NOTES AND CORRESPONDENCE

The Energetically Consistent Shallow-Water Equations

PETER R. GENT

National Center for Atmospheric Research, Boulder, Colorado

18 February 1992 and 26 June 1992

ABSTRACT

It is shown that the very frequently used form of the viscous, diabatic shallow-water equations are energetically inconsistent compared to the primitive equations. An energetically consistent form of the shallow-water equations is then given and justified in terms of isopycnal coordinates. Examples are given of the energetically inconsistent shallow-water equations used in low-order dynamical systems and simplified coupled models of tropical air–sea interaction and the El Niño–Southern Oscillation phenomena.

1. Introduction

The shallow-water equations (SWE) are used very frequently as an analogue of atmospheric and oceanic flows governed by the primitive equations (PRE). The reason is that the SWE contain several phenomena, such as gravity waves, that are present in the PRE but are simplified compared to the PRE in that they are two-dimensional, not three-dimensional. Often ideas and techniques are tried on the SWE before they are applied to the PRE. Good examples are nonlinear normal-mode initialization, low-order models and dynamical systems, and simplified coupled atmosphere–ocean models.

It is shown that the very frequently used form of the viscous, diabatic (i.e., mass source or sink) SWE are energetically inconsistent compared to the PRE. The inconsistencies are that the viscous term is not negative definite in the kinetic energy (KE) budget, and the diabatic term affects both the KE and potential energy (PE) budgets, whereas in the PRE it only affects the PE budget. The energy budgets in the PRE are reviewed briefly in section 2. In section 3 an energetically consistent form of the SWE is given and is justified by an analogy to the PRE transformed into isopycnal coordinates. Section 4 contains some examples of the use of the energetically inconsistent SWE in low-order dynamical systems and in simplified coupled models of tropical air–sea interaction and the El Niño–Southern Oscillation (ENSO) phenomena.

2. The primitive equations

The hydrostatic, Boussinesq primitive equations in physical coordinates for large-scale atmosphere and ocean circulations are frequently written as

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f}\times\mathbf{u} + \frac{1}{\rho_0}\nabla p = \mathbf{F} + \nu\nabla^2\mathbf{u}, \quad (1)$$

$$p_z - g\rho = 0, \quad (2)$$

$$\nabla\cdot\mathbf{u} + w_z = 0, \quad (3)$$

$$\frac{D\rho}{Dt} = Q, \quad (4)$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla + w\frac{\partial}{\partial z}. \quad (5)$$

For the ocean z, p, and ρ represent depth measured downwards, pressure, and density, respectively. For the atmosphere they represent a modified pressure coordinate, geopotential, and potential temperature, respectively [see Hoskins and Bretherton (1972)]. A Laplacian form for momentum dissipation has been assumed, and the momentum forcing term, $\mathbf{F}$, and the diabatic term in the density equation, Q, are left unspecified. These equations have a KE density of $\frac{1}{2}\mathbf{u}\cdot\mathbf{u}$ and a PE density of $-gz\rho/\rho_0$, and Eqs. (1) and (4) give

$$\frac{D}{Dt}\left(\frac{1}{2}\mathbf{u}\cdot\mathbf{u}\right) + \mathbf{u}\cdot\nabla p/\rho_0 = \mathbf{u}\cdot\mathbf{F} + \nu\mathbf{u}\cdot\nabla^2\mathbf{u}, \quad (6)$$

$$\frac{D}{Dt}(-gz\rho/\rho_0) = -gw\rho/\rho_0 - gzQ/\rho_0. \quad (7)$$

Integrating over a rectangular volume using the kinematic boundary conditions of $\mathbf{u}\cdot\mathbf{n} = 0$ at horizontal boundaries and $w = 0$ at vertical boundaries gives

$$\frac{d}{dt}\mathrm{KE} = \frac{g}{\rho_0}\langle\rho w\rangle + \langle\mathbf{u}\cdot\mathbf{F}\rangle + \nu\langle\mathbf{u}\cdot\nabla^2\mathbf{u}\rangle, \quad (8)$$

$$\frac{d}{dt}\mathrm{PE} = -\frac{g}{\rho_0}\langle\rho w\rangle - \frac{g}{\rho_0}\langle zQ\rangle, \quad (9)$$

