Variational and Optimization Methods in Meteorology: A Review

by

F. X. Le Dimet[1]

Université Blaise Pascal/(CLERMONT-2)
63177 Aubiere Cedex, France

and

I. M. Navon

Supercomputer Computations Research Institute
and Department of Mathematics
Florida State University
Tallahassee, FL 32306-4052

June 1988

[1] Work done while visiting SCRI.

Abstract

Recent advances in variational and optimization methods applied to increasingly complex numerical weather prediction models with larger numbers of degrees of freedom mandate to take a perspective view of past and recent developments in this field, and present a view of the state of art in the field.

Variational methods attempt to achieve a best fit between data and model subject to some '*a priori*' criteria – in view of resolving the undeterminancy problem between the size of the model and the respective number of data required for its satisfactory solution.

This review paper presents in a synthesized way the combined views of the authors as to the state of the art of variational and optimization methods in meteorology.

Issues discussed include topics of variational analysis, variational initialization, optimal control techniques, variational methods applied for numerical purposes and constrained adjustment, and finally how some of the variational and optimization methods discussed in the review relate to each other.

1. Introduction

In the last few years due to a constant increase in the need for more precise forecasting and nowcasting, several important developments have taken place in Meteorology directed mainly in two different directions:

a) Modelling at either large scale or at smaller scales. Recently, many models have been developed including an ever increasing number of physical processes and parametrization of subgrid phenomena.

b) Data: New sources of data such as satellite data, radar, profilers, and other remote sensing devices have led to an abundance of widely distributed data. However, a common characteristic of these data is to be heterogeneous either in their space or time density or in their quality.

Therefore, a cardinal problem is how to link together the model and the data. This problem induces several questions:

i) How to retrieve meteorological fields from sparse and/or noisy data in such a way that the retrieved fields are in agreement with the general behaviour of the atmosphere? (Data Analysis)

ii) How to insert pointwise data in a numerical forecasting model? This information is continuous in time, but localized in space (satellite data for instance)? (Data assimilation problem)

iii) How to validate or calibrate a model (or to invalidate it) from observational data? The dual question in this case being how to validate (invalidate) observed data when the behaviour of the atmosphere is predicted by a numerical weather prediction model.

For these questions a global approach can be defined by using a variational formalism.

1.1. Variational Methods in Meteorology: A Perspective

There are two main approaches employed when modeling a system described by a state variable, X. The first approach consists of finding a set of equations F such that X is the unique solution of the state equation

$$F(X) = 0. \tag{1}$$

In most cases system F must have as many equations as X has components in order to possess a unique solution – this is the problem of closure. In meteorology this problem has often been solved by using various artifacts such as adding supplementary equations. The second approach to the problem of closure is the variational one consisting in finding X as the solution of a problem of optimization *i.e.*by finding the extremum of some known functional J. Such an approach was proposed in theoretical mechanics some 245 years ago by Euler (1744, 1764) and by Lagrange (1760a, 1760b).

In the domain of numerical analysis Sobolev or Galerkin type methods are also based upon variational principles (Ritz (1908), Galerkin (1915)).

In meteorology, using the most general terms, we assume the state of the atmosphere to be described by a set of equations $G(X) = 0$.

If this system possesses fewer equations than unknowns, the system is said to be non-closed. However, one can still close it by introducing a variational approach.

If X_{obs} is an observation of a meteorological field, we will choose from among all the solutions of the system $G(X) = 0$ the solution closest to the observation X_{obs}. The resulting solution will be the optimal solution. In this manner a connection is established between the data and the observations.

In meteorology, the first application of variational methods has been pioneered by Sasaki (1955, 1958). Washington and Duquet (1963), Stephens (1966, 1968) and Sasaki (1969, 1970a, 1970b, 1970c) have given a great impetus towards the development of variational methods in meteorology.

In a series of basic papers Sasaki (1969, 1970a, 1970b, 1970c) generalized the application of variational methods in meteorology to include time variations and dynamical equations in order to filter high-frequency noise and to obtain dynamically acceptable initial values in data void areas.

In all these approaches, the Euler-Lagrange equations were used to calculate the optimal X.

Numerous other works applying these ideas appeared in the meteorological literature during the 1970's using the variational formulation. These works will be surveyed and classified in the later sections of this review. In parallel with the introduction of variational methods in meteorology, starting in the 1960's and 1970's, mathematicians in coordination with other scientific disciplines have achieved significant advances in optimization theory and optimal control, both from the theoretical viewpoint as well as from the computational one. In particular significant advances have been achieved in the development of optimization algorithms (Gill, *et al.* (1981), Fletcher (1980a, 1980b), Powell (1981), Bertsekas (1982), Lugenberger (1974) to cite but a few).

Optimal control methods have been introduced by Pontryagin, *et al.* (1960), and they have been generalized for systems governed by partial differential equations (Lions (1968)).

The application of an optimal control theory to meteorological problems has for the first time supplied the correct framework for a unified approach to analysis, data assimilation and initialization for meteorological problems.

Other techniques strongly related to variational and optimization theory, such as optimum interpolation, Kalman-Bucy filtering (Ghil, *et al.* (1981), smoothing splines Wahba (1975, 1981) Krieging, generalized cross-validation (GCV) Wahba and Wandelberger (1980) (for a unified approach see Lorenc (1986)) have also emerged during the last 10 years.

1.2. Variational Methods in Meteorology: The Optimization Theory View Point

Numerical weather prediction is based on the integration of a dynamic system of partial differential equations modeling the behavior of the atmosphere.

From a mathematical view point this approach is equivalent to the classical Cauchy problem. Therefore discrete initial conditions describing the state of the atmosphere have to be provided prior to the integration.

