Tellus (1984), *36A*, 166–171

Balanced models in isentropic coordinates and the shallow water equations

By PETER R. GENT and JAMES C. McWILLIAMS, *National Center for Atmospheric Research*[1], *Boulder, Colorado 80307, U.S.A.*

(Manuscript received March 30, in final form July 21, 1983)

ABSTRACT

It is shown that there is an appropriate family of three balanced models in isentropic coordinates and the shallow water equations. "Appropriate" means the equations include exact hydrostatic, continuity and heat equations and form a complete system; some family members are second-order accurate in a midlatitude Rossby number expansion; the models conserve at least one integral invariant analogous to the primitive equation energy and enstrophy invariants; and the terms retained in the vorticity and divergence equations are the minimum required for the accuracy of the model in an asymptotic scale analysis. We compare vorticity and divergence in physical and isentropic coordinates to illustrate an aspect of the difference in physical content between balanced models defined in the different coordinates. We show that the early balanced models of Bolin and Charney are similar to members of our family in isentropic coordinates, but fail some of our criteria for appropriateness. Finally, we discuss why we think the above criteria are the right ones for balanced model families, and compare and contrast similar criteria for other intermediate models.

1. Introduction

In a pair of companion papers, Gent and McWilliams (1983a, b) we have examined a family of balanced models in physical coordinates (PC) that conserve an energy, but not an enstrophy, integral invariant. Consistent initial, boundary value problems are examined in the first paper, and regimes of validity for balanced models in the second. The family of balanced models is characterized by the degree of accuracy retained in the Coriolis terms, and whether the truncated horizontal divergence equation is linear or nonlinear.

The primary purpose of this paper is to show that there is an alternative family of balanced models defined in isentropic coordinates (IC) or the shallow water equations (SWE) in forms analogous to the PC family members. In Section 2, we show that three members of the family of balanced models are "appropriate" in IC because the equations include exact hydrostatic, continuity and heat equations and form a complete system; some family members are second-order accurate in a midlatitude Rossby number expansion; the models conserve at least one integral invariant analogous to the primitive equation energy and enstrophy invariants; and the terms retained in the vorticity and divergence equations are the minimum required for the asymptotic accuracy of the model (see the scaling arguments in Gent and McWilliams, 1983b). The integral invariant in IC is the exact enstrophy, rather than an energy as in PC, and it is retained even with topography. The analogues of the remaining members of the PC family have no integral invariants in IC, and we term them "inappropriate" (see Section 4). The physical content of analogous balanced models is different in PC and IC in part because of the different meaning of vorticity and divergence, and this is also discussed briefly in Section 2.

A secondary purpose of this paper is to clarify the work in two early papers on balanced models, which incorporated ideas related to IC and potential vorticity conservation. In Section 3, we discuss Bolin's (1956) model in IC and Charney's (1962)

[1] The National Center for Atmospheric Research is sponsored by the National Science Foundation.

model in PC, but show that although they are similar to our models, neither is appropriate by the criteria listed above.

In Section 4, we discuss the rationale behind the appropriateness criteria listed above for families of balanced models. We then compare and contrast similar criteria for other intermediate models.

2. Balanced models in IC and the SWE

We use the notation of McWilliams and Gent (1980), and use capital letters for quantities in IC. IC are defined by

$$T = t, \quad X = x, \quad Y = y, \quad Z = \theta, \tag{1}$$

where θ is the potential temperature for the atmosphere and minus potential density for the ocean. The inviscid, adiabatic and Boussinesq momentum primitive equation (PE) in IC is

$$\textit{PE-IC} \quad \frac{\mathrm{D}\mathbf{u}}{\mathrm{D}T} + f\mathbf{k} \times \mathbf{u} + \nabla\Phi = 0, \tag{2}$$

where

$$\frac{\mathrm{D}}{\mathrm{D}T} = \frac{\partial}{\partial T} + u\frac{\partial}{\partial X} + v\frac{\partial}{\partial Y}. \tag{3}$$