Corresponding author address: Dr. Peter R. Gent, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000.

where angle brackets denote a volume integral. Imposing a viscous momentum boundary condition of no slip ($\mathbf{u}\cdot\mathbf{s} = 0$) or free slip [$\partial(\mathbf{u}\cdot\mathbf{s})/\partial n = 0$] gives

$$\frac{d}{dt}\mathrm{KE} = \frac{g}{\rho_0}\langle\rho w\rangle + \langle\mathbf{u}\cdot\mathbf{F}\rangle - \nu\langle\nabla\mathbf{u} : \nabla\mathbf{u}\rangle. \quad (10)$$

Equations (9) and (10) explicitly show three well-known and desirable properties of the energy forms. They are

(i) The conversion terms in the two equations are identical, so that the sum of KE and PE is conserved in adiabatic, inviscid flow.

(ii) The viscous term in the momentum equation only affects KE and the dissipation term is negative definite in the KE budget.

(iii) The diabatic term only affects PE.

It might be asked why the properties listed above only involve energy and not other quantities of interest to the PRE, such as momentum or potential vorticity. The momentum diffusion term in (1) is in the form of the divergence of a stress tensor, which should be retained in the SWE, see Schär and Smith (1993). However, the momentum budget is influenced by the Coriolis term, the pressure gradient term as well as the nonconservative term in (1), so it is less easy to interpret than the energy budget, which is not influenced by the Coriolis or pressure gradient terms. The potential vorticity conservation from (1)–(5) is influenced by the nonconservative terms in both (1) and (4). However, it is a Lagrangian conservation principle and its pointwise nonconservative terms will be difficult to mimic exactly in the SWE.

3. The shallow-water equations

The SWE are frequently used in the form

$$\frac{\partial\mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \mathbf{f}\times\mathbf{u} + g\nabla h = \mathbf{F}/h + \nu\nabla^2\mathbf{u}, \quad (11)$$

$$\frac{\partial h}{\partial t} + \nabla\cdot(h\mathbf{u}) = Q, \quad (12)$$

see, for example, Lorenz (1980). However, there are numerous other papers that could be cited, and some will be discussed in the next section. In (11) and (12) h is the fluid depth and Q is a mass source or sink that is analogous to the diabatic term in the PRE. The KE density of the SWE is $\frac{1}{2}h\mathbf{u}\cdot\mathbf{u}$, so that the KE equation is formed by adding the scalar product of $h\mathbf{u}$ and Eq. (11) to $\frac{1}{2}\mathbf{u}\cdot\mathbf{u}$ times Eq. (12). It is immediately apparent that, using either no-slip or free-slip boundary conditions to eliminate the boundary term, the dissipation term in the KE budget cannot be written in a form which ensures that it is negative definite. This is inconsistent with property (ii) of the PRE discussed above.

Gustafsson and Sundstrom (1978) used the momentum dissipation form

$$\nu\nabla\cdot(h\nabla\mathbf{u})/h, \quad (13)$$

and this is also mentioned as a possible form by Bernard and Pironneau (1991). These papers are concerned with the existence of solutions to the SWE using the energy method. The form (13) has several justifications. The term inside the brackets is a parameterization of the turbulent momentum flux and so should be proportional to the fluid depth. Equation (13) ensures that the diffusion of momentum is in the form of the divergence of a stress tensor as suggested by Schär and Smith (1993). Finally the form (13) has a negative definite effect on the KE budget if the viscous boundary condition is either no slip or free slip. However, the form (13) has been criticized by Schär and Smith (1993) because it does not have the same symmetries as the dissipation term in the PRE in (1). They propose an alternative, somewhat more complicated, form than (13) that has the symmetries and still has the form of divergence of a stress tensor for momentum and a negative definite effect on the KE budget. Note that the dissipation form (13) is the analogue of that used by Bleck and Boudra (1981) for the PRE in isopycnal coordinates where the thickness of the isopycnal layers replaces h. In fact, the inviscid SWE can be derived as a one-mode model from the inviscid PRE in isopycnal coordinates.

