Atmósfera (1991), **4**, pp. 145-164

Analysis of dynamic data assimilation for atmospheric phenomena. Effect of the model order

LUIS LE MOYNE HERNANDEZ

Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, 04510, México, D. F., México

JESUS ALVAREZ CALDERON

Departamento de Ingeniería de Procesos, Universidad Autónoma Metropolitana-Iztapalapa, Apdo. 55534, 09340, México, D. F., México

(Manuscript received Feb. 1st, 1990; accepted in final form Nov. 14, 1990)

RESUMEN

Se estudia el problema del pronóstico del tiempo basado en el uso de información *a priori* a través de modelos de principios físicos en conjunción con medidas obtenidas sobre la marcha. Nuestras derivaciones de un algoritmo de estimación ponen en evidencia que para entender y explotar las técnicas actuales de estimación se requieren ingredientes de conocimientos físicos de los procesos, técnicas para representar ecuaciones diferenciales parciales (como un conjunto de ecuaciones diferenciales ordinarias), propagación de errores numéricos y estimación condicional en estadística. Se concluye que la actualización del error estadístico de modelo así como del número y localización de las medidas, juegan un papel importante en el desempeño del estimador. Un tratamiento adecuado de los aspectos antes mencionados, genera un estimador cuya aplicación, tamaño, tratamiento numérico, esfuerzo computacional y medida de red deben ser tales que se garantice un cierto desempeño manteniendo simplicidad y facilidad de implementación.

ABSTRACT

The problem of weather forecasting based on the use of *a priori* information, from a first principles model, in conjunction with on-line measurements is addressed. Our derivation of an estimation algorithm evidences that understanding and exploitation of available estimation techniques require ingredients from physical knowledge of the process, tools to represent partial differential equations (as a reduced set of ordinary differential equations), numerical error propagation and conditional estimation in statistics. It is concluded that updating of the model error statistics and the number and location of measurements play an important role on the estimator performance. A suitable treatment of the preceding issues leads to an estimator whose accuracy, size, numerical treatment, computational effort and measurement mesh can be chosen so that performance is guaranteed while keeping simplicity and ease of implementation.

1. Introduction

Techniques for weather forecasting are an active area of research in atmospheric sciences. According to the deterministic or statistical nature of models used, those techniques can be classified in three classes: fully deterministic, fully statistical and deterministic-statistical.

The fully deterministic techniques (Charney *et al.*, 1950; Adem, 1962; Castro *et al.*, 1985), consist in the time integration of a model represented by a set of equations derived from first principles (mass, momentum and heat conservation as well as thermodynamics and constitutive equations). The time evolution of the model-reality missmatch is induced by the initial deviation, with respect to reality, of the model and by a persistent exogenous deviation produced by the imperfections associated to both the conceptualization of the phenomenon and the numerical approximations used to solve the model equations. It must be pointed out that the preceding forecasting techniques do not possess a self-correcting scheme based on measurements received as the estimator is running.

Fully statistical models (Le Moyne *et al.,* 1986) consist of difference equations that describe the temporal evolution of a set of statistical parameters that represent the state of the process. The statistical model is usually a linear auto regressive moving average model (ARMA) built, not from first principles, but from *a priori* study of the statistical behavior of discrete, in space and time, data. Once this is done, available recursive statistical tools are used to devise the estimation algorithm. Forecasting requires interpolation in space and extrapolation (to the future) in time. The advantage of the recursive statistical techniques is that, by incorporating data (as they become available), in a feedback scheme, the predictor acquires an error self-correcting feature. The disadvantage is that the fully statistical technique ignores the knowledge, about the process, available from research efforts on model construction based on first principles.

Deterministic-statistical models combine statistical and deterministic features (Miyakoda and Talagrand, 1971). Our work is based on that kind of models.

In the communication and control theory (Kalman, 1960), the estimation of unmeasured states based on incomplete noisy measurements has been addressed. Jones (1965) applied Kalman's, optimal estimation technique to meteorology forecasting using an ARMA-type model. Petersen (1968) applied Kalman's theory in meteorology to carry out objective synoptic observations. That is, he built a time-space statistical correlation to perform a space interpolation and a time prediction.

Epstein (1969) used a deterministic model where he acknowledged an erroneous initial condition that was modeled as a random variable. As a consequence, the deterministic evolution model became a stochastic process that provided, in principle, the evolution of the probability density funtion. However, it did not provide means to update the forecasting as new measurement information became available (Jazwinsky, 1970). Moreover, posing the problem as a full characterization of a non-Gaussian statistics leads to complicated numerical implementation problems.

Epstein and Pitcher (1972), faced the latter disadvantage by restricting themselves from the beginning to a linear system with Gaussian statistics. As a result, two moments suffice to yield a complete statistical characterization. Along the same line of work, Pitcher (1977) incorporated persistent forcing errors due to model truncation. This was accomplished by modeling the truncation error as an additive, exogenous stochastic process. Miyakoda and Talagrand (1971) used data over a period of several days with a deterministic estimator and obtained a forecasting model which assimilated in an optimum way the observations as they become available. Ghil *et al.* (1981) developed a shallow water model suitable for treatment with Kalman's theory. The estimator model used a finite difference approximation. The use of standard finite difference techniques to represent the model may lead to an inefficient numerical algorithm. This is important if we are interested in developing an optimal estimator, because the optimization of the filter consisting on reducing (or increasing), the number of active states and it is not very easy in finite differences.

Stochastic estimation based on combination of model prediction with actual measurements, is a rather mature area in systems theory and has undergone ample testing in various scientific and engineering disciplines.

In athmospheric science, prediction with data assimilation has been a subject of study for various decades. Only recently, that line of research has found a formalization and systematization in optimal stochastic estimation theory (Ghil, 1981), Ghil, (1989) gave an overview on the state of the art in prediction atmospheric studies. Although the estimation theory used is a standard one, the value of that work is the move towards systematization in atmospheric prediction.

In this work we address issues which are relevant in understanding the model-based estimation scheme and in implementing it efficiently. We derive the estimator model as one where three kinds of model approximations have been made: conceptual, representation in a finite dimension and numerical time integration. The additional feedback introduced by data assimilation is conceptualized as a devise to endow the predictor with ability to tolerate model and numerical time-integration errors while maintaining precision.

Ghil *et al.* (1988), gave an overview on the state of the art on the meteorology applications of sequential estimation with data assimilation. Their work was based on a Kalman estimator which eliminated fast undesirable model modes.

In this work, we focus on issues relevant to the model-based, with data assimilation, stochastic atmospheric estimator. Those issues are: effect and role of model order and sensors' allocation on estimator performance. Ultimately, recursive stochastic estimators combine and reconcile, in a suitable manner, information from actual measurements and from a model. The estimator model is posed as an approximation to reality which includes a sequence of approximations: conceptual, representation in a reduced finite dimension and time integration. With standard estimation techniques, we derive an estimator for a hemispheric barotropic model. Derivations stress the appearance of all model estimator approximations. By doing so, we establish a framework which clearly shows how the estimator is conformed and how it works. Connections as well as interpretations with physical aspects and numerical implementation issues are established. Regarding estimator performance, we address three issues: the effect of order size of the estimator model, the effect of updating the model error statistics and the influence of the number and location of measurements.

