Application of the Compressible, Nonhydrostatic, Balanced Omega Equation in Estimating Diabatic Forcing for Parameterization of Inertia–Gravity Waves: Case Study of Moist Baroclinic Waves Using WRF

MAHNOOSH HAGHIGHATNASAB, MOHAMMAD MIRZAEI, AND ALI R. MOHEBALHOJEH

Institute of Geophysics, University of Tehran, Tehran, Iran

CHRISTOPH ZÜLICKE

Leibniz Institute of Atmospheric Physics, University of Rostock, Kühlungsborn, Rostock, Germany

RIWAL PLOUGONVEN

Laboratoire de Météorologie Dynamique, IPSL, Ecole Polytechnique, Université Paris-Saclay, CNRS, Palaiseau, France

(Manuscript received 13 February 2019, in final form 31 August 2019)

ABSTRACT

The parameterization of inertia–gravity waves (IGWs) is of considerable importance in general circulation models. Among the challenging issues faced in studies concerned with parameterization of IGWs is the estimation of diabatic forcing in a way independent of the physics parameterization schemes, in particular, convection. The requirement is to estimate the diabatic heating associated with balanced motion. This can be done by comparing estimates of balanced vertical motion with and without diabatic effects. The omega equation provides the natural method of estimating balanced vertical motion without diabatic effects, and several methods for including diabatic effects are compared. To this end, the assumption of spatial-scale separation between IGWs and balanced flows is combined with a suitable form of the balanced omega equation. To test the methods constructed for estimating diabatic heating, an idealized numerical simulation of the moist baroclinic waves is performed using the Weather Research and Forecasting (WRF) Model in a channel on the $f$ plane. In overall agreement with the diabatic heating of the WRF Model, in the omega-equation-based estimates, the maxima of heating appear in the warm sector of the baroclinic wave and in the exit region of the upper-level jet. The omega-equation-based method with spatial smoothing for estimating balanced vertical motion is thus presented as the proper way to evaluate diabatic forcing for parameterization of IGWs.

1. Introduction

The notion of balance has played a major role in shaping our understanding of a great range of phenomena involving waves and vortical flows (Vallis 1996; McIntyre and Norton 2000; Ford et al. 2000; McIntyre 2009). Among the developments has been the introduction of balanced omega equation (Hoskins et al. 1985), in which the impact of diabatic processes is

DOI: 10.1175/JAS-D-19-0039.1
© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
IGWs as orography, convection, shear instability, jet streams, and fronts (e.g., Uccellini and Koch 1987; Eckermann and Vincent 1993; O’Sullivan and Dunkerton 1995; Guest et al. 2000; Plougonven and Snyder 2007). The IGWs affect the atmospheric general circulation through breaking and dissipation by exerting an acceleration on the mean flow (e.g., McLandress 2002; Alexander et al. 2010; Kim et al. 2003) and contributing significantly to remarkable phenomena of the quasi-biennial oscillation (QBO) (Dunkerton 1997; Evan et al. 2012) and semiannual oscillation (Ern et al. 2015) in the stratosphere. The general circulation models (GCMs) fail to capture a wide spectrum of the IGWs due to their limitation in resolving short spatial and temporal scales, which need to be parameterized in the GCMs based on information available on their sources and propagation characteristics. In parameterization, the IGWs are commonly classified into orographic waves forced by Earth’s topography and nonorographic waves forced by jets, fronts, and convection. Among the latter, convection plays an important role in generation of IGWs in the tropical region and during summer in midlatitudes. Three major mechanisms for generation of IGWs from convection has been proposed in previous studies including thermal forcing (Lin and Smith 1986; Chun and Baik 2002), obstacle effect (Clark et al. 1986; Beres et al. 2002), and mechanical oscillator effect (Pierce and Coroniti 1966; Lane et al. 2001).

There seems to be a close connection between the characteristics of the convection cells and the spectrum of IGWs generated by convection (e.g., Alexander and Holton 1997; Piani et al. 2000; Song et al. 2003; Beres and Alexander 2004; Choi and Chun 2014), which is the basis for the thermal forcing mechanism. The spectrum of IGWs generated by convection depends on the latent heat released, properties of convection in the lower troposphere and the environmental wind (Beres and Alexander 2004; Beres et al. 2005). There have been many studies devoted to finding a suitable method for parameterization of convectively generated IGWs in the GCMs mainly following two main approaches, based on either the momentum flux (e.g., Song et al. 2003; Beres and Alexander 2004) or the energy (e.g., Zülicke and Peters 2008; Mirzaei et al. 2014) transmitted by the waves. Beres and Alexander (2004) presented a method for estimating the spectrum of IGWs forced by convection based on the linear analysis of the two-dimensional primitive equations. Their analysis of the momentum flux carried by gravity waves showed generation of nonsteady gravity waves by the oscillatory component and stationary gravity waves by interaction of the environmental wind with the steady component of the heating. In the relations developed by Beres and Alexander (2004), the gravity wave spectrum depends on the depth of the convective heating, environmental wind in the troposphere, and the convective heating rate. Beres et al. (2005) implemented the method of Beres and Alexander (2004) to estimate the momentum flux of the gravity waves in the Whole Atmosphere Community Climate Model (WACCM2). Stephan and Alexander (2015) introduced a modeling approach based on retaining the large-scale resolution of the GCMs but incorporating the gravity waves generated by convection. In this way, they carried out a dry idealized simulation using the Weather Research and Forecasting (WRF) Model forced by diabatic heating field obtained by converting radar precipitation rates, and succeeded to reproduce the observed gravity waves associated with squall lines and mesoscale convective systems.

In the energy-based parameterization of IGWs generated by nonorographic sources, including convection, proposed by Mirzaei et al. (2014), a spatial averaging procedure is used to decompose the circulation into large- and small-scale components treated as representing balanced flow and IGWs, respectively. Underpinning this procedure is the assumption of a spatial-scale separation between IGWs and balanced flows (Aspden and Vanneste 2010). Parameterization of IGW generation is then reduced to finding relations between the small-scale IGWs and the large-scale adiabatic (jet–front) and diabatic (convection) sources. For the diabatic source, the problem becomes particularly challenging because of the coupling between latent heating and vertical motion in moist convection, which is itself parameterized in GCMs (Lane and Reeder 2001; Müller et al. 2018).