$\mathbf{u}$ is horizontal velocity, f the Coriolis frequency and Φ the Montgomery potential. The vorticity equation derivable from eq (2) is

$$\textit{PE, BE-IC} \quad \frac{\mathrm{D}}{\mathrm{D}T}(f + \nabla^2\Psi) + (f + \nabla^2\Psi)\nabla^2\Sigma = 0, \tag{4}$$

where

$$\mathbf{u} = \mathbf{k} \times \nabla\Psi + \nabla\Sigma. \tag{5}$$

Ψ and Σ are the stream function and velocity potential in IC, and the spatial differential operators are purely horizontal. In the balance equations in PC (BE-PC), only one term is deleted from the exact vorticity equation, and it arises from the vertical advection of the divergent component of velocity. Vertical advection is implicit in IC for adiabatic flows, so that the natural vorticity equation for BE-IC is the exact eq. (4). In IC the exact continuity and heat equations combine to give

$$\textit{PE, BE-IC} \quad \frac{\mathrm{D}h_Z}{\mathrm{D}T} + h_Z\nabla^2\Sigma = 0, \tag{6}$$

where $z = h(X, Y, Z, T)$ is the "height" of a θ surface. The vertical coordinate, z, in PC is the modified pressure coordinate first used in Batchelor (1953) and in Hoskins and Bretherton (1972), which is proportional to the Exner function, for the atmosphere, and is height for the ocean. h can be calculated from the hydrostatic equation, which in IC, is

$$\textit{PE, BE-IC} \quad h = -\frac{\theta_0}{g}\Phi_Z, \tag{7}$$

where g is gravity and θ_0 is a reference potential temperature or density. All balanced models in our family retain the exact heat, continuity and hydrostatic equations, so that (6) and (7) are also equations of BE-IC. From (4) and (6), it is trivial to show that the primitive equation potential vorticity, Q, is conserved pointwise in BE-IC; i.e.,

$$\frac{\mathrm{D}Q}{\mathrm{D}T} = 0, \tag{8}$$

where

$$Q = \frac{g}{\theta_0 f_0^2}(f + \nabla^2\Psi)h_Z^{-1}, \tag{9}$$

and f_0 is a reference value of f. It is then also easy to show that total enstrophy is conserved because

$$\text{enstrophy} = \iint_\Omega \int_{Z_s}^{Z_u} h_Z Q^2 \,\mathrm{d}X\,\mathrm{d}Y\,\mathrm{d}Z, \tag{10}$$

is independent of time in either a bounded or periodic domain. Here Ω is the horizontal domain, Z_u and Z_s the upper and lower boundaries, and this result remains true for flow over topography. The remaining equation of the BE-IC is the truncated divergence equation, and it is natural to use the same form of the equation as in BE-PC. Thus,

$$\textit{BE-IC} \quad \nabla^2\Phi = \nabla\cdot(f\nabla\Psi) + 2J(\Psi_X, \Psi_Y), \tag{11}$$

where J is the horizontal Jacobian operator in IC. It can be shown that the BE-IC equations, (4), (6), (7), and (11), do not conserve the integral of the obvious form of energy density with a kinetic energy density of $\frac{1}{2}\nabla\Psi\cdot\nabla\Psi$, and we have found no modified form of energy conservation at all for this model. The BE-IC model retains the minimum number of terms compatible with ensuring second-order accuracy in the usual midlatitude expansion in small Rossby number (see Gent and McWilliams, 1983b).

We can also define the global balance equations (gBE-IC), which consist of eqs. (4), (6) and (7), but the divergence equation is truncated to an analogue of the gBE-PC:

$$gBE\text{-}IC \quad \nabla^2 \Phi = \nabla \cdot (f \nabla \Psi) + 2J(\Psi_X, \Psi_Y) + J(f, \Sigma). \tag{12}$$

Thus gBE-IC treats all the Coriolis terms exactly, and hence is valid in very large domains (Gent and McWilliams, 1983b). We can also define the f-plane balance equations (fBE-IC), which consist of eqs. (6) and (7) and (4) and (11) with f set to f_0. These two models retain the pointwise and integral conservation of the full and f-plane PE potential vorticity, respectively.

One more nonlinear balanced model, called βBE, and four linear balanced models are included by Gent and McWilliams (1983a, b) in the PC balanced model family, because they all conserve the same integral energy. Analogues of these models in IC would truncate the exact vorticity equation (4) in a way which would not conserve any form of potential vorticity. We call these models inappropriate in IC because they do not conserve either a global energy or enstrophy, whereas all other balanced models we advocate conserve one or the other (see Section 4).

A possible method of solution for the BE-IC models is to form an equivalent of the omega equation for Σ from eqs. (4), (6), (7), and (11) or (12). This equation would form part of an iteration cycle that would also include solving eqs. (4), (7), and (11) or (12) for Ψ, h and Φ, respectively. This method is almost the same as that we discussed for PC in Gent and McWilliams (1983a). Also discussed in the appendix of that paper are the different choices for horizontal boundary conditions in a finite domain; similar choices would also have to be made in IC. Because these issues are so similar in PC and IC, we do not discuss them further.