To test the estimator under reality-model mismatch, estimator performance as a function of the estimator model order (number of harmonics) was studied. The results show that a reduced (simpler) estimator model with an adequate measurement mesh, suffices to yield adequate forecasting. It is concluded that the election of the number and location of measurements play an important role on the performance of the estimator.

2. The deterministic model

The present technique is intended for a wide class of atmospheric models. Without restricting the applicability, here the simplest barotropic vorticity model (BVM) is used to illustrate our points. The BVM describes approximately the horizontal flow of an atmosphere, which is vertically averaged, homogeneous, incompressible, inviscid and with vorticity conservation in all the flow points. The BVM is described by the following equation (Charney *et al.*, 1950):

$$\frac{\partial \varsigma}{\partial t} = -V \cdot \nabla(\varsigma = f) \qquad (1)$$

where ς is the vertical component of the relative vorticity. V is the velocity and f is the Coriolis parameter. In natural coordinates, the preceding equation can be expressed, as follows (Silberman, 1954):

$$\frac{\partial \varsigma}{\partial t} - \frac{1}{a}(V_\theta \frac{\partial}{\partial \theta} + \frac{V\lambda}{\sin\theta}\ \frac{\partial}{\partial \lambda})(\varsigma + 2\Omega\cos\theta) \qquad (2)$$

where α is the radius of the Earth. λ and θ are the longitude and the co-latitude, respectively. V_λ and V_θ are the components of the horizontal velocity and Ω is the Earth angular velocity. Furthermore,

$$\varsigma = \frac{1}{a\sin\theta}(\frac{\partial}{\partial\theta}(V_\lambda \sin\theta) - \frac{\partial V_\theta}{\partial \lambda}) \qquad (3)$$

Introduction of a stream function ψ

$$V_\lambda = \frac{1}{a}\frac{\partial\psi}{\partial\theta} \quad \text{and} \quad V_\theta = -\frac{1}{a\sin\theta}\frac{\partial\psi}{\partial\lambda} \qquad (4)$$

leads to the following representation of equation (2):

$$\nabla_h^2 \frac{\partial\psi}{\partial t} = \frac{1}{a^2\sin\theta}\left[\frac{\partial\psi}{\partial\lambda}\frac{\partial}{\partial\theta} - \frac{\partial\psi}{\partial\theta}\frac{\partial}{\partial\lambda}\right] \times \left[\nabla_h^2\psi + 2\Omega\cos\theta\right] \qquad (5)$$

This is a hyperbolic partial differential equation (PDE) with an initial value problem in time and a boundary value problem in a two-dimensional space. Boundary conditions are accounted by construction.

3. Reduced model for estimation

3.1 Discretization in space

Equation (5) describes the time evolution of an element [a function Ψ (λ, θ, t)] in an infinite dimensional space. As in a standard numerical technique, the estimator algorithm requires a suitable approximation of the solution to equation (5) in a reduced, finite dimensional space (Alvarez *et al.*, 1981). In other words, the PDE (5) must be approximated by a set of ordinary differential equations. In principle, model reduction can be achieved with any standard technique. Among them are, finite differences, finite elements, spectral methods or combination of them. Estimator performance is affected, in terms of accuracy and efficiency, by the selection of the model reduction technique. A detailed study of this matter is beyond the scope of the present work and here we circumscribe ourselves to work with one of the standard solution existing procedures for equation (5): a global approximation in terms of a double Fourier series solution with Orszag's basis functions (Orszag's, 1970). Orszag's approximation is a suitable an efficient model reduction technique which yields not only a small set of ODE's, but also a numerically well-conditioned problem. As we shall see, these features will be translated to the estimation algorithm performance.

The stream function is approximated by a truncated double series:

$$\psi(\lambda,\ \theta,\ t) = a^2\Omega\sum_{i=0}^{n}\sum_{j=i}^{1+m}(\cup_j^i(t)\cos i\lambda + V_j^i(t)\sin i\lambda)P_j^i(\sin\theta) \qquad (6)$$

where p^i are the associated Legendre normalized functions of the first kind of order i and degree j. $U_j^i(t)$ and $V_j^i(t)$ are time dependent Fourier coefficients. For simplicity, time dependency will be dropped from the notation. Substitution of (6) into (5) yields

$$\frac{dU_j^i}{dt} = \frac{2\Omega i}{j(j+1)} V_j^i - \frac{a^2\Omega}{j(j+i)} E_j^i$$

$$\frac{dV_j^i}{dt} = -\frac{2\Omega i}{j(j+1)} U_j^i - \frac{a^2\Omega}{j(j+i)} F_j^i \tag{7}$$

where E_j^i and F_j^i are Orszag's coefficients (1970). Equations (7) contitute a set of ODE's whose states dependent variables are the expansion coefficients U_j^i and V_j^i.

To facilitate manipulations, the above equation set is rewritten in vector notation:

$$\dot{x} = Ax + b, \quad x(0) = x, \tag{8}$$

where

$$x^T = \left[U_1^1,\ U_2^1 ... U_n^1,\ U_1^2,\ U_2^2 ... U_n^m,\ V_1^1, ... V_n^m\right]$$

$$b^T = \frac{a^2\Omega}{j(j+1)} \left[E_1^1, ..., E_n^m,\ F_1^1, ... F_n^m\right]$$

$$A = 2\Omega \begin{bmatrix} OD \\ -DO \end{bmatrix}$$

where $(\ )^T$ denotes vector transposition and $D =$ diagonal $\frac{i}{j(j+1)}$

3.2 Discretization in time

In principle, any standard numerical integration scheme (Euler, Euler modified, Runge-kutta, Gear) for ODE's can be used to carry out time discretization of equations (7). For a standard numerical integration (as the one used in fully deterministic predictors), election of the time integration technique is a crucial step that affects the performance and accuracy of numerical integration. On the other hand, it is a well known fact in estimation theory that introduction of measurements (as they become available), in a properly designed feedback, incorporates a self-correcting capability to the prediction. Therefore, a properly designed estimator should be able to yield adequate predictions in spite of using a simple, possibly less accurate, time integrator.

Following Alvarez *et al.* (1981), represent equation (8) as a difference equation:

$$x(t_{k+1}) = e^{A(t_{k+1}-t_k)} x_{tk} + \int_{t_k}^{t_{k+1}} e^{A(t_{k+1})} b(\tau) d\tau \tag{9}$$

Introduce an approximation: $b(t)$ is taken as a piece-wise (staircase) $b_k(t)$ constant function:

$$b(t) = b(t_k), \qquad t_k \leq t \leq t_{k+1}$$

Hence, the last equation can be approximated with the following first-order vector difference equation:

$$x_{k+1} = Fx_k + Gb_k \tag{10}$$

where

$$F = e^{A\Delta t} \qquad b_k = b(t_k)$$

$$G = Ae^{A\Delta t} \qquad \Delta t = t_{k+1} - t_k \tag{11}$$

Since (8) is a linear, time invariant equation, matrices F and G are not state or time dependent. Therefore, F and G are computed at once and stored. The aforementioned matrices are evaluated with a truncated serie expansion for an exponential matrix (Balakrishnan, 1954):

$$F = \sum_{n=0}^{m} \frac{A^n \Delta t^n}{n!} \text{ and } G = \sum_{n=0}^{m} \frac{A^n \Delta t^{n+1}}{(n+1)!} \tag{12}$$