Mirzaei et al. (2014) developed the relation between the short-scale IGW-related ageostrophic flow and the convection-induced ageostrophic motions of larger scales by scaling the linearized equations of motion, providing a parameterization relation for IGWs energy in terms of the diabatic heating released from the resolved and parameterized convection in numerical simulations by the WRF Model. The diabatic heating required for the parameterization was estimated by smoothing the latent heat released during condensation in the simulations over a 500-km interval and using a threshold for diabatic heating based on scaling considerations.

A quantification of IGWs generated in numerical simulations of the moist baroclinic waves (BCWs) was presented by Mirzaei et al. (2017, hereafter MMZP), who carried out the wave–vortex decomposition (WVD; see Viúdez 2007; Snyder et al. 2007; Mirzaei et al. 2012; Vanneste 2013; Bühler et al. 2014) by employing the nonlinear balance relations of $\delta-\gamma$ and Bolin–Charney
and the Rossby number expansion up to third order. MMZP applied the WVD methods to the Boussinesq primitive equations on the \( f \) plane to obtain the omega equations in compressible (incompressible) form with explicit (implicit) forcing by diabatic heating. They found out that the WVD method using second-order Rossby number expansion in incompressible form is the optimal method as far as the energy of IGWs is concerned. The difference between the results of the WVD in compressible and incompressible forms can be related to the accuracy of the diabatic heating estimation, which may significantly affect the unbalanced flow.

In the current work the WVD procedure of MMZP is further developed to incorporate in a different way the impact of forcing by diabatic heating on unbalanced flow in the idealized moist simulation of BCWs by the WRF Model. To this end, we derive a new WVD based on balance relations of the Rossby number expansion that both compressible and nonhydrostatic effects are explicitly taken into account. In particular, we explore the consequences of assuming both the spatial-scale separation between IGWs and balanced flows and the validity of the compressible, nonhydrostatic, balanced omega equation that is derived here. Through this assumption, different methods are presented and compared for the estimation of diabatic forcing; the methods involve spatial smoothing, running the model at low resolution and scaling relations following the work of Mirzaei et al. (2014). The estimates are compared with the WRF Model’s own diabatic heating and that of the slantwise convection proposed in Emanuel et al. (1987) as previously applied in MMZP.

This paper is organized as follows. Section 2 presents the omega equation based on Rossby number expansion as well as the methods of obtaining diabatic forcing followed by details of the numerical simulations carried out in section 3. The results of different methods used to determine the diabatic forcing are given in section 4. Finally, concluding remarks are presented in section 5.

### 2. Methodology

The objective of WVD as applied in previous studies (e.g., Vallis 1996; Mohebalhojeh 2002; MMZP) is to achieve optimal accuracy in representing IGWs, within the constraints set by the slow quasi manifold (Ford et al. 2000). MMZP presented three different sets of balance relations of the first-order \( \delta \gamma \) and Bolin–Charney and the first- to third-order Rossby number expansion, which is constructed by applying an asymptotic expansion to the primitive equations using the Rossby number (\( Ro \)) as the small parameter. Using each one of the three methods lead to an omega equation that can be solved to obtain balanced vertical velocity \( w_b \), and subsequently, the unbalanced vertical velocity from \( w = w_b + w_{\text{unb}} \), with \( w \) the actual vertical velocity. MMZP concluded that the second-order Rossby number expansion is the optimal method to determine balanced flow by solving the omega equation.

#### a. Formulation of omega equation

Our aim here is to develop a new omega equation similar to (2.24) in MMZP for Rossby number expansion method without making use of hydrostatic balance relation. For completeness, we rewrite the vorticity and divergence equations as in MMZP:

\[
-f \delta = \{ Ro \} \left[ \frac{\partial \zeta}{\partial t} + \nabla \cdot (V \zeta) \right] + \{ Ro \} k \cdot \nabla \times \left( \frac{\partial V}{\partial z} \right).
\]

(2.1)

\[
\gamma = \{ Ro \} \left[ -2J(u, v) + \nabla \cdot (V \delta) + \nabla \cdot \left( \frac{\partial V}{\partial z} \right) + \frac{\partial \delta}{\partial t} \right].
\]

(2.2)

where the Coriolis parameter \( f \) is taken to be constant, \( \delta \) is the horizontal divergence, \( \{ Ro \} \) denotes order in Rossby number, \( V = (u, v) \) is the horizontal velocity vector, \( \zeta \) is the vorticity, \( k \) is the unit vector in the local vertical direction, and \( w \) is the vertical velocity. The horizontal acceleration divergence \( \gamma = \nabla \cdot D\mathbf{V}/Dt \) is given by

\[
\gamma = -\frac{1}{\rho_0} \nabla^2 p.
\]

(2.3)

where \( p \) is pressure and \( \rho_0(z) \) denotes the reference density. Now we take the time derivative of (2.2),

\[
\frac{\partial \gamma}{\partial t} = \{ Ro \} \left[ -2J(u, v) + \nabla \cdot (V \delta) + \nabla \cdot \left( \frac{\partial V}{\partial z} \right) + \frac{\partial \delta}{\partial t} \right].
\]

(2.4)

and (2.3),

\[
\frac{\partial \gamma}{\partial t} = \{ Ro \} \left[ \frac{\partial \zeta}{\partial t} - \frac{1}{\rho_0} \frac{\partial \zeta^2}{\partial t} \right].
\]

(2.5)

Substituting (2.1) in (2.5) for the time tendency of \( \zeta \), the time tendency of \( \gamma \) becomes

\[
\frac{\partial \gamma}{\partial t} = -f^2 \delta - \{ Ro \} f \nabla \cdot (V \zeta) - \{ Ro \} f k \cdot \nabla \times \left( \frac{\partial V}{\partial z} \right) - \{ Ro \} \frac{1}{\rho_0} \nabla^2 \left( \frac{\partial p}{\partial t} \right).
\]

