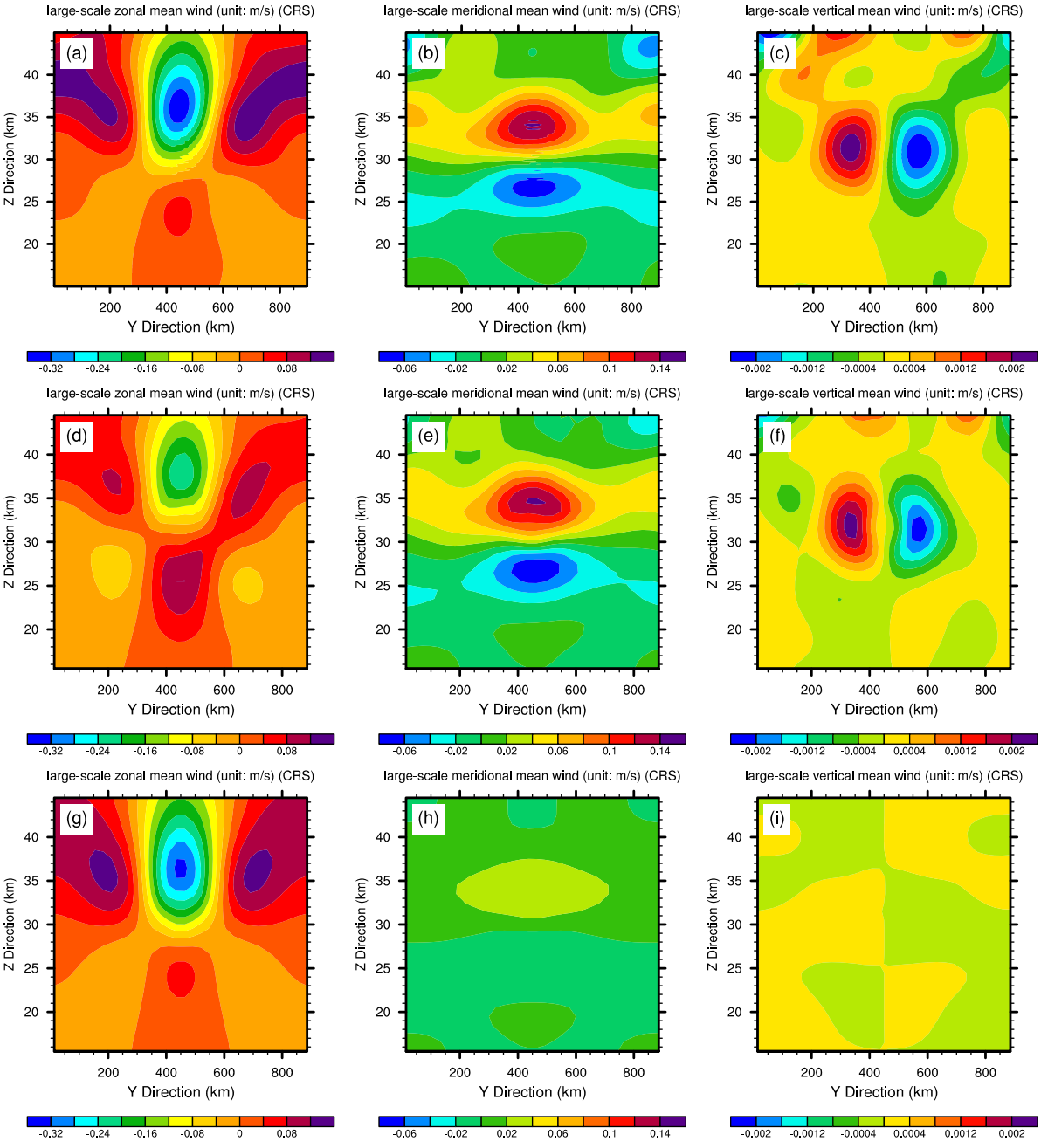
FIG. 7. Same as Fig. 6, but for the 2DWP90 case.



922 FIG. 8. Same as Fig. 6, but for the 2DWP90-HAMP case. Please refer to Table 3 for the details of the
923 sensitivity experiments.



924 FIG. 9. Same as the results from the direct scheme in Fig. 8, except that the sensitivity experiments of
925 (top row) 2DWP90-HAMP-NoET, (middle row) 2DWP90-HAMP-NoHeat, and (bottom row) 2DWP90-HAMP-
926 NoET-NoHeat are shown. In 2DWP90-HAMP-NoET (top row) the elastic term has been ignored in the sim-
927 ulations. In 2DWP90-HAMP-NoHeat (middle row) the wave-induced heating term has been ignored, and in
928 2DWP90-HAMP-NoET-NoHeat (bottom row) both elastic term and heating term have been neglected. Please
929 refer to Table 3 for the details of the sensitivity experiments.



