A Probabilistic Theory for Balance Dynamics

GREGORY J. HAKIM

University of Washington, Seattle, Washington

(Manuscript received 18 April 2007, in final form 16 November 2007)

ABSTRACT

Balance dynamics are proposed in a probabilistic framework, assuming that the state variables and the master, or control, variables are random variables described by continuous probability density functions. Balance inversion, defined as recovering the state variables from the control variables, is achieved through Bayes' theorem. Balance dynamics are defined by the propagation of the joint probability of the state and control variables through the Liouville equation. Assuming Gaussian statistics, balance inversion reduces to linear regression of the state variables onto the control variables, and assuming linear dynamics, balance dynamics reduces to a Kalman filter subject to perfect observations given by the control variables.

Example solutions are given for an elliptical vortex in shallow water having unity Rossby and Froude numbers, which produce an outward-propagating pulse of inertia–gravity wave activity. Applying balance inversion to the potential vorticity reveals that, because potential vorticity and divergence share well-defined patterns of covariability, the inertia–gravity wave field is recovered in addition to the vortical field. Solutions for a probabilistic balance dynamics model applied to the elliptical vortex reveal smaller errors (“imbalance”) for height control compared to potential vorticity control.

Important attributes of the probabilistic balance theory include quantification of the concept of balance manifold “fuzziness,” and clear state-independent definitions of balance and imbalance in terms of the range of the probabilistic inversion operators. Moreover, the theory provides a generalization of the notion of balance that may prove useful for problems involving moist physics, chemistry, and tropical circulations.

1. Introduction

Theoretical and practical motivation for balance dynamics is derived from the search for approximate dynamical systems that accurately reproduce solutions of the full system, but with fewer degrees of freedom. Put differently, the quest in balance dynamics is to accurately compress, or encode, information into fewer variables that I shall refer to as “control” variables. To retrieve the information compressed in these control variables, a decompression algorithm is needed to “invert” the state variables; I shall call this a balance relationship. For example, geostrophic and hydrostatic balance decompress the horizontal velocity and temperature fields, respectively, from the pressure field. Balance compression is lossy, and the error in the inversion defines the unbalanced part of the state. A complete balance model is obtained when the balance relationship is coupled to balance dynamics in the form of a temporal evolution equation for the control variables. Extant balance relationships and dynamics are defined deterministically, whereas here, a new theoretical framework is proposed based on conditional probabilities.

In the interest of brevity, the vast literature on balance models will not be summarized here; the interested reader may consult McIntyre and Norton (2000) and Muraki et al. (1999) for references. Potential vorticity (PV) figures prominently in the balance model literature, in part because in two well-known classical balance models, the barotropic vorticity equation and the quasigeostrophic (QG) system, the dynamics compress naturally on this variable. Furthermore, two basic facts elevate the status of PV among other control variables. The first fact is that linear modes of the primitive equations for states of rest are distinguished, exactly, by linearized PV,¹ and therefore linear inertia–gravity

¹ Inertia–gravity waves have zero linearized PV, whereas Rossby waves do not.

Corresponding author address: Gregory J. Hakim, Department of Atmospheric Sciences, University of Washington, Box 351640, Seattle, WA 98195-1640.
E-mail: hakim@atmos.washington.edu

DOI: 10.1175/2007JAS2499.1

© 2008 American Meteorological Society
waves are effectively filtered through a PV inversion decompression algorithm. Second, for adiabatic and inviscid conditions, PV is conserved on fluid particles for the full system, which naturally reduces dimensionality; for the hydrostatic Boussinesq equations, the compression ratio for the PV is 3.\(^2\)

There are, however, several problems with using PV as a control variable. One problem is that the PV is related nonlinearly to the state variables, which renders inversion problematic in the case where one would like to attribute parts of the state to parts of the PV field. A second problem, which derives from the first, is that multiple states may share the same PV distribution, so that in the absence of ancillary information, inversion is nonunique. Third, the PV has an unknown relationship to some state variables, particularly those that relate to moisture, and therefore inversion cannot even be defined in these cases. Put differently, it is unclear how well the “PV paradigm” generalizes the notion of balance beyond the classical confines of extratropical dry dynamics.