In order to retrieve a complete description of the atmosphere one can add information to the raw data using one of the following families of several methods:

a) Perform a simple interpolation, *i.e.*, no information is added to the data. This procedure is purely algorithmic.

b) Add as information the statistical structure of the fields *i.e.*, use an optimal interpolation type method. Unfortunately this information is not always available or may be inadequate for instance as is the case with a paroxysmal event.

c) A third way is the variational method. Variational methods are based on the fact that a given meteorological observation has not an intrinsic credibility. The same measurement of wind, to give just an example, may be used to study the flow around a hill, or may be inserted in a mesoscale model, or may be used in a global model of atmospheric circulation. According to the particular framework where the data will be used, variable trust will be attributed to the same data.

Variational methods try to achieve a best fit, with respect to some '*a priori*' criterion, of data to a model by placing the data into the most adequate framework where it should be used, and permits us to link the data and the model.

In the first part of the paper we will show how variational methods can be defined and which are the ingredients necessary to build a variational method, all this in the perspective of the surveyed accumulated work. Then we will show how to solve related variational problems in the framework of a systematic classification of the reviewed work. This classification will permit us to review different variational methods as well as the context in which they were performed.

The last section will be devoted to future developments and potential applications of variational methods in meteorology.

2. Ingredients of a Variational Method

2.1. Definition of a Variational Method

In the most condensed way a variational method may be defined as a search, amongst all the possible solutions of a model, of the solution closest to a given observation. Therefore a variational method will be defined by the following ingredients:

i) An atmospheric variable X – describing the state of the atmosphere.

ii) A model which may be mathematically written as:

$$B\frac{dX}{dt} + A(X) = 0 \tag{2}$$

where B is either the null-operator for a steady state model, or the identity operator for a dynamical model. A is a linear or non-linear operator.

We suppose that system (1) is not closed by which we mean that in order to obtain an unique solution to (1) some additional information has to be provided.

iii) U – a control variable. U may be comprised of the initial conditions, boundary conditions, or both, the vector X itself or a part of it. Once U is defined – a unique solution $X(U)$ of (2) will be associated with it.

The vector control variable U must belong to some set of admissible control U_{ad}. The definition of U_{ad} may include physical information which can be stated in the form of inequalities.

iv) J, a cost function measuring the difference between a solution of (1) associated with U and the observations X_{obs}.

v) An observation X_{obs} of the meteorological fields.

A variational problem is determined in terms of these last five items and it can be stated as problem (P) *i.e.*:

(P) Determine U^* which belongs to U_{ad} and minimizes the cost function J.

The second stage of the solution of the variational problem will be to determine, or at least to approximate U^* (and therefore the optimal associated state of the atmosphere $X(U^*)$.

In order to achieve this, we first have to set up an optimality condition and then to perform an algorithm for solving problem (P).

2.1.1. The optimality Condition

A general optimality condition is given by the variational inequality (see Lions (1968))

$$(\nabla J(U^*), V - U^*) \geq 0 \text{ for all } V \text{ belonging to } U_{\mathrm{ad}}, \tag{3}$$

where ∇J is the gradient of the functional J with respect to the variable U.

In the case where U_{ad} has the structure of a linear space, variational inequality (3) is reduced to the equality

$$\nabla J(U^*) = 0 \tag{4}$$

2.1.2. The Algorithm of Solution

As stated above – variational problems are problems of optimization with or without constraints. There exist standard procedures (Le Dimet and Talagrand (1986), Navon and Legler (1987)) to solve them.

A common requirement of these procedures is the need to explicitly supply the gradient of J with respect to U to the code.

Moreover, the basic problem to be solved is always a problem of unconstrained minimization for which the method of conjugate gradient may be used (see Navon and Legler (1987)).

2.1.3. Variational methods: for which purposes?

The first applications of variational methods were for objective analysis of meteorological fields, *i.e.*to retrieve fields from pointwise distributed data in space. In most of the important meteorological situations the temporal evolution of the fields is crucial, therefore, some attempts were carried out towards extending variational analysis to dynamic analysis. Introducing sparsity of data in time using variational tools has led to 4-D data assimilation for numerical weather prediction models. To perform a forecast a meteorological model requires an initial condition. This initial condition must be as close as possible to the observations while remaining compatible with the model. The problem of initialization may be stated as a variational problem and solved in this way.

A general formalism of variational problems has to deal with observations but these observations may not necessarily be physical ones. For instance they may result out of a numerical model (output of a numerical model). Furthermore, the constraints imposed upon the analysis may have no physical origin and could only have been introduced for numerical purposes.

Many applications were carried out in similar situations as mentioned above resulting in a global approach of variational methods, such as for instance enforcing conservation of integral invariants in numerical models (Navon (1981), Navon and de Villiers (1983)), or design of discretization schemes (Sasaki (1976)). A major difficulty for the classical approach to variational methods for meteorologically significant problems, in particular for those where dynamics play a prominent part, is the fact that the size of the discrete problem to be solved is prohibitive.

A way to circumvent this difficulty is to introduce optimal control methods permitting a significant reduction of the problem size. These techniques, upon which we will expand in a later section, introduce the adjoint of the numerical model. Knowledge of the adjoint of the model turns out to be particularly useful, because it can be applied towards a sensitivity analysis (Hall and Cacuci (1982, 1984)) or for environmental studies such as the estimation of the impact of industrial pollution upon the environment (see Marchuck (1982)).

In this review paper we will present the most important contributions concerning applications of variational methods using the general formalism of mathematical programming.

3. Variational Analysis

Basically, the problem of retrieving meteorological fields X from observations $\widetilde{X}$, in such a way that X verify some model:

$$F(X) = 0 \tag{5}$$

and are as close as possible, in the sense of a given functional J, to the observations $\widetilde{X}$, is a problem of optimization with constraints.