(2.6)
Equating (2.4) and (2.6) and taking partial derivative with respect to \( z \), we arrive at the following equation:

\[
- f^2 \frac{\partial \delta}{\partial z} = \{ \text{Ro} \} f \nabla \cdot \left( \frac{\partial}{\partial z} (V \zeta) \right) + \{ \text{Ro}^2 \} f \mathbf{k} \cdot \nabla \left( w \frac{\partial V}{\partial z} \right) + \{ \text{Ro} \} \frac{1}{\rho_0} \nabla \cdot \frac{\partial p}{\partial t} + \{ \text{Ro}^2 \} \frac{\partial}{\partial z} \left[ -2 J(u, v) + \nabla \cdot (V \delta) + \nabla \cdot \left( w \frac{\partial V}{\partial z} \right) + \frac{\partial^2 \delta}{\partial t^2} \right].
\]  

(2.7)

To reach an independent expression for \((\partial/\partial z)/(\partial p/\partial t)\), the thermodynamics energy equation as

\[
\frac{DT}{Dt} - \frac{1}{\rho c_p} \frac{Dp}{Dt} = Q
\]

(2.8)

is invoked in which \( T \) is the temperature, \( \rho \) is the density, \( c_p \) is the heat capacity per constant pressure, and \( Q \) is the heating rate per unit mass. Now using (2.8), the equation of state \( p = \rho RT \) where \( R \) denotes the specific gas constant, and the compressible continuity equation \( Dp/Dt = -\rho (\delta + \partial w/\partial z) \) yields

\[
\frac{\partial}{\partial z} \left( \frac{\partial p}{\partial t} \right) = \frac{\partial}{\partial z} \left( \frac{p}{1 - \kappa} \left[ w \frac{d}{dz} (\ln \rho_0) \right] \right) - \frac{\partial}{\partial z} (V \cdot \nabla p) - \left( \frac{\partial w}{\partial z} \frac{\partial p}{\partial z} \right) - w \frac{\partial^2 p}{\partial z^2} + \frac{\partial (\rho Q)}{\partial z} (1 - \kappa).
\]

(2.9)

Without loss of generality, to relate \( \delta \) to \( w \), the mass continuity equation can be approximated in the anelastic form [see (2.34) in Holton and Hakim (2013)]:

\[
\delta = -\frac{1}{\rho_0} \frac{\partial}{\partial z} (\rho_0 w).
\]

(2.10)

Taking partial derivative of (2.10) with respect to \( z \), the following equation is then obtained:

\[
\frac{\partial^2 \delta}{\partial z^2} = -\frac{\partial^2 w}{\partial z^2} - \frac{\partial w}{\partial z} \frac{\partial p}{\partial z} (\ln \rho_0) - w \frac{\partial^2}{\partial z^2} (\ln \rho_0).
\]

(2.11)

Substituting (2.11) and (2.9) into (2.7), and then setting \( p = \bar{p}(z) + p', \bar{p}(z) \) and \( p' \) being the reference and perturbation pressure fields, respectively, the omega equation is given by

\[
\mathcal{D} w = b \left\{ \{ \text{Ro} \} \left\{ \frac{\partial}{\partial z} [V \cdot (V \xi)] - \frac{1}{\rho_0} \nabla^2 \left( \nabla \cdot \frac{\partial p}{\partial z} \right) \right\} 
- \frac{1}{\rho_0} \nabla^3 \left( \frac{\partial V}{\partial z} \cdot \nabla p \right) + \frac{1}{\rho_0} \nabla^2 \left[ \frac{\partial (\rho Q)}{\partial z} (1 - \kappa) \right] \right\} 
+ \{ \text{Ro} \} \left\{ \frac{1}{\rho_0} \nabla^2 \left[ \frac{p w}{1 - \kappa} \frac{d}{dz} (\ln \rho_0) \right] \right\} 
- \frac{1}{\rho_0} \nabla^2 \left( \frac{\partial w}{\partial z} \frac{\partial p}{\partial z} \right) 
+ \{ \text{Ro}^2 \} \left\{ f \mathbf{k} \cdot \nabla \times \frac{\partial}{\partial z} \left( w \frac{\partial V}{\partial z} \right) - \frac{1}{\rho_0} \nabla^2 \left( \frac{\partial^2 p'}{\partial z^2} \right) \right\} 
+ \{ \text{Ro}^2 \} \left\{ \frac{\partial}{\partial z} \left[ -2 J(u, v), + \nabla \cdot (V \delta), + \nabla \cdot \left( w \frac{\partial V}{\partial z} \right) + \frac{\partial^2 \delta}{\partial t^2} \right] \right\}.
\]

(2.12)

where \( \mathcal{D} = \nabla^2 + \mathcal{L} \), with \( \mathcal{L} = a \partial^2/\partial z^2 + \alpha \partial / \partial z + \beta \), \( a = \rho_0 f^2/(d^2 \bar{p}/d z^2), \alpha = d \ln \rho_0 / d z, \beta = d^2 \ln \rho_0 / d z^2 \), and \( b = a f^2 \). The doubly underlined term is solely due to compressibility and the singly underlined terms are those due to taking nonhydrostatic effect into account. In terms of ordering in Rossby number, no prior assumption has been made here on the magnitude of the compressibility and nonhydrostatic effects. This and the consideration of
nonhydrostatic effects for estimating \( \frac{\partial}{\partial \eta}(\rho \frac{\partial \eta}{\partial t}) \) distinguishes the derivation here from that in MMZP.

**b. Diabatic forcing: Solving omega equation**

To make the numerical solution of (2.12) amenable, a vertical-mode decomposition is used to transform \( \mathcal{L} \) on the left-hand side of (2.12) into a set of two-dimensional modified Helmholtz operators, details of which can be found in MMZP. For WVD, a standard Rossby number expansion is used for \( w \):

\[
 w_b = w_b^{(0)} + \{ \mathcal{R}_0 \} w_b^{(1)} + \{ \mathcal{R}_0^2 \} w_b^{(2)},
\]

where \( w_b^{(0)} \), \( w_b^{(1)} \), and \( w_b^{(2)} \) denote, respectively, the zeroth, first, and second order in \( \mathcal{R}_0 \) contributions to balanced \( w \), that is, \( w_b \).