Here we break from the deterministic tradition of previous studies on balance dynamics and define probabilistic balance models. The theory is defined in section 2, where balance relationships are established by conditional probability density functions, and inversion through Bayes’ theorem for these functions. Balance dynamics are recovered by applying the Liouville equation to the conditional probabilities. Although this theory is exact, it is also computationally impractical except for very low-dimensional systems, motivating Gaussian statistics. Section 3 is devoted to shallow water experiments, which are used to provide a practical illustration of the Gaussian-approximated theory. Implications of these ideas for balance dynamics are discussed in section 4, and conclusions are given in section 5.

2. A probabilistic theory for balance

Define a state vector, \(\mathbf{x}\), as an \(N \times 1\) vector of all variables for the full system in a chosen set of coordinates. Let \(\mathcal{X}\) denote an \(N \times 1\) set of random variables that correspond to elements of the state vector. It will prove useful later to define an \(N \times M\) ensemble state matrix, \(\mathbf{X}\), consisting of \(M\) independent columns representing state vector realizations for \(M\) ensemble members. Similar to the state variables, a control vector is defined by an \(n \times 1\) vector of variables, \(\mathbf{q}\), an \(n \times 1\) vector of random variables, \(\mathbf{q}\), and an \(n \times M\) ensemble control matrix, \(\mathbf{Q}\).

a. Probabilistic balance inversion

The state and control random variables are described by joint probability density functions. For the state vector, the joint probability is given by \(f(x) = f(x_1, x_2, \ldots, x_N)\), and for the control vector, the joint probability is given by \(f(q) = f(q_1, q_2, \ldots, q_n)\).\(^3\) These probability density functions may be used to define a balance relationship as the conditional probability of the state vector variables given the control variables, \(f(x|q)\). This conditional probability may be inverted using a multidimensional version of Bayes’ theorem for continuous random variables,

\[
f(x|q) = \frac{f(x,q)}{f(q)}. \tag{1}
\]

Note that (1) reverses the conditional probability such that the right-hand side is a function of the conditional probability of the control variables given the state variables. Equivalently, using the definition for joint probability, (1) may be expressed in terms of the joint state–control probability density function as

\[
f(x|q) = \frac{f(x,q)}{f(q)}. \tag{2}
\]

b. Probabilistic balance dynamics

Balance dynamics are defined by the time evolution of the conditional probability of the state variables given the control variables. Letting \(\mathbf{x}=[x_1, x_2, \ldots, x_N, q_1, q_2, \ldots, q_n]^T\), an \((N+n) \times 1\) vector containing both the state and control variables, the dynamics may be described by

\[
\frac{d\mathbf{x}}{dt} = \mathcal{A}(\mathbf{x}), \tag{3}
\]

where \(\mathcal{A}\) is, in general, a nonlinear vector-valued function of the state and control variables.

Letting

\[
\rho = f(x|q) = f(x|\mathbf{x}),
\]

the Liouville equation defines the temporal evolution of the conditional probability density (Ehrendorfer 1994) as

\[\]
\[
\frac{\partial \rho}{\partial t} = \nabla \cdot [\rho \mathbf{A}(\mathbf{x})],
\]
(4)

where
\[
\nabla = \frac{\partial}{\partial x_k} \mathbf{e}_k
\]
(5)
is a generalized gradient operator with unit vectors \(\mathbf{e}_k\) in the coordinates of \(\mathbf{x}\). The full-balance model consists of the balance dynamics (4) constrained by the balance relationship (2). For any reasonable discretization of the probability density functions, solving this full-balance model is only practical for a low-order system having few degrees of freedom. Further progress, along with explicit results for high-dimensional problems, is possible assuming Gaussian statistics.