Sasaki (1970) in historical paper has introduced two formalisms:

a) The *weak constraint formalism* consists in minimizing without constraint the functional J defined by

$$J_1(X) = J(X) + K\|F(X)\|^2. \tag{6}$$

It is easily seen that for large values of K, $F(X)$ has to be small for minimizing J_1, therefore, for a specified value of K, constraint (5) is only approximately verified. In what follows K is a generic constant used as a coefficient of a weak constraint. This is justified by the fact that equation (5) is not a perfect representation for the atmosphere and therefore should not be satisfied with a greater precision than its own accuracy.

The optimal condition, which in the Euler-Lagrange equation gives the optimal analyzed field X^*, is the solution of the equation

$$\nabla J_1(X^*) = \nabla J(X^*) + 2K\ F'(X^*) \cdot X^* = 0. \tag{7}$$

In this equation ∇J_1(respectively ∇J) is the gradient of J_1(respectively ∇J_1) with respect to X, while F' is the Jacobian matrix of F. No standard method exists for solving (7). As such a method of solution has to be chosen in agreement with the particular expressions for J and F. In the majority of cases, and even always when F is non-linear, an iterative algorithm has to be carried out.

b) The *strong constraint formalism* imposes upon the optimally analysed field X^* to exactly verify equation (5) (in fact only up to discretization and round-off errors). To implement this condition the Lagrangian $L(X, \Lambda)$ is introduced, given by

$$L(X, \Lambda) = J(X) + (F(X), \Lambda) \tag{8}$$

where Λ, the Lagrange multiplier, has the same dimension as F. The Euler-Lagrange optimality condition, gives X^* and Λ^* and may be written as

$$\frac{\partial L}{\partial X}(X^*, \Lambda^*) = 0 \tag{9}$$

$$\frac{\partial L}{\partial \Lambda}(X^*, \Lambda^*) = 0 \tag{10}$$

As before, no standard method exists for solving system (5), and in the majority of cases, X^* is eliminated between (9) and (10) leading to a unique system for Λ^*. X^* is then computed using equation (10).

Therefore, a variational analysis is defined by different choices of the cost function J of the model F, and by the method of resolution. We shall now briefly survey the main choices for these principal ingredients.

3.1. Choices of the Cost Function J

The prime objective of J is to measure the proximity between an observation and a solution of the model. So J must have the property of a norm. Variational methods are based upon the computation of the gradient, therefore, the functional J has to be differentiable, which is easily implementable as most of the time J is the square of a norm. If for instance wind and geopotential fields are observed, the functional J assumes the form

$$\begin{aligned} J(u, v, \phi) = \sum_i \sum_j & \alpha(i,j)\,(u(i,j) - \tilde{u}(i.j))^2 + \beta(i,j)\,(v(i,j) - \tilde{v}(i,j))^2 \\ & + \gamma(i,j)\left(\phi(i,j) - \tilde{\phi}(i,j)\right)^2 \end{aligned} \tag{11}$$

where u, v are the wind components, ϕ is the geopotential, the summation is extended on the whole domain, and α, β, and γ are weights. These weights have a dual purpose:

i) make J a non-dimensional quantity

ii) reflect the confidence we have in the quality of the observed data.

Several choices are possible for these coefficients. One of the most often used weights is the Gauss precision moduli, which are defined as the reciprocal of twice the variance of the errors of observation for the respective observed elements.

Sasaki (1971) studied a theoretical interpretation of anisotropically weighted smoothing on the basis of numerical variational analysis. The results he obtained suggest that the weights for the upstream and downstream observations should be of the same magnitude and as much as three times larger than the respective weight for the crosswind direction. This work is also related to the anisotropic weighting factors for Cressman objective analysis scheme (Endlich and Mancuso (1968)).

In almost all analysis applications these weights are taken to be constant, except in the vertical coordinate. Another criterion for choosing the weight functions is to render the numerical methods used to solve the optimality system, convergent. If the problem of analysis is stated as a problem of mathematical programming then the standard codes of unconstrained optimization include an automatic scaling of the variables.

The cost function may also include terms which act as filters in time or in space (see Sasaki (1970b)). For a given analyzed field φ, the addition of a term in the form

$$K\left(\frac{\partial^2\varphi}{dx^2}+\frac{\partial^2\varphi}{dy^2}\right) \tag{12}$$

will tend to smooth the curvature of the φ-field. Wahba and Wandelberger (1980) used a functional in the form

$$J_m(\phi)=\sum_{\alpha_1+\alpha_2+\alpha_3+\alpha_4=m}\frac{m!}{\alpha_1!\alpha_2!\alpha_3!\alpha_4}\int\int\int\int\left(\frac{\partial^m(\varphi)}{\partial x^{\alpha_1}\partial y^{\alpha_2}\partial p^{\alpha_3}\partial t^{\alpha_4}}\right)dx\,dy\,dp\,dt. \tag{13}$$

where x, y are space variables for the analysis of a field φ, where p the atmospheric pressure is used as the vertical coordinate, m is an integer to be determined, and the coefficients are given in such a way that an explicit representation for the minimizer can be found.

Sasaki (1970a) also used a so called timewise localized formalism with one of the constraints assuming the form:

$$K\left(\frac{\partial\varphi}{dt}\right)^2 \tag{14}$$

where the local time derivative is usually replaced by the other terms of a conservation law. This term acts as a penalty term and allows only a slowly varying evolution of the field. This approximation is valid only for quasisteady state events.

If more statistical information is available it may be inserted in the weight function to enhance the static consistency of the fields X. Such a term may have the form (Lorenc (1986)):

$$K(\varphi - \varphi_c)C^{-1}(\varphi - \varphi_c) \tag{15}$$

where φ_c is a climatological value for the variables φ and C is the covariance matrix for φ.