The solution to (2.12) provides us with a method to estimate diabatic heating as follows. Let us denote the diabatic forcing term in (2.12), that is, \( \{1/\rho_0\} \nabla^2 [\rho \rho R_0 / (1 - \kappa)] / \partial z \), by \( F_d \) and the balanced vertical velocities arising from adiabatic and diabatic processes by, respectively, \( w_{ba} \) and \( w_{bd} = w_b - w_{ba} \). If we know \( F_d \), the diabatic heating \( Q \) can then be computed by

\[
 Q = \frac{1 - \kappa}{\rho R} \int_{z}^{\zeta_{top}} \nabla^2 (\rho F_d) \, dz. \tag{2.14}
\]

From (2.12), by definition, \( F_d \) is related to \( w_{bd} \) through

\[
 F_d = b^{-1} \mathcal{L} w_{bd}. \tag{2.15}
\]

If \( w_b \) can be estimated directly from model results, which will be discussed in section 4, then the omega equation can be used to estimate \( w_{ba} \) and thus \( w_{bd} \) is determined. Let us denote the averaging operator giving us the estimate for large-scale, balanced circulation by \( \mathcal{A} \).

The direct estimate of \( w_b \) can then be taken as \( \mathcal{A} \) \( w_{mod} \), with \( w_{mod} \) being the WRF Model’s actual vertical velocity. In this way, the diabatic forcing becomes

\[
 F_d = b^{-1} \mathcal{L} (\mathcal{A} w_{mod} - w_{ba}). \tag{2.16}
\]

It is worth considering two limits here. First, when there is no contribution from jets and fronts so \( w_{ba} = 0 \), (2.16) attributes the large-scale circulation to diabatic heating. Second, when there is no contribution from actual diabatic sources so \( w_{bd} = 0 \) and \( w_{ba} = w_b \), (2.16) attributes the imbalance in the large-scale circulation, that is \( \mathcal{A} w_{mod} - w_b \), to a diabatic forcing and through (2.14) to an equivalent \( Q \).

From (2.16), the estimate for diabatic forcing depends on the estimate for the adiabatic balanced vertical velocity. Given the optimality of the second-order Rossby number expansion (MMZP), two possible second-order estimation of \( w_{ba} \) are explored here. Writing \( w_{ba} = w_{ba}^{(1)} + w_{ba}^{(2)} \) with \( w_{ba}^{(1)} \) and \( w_{ba}^{(2)} \) respectively, the first and second order in Rossby number contributions to adiabatic balanced vertical velocity, in the first estimation the terms due to compressibility and nonhydrostatic effect that arise from vertical motion are dropped in (2.12) leading to the equations

\[
 \mathcal{L} w_{ba}^{(1)} = b \left\{ \frac{\partial}{\partial z} \left[ \nabla \cdot \left( (V^{(0)}(\xi_{top})) \right) \right] - \frac{1}{\rho_0} \nabla^2 \left[ V^{(0)} \cdot \nabla \frac{\partial p^{(0)}}{\partial z} + V^{(1)} \cdot \nabla \frac{\partial p^{(0)}}{\partial z} \right] \right. \\
- \frac{1}{\rho_0} \nabla^2 \left[ \frac{\partial V^{(0)}}{\partial z} \cdot \nabla p^{(1)} + \frac{\partial V^{(1)}}{\partial z} \cdot \nabla p^{(0)} \right] + \frac{\partial}{\partial z} [-2J(u^{(0)}, v^{(0)})] \right\}, \tag{2.17}
\]

In the second estimation, the terms involving vertical velocity coming from compressibility and nonhydrostatic effects are retained in (2.12) and thus a further contribution to the second-order solution denoted by \( w_{ba}^{(2)} \) is obtained by solving the equation

\[
 \mathcal{L} w_{ba}^{(2)} = b \left\{ \frac{1}{\rho_0} \nabla^2 \left[ \frac{\partial}{\partial z} \left[ \rho_0 \frac{\partial p^{(1)}}{\partial z} \right] \right] \right. \\
- \frac{1}{\rho_0} \nabla^2 \left[ \frac{\partial w^{(1)}}{\partial z} - \frac{\partial p^{(0)}}{\partial z} \right] \right\}, \tag{2.19}
\]

Throughout, the same vorticity inversion procedure of MMZP is carried out. For completeness, a brief account is given here. Considering the leading-order geostrophic relations, \( w_b^{(0)} = 0 \) and \( p^{(0)} \) is computed from \( \nabla^2 p^{(0)} = f \rho_0 \kappa^{(1)} \) in (2.3). Further, \( \kappa^{(1)} = \xi, \xi \) being the actual velocity field. The zeroth-order velocity field \( (u^{(0)}, v^{(0)}) \) is obtained by inverting \( \kappa^{(0)} \) and \( \delta^{(0)} = 0 \) using (2.17) and (2.18) of MMZP, respectively. Inverting (2.17), \( w_{ba}^{(1)} \) is computed, which is followed by the computation of \( \delta^{(1)} \) from (2.10), and thus \( (u^{(1)}, v^{(1)}) \) from inverting.
\( \delta^{(1)}, \xi^{(1)} = 0 \). Likewise, \((u^{(0)}, v^{(0)})\) field is obtained from \( \partial \xi^{(0)} / \partial t = -\nabla \cdot \left[ \nabla (\delta^{(0)}) \right] \) and \( \delta^{(0)} / \partial t = 0 \). The first-order pressure field \( p^{(1)} \) is computed from \( \nabla^2 p^{(1)} = 2 \rho_0 j(u^{(0)}, v^{(0)}) \), which comes from substituting (2.3) into (2.2). For \( w_{b} \), (2.18) is inverted, which is followed by solving \( \delta^{(2)} \) from (2.10) and subsequently, the second-order velocity field \((u^{(2)}, v^{(2)})\) from \((\delta^{(2)}, \xi^{(2)} = 0)\).

c. Diabatic forcing: Scaling relations

As a simpler, less costly way of estimating diabatic heating, a scaling procedure similar to that of Mirzaei et al. (2014) can be adopted. To facilitate scaling analysis, based on (2.12), the following approximate equation for the balanced diabatic vertical velocity can be written in two-dimensional form in the \( x-z \) plane as
\[
N^2 \frac{\partial^2 w_{bd}}{\partial x^2} + f^2 \frac{\partial^2 w_{bd}}{\partial z^2} = \frac{1}{\rho_0} \frac{\partial^2}{\partial z^2} \left[ \frac{\partial (\rho R Q)}{\partial z} \left( 1 - \kappa \right) \right],
\]
(2.20)
in which \( N^2 = (1/\rho_0) \partial \rho / \partial z \) is related to the rate of density decrease with height. Making use of the non-divergence relation \( \partial u_{bd}/\partial x + \partial w_{bd}/\partial z = 0 \) for the circulation induced by diabatic heating, one can then write
\[
N^2 \frac{\partial w_{bd}}{\partial x} - f \frac{\partial u_{bd}}{\partial z} = \frac{1}{\rho_0} \frac{\partial}{\partial z} \left[ \frac{\partial (\rho R Q)}{\partial z} \left( 1 - \kappa \right) \right],
\]
(2.21)
from which the scaling relations for the ageostrophic velocity components can be deduced as
\[
u_{bd} = \frac{1}{f^2 L_h} \frac{QR}{1 - \kappa},
\]
(2.22)
\[
w_{bd} = \frac{1}{N^2 L_z} \frac{QR}{1 - \kappa}.
\]
(2.23)