c. Gaussian statistics

Gaussian statistics imply that the state and control random variables are completely determined by a mean vector, denoted here by an overbar, and a covariance matrix:
\[
X \sim N(\bar{x}, \mathbf{B})
\]
(6)
\[
Q \sim N(\bar{q}, \mathbf{P}),
\]
(7)
where \(\mathbf{B}\) and \(\mathbf{P}\) are the state and control variable covariance matrices, respectively. The specific functional form of the probability density for the state variables is given by
\[
N(\bar{x}, \mathbf{B}) = \frac{1}{\sqrt{2\pi}^n \det{\mathbf{B}}} \exp\left\{-\frac{1}{2} (x - \bar{x})^T \mathbf{B}^{-1} (x - \bar{x}) \right\},
\]
(8)
with a similar expression for the control variables. For the joint state–control vector, \(\mathbf{x}\), the covariance matrix is given by
\[
\mathbf{E} = \langle \bar{x} \bar{x}^T \rangle = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{P} \end{bmatrix},
\]
(9)
where \(\langle \cdot \rangle\) indicates the expected value. The submatrices are defined as follows: \(\mathbf{B} = \langle xx^T \rangle\) is the \(N \times N\) state variable covariance matrix, \(\mathbf{P} = \langle qq^T \rangle\) is the \(n \times n\) control variable covariance matrix, and \(\mathbf{C} = \langle qx^T \rangle\) is the \(N \times n\) state–control variable covariance matrix.

For these Gaussian statistics, one can show that the Bayesian balance inversion (1) is given by
\[
f_{X|\Omega}(x|q) \sim N(\bar{x}', \mathbf{B}').
\]
(10)
The mean value of the inverted state variables is given by
\[
\bar{x}' = \bar{x} + \mathbf{L} (q - \bar{q})
\]
(11)
and the covariance matrix of the inverted state variables is given by
\[
\mathbf{B}' = \mathbf{B} - \mathbf{L} \mathbf{C}^T,
\]
(12)
where
\[
\mathbf{L} = \mathbf{C} \mathbf{P}^{-1}.
\]
(13)
Balance dynamics require the propagation of \(\bar{x}, \bar{q}, \bar{q}, \mathbf{B}, \mathbf{C}, \) and \(\mathbf{P}\). These are formidable calculations through (3), and since the details depend on the form of the nonlinear operator \(\mathbf{H}\), this will not be pursued here. For linear (or tangent linear) dynamics, solutions can be written in terms of a propagator matrix for both the state and control variables:
\[
\mathbf{x}(t) = \mathbf{M}(t, t_0) \mathbf{x}(t_0)
\]
\[
\mathbf{q}(t) = \mathbf{R}(t, t_0) \mathbf{q}(t_0).
\]
(14)
Similarly, the covariance matrices needed by (11) and (12) evolve according to
\[
\mathbf{B}(t) = \mathbf{M}(t, t_0) \mathbf{B}(t_0) \mathbf{M}(t, t_0)^T
\]
\[
\mathbf{P}(t) = \mathbf{R}(t, t_0) \mathbf{P}(t_0) \mathbf{R}(t, t_0)^T
\]
\[
\mathbf{C}(t) = \mathbf{M}(t, t_0) \mathbf{C}(t_0) \mathbf{R}(t, t_0)^T.
\]
(15)
Probabilistic balance dynamics are given by solutions of (14) and (15) that satisfy (11) and (12) for all times. In practice, models are integrated in discrete time steps, and the following probabilistic balance-dynamics algorithm applies at each time step:

1) Given mean values for the state and control variables, \(\bar{x}\) and \(\bar{q}\), and the covariance matrices \(\mathbf{C}, \mathbf{B}, \) and \(\mathbf{P}\), update the state mean to \(\bar{x}'\) and covariance to \(\mathbf{B}'\) using (11) and (12), respectively.

2) Propagate the updated \(\bar{x}\) and \(\bar{q}\) by (14) and updated \(\mathbf{C}, \mathbf{B}, \) and \(\mathbf{P}\) by (15) forward in time by one time step. Return to step 1).

This sequential algorithm, where means and covariances are propagated and then updated with new information, is reminiscent of state estimation, where new observations serve the role that control variables do here. In fact, the probabilistic balance-dynamics algorithm is a form of Kalman filter in the limit of vanishing observation error covariance,\(^4\) provided that \(\mathbf{q} = \mathbf{Hx}\), where \(\mathbf{H}\) is the “observation operator.” Furthermore, in this limit the Kalman gain matrix becomes
\[
\mathbf{K} = \mathbf{B} \mathbf{H}^T [\mathbf{H} \mathbf{B} \mathbf{H}^T]^T,\]
\(^4\) The control variables are known to take deterministic values \(\mathbf{q}\).
or, in the notation of this paper, 

$$K = CP^{-1},$$

which is simply the inversion operator. $L = H^{-1}$. This means that state variables are constrained by linear regression onto the control variables.