A classical method of mathematical programming for solving the familiar problem of constrained optimization given by

$$\min J(X) = 0 \tag{16}$$

subject to the equality constraints:

$$F(X) = 0 \tag{17}$$

is to introduce the penalized functional:

$$J_\epsilon(X) = J(X) + \frac{1}{\epsilon}\|F(X)\|^2 \tag{18}$$

and to minimize $J_\epsilon(X)$ without constraint giving the optimal solution X_ϵ^*. Here ϵ represents a sequence of real numbers used as penalty parameters and tending to zero. It is very easy to see the similarity of this approach with Sasaki's weak constraint. With some additional hypothesis it can be shown that X_ϵ^* tends to X^*, solution of the constraint problem at the limit. Let us make three remarks:

a) With the weak constraint formalism only one step of a penalty method is performed. Therefore, the dependence of the solution of the weak constraint method upon the coefficient K cannot be exhibited. This dependence is more easily shown using a penalty term and different values of ϵ.

b) In the weak constraint formulation the constraint does not have to be exactly satisfied. Also, it is not possible to control to what extent the value of the deviation of the constraint at the final stage from zero is due to physical or numerical reasons. Theoretically, for penalty type methods at the optimum, the constraint is exactly satisfied. Nevertheless, the method of optimization without constraint which has to be used to solve the minimization problem needs to specify some stopping criteria which permit control to the satisfaction of the constraints. Choices of the constants used in these stopping criteria have to be carried out based on physical considerations.

c) A well known fact is that for small values of ϵ, *i.e.*for values of the variable close to the optimum, the method of unconstrained optimization may turn out to be ill-conditioned, leading to serious numerical problems. A way to deal with this difficulty is to introduce the Augmented Lagrangian approach for the problem of optimization with constraints. This is done in Navon and de Villiers (1983) and in Le Dimet and Segot (1986). These methods used a so called Augmented Lagrangian L defined by:

$$L(X, \Lambda) = J(X) + \frac{1}{\epsilon}\|F(X)\|^2 + (\Lambda, F(X)) \tag{19}$$

where the vector Λ is the Lagrange multiplier of the constraint F. In Sasaki's terminology F is considered both as a weak constraint and a strong one. Standard methods (Bertsekas (1975, 1982), Fortin and Glowinski (1983)) exist for solving this problem. A major advantage of this method is its ability to prevent numerical instabilities. The gradient of the Augmented Lagrangian has to be computed for both variables, but this task does not need more computation and/or storage than either the penalty or the duality methods.

3.2. Choices of Models

In variational analysis, models are used as constraints to fit the analysis to the data. Of course, the quality of the analysis depends upon the quality of the model which is used for the adjustment. A wrong model cannot give a good analysis. Therefore, during the practical realization of a variational method, the quality of the model, has to be kept in mind.

Applications of variational analysis were conducted for a multitude of case studies. Some operational uses have been done by Lewis (1972) for the upper air analysis on the Pacific Ocean and also by Seaman, *et al.* (1977) for the analysis in the Australian Region.

Variational methods are particularly well adapted for events with sparse data and irregular fields. Many studies were carried out on squall lines, for instance by Charba and

Sasaki (1981), who use a gravity current model for studying a squall line. Sasaki and Lewis (1970), Lewis (1972) used a variational formalism for the analysis of squall line and severe storms. Sheets (1973) has developed a variational optimization technique for an analysis scheme and applied it to the high portion of a hurricane, as well as to study the presumed effect of seeding on the hurricane.

Many applications have been performed using the divergence free constraint (Sherman (1978)) and especially for environmental studies (Wilkins (1971), McCraken*et al.* (1978)). See also O'Brien(1970).

Soliz and Fein (1980) have developed a so called Pattern Conserving Technique (PCT) which is used to obtain 3 dimensional grid fields of wind, temperature and height.

Variational methods have never been implemented using an operational model as a constraint. The main reason for this is the lack of systematic approach for the algorithmic side of the variational method. For each problem an adequate numerical method was designed. A major contribution was the mathematical programming formulation of variational methods which consist of considering the problem in its algorithmic perspective where standard and high performance codes of optimization can be used.

More and more data are available from remote sensors, being provided mainly by radars or by satellites. Classical meteorological fields such as wind, temperature, and humidity are not directly measured by remote sensing but they can be estimated using mathematical inversion methods in combination with empirical laws for measured quantities such as reflectivity for radar or spectral bands for satellites. As such these data may be of poor accuracy. Another common property of remote sensors is to provide data with very heterogeneous resolution, for instance radars give information only in regions having reflectivity.

Interpolation has no sense in this context, statistical methods cannot be carried out in the absence of elementary information data for statistics. Therefore, variational methods are well suited for this type of data because they impose upon the retrieved fields a physical consistency through the model equations which are used as constraints.

Ray *et al.* (1980), and Ziegler (1986) studied air flow in convective storms using Doppler radar observations. In their application, a variational analysis simultaneously imposes two kinematic boundary conditions and the mass continuity equation on Doppler velocities to derive the three-dimensional thunderstorm air motion.

The analysis fields are obtained by minimizing

$$E = \int\int \left\{ \int \left[\alpha^2 (u - \tilde{u})^2 + \beta^2 (v - \tilde{v})^2 \right] dz + \Lambda \int \rho \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) dz - C \right\} dx\, dy \qquad (20)$$

where $\tilde{u}$, $\tilde{v}$ are the observations, and u, v are the final fields, ρ the air density, α^2 and β^2 are Gauss precision moduli, and Λ a Lagrange multiplier.

The adjustment requires the integrated density weighted horizontal divergence from the surface to a height Z_t to be a constant defined by

$$C = \int_0^{Z_t} \rho \left\{ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right\} dz = - \int_0^{Z_t} \frac{\partial \rho w}{\partial z} dz \qquad (21)$$

Testud and Chong (1983) and Chong *et al.* (1983) also use a variational procedure to retrieve data from Doppler radar observations. In a first step, the retrieved fields are computed in such a way that they minimize the distance to the observation and verify some regularity condition (finite curvature for the retrieved fields). In a second step, (Chong *et al.* (1983)), the error in the retrieved 3-dimensional wind field due to temporal variation is minimized.