Here, \( L_h \) and \( L_z \) are, respectively, the horizontal and vertical length scales of the diabatically induced ageostrophic motion. For IGW generation, as in Mirzaei et al. (2014), the threshold \( u_{th} = f L_h/(2\pi) \) is taken, which using the quasigeostrophic scaling \( f^2/N^2 = L_z^2/L_h^2 \) leads to the following threshold for diabatic heating:
\[
Q_{th} = \frac{f N^2 L_z^2 (1 - \kappa)}{2 \pi R}.
\]
(2.24)

Assuming a vertical scale of \( L_z = 3 \) km for the convection-induced circulation, \( \kappa = 0.3 \) and \( R = 287 \) J kg\(^{-1}\) K\(^{-1}\), a threshold of \( Q_{th} = 1.26 \) K h\(^{-1}\) is obtained for diabatic heating. The relation (2.23) can be used to obtain a bulk estimate for \( Q \) provided that \( w_{bd} \) is known. The methods described for \( w_{b} \) in sections 4a and 4b and the solution to the adiabatic form of the omega equation for \( w_{bd} \) can provide the estimate needed for \( w_{bd} \).

For every grid point, the bulk balanced diabatic vertical velocity is computed by volume-averaging \( w_{bd} \) in a rectangle cube with dimensions of 100, 100, and 1 km in \( x, y, \) and \( z \) directions, respectively.

d. Diabatic heating: SC method

The equation introduced by Emanuel et al. (1987) for diabatic heating produced in slantwise convection, called SC here for reference, reads in \( z \) coordinates as
\[
Q = w_b^{(1)} \left\{ \frac{\partial}{\partial z} \left( \frac{\partial \theta}{\partial \theta} \delta \Gamma_m \frac{\partial \theta}{\partial \theta} \right) \right\}^{(0)}_b,
\]
(2.25)

with the term in the curly brackets being computed using the zeroth-order balanced thermodynamical fields. It should be noted that (2.26) is used only in the grid points with relative humidity greater than 99% and \( Q \) is set to zero elsewhere. The same criterion is applied to \( Q \) given by (2.14) and (2.23), but no restriction is made on vertical velocity being positive.

3. WRF simulations

The numerical simulations were carried out using the Advanced Research version of WRF (ARW) (Skamarock et al. 2008), which integrates the fully compressible, non-hydrostatic form of the primitive equations. The simulations were performed in a channel of \( L_x = 4000 \) km length, \( L_y = 10,000 \) km width, and \( L_z = 30 \) km height on an \( f \) plane with boundary conditions of periodic in the \( x \) direction, symmetric in the \( y \) direction, and free slip at the bottom (see MMZP for more details). To prevent the reflection of upward-propagating gravity waves from the top boundary, a sponge layer with \( w \) Rayleigh damping (Klemp et al. 2008) of \( 0.2 \) s\(^{-1}\) coefficient is used in the upper 8 km of the model. To produce a two-dimensional jet suitable for baroclinic waves, the initial conditions are constructed in a manner similar to MMZP (see their Fig. 1a) by inverting an ideal two-dimensional uniform distribution of potential vorticity with 0.4 and 4 PV units (1 PVU = \( 10^{-6} \) K kg\(^{-1}\) m\(^2\) s\(^{-1}\)) values in the troposphere and stratosphere, respectively. The most unstable normal mode is then superposed on the two-dimensional jet.
[see Rotunno et al. (1994) and Plougonven and Snyder (2007) for details]. The model was run with a horizontal resolution of 25 km and a vertical resolution of 250 m for reference run with a triharmonic horizontal hyperdiffusion $\nu_h \nabla^3$, described by Kneivelet et al. (2007), where $\nu_h = 2.5 \times 10^{21} \text{m}^6\text{s}^{-1}$, to filter nonphysical structures at the smallest scales of the flow. The model was also run with a horizontal resolution of 100 km and a vertical resolution of 1000 m. The latter low-resolution run is referred to as LOW. To initialize the model, an ideal profile of relative humidity [see Tan et al. (2004) and MMZP for details] is added to the jet. For the amplitude of the initial moisture, in (5.1) of MMZP, RH$_0$ is set to 0.85. The same setup for humidity, as well as the cumulus and microphysics parameterization schemes, is used in both the reference run, called M85, and the LOW run (see Table 1). It should be emphasized here that the main arguments are valid irrespective of the parameterizations used for convection, microphysics and boundary layer. To show that, in the supplementary material, the corresponding results for the WRF single-moment 6-class microphysics scheme (WSM6), which leads to rather marked differences in diabatic heating compared to the Kessler scheme used here, are provided.

### 4. Results

Central to our analysis are estimates of the balanced vertical velocity $w_b$, and the adiabatic contribution to $w_b$, that is, $w_{ba}$. To begin with, results are presented for $w_{ba}$ using the second-order Rossby number expansion introduced in section 2b. The solutions to the omega equations, (2.17) and (2.18), which exclude the terms due to compressibility, constitute the second-order accurate balanced adiabatic vertical velocity at day 8 and $z = 11$ km (Fig. 1a). Inclusion of the compressibility, that is the contribution of $w^{(2,C)}$ from (2.19) in $w_{ba}$, leads to the solution shown in Fig. 1c. Once $w_{ba}$ in used for WVD, the corresponding approximations to unbalanced vertical velocities ($w_{unb} = w - w_{ba}$) shown in Figs. 1b and 1d contain balanced contributions from $w_{ba}$. Referring to Fig. 1a, a familiar large-scale quadrupole structure can be seen for $w_{ba}$. The marked negative and positive regions of $w_{ba}$ coincide with the entrance and exit regions of the jet stream. The corresponding unbalanced vertical velocity in Fig. 1b exhibits wave structures reminiscent of IGWs. Regarding the results for the inclusion of compressibility, contamination by small-scale noise, particularly in $w_{ba}$ (Fig. 1c), is apparent. Further, compared to the incompressible solutions, a shift to larger values can be noted. Therefore, there is a marked scale effect discussed below, which is the root cause of departure from the optimal WVD by inclusion of compressibility.