### 3. Experiments for rotating shallow water

The Gaussian theory is illustrated here for the rotating shallow water equations. An elliptical-vortex example is chosen for simplicity and because it is not sensitive to the initial conditions, allowing a small ensemble to be used for the calculations. Balance inversion is illustrated first for solutions to a random initial condition that are not constrained by the control variables. Balance dynamics are illustrated in the subsequent subsection, where the state variables are constrained in time by the control variables.

The nondimensional shallow water equations are given by

$$\frac{\partial u}{\partial t} = -\epsilon \mathbf{V} \cdot \nabla u - \epsilon \frac{\partial h}{F^2 \partial x} + v,$$

$$\frac{\partial v}{\partial t} = -\epsilon \mathbf{V} \cdot \nabla v - \epsilon \frac{\partial h}{F^2 \partial y} - v,$$

and

$$\frac{\partial h}{\partial t} = -\nabla \cdot (\mathbf{V} h).$$

Here $\mathbf{V} = (u, v)$ is the horizontal velocity vector, $h$ is the fluid depth, $F = (U/\sqrt{gh^*})$ is the Froude number, and $\epsilon = (U/L)$ is the Rossby number. Constants $g$ and $f$ represent the gravitational and Coriolis parameters, respectively. Nondimensional scaling is as follows:

$$(x, y) \sim L, \quad (u, v) \sim U, \quad t \sim L/V, \quad \text{and} \quad h \sim h^* = \frac{U^2}{F^2 g}.$$

Potential vorticity is conserved according to

$$\frac{\partial Q}{\partial t} = -\mathbf{V} \cdot \nabla Q,$$

where $Q = [(1 + \epsilon\xi/h)]$ is the potential vorticity, which scales as $(f/h^*)$. For all experiments, $\epsilon = F = 1$, which is a parameter regime far from those typically associated with balance.

The ensemble initial condition selected to illustrate the theory is an elliptical vortex, with small random changes in position, intensity, and ellipticity. This initial condition produces a gravity wave field that radiates from the elliptical vortex, allowing for a test of the inversion method in the presence of a field traditionally considered unbalanced. Moreover, this initial condition has weak error growth, so that errors are limited to a small subspace (initially 5 degrees of freedom) that is spanned well by a small ensemble. The divergence field is initialized to zero, and therefore the covariance relationship between divergence and PV is determined by the dynamics.

Specifically, the elliptical cyclonic vortex is defined by a streamfunction

$$\psi = r_A \exp \left\{ - \left[ \left( \frac{x - r_2}{r_3} \right)^2 + \left( \frac{y - r_4}{sr_5} \right)^2 \right] \right\},$$

(20)

where $A = -0.25$ is an amplitude parameter, $\sigma = 0.7$, defines the vortex ellipticity, $r_1$, $r_2$, $r_3$, $r_4$, $r_5 \sim N(1, 0.01)$, and

$$\left( u, v \right) = \left( -\frac{\partial \psi}{\partial y}, \frac{\partial \psi}{\partial x} \right).$$

(21)

The disturbance height field is initialized by solving the nonlinear balance equation

$$\nabla^2 h' = \frac{F^2}{\epsilon} \nabla^2 \psi + 2F^2 \left( \frac{\partial \psi}{\partial x}, \frac{\partial \psi}{\partial y} \right).$$

(22)

A 50-member ensemble is populated using random draws for the $r$ parameters in (20). Solutions are determined using the pseudospectral method for a spectral resolution of $64 \times 64$ on a domain having dimensions $2\pi \times 2\pi$.

The solution given in Fig. 1 shows that the elliptical cyclonic vortex rotates counterclockwise with a quadrupole of divergence near the vortex. A ring of divergence radiates away from the vortex, reflecting a pulse of gravity waves generated by the initial condition.