In addition, it is also immediately apparent from forming the SWE KE equation from (11) and (12) that the diabatic term, Q, will affect the KE budget, which is inconsistent with property (iii) of the PRE discussed above. Again the resolution of this inconsistency comes from the PRE in isopycnal coordinates. Transforming the diabatic PRE given in (1)–(5) to isopycnal coordinates will result in a substantial derivative of the form

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla + Q\frac{\partial}{\partial\rho}, \quad (14)$$

where the differentiation by t, x, and y is now with respect to constant ρ surfaces. Thus, the momentum equation will have a term $Q\mathbf{u}_\rho$ that combines with the Q term from the thickness equation so that Q does not affect the momentum or KE budgets. This suggests that a term proportional to Q should be added to the SWE velocity equation (11), which represents the changes to momentum or KE as mass is added to the system.

It turns out that for the SWE, a single term cannot be found such that Q does not affect both momentum and KE. If $\mathbf{u}Q/h$ is added to the left-hand side of (11) then Q does not affect momentum, whereas $\mathbf{u}Q/2h$ needs to be added so that Q does not affect KE. Thus, a choice has to be made whether the mass source or sink should not affect either momentum or KE. If Q is meant to represent a diabatic term in analogy to the PRE, then for the reasons discussed at the end of section 2, the choice should probably be not to affect the KE. The form of the SWE that is energetically consistent with properties (i)–(iii) of the PRE listed above is

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \mathbf{f}\times\mathbf{u} + g\nabla h + \mathbf{u}Q/2h = \mathbf{F}/h + \nu\nabla\cdot(h\nabla\mathbf{u})/h, \quad (15)$$

$$\frac{\partial h}{\partial t} + \nabla\cdot(h\mathbf{u}) = Q. \quad (16)$$

The KE and PE densities associated with (15) and (16) are $\frac{1}{2}h\mathbf{u}\cdot\mathbf{u}$ and $\frac{1}{2}gh^2$, and the equations for the energy budgets are

$$\frac{d}{dt}\mathrm{KE} = -g\langle h\mathbf{u}\cdot\nabla h\rangle + \langle \mathbf{u}\cdot\mathbf{F}\rangle - \nu\langle h\nabla\mathbf{u} : \nabla\mathbf{u}\rangle, \quad (17)$$

$$\frac{d}{dt}\mathrm{PE} = g\langle h\mathbf{u}\cdot\nabla h\rangle + g\langle hQ\rangle, \quad (18)$$

where angle brackets now indicate a horizontal integral. These should be compared to the PRE energy budgets given in (9) and (10).

4. Discussion and conclusions

The form of the SWE given in (11) and (12) can be traced back to the seminal paper by Lorenz (1980). In Lorenz' equations **F** is zero and the forcing is in the diabatic, or mass source or sink, term Q. Lorenz used a low-order truncation of the SWE that gave a set of nine ordinary differential equations, which contain strange attractors, limit cycles, period doubling bifurcations, and other interesting phenomena. The Lorenz paper has inspired many further papers, for example, Gent and McWilliams (1982) and Curry and Winsand (1986). Most, if not all, of these subsequent papers have used the same form of the SWE. Following the argument given in section 3, all these papers use a form of the SWE inconsistent with properties (ii) and (iii) of the PRE given in section 2.