The extent to which a set of balanced relation is near to optimal can be assessed by looking into the time evolution of the eddy kinetic energy of the unbalanced solution over the entire domain as defined by

$$K_{unb} = \frac{1}{L_x L_y} \int \left( \frac{\nu_{unb}^2 + \nu_{unb}^2}{2} \right) dx \ dy \ dp$$

where $u_{unb}$ and $v_{unb}$ are the unbalanced horizontal velocity components whose zonal mean values have been extracted. The measure of optimality is determined by the smallness of $K_{unb}$ during the main wave generation events at the peak of baroclinic activity. For reference, the $K_{unb}$ measures for the second-order accurate WVDs introduced here are denoted by $K_{unb}^{NH}$ and $K_{unb}^{NH,C}$ for the exclusion and inclusion of compressibility effects through $w^{(2,C)}$ term, respectively. For comparison, results are also presented for the second-order hydrostatic WVDs introduced in MMZP, which are denoted by $K_{unb}^{NH}$ and $K_{unb}^{NH,C}$ for the exclusion and inclusion of compressibility effects, respectively. Similar measures are defined for the balanced solution including $K_b$ for the global eddy kinetic energy of the balanced flow. For the four sets of WVDs discussed, the time evolution of $K_b$ and $K_{unb}$ over the first 15 days of baroclinic wave is shown in Fig. 2. For $K_b$ as given by each of the WVD procedures, the evolution of the baroclinic wave can be summed up as a rapid increase around day 4 and a double peak between days 6 and 9, followed by a decline in a manner reminiscent of the baroclinic life cycles of Thorncroft et al. (1993). For $K_{unb}$, a noticeable increase is seen for the solution with nonhydrostatic and compressible effects included $K_{unb}^{NH,C}$ (black line) relative to that with compressibility excluded $K_{unb}^{NH}$ (green line). The increase is such that the peak of $K_{unb}$ over the first 15 days of baroclinic life cycle, denoted by $K_{unb}$, for the inclusion of compressibility $K_{unb}^{NH,C}$ is about 1.4 times that of the
incompressible estimate $K_{NH}^{unb}$. A similar increase is also seen in $K_{unb}$ for the hydrostatic WVD, once compressibility effects are taken into account with $K_{NH}^{HC} \approx 1.2 K_{NH}^{HJ}$. Therefore, in agreement with Fig. 1, the $K_{unb}$ measures presented further corroborate the conclusion in MMZP on the adverse effect of compressibility on balanced solutions. As noted by one of the reviewers, this departure would be expected if the scales affected by compressibility did not satisfy balance assumptions. At the same time, $K_{NH}^{unb} \approx 1.9 K_{NH}^{HJ}$, which means that, in tandem with

Fig. 1. (a) The adiabatic part of the balanced vertical velocity $w_{ba}$ and (b) the unbalanced vertical velocity $w_{unb}$ at 11 km, obtained by solving (2.17) and (2.18). (c),(d) As in (a) and (b), respectively, but with the contribution of $w_{unb}^{(2.19)}$ included by solving (2.19) (see the text for more explanation). The unit for $w_{ba}$ and $w_{unb}$ (color shaded) is $10^{-2} \text{ m s}^{-1}$, the thick magenta contour is 3 PVU, and the thick black contours show wind speeds greater than 30, with a 10 m s$^{-1}$ interval. The thick dashed contour in (a) is the isoline of 500-hPa pressure level at $z = 5 \text{ km}$. 
compressibility, consideration of nonhydrostatic effects in WVD leads to further departure from optimal balance. That is, both compressibility and nonhydrostatic effects add considerable scale effects to the WVD procedure. Despite this, the current nonhydrostatic WVD has the advantage of retaining an explicit representation of diabatic heating, which is essential for our purpose.

For a grasp of the structure of baroclinic wave and the associated adiabatic and diabatic sources of unbalanced flow, Fig. 3 is presented at day 8 near to the peak of baroclinic activity based on the results from the WRF’s own fields. The position of the low-level fronts is marked by the strong potential temperature gradients present at \( z = 4 \) km. Two distinct regions of positive values of diabatic heating are seen in Fig. 3a: a crescent shape in the warm sector of the BCW and two nearby compact regions in the jet exit. Signatures of both regions are seen in the mass-weighted vertical average of heating in Fig. 3b. The cross sections in Figs. 3c and 3d plotted through the two main regions of diabatic heating provide further information on the relation between the BCW and unbalanced flow. Two prominent upward- and downward-propagating wave signatures are present in Fig. 3c emanating from a region in which both jet and convective sources are active. The downward-propagating wave is particularly interesting as it is modulated by the presence of convection and an active warm front with the associated upgliding of warm air and diabatic heating. Compared to Fig. 3c, the cross section in Fig. 3d is rather featureless, but there is an active region of diabatic heating far upstream of the jet region and in connection with the crescent-shape heating in Fig. 3a.

The other central information in our analysis comes from an averaging procedure to determine \( w_b \) from model results. Here, two procedures are used for \( w_b \) based on (i) spatial smoothing of the actual vertical velocity of the M85 run and (ii) the actual vertical velocity of the LOW run. It is also worth mentioning here that in presenting the results no restriction to positive values associated with condensational heating is made on \( Q \). In this way, a whole set of complex resolved and unresolved processes of cloud and convection is allowed to contribute to both positive and negative values of \( Q \). Nevertheless, the main comparisons among the methods remain valid, in case the negative values of \( Q \) are filtered.