The truncation is such that, for a given time increment Δt, the contribution of an additional term in the series is less than a certain prescribed tolerance:

$$\epsilon =_{\substack{\max \\ ij}} | f_{ij}^{m} - f_{ij}^{m-1} |, \quad \left[f_{ij}^{m} \right] = F = \sum_{n=0}^{m} \frac{A^n \Delta t^n}{n!} \tag{13}$$

3.3 Relationship between measurements and model states

For the present application model, the sensors are assumed to be m geopotential measurements at discrete locations in the two-dimensional longitude-latitude space. Measurements are available at discrete points in time. The relationship between the m-dimensional measurement vector, at time t_k, and the model states (Fourier coefficients) is given by:

$$y_k^T = \Big[p(\lambda_1,\ \theta_1,\ t_k),\ p(\lambda_1,\ \theta_2,\ t_k), \ldots;\ p(\lambda_2,\ \theta_1,\ t_k),\ p(\lambda_2,\ \theta_2,\ t_k) \ldots;$$

$$\ldots; p(\lambda_m,\ \theta_{m-1},\ t_k), \ldots, p(\lambda_m,\ \theta_m,\ t_k) \Big]$$

where $p(\lambda_i,\ \theta_i,\ t_k)$ is the measured pressure at coordinates $(\lambda_i,\ \theta_j)$ at time t_k, expressed as Fourier's coeficients.

By definition, the evolution of the state x describes the process dynamics. In general, the measurements and the states do not have to coincide, the state vector x (in this case, the expansion coefficients) and the measurement vector y_k are related by: (see appendix A)

$$y_k = Cx_k \tag{14}$$

$$C_{im} = \frac{a^2\Omega^2}{g}\left[\frac{2i}{i_2}\sqrt{\frac{(i+2)^2-m^2}{4(i+2)^2-1}}P_{i+1}^m + \right.$$

$$\left.\frac{2(i+1)}{i-1}\sqrt{\frac{i^2-m^2}{4i^2-1}}P_{i-1}^m\right]\cos m\lambda$$

Where C is the transformation matrix between states and observations.

4. The Stochastic Model

In a realistic situation, the reduced model (10) is only an approximation to reality. The reduced model contains errors due to: the conceptual simplifications underlying model (1), model reduction via a truncation of an infinite Fourier double series (6), the assumption of piece-wise constant exogenous inputs, and the truncation involved in the evaluation of the exponential matrix (11). In addition, and due to model missmatch between expansion coefficients and actual observations as well as to inherent uncertainties associated to the measuring devises, the measurement equation (14) is not exact. To face these facts, the above model and measurement imperfections are accounted for as additive random errors that enter the dynamic process at discrete points in time. That is,

$$x_{k+1/k} = Fx_k + Gb_k + w_k \tag{15}$$

$$y_k = Cx_k + v_k \tag{16}$$

where w_k and v_k are vectors assumed to be stochastic time sequences with a characterized joint statistics. w_k and v_k are zero-mean, independent, Gaussian sequences. That is,

$$E\left[w_k\right] = E\left[v_k\right] = 0, \quad E\left[w_k^T v_i\right] = 0; \quad k,\ i = 1,\ 2,$$

where E [.] is the expectation operator. Model and measurement errors are uncorrelated (white) sequences. That is,

$$E\left[w_k^T w_i\right] = \delta_{ki}Q_k, \; E\left[v_k^T v_k\right] = \delta_{ki}R_k$$

Where Q_k and R_k are error covariance matrices. δ_{ki} is a Kronecker delta ($\delta_{ki} = 1$ if $k = i$, and 0 otherwise).

The presence of the stochastic input sequences induces statistical properties to the state and measurement sequences (x_k) and (y_k). Hence, model (15 and 16) induces a joint statistics between the state and measurement sequences, and therefore, it is possible to pose the estimation problem as a conditional estimation: "given the past and the present measurements, obtain an estimate of the states (expansion coefficients) n steps ahead in time".

Before further pursuing, let us state a final assumption: the state and the error sequences are independent. That is,

$$E\left[(x_k - \overline{x}_k)^T w_i\right] = 0, \ E\left[(x_k - \overline{x}_k)^T v_i\right] = 0; \ k, \ i = 1, \ 2, \ldots$$

where $\overline{x}_k = E\left[x_k\right]$ is the mean of x_k. Finally, an initial guess for the state statistics must be assessed:

$$x_o = E\left[x_o\right], \ P_o = E\left[(x - x_o)(x - x_o)^T\right]$$

To conclude, the estimation scheme requires: a model built from first principles (F, G, C), a sequence of measurements (y_k), an initial state guess x_o, and the value of the covariances P_o, Q_k and R_k.

The required initial state value and the error covariances can be seen as design parameters for the estimation scheme. The election of those parameters should be guided, as much as possible, by designer knowledge on: the expected errors of the one-step prediction model, information about the uncertainty of the measurement, and the value of the initial state error. A great deal of this information can be gathered before the estimation scheme is implemented. As we shall see, the statistics design parameters can be adjusted either from *a priori* test or when the estimation scheme is running. Regarding implementation on a digital computer, we are aiming to end-up with a recursive estimator. In other words, the forecasting should be based on the storage of a "moving initial condition" that contains all past information.

5. Construction of the estimator

The solution to the estimator problem can be obtained with various procedures (Meditch, 1969): weighted least squares, minimum variance or maximum likelihood (Jazeinsky, 1970). Since equations (15) and (16) are linear and the statistics is Gaussian all procedures lead to the same result.

The time discrete version of the Kalman estimator has two advantages: the resulting estimator is well suited for on-line digital computations and the estimator construction procedure shows, in a transparent manner, the nature of the estimation strategy. Because the exogenous and initial statistics are Gaussian and the system is linear, the $(x_k) - (y_k)$ joint statistics is also Gaussian and therefore, the description of state estimate mean and covariance constitute a complete statistical characterization.

Suppose, for a moment, that at time t_k the mean and covariance of the state are known and that a prediction at time t_{k+1} (before receiving a new measurement y_{k+1}) is required. This prediction can be posed as a conditional expectation:

$$\hat{x}_{k+1|k} = E\left[x_{k+1|y_k},\ y_{k-1} \dots, y_o\right]$$

For the measurement update, we shall need the joint model-based statistics $f_{x|y}(x, y)$ at time t_{k+1}. Applying the conditional expectation operator $E[(.) \mid y]$ to both sides of equation (15) yields

$$\hat{x}_{k+1|k} = Fx_{k|k} + Gb_k \tag{17}$$

The assumption of (w_k) with zero mean has been used and hence, the noise term in equation (15) has vanished. From equations (15) and (16) valuation of $E\ [(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^T]$ gives a one-step prediction of the estimate error covariance:

$$P_{k+1|k} = P_{k|k} F P_{k|k}^T + Q_k \tag{18}$$

$$E\left[(x_{k+1} - \hat{x}_{k+1|k})(y_{k+1} - C\hat{x}_{k+1|k})^T\right] = CP_{k+1|k}C^T \tag{19}$$

$$E\left[(y_{k+1} - \hat{C}x_{k+1|k})(y_{k+1} - C\hat{x}_{k+1|k})^T\right] = FP_{k+1|k}C^T + R_{k+1} \tag{20}$$

Once the model-based prediction has been done, we proceed to evaluate the minimum variance estimate of x_{k+1} when a realization (measurement) of y_{k+1} is available.