In satellite meteorology, variational methods were used by Ghil and Mosebach (1978) for asynoptic data assimilation. Hoffman (1982, 1984) removed ambiguity from Seasat-A satellite scatterometer by choosing the alias closest to the analyzed field. The application of his method was illustrated for a limited region in the North Atlantic. The method involved direct minimization of a functional. His analysis uses both satellite and conventional data applied to the study of a storm. Cram and Kaplan (1985) developed a variational method to assimilate VAS temperature and moisture gradient information into a mesoscale model. The constraint they used is to match the VAS data gradient and the model first guess absolute value. The functional which is minimized is given by

$$J = \int\int \left[A(v - \tilde{v})^2 + B \left(\frac{\partial v}{\partial x} - \frac{\partial \hat{v}}{\partial x} \right)^2 + C \left(\frac{\partial v}{\partial y} - \frac{\partial \hat{v}}{\partial y} \right)^2 \right] dx\, dy. \qquad (22)$$

Here $\tilde{v}$ is the model first guess variable, v is the model adjusted variable, while $\hat{v}$ is the VAS data variable. A, B, and C are matrices of weights. A is defined at each grid point where

VAS data were analyzed and set to zero elsewhere, while B and C are defined between grid points and set to zero whenever VAS data is not defined at both surrounding grid points.

VAS data together with geopotential derived from rawinsondes (RAOB) are analyzed by Lewis *et al.* (1983). Two methods are developed. The first incorporates the statistics of RAOB-derived potential vorticity into the VAS vorticity analysis by making a least square adjustment with the constraint to have the first and second moments identical to the RAOB analysis. The second method implements a mutual least square adjustment to RAOB and VAS vorticity with the constraint being that forecast and hindcast of potential vorticity to the time midway between the analyses are equal.

Sasaki and Goerss (1982) developed a method for analyzing and assimilating data into a baroclinic primitive equation model. The model is discretized using a staggered grid system with centered space and time differences. At each synoptic time a variational procedure is performed to combine the newly acquired upper air and surface observations with the first guess fields. The functional which is minimized is given by

$$\begin{aligned} J = \int\int\int \Bigg\{ & A(\phi - \hat{\phi})^2 + B\left(\frac{1}{R}\cdot\frac{\partial\phi}{\partial\pi} - \widehat{T}\right)^2 + C_u(u-\hat{u})^2 + C_v(v-\hat{v})^2 \\ & d\left(\nabla\phi - \nabla\tilde{\phi}\right)^2 + e\left(\nabla^2\phi - \nabla^2\tilde{\phi}\right) + \ell(\zeta - \hat{\zeta}) + g(D - \widetilde{D})^2 \\ & + h\left(\frac{1}{a\cos\theta}\cdot\frac{\partial\phi}{\partial\lambda} - fv\right)^2 + i\left(\frac{1}{a}\cdot\frac{\partial\phi}{\partial\theta} + fu\right)\Bigg\}\, d\lambda\, d\theta\, d\pi \end{aligned} \tag{23}$$

here θ and λ are spherical coordinates, π the vertical coordinate, f is the coriolis parameter, A, B, C_u, and C_v are three dimensional weight matrices. $\tilde{\phi}$, $\tilde{u}$, and $\tilde{v}$ are first guess field values while the observations are given by $\hat{\phi}$, $\hat{u}$, and $\hat{v}$. The weights d, e, ℓ, g, h, and i are constant on the domain. Therefore, the resulting analysis will adjust at best the observation with the constraint of having the same horizontal gradient, Laplacian, vorticity and divergence as the first guess field. When considered as penalty terms, the two last terms in equation (23) will enforce the geostrophic equilibrium between analyzed fields. The quality of the adjustment to the geostrophic equilibrium will depend upon the relative values of the coefficients of the cost function.

3.3. Analysis of Dynamic Data

We assume that a model of the atmosphere is given by the differential equation

$$\frac{dX}{dt} = F(X), \quad 0 \le t \le T. \tag{24}$$

and let $X_{\text{obs}}(t)$ be an observation of the atmospheric fields during the same period of time $[0, T]$. There are several ways to endow the analyzed fields with dynamical consistency.

The first method is the time-wise localized formalism in Sasaki's terminology. It consists of considering that at any moment the fields are only slowly evolving, *i.e.* $\frac{dX}{dt}$ has to be small. At a given moment, the optimal analysis will be the closest to the observation, subject to the weak constraint that $\frac{dX}{dt}$ remains small, therefore, the time derivatives are introduced as penalty terms in the cost function. J, the cost function, will be defined as:

$$J(X) = \int \left((X - X_{\text{obs}})^2 + C \left(\frac{dX}{dt} \right)^2 \right) d\Omega \tag{25}$$

where C is the penalty coefficient.

Some remaining questions are:

a) What is the dependence of the optimal analysis upon the choice of constant C?

b) If $\frac{dX}{dt}$ has to be small, then C has to be a large value, therefore, causing the numerical solution to be ill-conditioned.

c) How to link together two successive analyses? Relative values of $\frac{dX}{dt}$ at times T and $T + \Delta t$, could be different, leading to numerical noise in the analysis.

d) The norms, which are chosen to measure the proximity between the analysis and the observation and to impose the constraints, are of the L_2 type, therefore, permitting localized high values for the constraints. This situation may arise in limited area analysis if the boundary terms are not carefully discretized.