### a. Omega equation: Smoothing

Spatial smoothing is commonly used as a way of filtering the signature of IGWs and reaching an approximate balanced state (e.g., Zülicke and Peters 2006; Wei and Zhang 2014). Here, the actual vertical velocity of the M85 run is smoothed by a Cressman filter to estimate \( w_b \) (see Fig. 4a). With \( w_b \) and \( w_{\text{ba}} \) known from solving (2.17) and (2.18), the diabatic heating is determined from (2.14). The results have been shown in Fig. 5 for day 8 at the peak of IGW activity (see Fig. 2) during the evolution of the BCW. A peak value of \( 2.4 \) K h\(^{-1}\) is seen for \( Q \) at the exit region of the upper-level jet–front system (Fig. 5a). The warm sector of the BCW in the lower troposphere is covered by a positive heating (warming) region, which is clearly seen in the mass-weighted vertical average of heating presented in Fig. 5b. With a maximum magnitude of \( 1.8 \) K h\(^{-1}\), the
FIG. 3. The diabatic heating $Q$ (contours; interval: 0.2 K h$^{-1}$) as given by the microphysics and convection parameterizations of the WRF Model: (a) at $z = 4$ km and (b) the mass-weighted vertical average. (c),(d) Cross sections marked by, respectively, the solid and dashed lines in (a) for day 8. The purple contours are potential temperature at $z = 0.25$ km, with a contour interval of 10 K; the thick magenta contour is 3 PVU at $z = 8$ km, and the thick black contours indicate wind speed at $z = 8$ km for values greater than 30, with a 20 m s$^{-1}$ interval. Also shown in (a), (c), and (d) are the unbalanced vertical velocity (color shaded), with a unit of 10$^{-2}$ m s$^{-1}$, as obtained by solving (2.17) and (2.18). The thick blue contour in the region around $x \in [1500, 2000]$ km and around $y = 4000$ km in (a) denotes the 0.2 K h$^{-1}$ isoline of the 500-km-averaged diabatic heating of the WRF Model.
heating region in the warm sector is less intense but spatially larger than that of the exit region of the jet stream. The latter maximum coincides well with a similar maximum of 0.7 K h\(^{-1}\) magnitude in the northeast of the warm sector in Fig. 5b. The cooling region west of the warm sector can be related to the cold air descending in the entrance region of the upper-level jet stream, extending to the base of the trough where the maximum of cooling is estimated to be about \(-1.5\) K h\(^{-1}\). Referring to Fig. 5b, a cooling region in the entrance region of the upper-level jet stream and upstream of the tropopause folding is manifested. Comparing Figs. 5a and 5b, despite the quantitative differences, an asymmetric quadrupole pattern for \(Q\) consisting of generally well-defined centers of warming and cooling can be inferred. Overall, while giving spatially broader structures than those seen in Figs. 3a and 3b, the estimates provided for diabatic heating are much richer than the spatially averaged heating of WRF Model’s diabatic heating carried out in Mirzaei et al. (2014). To compare with, the 0.2 K h\(^{-1}\) isoline of Cressman-filtered \(Q\) of WRF with a radius of 250 km, the same filtering operation used for \(w\), has been presented in Fig. 3a with thick blue line.

Plotted in Fig. 5c is the cross section marked by the solid line in Fig. 5a. A positive anomaly of potential vorticity is present, associated with the tropopause folding in the exit region of the jet stream, which is marked by the 3-PVU contour indicating the tropopause level. The

\[\text{tropopause folding and the associated upper-level jet–front system intensify the vertical motion in the lower troposphere with ascending (descending) motion and warming (cooling) appearing downstream (upstream).}\]

The signature of downward-propagating IGWs emanating from the jet region is manifested well in the estimates for \(Q\). This is indeed a good candidate for the amplification or modulation of IGWs by convection (Koch and O’Handley 1997). The maximum warming of about 2.4 K h\(^{-1}\) is observed at the 3–4-km layer in the exit region of the upper-level jet–front system. The other cross section plotted in Fig. 5d is that of the dashed line in Fig. 5a. With the 3-PVU contour stretched down to 3-km level, the tropopause folding in Fig. 5d is stronger than that seen in Fig. 5c. It should be noted that, because of cyclonic Rossby wave breaking, the flow direction in Fig. 5c is opposite that of Fig. 5b. Consequently, the ascending and warming is in the downstream of the tropopause-folding region with the reverse in the upstream.

b. Omega equation: Low resolution

At sufficiently low resolutions, WRF strongly damps the short-scale vertical velocity and thus IGWs. On this basis, the LOW run with horizontal and vertical resolutions of, respectively, 100 and 1 km, is used here to obtain another estimate for balanced vertical velocity (see Fig. 4b). For this purpose, the actual vertical
velocities of the LOW are spatially interpolated to 25-km horizontal resolution and 250-m vertical resolution of the M85 using the cubic Hermite interpolation. The rest of the procedure to estimate $Q$ is the same as that carried out for spatial smoothing in section 4a. Shown in Fig. 6 are the results for the horizontal distribution at $z = 4$ km and the mass-weighted vertical average of $Q$ at day 8 of the life cycle of the baroclinic wave. A qualitative agreement is apparent with the corresponding results for the smoothing presented in Figs. 5a and 5b, which is encouraging. Quantitatively, however, the $Q$ extrema are
quite sensitive to the way balanced vertical velocity is estimated. In Fig. 6a, the maximum warming of about 2.1 K h\(^{-1}\) in the exit region of the upper-level jet stream and 1.2 K h\(^{-1}\) in the warm sector of the BCW are in the same positions but smaller than the corresponding maxima in the smoothing. Moreover, compared to Fig. 5a, the main cooling of \(-1.2\) K h\(^{-1}\) is weaker and shifts slightly to the east. Referring to Fig. 6b, the warming maximum of about 0.3 K h\(^{-1}\) in the northeast of the warm sector of the BCW and the cooling...
maximum of about $-0.4 \text{ K h}^{-1}$ in the entrance region of the upper-level jet stream are similarly weaker than the corresponding maxima of the smoothing.