To facilitate manipulations, we introduce the following notation:

$\overline{x} = \overline{x}_{k+1\|k}$	: model-based prediction of the state
$x = \hat{x}_{k+1\|k+1}$	: estimate after measurement update
$y = y_{k+1}$	: measurement
$\overline{y} = C\overline{x}$	: model-based prediction of the measurement
$\overline{P}_{xx} = P_{k+1\|k}$	: error covariance before measurement update
$\overline{P}_{xx} = P_{k+1\|k+1}$	: error covariance after measurement update
$P_{yy} = CP_{k+1\|k}C^T$	
$P_{xy} = FP_{k+1\|k}C^T$ $+R_{k+1}$	
$K = K_k$	: gain matrix

Since the injected statistics is Gaussian and the model is linear, the minimum variance estimator is a linear one (Jazwinski, 1970).

$$x = \overline{x} + K(y - \overline{y}) \tag{21}$$

The value of the gain matrix K and the *a posteriori* covariance P can be obtained directly with the orthogonal projection theorem (Anderson and Moore, 1979). This theorem can be thought of as a corollary to the conditions for a minimum variance (MV) estimate which has replaced the joint probability density function $f_{xy}(x,\ y)$ by the conditional one $f_{x|y}(x,\ y)$, by means of theorem of Bayes (Anderson and Moore, 1970). The necessary and sufficient condition to have a MV estimate is given by the orthogonality of the state error $(x - \hat{x})$ with respect to the innovation $(y - \overline{y})$.

That is,

$$E\Big[(x - \overline{x})(y - \overline{y})^T\Big] = 0$$

where 0 is a $n\ x\ m$ matrix with zero entries. Substituting (21) in the above condition and solving for K yields

$$K = \overline{P}_{xy}\overline{P}_{yy}^{-1} \tag{22}$$

The *a posteriori* covariance is obtained from the Pythagorean identity in the probability space:

$$P_{xx} = \overline{P}_{xx} - \overline{P}_{xy}\overline{P}_{yy}^{-1}\overline{P}_{xy} \tag{23}$$

Substitution of the *a priori* statistics (17, 18, 19, and 20) in expressions (21, 22, and 23) yields the equations for the measurement update. The resulting equations for the recursive estimator are:

Initialization

$$x_{o|o} = x_o, \qquad P_{o|o} = P_o \tag{24}$$

Propagation

$$\hat{x}_{k+1|k} = F_k \hat{x}_{k|k}, \tag{25}$$

$$P_{k+1|k} = F_k P_{k|k} F_k^T + Q_k. \tag{26}$$

Update

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + k_{k+1}(y_k - C_k^T \hat{x}_{k|k-1}) \tag{27}$$

$$P_{k|k} = P_{k|k-1} - K_{k+1} C_k^T P_{k|k-1} \tag{28}$$

where

$$k_{k+1} = P_{k|k-1} C_k \left[C_k^T P_{k|k-1} C_k + R_k \right]^{-1} \tag{29}$$

The fundamental idea behind the construction of the estimator is to use data, as they arrive, to reconcile, in the minimum variance sense, the model-based state prediction with the value of the measurement. Note that if the prediction of the measurement coincides with the actual measurement ($x = \overline{x}$), the state prediction is left unchanged by the measurement update.

Adaptation of the model error statistics

The matrix sequece Q_k models the uncertainty associated to the model prediction. That uncertainty is due to a model which approximately describes a physical phenomenon and to numeric approximations to that model. Q_k can be assigned large enough so that all situations are considered. However, it seems natural to use the model-based prediction of the measurement and the actual measurement to assess, in an adaptive fashion, the uncertainty Q_k.

Since uncertainty (Q_k) is injected to be model-based prediction step, the propagated uncertainty ($P_{k=1}$) is larger than the departure uncertainty (P_k). Mathematically this can be easily seen: equation (16) shows that the matrix P_{k+1} is result of adding two positive-definite symetric matrices. As a consequence, any scalar measure of the error covariance matrix (trace, determinant, maximum eigenvalue) increases. Usualy, Q_k can be pre-assigned and kept constant. Because past model error are available, it is natural to update the model matrix error covariance Q_k. Update of Q_k is a standard tool in estimation a theory Chin (1979) Mehra (1972). Phillips (1986) claims that in the atmospheric practice, use of update matrix Q_k is precluded because an excessive computing time is required. However, the excessive computational effort may have its main cause on the use of detailed models, large finite models and unnecessary accurate integration technics. In the present work, we use simple reduced models and, accordingly, a simple time integration technique. This alleviates the computational effort to store and update the matrix Q_k. To retain only the essence of a model error adaptive scheme, we use the following scheme:

$$Q_k = a_k I \tag{30}$$

where a_k is the mean squared measurement prediction error. That is,

$$a_k = \frac{1}{m} \left[(y_k - C_k x_{k+1|k})(y_k - C_k x_{k+1|k})^T \right] \tag{31}$$

6. Results

Numerical simulation tests were devised to illustrate the performance and to exhibit some features of the estimation scheme based on data assimilation. The initial states (harmonic coefficient), and their initial errors statistics were obtained as follows. Using isohipsa data, over the a 5° lat, long mesh, taken from Silverman (1954), the harmonic coefficients were calculated with a 100 harmonic series expansion. With the difference between Silverman's isohipsae data and ours, a mean squared average error for our harmonic coefficients were evaluated by means of expression (30). For the n-th order ($\leq$ 100) estimator, the initial state matrix covariance uncertainty is obtained from (31). The renewal isohipsa data were assimilated at 2 hour time intervals. In our simulation, the measurement data were provided by the numerical integration of the barotropic model with 100 harmonics. In other words, the numerical simulation (with an explicit Euler method with integration step of two hours) with 100 harmonics acted as "reality". At each time interval, as in the initial assessment of uncertainty, the difference between the "reality" measurements and the n-th order estimator predicted measurements, was used to assign the state model uncertainty matrix Q_k (eqs. 30 and 31). With one time-interval lag, the update of Q_k provides the adaptive feature of the estimator with respect to modeling errors.

As we say in section 3 (eq. 12), the time-integration for the n-th order estimator was done with a further numerical approximation. While in the "reality continuous model", an Euler (2 hours intervals) was used, here, the time advance was done with the discrete model scheme (eqs. 10 to 13).