Analogous models in PC and IC have different physical content, not only because of differences in model accuracy at order Rossby number squared, but also because of the different meanings of fundamental quantities in the two coordinate systems. For example, vorticity and divergence in the two coordinate systems are related as follows:

$$\begin{aligned} \nabla^2 \Psi &= \nabla^2 \psi + \nabla\theta \cdot \mathbf{k} \times \mathbf{u}_z \theta_z^{-1}, \\ &= \nabla^2 \psi - [\nabla\theta \cdot \nabla\psi_z + J(\theta, \chi_z)]\, \theta_z^{-1}, \end{aligned} \tag{13}$$

$$\begin{aligned} \nabla^2 \Sigma &= \nabla^2 \chi - \nabla\theta \cdot \mathbf{u}_z \theta_z^{-1}, \\ &= \nabla^2 \chi - [\nabla\theta \cdot \nabla\chi_z - J(\theta, \psi_z)]\, \theta_z^{-1}, \end{aligned} \tag{14}$$

where left-hand side derivatives are in IC, and those on the right-hand side are in PC. ψ and χ are the stream function and velocity potential in PC. Thus there may be circumstances where the geometry of the flow is such that the Laplacian of the stream function (or of the velocity potential) is markedly different in the two systems. For example, differences in divergence would involve strongly sloping θ surfaces with highly curved flow, so that θ and ψ_z are sufficiently distinct.

All the appropriate IC models can also be defined in the context of the SWE when eq. (7) drops out and h_Z in eq. (6) and Φ in eq. (11) or (12) are both replaced by the total depth of the fluid. Like their counterparts in IC, these shallow water balanced models conserve the exact potential vorticity, which was originally shown by Charney (1962), and do not conserve an energy, which was shown by Gent and McWilliams (1982). This consistency between balanced models in IC and the SWE is not surprising since a one-mode PE model in IC yields the SWE. This follows from separating the equations by linearizing about a state of rest using the same vertical structure functions as in PC, but now assumed to be functions of Z rather than z. A projection of the PE in IC onto a single Z mode yields the exact SWE.

3. The balanced models of Bolin and Charney

The first person to consider a balanced model in IC was Bolin (1956). Bolin's equations are the exact continuity and hydrostatic equations (6) and (7), a truncated divergence equation that is somewhat more complicated than our gBE-IC eq. (12), and the exact equation for DG/DT, the substantial derivative of the right-hand side of (15) below. The last equation can be derived from the primitive momentum and vorticity equations (2) and (4). Thus Bolin's model differs from ours in the absence of an explicit vorticity equation, and hence does not have an enstrophy integral invariant. For varying f, Bolin's truncation of the divergence equation is

$$\begin{aligned} Bolin \quad \nabla^2 \Phi + \tfrac{1}{2} f^2 &= \tfrac{1}{2}[(f + \nabla^2 \Psi)^2 - A^2 - B^2] \\ &+ \nabla f \cdot \nabla \Psi + J(f, \Sigma) \equiv G, \end{aligned} \tag{15}$$

where

$$A = u_X - v_Y, \quad B = u_Y + v_X. \tag{16}$$

This equation reduces to eq. (12) if only the Ψ components of A and B are substituted into (15). Thus, (15) retains more than the minimum number of terms required for second-order accuracy in a Rossby number expansion. For these reasons, we have not included it as an appropriate balanced model in IC.

As we already mentioned, Charney (1962) showed that the SW-BE conserve the exact potential vorticity. Guided by this, Charney proposed a balanced model in PC where the usual BE-PC truncated vorticity equation is replaced by the equation of conservation of exact PC potential vorticity. This means that Charney's (1962) model has an enstrophy but not an energy integral, and thus has different physics than any of the PC family of balanced models discussed in Gent and McWilliams (1983a, b). If Charney's equations are transformed into IC, then eqs. (6), (7) and (8) are obtained, from which eq. (4) can be deduced. However, if Charney's BE-PC truncated divergence equation is transformed to IC, it contains all the terms in eq. (11) plus many more due to the transformation. Thus Charney's (1962) model in IC is very close to our BE-IC model, but is not minimal because more terms are retained than is necessary to ensure second-order accuracy in the Rossby number.