c. Scaling relations

As computationally less expensive, the scaling relations of section 2c are used here to provide a bulk estimate for $Q$. Except using (2.23), the procedure is exactly the same as that in the smoothing. The main objective is to assess the performance of the scaling relations by examining the way $Q$ is altered when compared with a more complete procedure. This can help to clarify the accuracy of the approximations used in the scaling relations and provide justification for their use in parameterization of diabatically generated IGWs. Referring to Fig. 7a, with no significant change in horizontal scale and position, the maximum warming of about $2.7 \text{ K h}^{-1}$ in the exit region of the upper-level jet stream is stronger than that in the smoothing (Fig. 5a) by about $0.3 \text{ K h}^{-1}$. Moreover, the warming center in the northeast of the warm sector of the BCW (Fig. 7a) is also present but has shrunk and shifted westward compared to the results for the smoothing. It is only in the exit region of the upper-level jet stream that the bulk diabatic heating exceeds the threshold of $1.26 \text{ K h}^{-1}$ obtained in section 2c. Comparing Figs. 5b and 7b, it can be seen that the warming inside and northeast of the warm sector is of smaller scale than the smoothing. The maximum warming and cooling of about $0.2$ and $-0.3 \text{ K h}^{-1}$ for the mass-weighted vertical average of $Q$ in, respectively, the exit and entrance regions of the upper-level jet stream, are also weaker than the corresponding maxima in the smoothing. In general, it can be deduced that the major adverse impact of the scaling relations is on the heating inside and northeast of the warm sector of the BCW.

d. Slantwise convection

The results for $Q$ obtained by the SC method (section 2d) at day 8 of the BCW life cycle presented in Fig. 8 depart radically from the estimate for diabatic heating shown so far. Quantitatively, $Q$ values are an order of magnitude smaller and tend to align with the position of the upper-level wave as shown with the thick dashed contour in Fig. 1a. While the jet exit region is dominant in producing maxima of warming, there is only a fragmentary sign of the crescent-shape warming in the warm sector of the BCW. Overall, the presence of unbalanced processes is essential in shaping the distribution of diabatic heating.

5. Concluding remarks

The main objective of this study was to develop and examine omega-equation-based methods to determine the diabatic forcing required for parameterization of IGWs in the GCMs, which generally are not capable of resolving IGWs and convection cells due to

![Fig. 8](image-url)
the limitation of their spatial scales. This is an important issue in parameterization of the diabatically generated IGWs but so far we have been relying on the diabatic heating provided by the GCM itself in which the physics part is parameterized. The use of the omega equation in our estimates of diabatic heating is generally in line, but goes beyond, the recent works on the diagnosis of jet–front systems by the quasigeostrophic omega equation (e.g., Danioux et al. 2012; Martin 2014; Hien et al. 2018).

To disentangle parameterization of IGWs from the physics parameterization schemes, and in particular that of convection, a suitable form of omega-equation-based WVD procedure was designed. To this end, first the WVD constructed based on the balance relations of the Rossby number expansion in MMZP was extended to obtain a new omega equation that includes an explicit expression of diabatic heating applicable in fully compressible and nonhydrostatic conditions. The WVD thus constructed was applied to an ideal moist simulation of BCW life cycle using the WRF Model data for determining the balanced vertical velocity. The omega equation was first solved excluding diabatic heating forcing to find the adiabatic part of the balanced vertical velocity $w_{ba}$. Based on the peak values of the estimates for global kinetic energies of IGWs obtained by solving compressible and incompressible forms of the omega equation constructed using Rossby number expansion, it was shown that the WVD method under incompressibility assumption is indeed the optimal WVD, confirming the earlier finding of MMZP.

Furthermore, comparing current results with those in MMZP demonstrated that the incompressible WVD built in MMZP is indeed the best method to represent the energy of unbalanced flow or IGWs for two reasons. First, the diabatic heating forcing implicitly appeared in the omega equation cannot be accurately estimated using Emanuel et al. (1987) relation for slantwise convection. Second, the hydrostatic balance assumption filters finescale structures arising from the nonhydrostatic effects, leading to a reduction in the energy of IGWs. Nevertheless, for our purpose here, the compressible and nonhydrostatic form of the omega equation was used for the explicit representation of diabatic heating. The $w_{ba}$ obtained from the omega equation was then applied to determine the diabatic part of the balanced vertical velocity $w_{bd}$. Exploiting spatial-scale separation between IGWs and balanced flows enabled us to obtain estimates of the balanced vertical velocity $w_b$ and thus $w_{bd}$ from (i) smoothing the vertical velocity in M85 and (ii) the vertical velocity of LOW run.

The resulting estimates for $w_{bd}$ provided the two corresponding estimates of $Q$. The two methods employed to estimate $Q$ differ in the way they approximate the balanced vertical velocity. Both methods agreed on the location of maximum heating being in the exit region of the upper-level jet stream associated with upper-level fronts and tropopause folding, and the maximum cooling being in the entrance region of the upper-level jet stream at peak of the wave activity during the evolution of the BCW life cycle. On the other hand, the mass-weighted vertical average of $Q$ as estimated by the two methods differed in terms of magnitudes and occurrence region. The scaling relations were developed in the same manner of Mirzaei et al. (2014) as a way of reducing the computational cost by giving a direct relation between $Q$ and $w_{bd}$. The procedure to determine $w_{bd}$ remains the same as the smoothing method. As the bulk diabatic heating exceeded the threshold of 1.26 K h$^{-1}$, which was indicated by Mirzaei et al. (2014) as the convective forcing in parameterization relation of the IGWs, a heating center was detected in the exit region of the upper-level jet stream that had a similar pattern with but weaker than that of the smoothing method. Overall, as an approximation method, the scaling relations can be useful to determine the diabatic heating by reducing the computational cost especially in parameterizing IGWs in the GCMs. For the given dataset, the working of the scaling relation is also a sign for the dominance of hydrostatic and incompressible dynamics. Comparison was also made with results when $Q$ is given solely by balanced processes through slantwise convection as in the procedure introduced by Emanuel et al. (1987). The slantwise convection led to substantially different patterns of cooling and heating, which fall sharply smaller than the threshold diabatic heating.

Finally, it should be remarked that our (2.16) for diabatic forcing has the potential to unify the jet–front–convection sources of IGWs, as it also produces an equivalent adiabatic forcing arising from imbalances due to jets and fronts. To explore this possibility and the extension to spherical geometry warrant further studies, which we plan to do in future research.

Acknowledgments. We thank the University of Tehran and IAP for providing support during this research. The studies of CZ were partly funded by the Deutsche Forschungsgemeinschaft through Grant ZU 120/2-2 for the Research Unit FOR 1898 (Multiscale dynamics of gravity waves/spontaneous imbalance). We thank two anonymous reviewers for constructive criticism and helpful suggestions that have helped to improve the manuscript.