6.1 Effect of the approximation induced by the discrete nature of the estimator model

Figure 1 displays the estimator performance at a particular coordinate (15° north, 65° west). This coordinate coincides with a measurement point which has measurement error. To exhibit the effect of the discretization error, a full mesh data is used. The mesh consist of 1297 equidistributed (5° apart) measurements. In this case, the estimator acts as a model-based filter that reconciles, in a minimum variance sense, model and mesh data errors. If the measurement at that point is erroneous (bias, mishandling) the estimator is able to attenuate the effect by using data from all the mesh. In Fig. 1, Curve R (continuous) represents the "real" time evolution associated to the 15°n, 65°w isohipsa. Curve M represents the predicted value from an estimator without

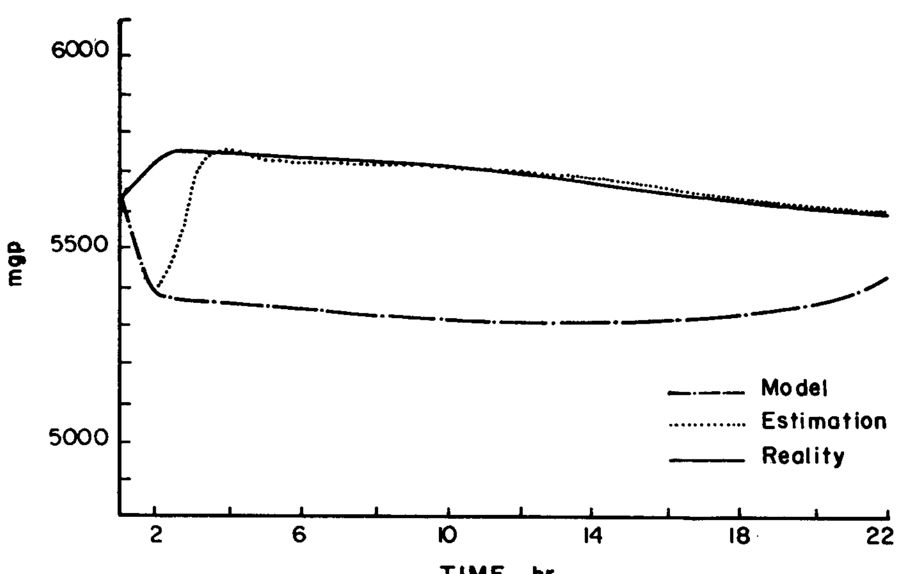

Fig. 1. Estimate of the (15° N, 65° W) geopotential hight from a full equidistributed (5° appart) mesh.

data assimilation. As mentioned before, the difference between R and M is due to the integration scheme. R used an explicit scheme based on an exponential transition matrix (see expression 11). Both R and M use 100 harmonics. Curve E is obtained when data are assimilated with the estimation scheme. As expected, the estimator is able to track the "reality" in spite of modeling inaccuracies.

6.2 The filtering-interpolation capability and the influence of the data location

In numerical approximation of partial differential equations (Voigt *et al.*, 1984), it is known that global methods (as expansions in harmonic series), lead to low dimension (number of expansion coefficients) approximation models (ordinary differential equations). In general, those methods suffer from a lack of robustness (ability to attenuate error propagation from equation residual error to solution profile error). The later limitation is aggravated as the approximation dimensionality increases. Adequate local methods (finite differences and finite elements) with adaptive mesh are better suited to cope with the robustness problem. However, the later methods require more elaborated implementations and lead to large dimensions in the approximated models. The implementation limitations are further enhanced when dealing with models in two and three dimensions (when altitude is considered). Orszag's (1970) global approximation is an ingenious global method which retains simplicity and produces acceptable robustness because uses tailored expansion functions. In this work, to obtain a small dimensionality in the estimation scheme, we have chosen Orzag's (1970) method. This should not be taken as a restriction because the estimation scheme can, in principle, be constructed from any model numerical approximation technique.

Because of the presence of feedback in integration, numerical integration is more robust than numerical differentiation. The presence of additional feedback is due to measurement corrections in the estimation algorithm. This fact offers the possibility of achieving robust predictions based on global methods while keeping the algorithms simple and the dimensionality small.

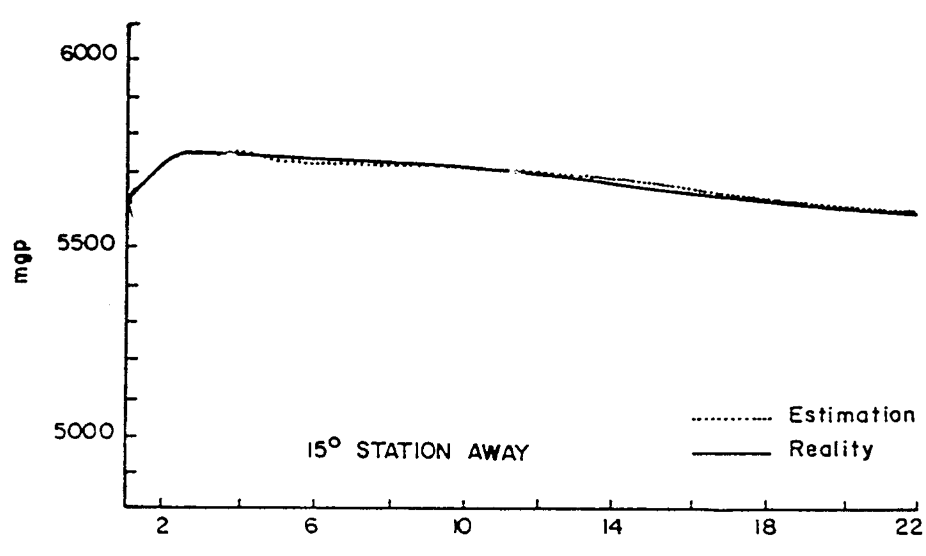

Fig. 2. Estimate of the (15° N, 65° W) geopotential hight from four [(0° N, 65° W), (30° N, 65° W), (15° N, 50° W), (15° N, 80° W)] surrounding measurements, which are 15° away.

Figure 2 shows the estimator performance when only four data are used to estimate an isohipsa (15° N, 65° W) which is not measured. The measurements [(15° N, 50° W), (15° N, 80° W), (0° N, 65° W), (30° N, 65° W)] surround, 15° apart, the location of the estimated isohipsa. The initial condition was taken with no error. In this case, the estimator produces a small error (about 50 m). This simulation exhibits the combined filtering-interpolation capability of the estimator. Figure 2a shows the performance when three measurement points [(15° N, 20° W), (15° N, 110° W), (60° N, 65° W)] , are used to estimate the isohipsa at (15° N, 65° W), in this case the information is generated 45° away, and estimate convergence is slower. Figure 3 shows the performance when three measurement points, 60° away [(15° N, 5° W), (15° N, 130° W), (75° N, 65° W)] are used to estimate the isohipsa at (15° N, 65° W). In this case, the measurements are so poor that the the estimator fails to track "reality", and a limitation of the estimator is brought up.