930 FIG. 10. Same as Fig. 6, but for the 2DWP90-HAMP-HF case at t=471.15 *min* (approximately four and a half
931 inertial periods). Please refer to Table 3 for the details of the sensitivity experiments.

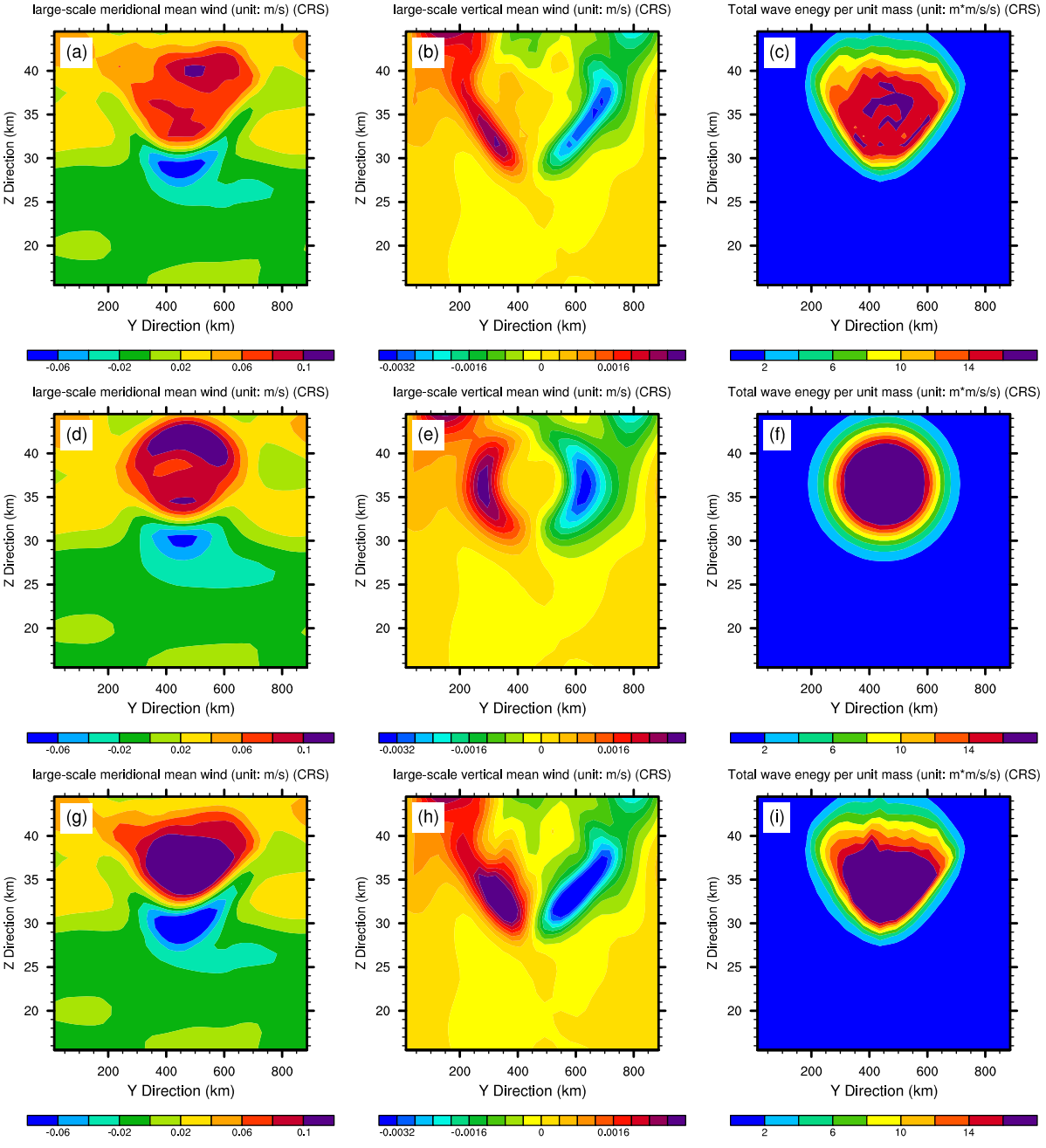


FIG. 11. All the subplots in this figure are based on the PincFloit/MS-GWaM. To be be specific, same as the PincFloit/MS-GWaM simulation in Fig. 10 using the direct approach (top row), now however at a much later stage (t=1832.25 min, approximately seventeen and a half inertial periods after initialization), and compared to results from a corresponding simulation where the mean-flow impact on the waves has been switched off (middle row), and a simulation where the saturation scheme, describing the effect of turbulent wave breaking, has been put out of use (bottom row). The three compared variables include the large-scale mean y-wind (left column, in m/s), large-scale mean vertical wind (middle column, in m/s), and the wave energy $E_w/\bar{\rho}$ per unit mass (right column, in $\mathrm{m}^2/\mathrm{s}^2$).