Bleck (1973, 1974) has solved models in IC that are more complicated than quasigeostrophy, but neither retain both the exact heat and continuity equations, nor are accurate at second-order in Rossby number, and hence are not close to our balanced model family.

4. The rationale for selecting appropriate intermediate models

In the introduction, we list four criteria that we have applied in selecting families of balanced models in PC and IC. In this section, we discuss why they, or their close counterparts, are likely to be of value in selecting intermediate models (i.e., models intermediate in physical accuracy between quasigeostrophy and the PE).

The first criterion is that they include the exact hydrostatic, continuity, and heat equations. This is important because many large-scale geophysical phenomena have their largest departures from quasigeostrophy in the heat equation (i.e., the Burger number is small; see Gent and McWilliams, 1983b). This allows the thermodynamics to be less constrained than in quasigeostrophy (e.g., the static stability can vary horizontally and evolve in time) and integrals of all powers of the temperature to be adiabatically invariant. Since the criterion of second-order accuracy (see below) implicates all terms in these equations, and no satisfactory form of them, less than exact but more than quasigeostrophic, has been discovered to date, we have made our criterion the exact equations.

The second criterion is that at least some family members have second-order accuracy in a mid-latitude Rossby number expansion. This requires, in particular, that the divergence equation represents gradient-wind (or cyclostrophic) balance as an improvement over geostrophic balance. Gradient-wind balance is the essential physical approximation which originally motivated interest in balanced models.

The third criterion is that the models have at least one integral invariant analogous to the energy and enstrophy invariants of the PE. Both of these invariants are positive definite functionals of their constituent flow quantities, and as such serve to bound these functionals of the solutions in time. In the case of the balanced model families in PC and IC, only one such invariant is available. The precise value of integral invariants remains unknown, but since our initial discussion of this issue (Section 8d, McWilliams and Gent, 1980), we have come to believe that their rôle in bounding solutions is an important one. In the two intermediate model intercomparisons we have made (the steady circular vortex (Appendix D, McWilliams and Gent, 1980) and the Lorenz triad model (Gent and McWilliams, 1982)) models with integral invariants have performed much better at large Rossby number than those without. Specifically, one intermediate model we proposed in McWilliams and Gent (1980), hypogeostrophy, which is an expansion of the PE valid through two orders in Rossby number, has gross errors at even intermediate Rossby numbers in both comparisons. We do not now advocate hypogeostrophy as an appropriate intermediate model. We prefer truncations of the PE which preserve at least one integral invariant, because expansions beyond leading order in some small parameter only have this property asymptotically.

The fourth criterion for families of balanced models is that they are minimal in terms retained in the truncations of the vorticity and divergence equations, consistent with the preservation of the integral invariant and the achievement of a given order of model accuracy. In Gent and McWilliams (1983b), regimes of validity were analyzed in terms of three parameters (Rossby number, Burger number (a measure of the strength of stratification), and a measure of the magnitude of variations in the Coriolis frequency) and, for many regimes, members of the balanced family were found to have second-order accuracy. The minimality criterion is primarily one to select the simplest model among other possible models which satisfy the other criteria. If extra terms do not demonstrably add to the virtues of a candidate model, then we feel they should be discarded.

The preceding criteria have been developed in the course of our studies of balanced models. However, there is a close analogy between the above criteria and ones which would select the semigeostrophic model of Hoskins (1975). There, the fundamental approximation is the geostrophic momentum approximation, rather than the gradient-wind approximation which formed the basis of our second balanced-model criterion. Semigeostrophy has second-order accuracy in frontal flow configurations with small curvature, although it has only leading-order accuracy in a Rossby number expansion (McWilliams and Gent, 1980). Semigeostrophy is based upon exact hydrostatic, continuity, and heat equations (criterion 1), and it has both energy and enstrophy invariants on the f-plane, but only an energy invariant for variable Coriolis frequency (criterion 3). Finally, semigeostrophy retains the minimum number of terms in the momentum equation compatible with its frontal geometry accuracy and integral invariant properties listed above (criterion 4).

A particular property of semigeostrophy is that the geostrophic momentum approximation entails analogous truncations of the PE in PC and IC, so that the physical content of the model is the same in the two coordinate systems. In contrast, analogous truncations of the PE in PC and IC for the balanced model families (and quasigeostrophy) result in models with different physical content, as discussed above.