#### a. Balance inversion

Here we apply the Bayesian definition for balance inversion of PV for the elliptical vortex solution at time $t = 3$. From (11),

$$\mathbf{x} - \bar{\mathbf{x}} = \mathbf{L} (\mathbf{q} - \bar{\mathbf{q}}),$$

(23)

where the $N \times n$ linear inversion operator $\mathbf{L}$ is defined by $\mathbf{L} = CP^{-1}$. The full 50-member ensemble is split into a 49-member ensemble to estimate the inversion operator, $\mathbf{L}$, and an independent member to test the performance of the operator. Calculations are performed in a spectral subspace by retaining the leading 32 Fourier harmonics in $x$ and $y$, leaving a total of 1024 degrees of freedom for the PV and 3072 for the state variables.
The dominant $p$ independent degrees of freedom are isolated from sampling noise using the singular value decomposition (svd)

$$C \approx U S V^T, \quad (24)$$

where $U$ and $V$ are orthogonal matrices of size $N \times p$ and $n \times p$, respectively, and $S$ is a $p \times p$ diagonal matrix containing the leading $p$ singular values of $C$. The linear inversion operator is approximated by

$$L \approx U S V^T (\bar{q} q^T)^{-1}, \quad (25)$$

where $\bar{q} = V^T q$, and the inverse is achieved by a svd-pseudoinverse technique (Golub and van Loan 1996).

Figure 2 shows results for the inversion of height and divergence from potential vorticity for the independent ensemble member. Root-mean-square (RMS) errors in the recovered height field are about 2% of the magnitude of the disturbance field (Fig. 2a). Although linearization (23) applies to disturbances from the mean, an inversion of the full PV field, including the mean, has similar errors (Fig. 2b), which are smaller than for the nonlinear balance solution (Fig. 2c). A noteworthy as-
pect of the inversion results is the recovery of divergence associated with the gravity wave field (Fig. 2d). This result derives from the fact the PV correlates with the gravity wave field through the dynamics; specifically, small differences in the initial ensemble result in small changes in the gravity wave field at later times. One way to quantify this relationship is through the leading singular vectors of the inversion result. The leading two singular-vector pairs show that the gravity wave divergence field relates primarily to changes in the ellipticity and strength of the vortex, as measured by the PV (Fig. 3).

Summary statistics of the accuracy of the PV inversion results are derived by constructing the inversion operator for a 25-member ensemble, leaving an ensemble of 25 independent members for error analysis. Performance is measured by spatial error variance in the inversion fields, normalized by the error in the original ensemble fields, as measured by the ensemble variance. For the height field, the ensemble-mean normalized error variance is 0.037 ± 0.031, where the error bars give the ensemble standard deviation of the normalized error variance. This result indicates that the
average error in the inverted height is about 3% of the error in the original data. For the divergence field, the ensemble-mean normalized error variance is 0.16 ± 0.12.

“Piecewise” PV inversion involves inverting a localized region of PV and attributing the inverted fields to the localized region (e.g., Davis and Emanuel 1991; Hakim et al. 1996). A natural limiting solution for piecewise inversion is given by Green’s function for the inversion operator, which represents the solution for a δ function of PV (Bishop and Thorpe 1994). In the context of the discrete solutions considered here, numerical estimates of Green’s function are obtained by one unity entry in the PV column vector with zeros for all other entries, which selects a single column of L. This provides a more thorough check of the inversion operator, because errors due to multicollinearity in the regression calculation can be masked in the inversion of the global field; for example, each point could have the same pattern, contributing a small amount to the global field.

Unlike Green’s functions for classical linear balance operators, such as the Laplacian, which are isotropic and independent of horizontal position, Green’s functions for the statistical inversion operator are anisotropic and depend on location (Fig. 4). However, an average over all points (Fig. 4c) reveals a nearly axi-symmetric solution with a decay pattern that is unlike that for the shallow water quasigeostrophic Green’s function (modified Bessel function of the second kind; not shown).

b. Balance dynamics

Probabilistic balance dynamics requires propagation of the mean and covariance for the state and control variables. As discussed in section 2c, probabilistic balance dynamics can be viewed as a form of Kalman filter, with the control variables serving the role of “perfect” observations. Solutions of (15) by direction calculation are impractical for most real systems, and therefore the propagation of these matrices is approximated here by an ensemble Kalman filter (EnKF; Evensen 1994). A 50-member ensemble is employed, implemented as a square root filter (Whitaker and Hamill 2002). Values for the control variables are determined by an independent integration of the shallow water equations for a 51st ensemble member. Control variables are assimilated at every eighth grid point for every fourth model time step. Results depend on these choices, especially if the number of control variables is changed significantly. A computational sponge layer is added near the lateral boundaries of the domain to prevent outgoing waves from reentering the periodic domain.