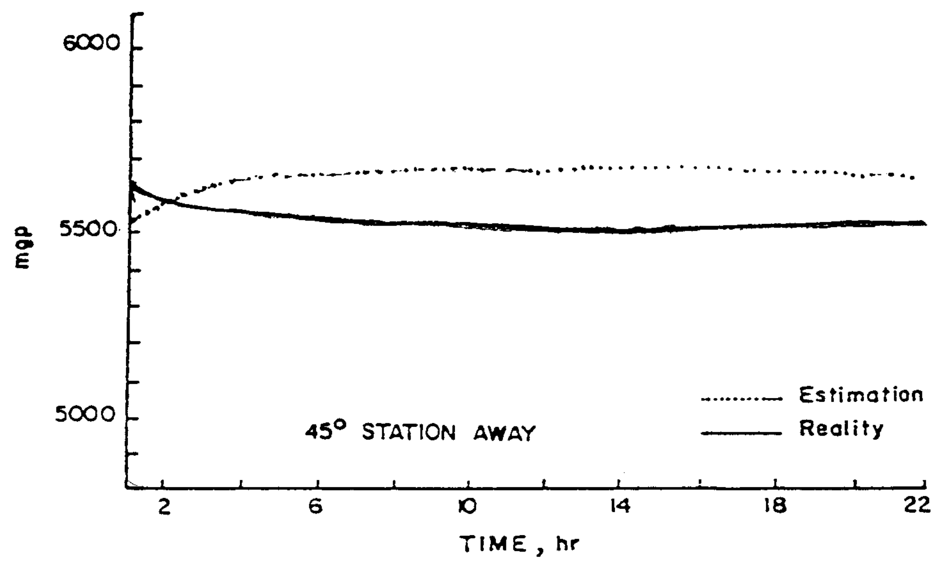

Fig. 2a. Estimate of the (15° N, 65° W) geopotential hight from three [(15° N, 110° W), (15° N, 20° W), (60° N, 65° W)] surrounding measurements, are 45° away.

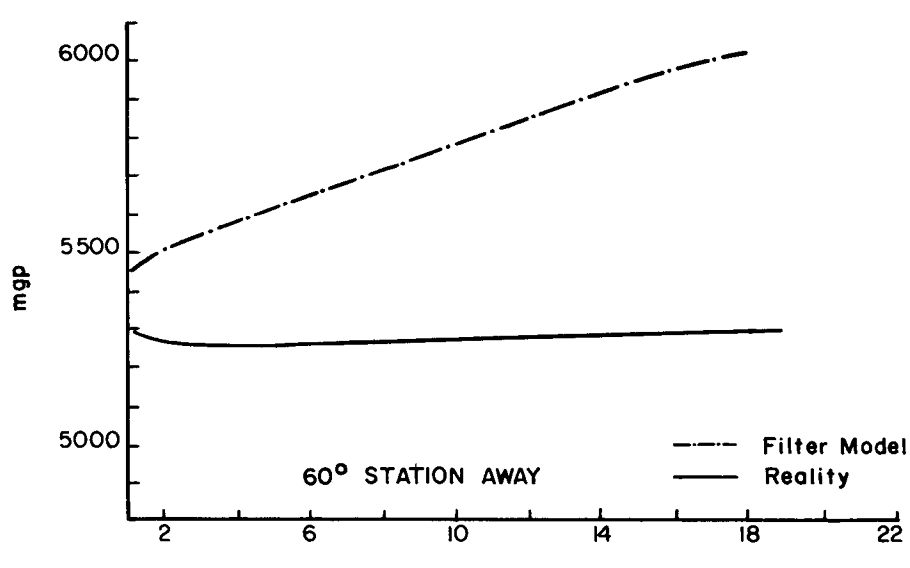

Fig. 3. Unsuccessful estimate of the (15° N, 65° W), geopotential hight from [(75° N, 65° W), (15° N, 5° W), (15° N, 125° W)] surrounding measurements which are 60° away.

The preceding simulations suggest that, for a certain region in the longitude-latitude space the design of an estimator includes the establishment of an adequate measurement mesh. A measurement mesh is characterized by the number and location of its measurement points. By adequate we mean a measurement mesh with a reduced number of elements which yields acceptable estimate performance (uncertainty and speed of convergence). A systematic treatment of the mesh election goes beyond the scope of this work. At this stage, it suffices to conclude that few measurements (15° away) are sufficient to provide acceptable estimates. The computational effort required by the estimation scheme is significantly reduced when few measurements are used. With few measurements, the quality of the estimate degrades as the measurements are scattered. In fact, in our case, at 60 away, the estimator breaks down.

6.3 Model order and robustness

In comparison to a predictor without data assimilation, the model-measurement estimation scheme should function with a rougher model approximation. This is so because the estimation feedback enhances numerical robustness (ability to tolerate numerical error propagation). Here we assess estimator performance in terms of the number of harmonics retained in the model approximation. To do so, we carried out a study of a process where 4 measurement stations [(10° N, 65° W), (20° N, 65° W), (15° N, 60° W), (15° N, 70° W)], forming a 5° side square, are used to estimate the isohipsa at point (15° N, 65° W). Fig. 4 shows the estimator performance

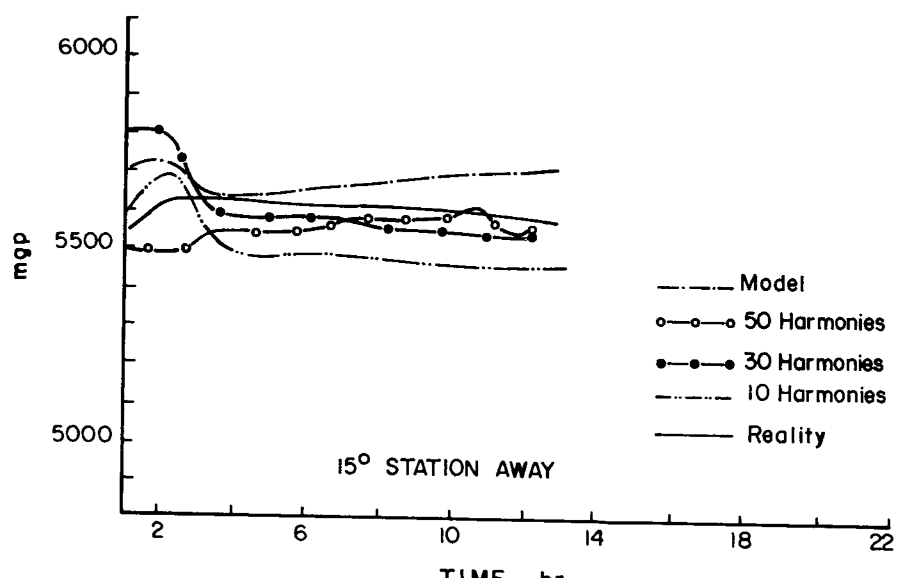

Fig. 4. Estimator performance as a function of the number of harmonics retained in the estimator. The (15° N, 65° W) geopotential is estimated from four surrounding measurements [(10° N, 65° W), (20° N, 65° W), (15° N, 60° W), (15° N, 70° W)] which are 5° away

when 100, 50, 30 and 10 harmonics are used. As it can be observed, the estimator is able to work, within acceptable precision, with only 10 harmonics. With less than 10 harmonics, the estimator diverges. Figure 5 shows the results when the measurement mesh is moved away from the prediction point. Now the measurement stations [(15° N, 50° W),(15° N, 80° W),(0° N, 65° W), (30° N, 65° W)], are 15° away. Here, the filter works satisfactory with 30 harmonics. When the measurements are moved 90° away the estimator requires at last 50 harmonics, this can be seen in Figure 6.