A way to deal with these difficulties would be to introduce (24) as a full constraint and to use the Augmented Lagrangian formulation:

$$\begin{aligned} L\left(X(t), \Lambda(t), \epsilon\right) = \int_0^T \int_\Omega & (X(t) - X_{\text{obs}}(t))^2 + \epsilon \left(\frac{dX}{dt} - F(X) \right)^2 \\ & + \left(\Lambda, \frac{dX}{dt} - F(X) \right) d\Omega \, dt \end{aligned} \tag{26}$$

where ϵ is a penalty coefficient tending to zero, Λ is a Lagrange multiplier function of time and space and Ω is the spatial domain of interest.

As pointed out above, the optimal analyses X^* associated with the optimal Lagrange multipliers are solutions of the system

$$\begin{aligned} \frac{\partial L(X^*, \Lambda^*)}{\partial X} &= 0 \\ \frac{\partial L(X^*, \Lambda^*)}{\partial \Lambda} &= 0 \end{aligned} \tag{27}$$

which can be solved by using Bertsekas' type of Augmented Lagrangian algorithm (Bertsekas (1982), Navon and de Villiers (1983)).

System (25) has to be discretized together with equation (24) using a time differencing scheme. The dimension of system (27) to be solved is equal to the dimension of X multiplied by the number of time-steps contained in the time interval $[0, T]$. For non-trivial problems we will obtain problems of very high dimensionality which may not be practical even for present large mainframe computers.

A method for imposing dynamical consistency to the analyzed fields has been proposed by Thompson (1969): the observations of the geostrophic vorticity at two successive observation times are adjusted at these two times subject to the constraint that they satisfy the barotropic vorticity equation. This method has been extended and applied by Lewis (1980) for the adjustment of vorticity at the level of non-divergence. Analyses were performed on the hemispheric scale in such a way that more than one disturbance center can be identified and that the region encompasses both rich and sparse data areas. The results show that time continuity, as well as the order of magnitude of vorticity adjustment are improved by this technique. This technique can be applied for the more data-void regions and can be patched with observations from a neighboring time.

Lewis and Bloom (1978) described a method for coupling two observations in time by using the forecast equations of horizontal momentum as dynamical constraints. The technique can be extended, always working on pairs of either observations or analyzed fields to enforce the dynamical consistency of the analysis. The scheme is tested on a squall line case, using the hourly surface observations from the Aviation network as data. The results of this study show that the build up of the convergence zone is better depicted by the variationally adjusted patterns and the correlation between surface convergence and the radar echo is enforced. differences between observed and analyzed fields are of the order of m sec^{-1}. The same approach is used by Bloom (1983) for the analysis of mesoscale rawinsonde data, the dynamical constraint being a set of forecast equations of horizontal momentum. This type of analysis permits computing the vertical velocity, which is shown to be more consistent with the weather events that occurred during the unconstrained case study.

3.4. Discretized Variational Analysis

To obtain the solution of a variational problem, a discretization has to be carried out on the constraint and on the cost function. In many papers, the Euler-Lagrange equations are written in continuous form and only then discretized. It has been pointed out by many authors that to discretize a variational problem and then to write the Euler-Lagrange equations of the discrete problem is not equivalent to writing first the Euler-Lagrange equations and then discretizing them. If the model, used as a constraint, has a non-linear term, then the Euler-Lagrange equations will have non-standard boundary conditions, which have to be simplified in order to solve the system of optimality. In many papers, simplifications applied on the boundary terms have no physical justification. Therefore, it is simpler to write the Euler-Lagrange equations, or to compute the gradient of the cost function on the discrete problem. Nevertheless, the boundary terms have to be carefully discretized in order to impose consistency between the discrete problem and the continuous one.

In many applications the discretizations were carried out using a finite difference approximation scheme in space. Some finite element discretizations were performed in several cases for studying wind field adjustment over complex terrains (Tuerpe, Gresho and Sani (1978)). Racher and Roset (1985) carried out a three-dimensional analysis of wind field over Hawaii, using a constraint of free divergence. Le Dimet and Segot (1987) proposed an algorithm for the adjustment of the 500 mb wind field. Their method implemented an Augmented Lagrangian algorithm discretized with the finite-element method using the observation stations as nodes of the finite elements. An advantage of the finite element method is that it does not necessitate a preliminary interpolation of the fields. A shortcoming of the finite element method is its higher computational cost, when compared to the finite difference discretization method.

3.5. Inequality Constraints

The general theory of optimization permits to introduce inequality constraints in the variational formalism. Real physical situations have to be simulated using inequality constraints. Sasaki and Goerss (1980) employed such a method for the adjustment of absolutely unstable atmospheric layers. The vertical temperature profile is adjusted in such a way that the vertical gradient of temperature is larger than some given value. The numerical method used for solving this problem introduce the so called slack variables. Modern theory of mathematical programming (Gill, Murray, and Wright (1981)) provides a wealth of mathematical tools for working with inequality constraints. In our opinion, important progress can be achieved by using this type of constraints with adequate mathematic tools such as the Augmented-Lagrangian method (see Bertsekas (1982), Navon and de Villiers (1983)).

As a partial conclusion, one can say that variational methods for the analysis of static fields are well suited for many meteorological situations, and by bringing the information contained in the physics of the atmosphere to the data they permit a coherent retrieval of meteorological fields. A synoptic view of various research works using variational methods is presented in Table 1.

4. Variational Initialization

The aim of initialization in meteorology is to prepare objectively analysed gridpoint data with a minimum of spurious high-frequency inertia-gravity noise while retaining accuracy of the forecasts for the meteorological scales of interest in the model.

Early variational initialization applied "dynamic" constraints such as the balance equation, hydrostatic relation or geostrophic balance (see Barker, Haltiner and Sasaki (1977), Haltiner, Sasaki, and Barker (1975), Haltiner and Barker (1976), and Stephens (1970)).

Lamb-waves were eliminated by enforcing as strong constraint the vanishing of the integrated mass divergence (Barker *et al.* (1977)).