Two control variables are considered, PV and height, for the elliptical-vortex initial condition defined in section 3. Errors are measured in terms of the vorticity and divergence fields, which are not explicitly constrained by the control variables. Results show that height control yields smaller errors in these fields compared to PV control (Fig. 5). For both height and PV, vorticity errors decline and reach an equilibrium level after about 30
assimilation cycles, whereas the divergence errors require at least 40 cycles to approach equilibrium because the divergence is initialized to zero.

Spatial patterns of unbalanced vorticity after 15 assimilation cycles ($t = 3$) exhibit a dipole pattern oriented along the major axis of the elliptical vortex for both height and PV control, although the magnitude of the imbalance is larger for PV control (Figs. 6a,b). The difference between height and PV control is more substantial after 50 assimilation cycles ($t = 10$), when unbalanced vorticity for height control exhibits a weaker reflection of the dipole pattern observed after 15 cycles (Figs. 6c,d). For PV control after 50 cycles, the unbalance vorticity is significantly larger than for $h$ control, with a spatial pattern that reflects a vortex that is too weak, with lower vorticity in the core, and misplaced spiral arms. Differences in the unbalanced divergence field are more dramatic (Fig. 7). After 15 assimilation cycles, a four-cell divergence pattern is apparent for the case of height control, but for PV control this field is dominated by noise with no evidence of the near-vortex four-cell pattern. Unbalanced divergence for height control after 50 assimilation cycles shows a small-scale wave pattern located mainly on the left side of the vortex (Fig. 7c). For PV control, the divergence field remains noisy, and the unbalanced divergence is much larger in magnitude (Fig. 7).

These results suggest that, after a period of “spinup,” probabilistic balance dynamics yields states that have maximum errors of about 2% of the true values in vor-
Fig. 5. Nondimensional RMS errors in (a) vorticity and (b) divergence for statistical balance dynamics controlled by PV (open circles) and height (filled circles). Time increases along the abscissa according to number of assimilation times, where one assimilation time equals 0.2 nondimensional time units.

Fig. 6. Vorticity (colors) and vorticity errors (lines) for statistical balance dynamics after (a),(b) 15 and (c),(d) 50 assimilation times. Results for height control are given in (a),(c) and PV control in (b),(d). Contours are every 0.1 units in (a),(b) and every 0.04 units in (c),(d). The zero contour is suppressed and all values are nondimensional.
ticity and 8% in divergence. This error reflects the unbalanced part of the state that is not constrained by information in the control variables. One caveat on these results is that the probabilistic balance model in these calculations involves assimilation of the control variables from an independent ensemble member that is unaffected by assimilation. In general, the winds derived from the assimilation step could be used to propagate the control variables [e.g., in (18) for height control and in (19) for PV control]. This is the standard procedure for balance models based on PV: winds derived from PV inversion are used to advect the PV. Many possibilities exist for probabilistic balance dynamics of this form, including pairing state and control variable ensemble members to share the same velocity field, and replacing ensemble propagation altogether by using fixed covariance matrices, such as time averages. The latter case may reduce probabilistic dynamics to classical balance models (e.g., quasigeostrophy) for appropriately chosen fixed covariance matrices. Evaluating the merit of these approximations, and the relationship of the probabilistic approach to existing balance models, are challenges for future research.