The consistent form of the SWE, (15) and (16), can be written as prognostic equations for the momentum, $h\mathbf{u}$ and the PE density, $P = \frac{1}{2}gh^2$. They are

$$\frac{\partial}{\partial t}(h\mathbf{u}) + \nabla\cdot(\mathbf{u}h\mathbf{u}) + \mathbf{f}\times h\mathbf{u} + \nabla P - \mathbf{u}Q/2 = \mathbf{F} + \nu\nabla\cdot(h\nabla\mathbf{u}), \quad (19)$$

$$\frac{\partial P}{\partial t} + \nabla\cdot(\mathbf{u}P) + P\nabla\cdot\mathbf{u} = ghQ. \quad (20)$$

These equations can be compared to those of Anderson and McCreary (1985). Setting the temperature to be constant in the Anderson and McCreary equations yields the SWE, but the dissipation term is of a different form, and the term involving Q is omitted, compared to (19). Thus, the equations of Anderson and McCreary (1985) have Q not affecting momentum but are also inconsistent with properties (ii) and (iii) of the PRE given in section 2. Again these equations have been used by several subsequent investigators, for example, Budin and Davey (1990) and Masumoto and Yamagata (1990, 1991), as the ocean component of simplified coupled models of tropical air–sea interaction and ENSO.

The SWE are only an analogue of the PRE. Nevertheless, they should be as consistent as possible. The KE and PE budgets in the two equation sets can be made consistent by using the form of the SWE presented in (15) and (16). The form of the dissipation term in (15) is not new. What is new is the proposal that if there is a diabatic, or a mass source or sink, term Q in (16), then there should be a corresponding term in the momentum equation (15) so that Q does not influence the KE budget.

Acknowledgments. Thanks to Linda Sundbye for inspiration and Joe Tribbia for information. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

REFERENCES

Anderson, D. L. T., and J. P. McCreary, 1985: Slowly propagating disturbances in a coupled ocean–atmosphere model. *J. Atmos. Sci.*, **42**, 615–629.
Bernard, C., and O. Pironneau, 1991: On the shallow-water equations at low Reynolds number. *Commun. Partial Differential Equations*, **16**, 59–104.
Bleck, R., and D. B. Boudra, 1981: Initial testing of a numerical ocean circulation model using a hybrid (quasi-isopycnic) vertical coordinate. *J. Phys. Oceanogr.*, **11**, 755–770.
Budin, G. R., and M. K. Davey, 1990: Intermediate models of intraseasonal and interannual waves in the tropics. *J. Mar. Syst.*, **1**, 39–50.
Curry, J. H., and D. Winsand, 1986: Low-order intermediate models: Bifurcation, recurrence, and solvability. *J. Atmos. Sci.*, **43**, 2360–2373.
Gent, P. R., and J. C. McWilliams, 1982: Intermediate model solutions to the Lorenz equations: Strange attractors and other phenomena. *J. Atmos. Sci.*, **39**, 3–13.
Gustafsson, B., and A. Sundstrom, 1978: Incompletely parabolic problems in fluid dynamics. *SIAM J. Appl. Math.*, **35**, 343–357.
Hoskins, B. J., and F. P. Bretherton, 1972: Atmospheric frontogenesis models: Mathematical formulation and solution. *J. Atmos. Sci.*, **29**, 11–37.
Lorenz, E. N., 1980: Attractor sets and quasigeostrophic equilibrium. *J. Atmos. Sci.*, **37**, 1685–1699.
Masumoto, Y., and T. Yamagata, 1990: The birth and evolution of an eastward-propagating air–sea coupled disturbance in an aquaplanet. *Meteor. Atmos. Phys.*, **44**, 1–9.
——, and ——, 1991: On the origin of a model ENSO in the western Pacific. *J. Meteor. Soc. Japan*, **69**, 197–207.
Schär, C., and R. B. Smith, 1993: Shallow-water flow past isolated topography. Part I: Vorticity production and wake formation. *J. Atmos. Sci.*, **50, 1373–1400.**