Typically the cost function for the balance equation constraint takes the form

$$I = \int \int \left[\alpha \left(\phi - \tilde{\phi} \right)^2 + \beta \left(\nabla \Psi + k \times \widetilde{V} \right)^2 + 2\lambda M \right] ds, \qquad (28)$$

where Ψ is the stream function to be determined, α and β are confidence weights, ϕ is the geopotential, J is the Jacobian operator, λ is a Lagrange multiplier, ds is an area element on the sphere, and V is the velocity field, while the tilde $\tilde{()}$ denotes objectively analyzed fields.

$M = 0$ is the equality constraint of the vanishing of the non-linear balance equation.

$$M(\phi, \Psi) - f\nabla^2\Psi + \nabla f \cdot \nabla\Psi + 2J\left(U_\Psi, V_\Psi\right) - \nabla^2\phi = 0 \tag{29}$$

In this approach a mutual adjustment of the wind and mass fields is achieved while attempting to satisfy exactly or approximately the classical non-linear balance equation constraint. More recently with the advent of more general balance relationships such as the non-linear normal mode initialization (N.M.I.), variational initialization procedures have been put forward attempting to achieve an adjusted state that

a) is on a presumed slow manifold,

b) fits the good data (high accuracy or high confidence data) as well as possible and fits poor data (low confidence data) less exactly.

Daley (1978) proposed a variational formalism for the constrained normal mode initialization of the shallow water equations and posed it as the minimization of the functional

$$I = \int\int_z \left[(V_0 - V_c)^2 W_v + (\phi_0 - \phi_c)^2 W_\phi\right] dA \tag{30}$$

where $\int\int dA$ is an integral over the atmosphere, V_0 and ϕ_0 indicated the observed values of velocity and geopotential fields, V_c and ϕ_c the values after constrained initialization, and W_v, W_ϕ the confidence weights.

In the variational formulation, Daley (1978), minimizes (30) subject to the constraint that the final state lies on the "slow" manifold. This constraint is approximated by requiring the satisfaction of the Machenhauer NMI balance condition.

$$Z = \frac{R_z(Z_c, Y_c)}{2\Omega_i \Lambda_z} \tag{31}$$

where Y_c and Z_c are the projections of the Rossby and gravity modes respectively, after adjustment has taken place, and Λ_z are the high frequencies. The later constraint is applied to the functional (31) by means of Lagrange's multipliers and the augmented functional is then minimized, leading to a set of Euler-Lagrange equations. One can show that unconstrained

N.M.I. initialization is a special case of constrained initialization using a particular choice of confidence weights W_ϕ and W_v.

The Daley (1978) procedure was iterative, because, once the rotational Rossby modes are changed by the Euler-Lagrange equations the gravitational manifold is no longer in balance. One has to iterate the procedure, replacing the gravitational manifold projection with the balance condition for the current rotational manifold projection.

Tribbia (1982a, 1982b) generalized the Daley approach by reducing the minimization to a series of linear least-square problems through an asymptotic expansion. He also allowed for longitude/latitude variable weights for a simple barotropic model. Using variational N.M.I. for four-dimensional data assimilation (Daley and Puri (1980)) noticed that the problem of minimization becomes difficult once the weights have full spatial variability.

Puri (1982) used constrained non-linear N.M.I. in an attempt to minimize mass loss in the ANMRC (Australian Numerical Meteorology Research Center) data assimilation scheme. Algorithmically his procedure followed the Daley (1978) approach and only the specification of weights was determined following the relative importance given to a particular field.

Phillips (1981) pointed out that large-scale meteorological analysis programs should concentrate on analyzing slow mode fields and to this purpose, observations must have their fast mode fields subtracted from them before they are used in the analysis.

Bourke and McGregor (1983) applied a variational constraint method involving constrained minimization of the changes in surface pressure while satisfying balance conditions of a non-linear vertical mode initialization scheme for a limited area baroclinic primitive equations prediction model. The variational constraint was introduced in order to prevent any changes in surface pressure occurring during initialization.

A pressure maintaining scheme formulated as a variational Euler-Lagrange problem has also been tested by Bourke and McGregor (1983).

A variational normal mode initialization using the variational formalism of Daley (1978) was employed by Puri (1983) with variable weights as a function of latitude. He then applied this scheme to a multi-level model by treating each vertical mode independently. Different weights were specified for each vertical mode.

It was found that unless the spatial variation of the weights is artificially simple, the solution of the variational problem is difficult for a model with a realistic number of degrees of freedom. This is because while the initialization itself is performed in normal mode space, the variation of the weights is carried out in physical space.

Temperton (1984) applied normal mode initialization to the European Center for Medium Range WeatherForecasting (ECMWF) multi-level grid point model minimizing an integral of the changes made by the initialization to the analyzed mass and wind fields, suitably weighted to control relative magnitudes of the adjustment to these fields.

The 3-D variational problem takes the form

$$I_v = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \int_0^{2\pi} \int_0^1 W_v \left[(\Delta u)^2 + (\Delta v)^2 \right] dp\, d\lambda \cos\theta\, d\theta \tag{32}$$

where Δu and Δv are changes made to the u and v fields and W_v is the confidence weight for the analyzed wind.

If the vertical modes have been chosen to be orthogonal then the variational integral for each vertical model can be expressed as

$$\begin{aligned} I(l) = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \int_0^{2\pi} & \Big\{ W_v \left[(\Delta \overline{u})_l^2 + (\Delta \overline{v})_l^2 \right] \\ & + (gD)_l^{-1} W_h \left(\Delta \overline{P} \right)_l^2 \Big\}\, d\lambda \cos\theta\, d\theta \end{aligned} \tag{33}$$

where D_l is the equivalent depth of the vertical model l, and where $\Delta\overline{u}$ and $\Delta\overline{v}$ are changes made to the vertical normal mode coefficients and $\Delta\overline{P}$ is the change made to the vertical normal mode coefficient of the auxilliary variable P.