4. Implications

In general, the inversion operator (13) is not square, and the compression ratio is given by the row rank divided by the column rank. For the shallow water equations, or hydrostatic primitive equations, PV yields a compression ratio of three; there are \( n \) model grid points for PV and \( N = 3n \) total prognostic variables. The dimension of the range of \( L \) in this case is \( n \). For states close to rest, the range corresponds to \( n \) Rossby waves and the \( 2n \) inertia–gravity waves constitute a null space of the operator because all inertia–gravity waves have zero linearized PV. While the statistical and modal definitions for balance yield equivalent results...
for linear dynamics close to a state of rest, the statistical approach provides a clear generalization to states for which the modal framework becomes less clear: the unbalanced parts of the state are not in the range of the operator \( L \). In should be noted that, for ensembles of size less than \( n \), ensemble sampling can affect the estimate of the balanced part of the state. Undersampling reduces the rank of \( L \) and increases the rank of the orthogonal complement of the range; that is, parts of the state unresolved by the sample are interpreted as unbalanced. Finally, “optimal” balance control variables may be defined as those that, for a given compression ratio, have minimum error (unbalanced component) in a chosen norm, or, the maximum compression ratio for a chosen error bound.

Experiments conducted here suggest that layer depth is a better control variable than PV for the shallow water equations. Because it is a nonlinear function of the state variables, PV could be problematic for the linear statistical approach defined here because the PV is not a unique function of the state variables. A pathological limit occurs when all independent members of an ensemble, each having unique values for the state variables, have identical PV, in which case the perturbations from the mean, and the balance state, are equal to zero. Such problems do not occur for linear functions of the state variables, such as layer depth, provided that the ensemble members are linearly independent.

As the results of section 3 indicate, the statistical inversion approach defined here allows for the recovery of gravity waves from the PV field. This is due to the fact that the gravity wave field correlates with the balance field that generates it. As with divergent ageostrophic circulations due to frontogenesis in rotating stratified fluids, these motions may be viewed as part of an adjustment process to restore balance. Furthermore, like the secondary frontogenetical circulations, these generalized secondary circulations (gravity waves) may be recovered from the balance control variables. Given sufficient time, the gravity waves propagate far from the source generation region such that they decorrelate with the PV field, and fall into the unbalanced part of the state.

5. Conclusions

A probabilistic theory for balance relations and balance dynamics is defined. The essential aspect of the theory is derived from the fact that the state of real physical systems is uncertain, and balance may be defined in terms of conditional probability density functions that account for this uncertainty. Bayes’ theorem provides the definition for balance relationships, where the conditional probability of the state variables given the control variables is inverted. Probabilistic balance dynamics involve the temporal evolution of the conditional probability density functions, constrained by the balance relationship for all time. This general theory is fully nonlinear and non-Gaussian, but impractical for computational implementation in real problems. Assuming Gaussian statistics reduces the inversion relationship to linear regression of the state variables on the control variables, in which case balance dynamics can be viewed as a Kalman filter that is cycled on error-free observations (the control variables).

This probabilistic approach provides a direct estimate of the “fuzziness” of a balance manifold in terms of probability density functions. From a state estimation perspective, the fuzziness is determined by the information contained in the control variables, which constrain the balance manifold. The amount of fuzziness depends on the choice of control variables and the metric used for measuring error. Again, from a state estimation perspective, the best control variables (“observations”) are those that contain the most information about the full state of the system. The limited shallow water experiments conducted here suggest that height is a better control variable compared to potential vorticity because it yields smaller errors. Research into optimal control variables, which yield the smallest errors, or the maximum compression ratio for a given error bound, would be a fruitful direction for future research.

Further research is also needed to test the probabilistic balance theory in more general systems, such as the hydrostatic primitive equations. An attractive property of the probabilistic approach is that it liberates balance dynamics from the confines of extratropical dry dynamics. Moist physics can easily be incorporated, so that “balanced precipitation” can be determined from the PV, for example (see, e.g., Gombos and Hansen 2008). Moreover, the theory may provide a generalization of the notion of balance to problems such as those involving chemistry and tropical circulations.

Acknowledgments. David Dritschel suggested the pathological PV thought experiment discussed in section 4, where all ensemble members share the same potential vorticity. This research was supported by the National Science Foundation through Grants 0205648, 0327658, and 0552004.

---

5 Such as unsteady states far from rest, for which linear mode solutions are difficult to define.
6 Except when the PV is approximated linearly for states close to rest, in which case the PV is zero for inertia-gravity waves (not shown).
REFERENCES