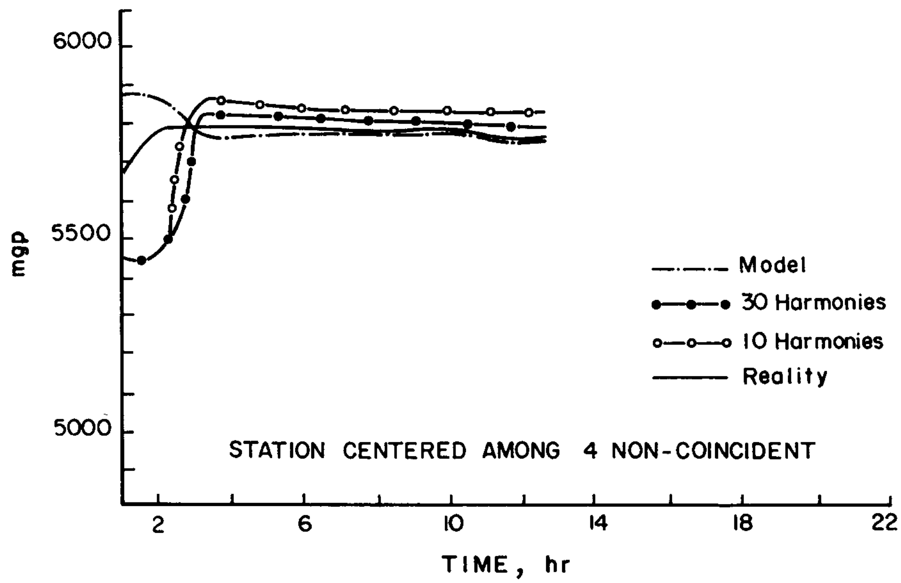

Fig. 5. Estimator performance as a function of the number of harmonics retained in the estimator. The (15° N, 65° W) geopotential is estimated from four surrounding measurements [(0° N, 65° W), (30° N, 65° W), (15° N, 50° W), (15° N, 80° W)] which are 15° away.

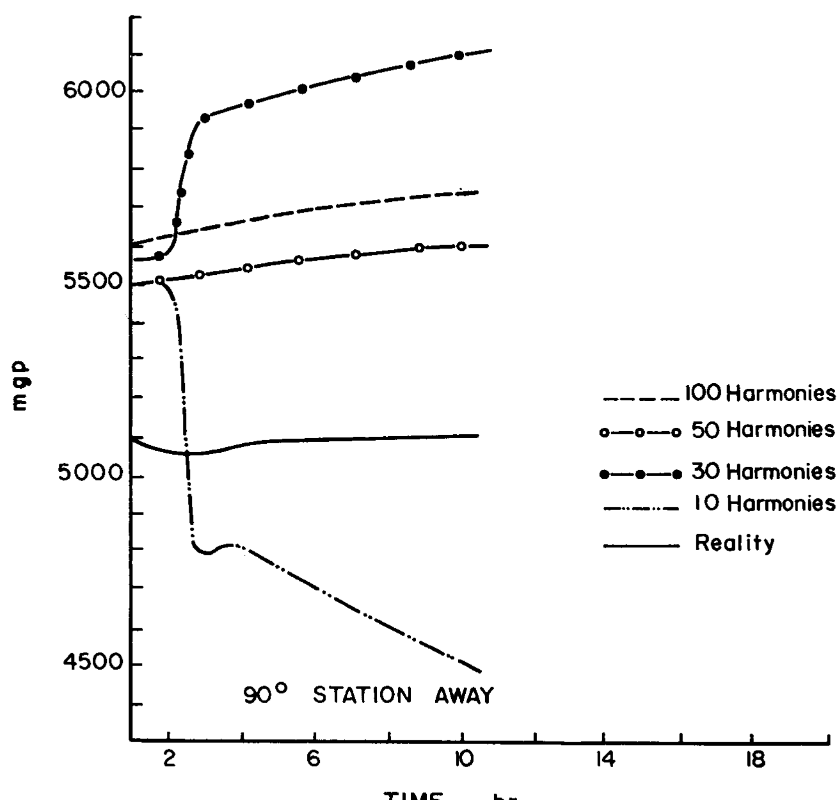

Fig. 6. Estimator performance as a function of the number of harmonics retained in the estimator. The (15° N, 65° W) geopotential is estimated from surrounding measurements [(75° N, 245° W), (15° N, 155° W), (15° N, 25° E)] which are 90° away.

Summarizing, the size of the estimator model depends on the size and location of the measurement mesh. As the measurement points are closer to the prediction point, the estimator error decreases. For example if measurements are 5° appart, there is a drastic (to 10 harmonics) estimator order reduction

7. Conclusions

Every modeling process involves a deviation from reality where every simplification, calculus or conception induce to more or less important errors, to avoid the real values. Only through a feedback process one may correct the estimations of any model so to do with any other model of greater efficiency, like our results show it.

As a part of a modeling process, the reduction in harmonics is able to yield a drastic reduction of computer times at developed models. In series like those used in the present work, this is not able of fact without the use of feedback processes because of the lost in preciseness.

Our results show the possibility to rebuilt useful preciseness, even, when the reduction in harmonics is so drastic as the use of only 10 harmonics (only 10% of the original design used by Orszag).

Data selection incorporated to a forecasting net, usually requires an objective process in computer, for the initialization of the process. However, the dynamic initialization process like the present is of a greater preciseness because, to every initialization, it incorporates the experience in the former estimations, generally not taken in account in modeling of static initial conditions.

As it has been shown it is possible to reduce significatively the density in data necessary to the initialization of the process because the sensibility in the equations to the data which is incorporated.

On the other side the dynamic covariance allows the sensibility of the initial condition to be adjusted giving as a result that corrections on introducing the feedback be as big as necessary and therefore permits a dynamic incorporation.

We consider the formulation in Kalman's theory is still able to give good contributions to the optimization in the atmosphere estimation and its frame of applicability is still wide and requires the most high finished attention.

Acknowledgements

The authors are indebted to Telma Castro Romero who contributed with many valuable suggestions, to Rafael Patiño Mercado for his aid in the computer programs. To Thelma del Cid for her help in the preparation of the final form of this article and to Ma. Esther Grijalva for the text processing.

APPENDIX A

THE MEASUREMENT MATRIX C

The stream and surface functions are approximated by a truncated double series:

$$\psi(\lambda,\ e,\ t) = a^2\Omega \sum_{i=0}^{10} \sum_{j=i}^{i+10} (\cup_j^i(t) \cos i\lambda + V_j^i(t) \sin i\lambda) P_j^i(\sin\theta) \tag{1}$$

$$\rho = \frac{a^2\Omega^2}{g} \sum_{m=0}^{10} \sum_{l=m}^{m+10} \left[A_l^m \cos m\lambda + B_l^m \sin m\lambda \right] P_l^m \sin\theta \tag{2}$$

The relation between geopotential and the stream function are given by Holton (1979):

$$\nabla^2 g\rho = 2\Omega \sin\theta \nabla^2 \psi \tag{3}$$

where g (gravity) is constant. From (1) and (2) the following laplacians are obtained:

$$\nabla^2\psi = -\Omega \left[\sum_{m=0}^{10} \sum_{n=m}^{m+10} (\cup_n^m \cos m\lambda + V_n^m \sin m\lambda) P_n^m n(n+1) \right]$$

$$\nabla^2\rho = -\frac{\Omega^2}{g} \sum_{m=0}^{10} \sum_{n=m}^{m+10} n(n+1)(A_n^m \cos m\lambda + B_n^m \sin m\lambda) P_n^m (\sin\theta)$$