A different, but similarly motivated, approach has been taken by Salmon (1983). He proposes a truncation of the Hamiltonian for the SWE, replacing total velocity by geostrophic velocity in the kinetic energy density. This has the effect of eliminating some terms in the momentum equation, compared to PE, but some new and unfamiliar terms are introduced. The model equations contain all the terms in semigeostrophy, plus others, but the asymptotic accuracy of his model is the same as that of semigeostrophy. The advantage he obtains over all other existing intermediate models, and semigeostrophy in particular, is two integral invariants for variable Coriolis frequency. Thus, Salmon's criteria can also be described as analogous to those listed above. His emphasis is different, however; to gain the second invariant (criterion 3), he accepts a weaker minimality criterion 4. Solutions are requird to assess the relative merits of the two models.

Thus, there are very similar criteria for selecting all the intermediate models that we now advocate.

5. Discussion

The fBE, BE and gBE models discussed in this paper have different physics depending upon whether they are defined in PC or IC. In the absence of explicit solutions, it is very difficult to say which coordinate system has physically better balanced models. The reason is that in an asymptotic analysis both models have formally equivalent second-order accuracy in a small Rossby number expansion, although Shapiro (1970) suggests that the equivalent Rossby number in IC will often be smaller than that in PC near the jet stream. The same comments apply to quasigeostrophy in PC and IC, but only to first-order in Rossby number, and any search for superior accuracy should probably begin with the quasigeostrophic models. Almost certainly, the question of superiority will depend upon the physical phenomenon being modeled.

A circumstance where IC might be preferable is if the upper and lower domain boundaries are surfaces of constant Z. In general, of course, this is not the case, and dealing with the difficult vertical boundary conditions has always been a problem in IC, and even in the atmospheric PC because the surface pressure is variable (see Kasahara, 1974). The simplest choice of horizontal boundary conditions in IC, namely periodic boundary conditions, is only appropriate for the fBE-IC model, and not the other IC balanced models, and hence for only a very restricted set of physical phenomena. Thus, we feel that difficulties in building a reasonably general numerical balanced model in IC are more formidable than in PC because of the known vertical boundary condition problem and the unfamiliar horizontal boundary conditions in a finite domain. This suggests that progress in obtaining solutions for the PC models will be more rapid than for the IC models, even though the conceptual virtues of the two model types are comparable at our present state of knowledge.

REFERENCES

Batchelor, G. K. 1953. Dynamical similarity of motions of a perfect gas atmosphere. *Q.J.R. Meteorol. Soc. 79*, 224–235.

Bleck, R. 1973. Numerical forecasting experiments based on the conservation of potential vorticity on isentropic surfaces. *J. Appl. Meterol. 12*, 737–752.

Bleck, R. 1974. Short-range prediction in isentropic coordinates with filtered and unfiltered numerical models. *Mon. Wea. Rev. 102*, 813–829.

Bolin, B. 1956. An improved barotropic model and some aspects of using the balance equation for three-dimensional flow. *Tellus 8*, 61–75.

Charney, J. G. 1962. Integration of the primitive and balance equations. Proc. Int. Symp. Numerical Weather Prediction, Tokyo, *Meteorol. Soc. Japan*, 131–152.

Gent, P. R. and McWilliams, J. C. 1982. Intermediate model solutions to the Lorenz equations: strange attractors and other phenomena. *J. Atmos. Sci. 39*, 3–13.

Gent, P. R. and McWilliams, J. C. 1983a. Consistent balanced models in bounded and periodic domains. *Dyn. Atmos. Oceans 7*, 67–93.

Gent, P. R. and McWilliams, J. C. 1983b. Regimes of validity for balanced models. *Dyn. Atmos. Oceans. 7*, 167–183.

Hoskins, B. J. 1975. The geostrophic momentum approximation and the semigeostrophic equations. *J. Atmos. Sci. 32*, 233–242.

Hoskins, B. J. and Bretherton, F. P. 1972. Atmospheric frontogenesis models: mathematical formulation and solution. *J. Atmos. Sci. 29*, 11–37.

Kasahara, A. 1974. Various vertical coordinate systems used for numerical weather prediction. *Mon. Wea. Rev. 102*, 509–522.

McWilliams, J. C. and Gent, P. R. 1980. Intermediate models of planetary circulations in the atmosphere and ocean. *J. Atmos. Sci. 37*, 1657–1678.

Salmon, R. 1983. Practical use of Hamilton's principle. *J. Fluid Mech. 132*, 431–444.

Shapiro, M. A. 1970. On the applicability of the geostrophic approximation to upper-level frontal-scale motions. *J. Atmos. Sci. 27*, 408–420.