Substitution of the last two expressions into equation (3) yields:

$$\Omega^2 \sum_{m=0}^{n} \sum_{l+m}^{m+n} (A_l^m \cos m\lambda + B_l^m \sin m\lambda) P_l^m (\sin\theta) l(l+1) =$$

$$2\Omega^2 \sin\theta \sum_{m=0}^{n} \sum_{l+m}^{m+n} l(l+1)(\cup_l^m \cos m\lambda + V_l^m \sin m\lambda) P_l^m \sin\theta$$

Recall the recurrence formula for associated Legendre functions:

$$\sin\theta P_l^m = \frac{1}{2l+1} \left\{ (l+m) P_{l-1}^m + (l-m+1) p_{l+1}^m \right\}$$

After comparison of power coefficients, one obtains the relationships between the two pairs of expansion coefficients:

$$\begin{pmatrix} U_{l-1}^m \\ V_{l-1}^m \end{pmatrix} = \left\{ \frac{1}{2} \begin{pmatrix} A_l^m \\ B_l^m \end{pmatrix} - \frac{l+2}{l} \left[\frac{(l+1)^2 - m^2}{4(l+1)^2 - 1} \right]^{1/2} * \right.$$

$$\left. \begin{pmatrix} U_{l+1}^m \\ V_{l+1}^m \end{pmatrix} \right\} \frac{l+1}{l-1} \left[\frac{4l^2 - 1}{l^2 - m^2} \right]^{1/2}$$

$$\begin{pmatrix} U_{l+1}^m \\ V_{l+1}^m \end{pmatrix} = \left\{ \frac{1}{2} \begin{pmatrix} A_l^m \\ B_l^m \end{pmatrix} - \frac{l-1}{l+1} \left[\frac{l^2 - m^2}{4l^2 - 1} \right]^{1/2} * \right.$$

$$\left. \begin{pmatrix} U_{l-1}^m \\ V_{l-1}^m \end{pmatrix} \right\} \frac{l}{l-2} \left[\frac{4(l+1)^2 - 1}{(l+1)^2 - m^2} \right]^{1/2}$$

This equation can be solved for the $\{A_i, B_i\}$ set: Then, the measurement equation is given by

$$y = Cx$$

where

$$C_{im} = \frac{a^2 \Omega^2}{g} \left[\frac{2i}{i+2} \sqrt{\frac{(i+2)^2 - m^2}{4(i+2)^2 - 1}} P_{i+1}^m + \right.$$

$$\left. \frac{2(i+1)}{i-1} \sqrt{\frac{i^2 - m^2}{4i^2 - 1}} P_{i-1}^m \right] \cos m\lambda.$$

REFERENCES

Alvarez, J., J. A. Romagnoli and G. Stephanopoulos, 1981. Variable measurements structure for the control of a chemical reactor. *Chem. Eng. Sci.*, **36**(5), pp. 787-902.

Balakrishnan, A.V., 1954. Elements of state space theory of systems. Optimization Software, Inc.

Bengtsson, L., 1971. An experiment in the Assimilation of Data in Dynamics Analysis. *Tellus* XXIII. **4-5**, pp. 328-336.

Castro, T., L. Le Moyne, and E. Villanueva. Predicción del campo de isohipsas a partir de la solución de una ecuación tipo elíptica por series de Fourier. *Geofís. Int.* **24**, pp 245-264.

Charney, J. G., R. Fjorboft and Von Neumann, 1950. Numerical integration of the barotropic vorticity equation. *Tellus* **2**, number 4, pp. 237-254.

Epstein, E., 1969. Stochastic dynamic prediction. *Tellus* XXI, **6**, pp. 739-759.

Epstein, E. S. and E. J. Pitcher, 1972. Stochastic analysis of meteorological fields. *J. Atmos. Sci.* **29**, pp. 244-257.

Espinoza, A., 1986. Solución de la ecuación de vorticidad en términos de series de esféricos Armónicos. Tesis, Facultad de Ciencias, UNAM.

Fleming, Rex J., 1972. Predictability with and without the Influence of Random External Forces. *J. Appl. Meteor.* **11** No. 8, pp. 1155-1163.

Ghil, M., S. Cohn, J. Tavantzis, K. Bube, and E. Isaacson, 1981. Application of estimation theory to numerical weather prediction. In Dynamic meteorology, data assimilation methods. Bengtsson, Ghil, and Kallen, eds. New York, Springer Berlang, 330 pp.

Ghil, M., 1989. Meteorological data assimilation for oceanographers. Part I: description and theoretical framework. *Dynamics of Atmospheres and Oceans,* **13**, pp. 171-218.

Holton, J. R., 1979. An Introduction to Dynamic Meteorology. 2o. Academic Press Inc.

Jazwinski, Andrew II, 1970. Stochastic Processes and filtering Theory. Academic Press. New York and London.

Jones, R, 1965. Optimal estimation of initial conditions for numerical prediction. *J. Atmos. Sci.*, **22**, pp. 658-663.

Kalman, R. E., 1960. A new approach to linear filtering and prediction problems. ASME Transactions, Part D (Journal of Basic Engineering), 82, 35-45.

Le Moyne, L., 1986. Modelo probabilístico de Pronóstico por Punto. *Geofís. Int.*, **25-3**, pp. 443-454.

Le Moyne, L., T. Castro, 1987. Uso de datos del pasado análisis dinámico. (Condiciones iniciales dinámicas en pronóstico numérico) Resúmenes U. G. M. Reunión 1987.

Lorenz, E., 1963. The Mechanics of Vacillation. *J. Atmos. Sci.* **20**, pp 448-464.

Meditch, J. S., 1969. Stochastic Optimal linear Estimation and Control. McGraw-Hill Book Company.

Miyakoda, K. and O. Talagrand, 1971. The assimilation of past data in dynamical analysis. I. *Tellus* XXIII, 4-5, pp. 310-317.

Miyakoda, K. and O. Talagrand, 1971. The assimilation of past data in dynamical analysis. II. *Tellus* XXIII, 4-5., pp. 318-327.

Orszag, S. A., 1970. Transform Method for the Calculation of Vector-Coupled Sums: Application to the Spectral form of the Vorticity Equation. *J. Atmos. Sci.*, *27*, pp. 890-895.

Petersen, D., 1968. On the concept and implementation of sequential analysis for linear random fields. *Tellus* XX, 4, pp. 673-686.

Phillips, N. A., 1986. The spatial statistics of random, geostrophic modes and first-guess errors. *Tellus*, **38**A, 314-332.

Pitcher, E. J., 1977. Application of stochastic dynamic, prediction to real data. *J. Atmos. Sci.*, **34** No. 1 pp. 3-21.

Silberman, I., 1954. Planetary waves in the atmosphere. *J. Meteor.* **11** pp. 27-34.

Voigt, R., Gottlieb, Y. Hussaini, 1984. Spectral methods for partial differential equations. SIAM, Philadelphia.