Temperton (1984) has shown that the problem of minimizing the cost functional can be reduced to a linear least squares problem solved in a very efficient way. Temperton (1984) tried different weight combinations for W_v and W_λ as well as assessing the impact on a subsequent 5 days forecast of both unconstrained non-linear N. M. I. as well as variational N. M. I. with different W_λ and W_v confidence weights. The forecast results turned out to be extremely similar and all rather successful in predicting major changes over 5 days.

In a later paper Temperton (1985) presented the concept of a variational implicit non-linear N.M.I. which can considerably simplify the application of variational constraints to the initialization procedure. The implicit N.M.I. method proposed by Temperton (1985, 1988) allows the performance of non-linear N.M.I. without explicitly knowing the normal modes at all.

Tseng (1985) presented a classical variational initialization where the linear balance equation was used as a constraint to adjust the mass and wind fields simultaneously while the determination of confidence weights is done by the formulas

$$\alpha = \frac{2}{\left(\sigma_u^2 + \sigma_v^2\right)}, \quad \beta = \frac{1}{\sigma_\varphi^2} \tag{34}$$

where

$$\sigma_u^2 = \overline{\left(u - \tilde{u}^2\right)^2}, \quad \sigma_v^2 = \overline{\left(v - \tilde{v}^2\right)^2}, \quad \sigma_\varphi^2 = \overline{\left(\varphi - \tilde{\varphi}^2\right)^2}$$

where $\tilde{u}$, $\tilde{v}$ are the observed horizontal wind components and $\tilde{\varphi}$ is the observed geopotential, and the overbar represents the average over all grid points.

A variational initialization procedure based on the bounded derivative method has been proposed by Navon and Semazzi (1986, 1987). It concerns the application of the bounded derivative method for initializing the exterior vertical mode of the GLAS barotropic model (Takacs, 1986). The minimization of a cost functional including full variability of weights as a function of longitude and latitude is carried out using an Augmented Lagrangian method (Navon and de Villiers (1983)). The cost functional of the constrained bounded derivative initialization includes as constraints the bounded derivative method height and divergence constraints (see Kasahara (1982), Semazzi and Navon (1986)).

5. Optimal Control Techniques

5.1. General Results

Optimal control methods for distributed systems have been extensively studied and applied in many areas such as mechanics, economics, engineering, oceanography, etc.

Due to the fact that the formalism of optimal control problems includes the minimization of a functional, the cost function, they are variational methods and as such their numerical solution requires the computation of the gradient of the cost functional with respect to the state variable.

In many cases, the cost function is only an implicit function of the state variable which may be an initial condition or a boundary condition. Therefore, more sophisticated mathematical techniques must be used for estimating the gradient. One such particular method, the adjoint model technique, was specially developed for this purpose. A difficulty of this approach is the necessity to write well-posed problems and to carefully specify the functional framework of the variational problem.

We assume that the state of the atmosphere is described by a variable X belonging to some Hilbert space $\mathcal{H}$ (of finite or infinite dimension) and by a model written as

$$F(X) = 0 \tag{35}$$

We suppose that X may be split into two parts, Y and U, each part belonging to the Hilbert spaces $\mathcal{Y}$ and $\mathcal{U}$, respectively.

Therefore, (35) may be written as

$$F(Y, U) = 0 \tag{36}$$

where U is the control variable, chosen in such a way that for each given U, Eq. (36) has a unique solution $Y(U)$.

In this way we may define G by

$$G : \mathcal{Y} \rightarrow \mathcal{U}$$

and for each U belonging to $\mathcal{U}$. Then

$$G(Y) = U \tag{37}$$

has a unique solution in $\mathcal{Y}$.

Furthermore, we will assume that for each Y belonging to $\mathcal{Y}$, $\frac{\partial F}{\partial Y}(Y)$ is an isomorphism from $\mathcal{Y}$ to $\mathcal{U}$.

Therefore, it is possible to define an inverse function Φ such that:

$$\begin{aligned} \Phi : \quad & \mathcal{U} \rightarrow \mathcal{Y} \\ & U \rightarrow \Phi(U) = Y \end{aligned}$$

verifying

$$\begin{aligned} \Phi(G(Y)) &= Y \\ \Phi'(U) &= \left(\frac{\partial F}{\partial Y}(\Phi(U) \right)^{-1} \end{aligned} \tag{38}$$

Another Hilbert space has to be defined: the space of observations Θ in which an observation Z_{obs} is given. As pointed out, the observation is not necessarily a physical one, and it is not supposed to verify the equations of the model.

Let C be a linear operator from the space of the state variable to the space of observations; for each value of the control U we associate a state of the atmosphere $Y(U)$ and an observation

$$Z(U) = CY(U). \tag{39}$$

The cost function $J(U)$ is a measure of the distance between the state associated to the control U and the observation. It is defined by:

$$J(U) = \frac{1}{2} \|CY(U) - Z_{\text{obs}}\|_{\Theta}^2 \tag{40}$$

Therefore, the problem is to determine the optimal control variable U^* defined by

$$J(U^*) = \min_{U} J(U). \tag{41}$$

From a theoretical viewpoint, the system of optimality giving U^* is dependent upon the gradient of J with respect to U.

From a numerical viewpoint, U^* may be estimated by an iterative method starting from a first given U_0. In the same way, the numerical implementation of the iterative method requires the computation of the gradient of J with respect to U.

For deriving the gradient, a systematic method is the following:

i) Let V be some variable belonging to $\mathcal{U}$; then the directional derivative of J in direction V will verify

$$\begin{aligned} J'(U,V) = \nabla J(U) \cdot V &= \big(C'(Y) \cdot V \cdot C(Y) - Z_{\text{obs}}\big)_{\Theta} \\ &= \big\langle C'(Y)V, \Lambda_{\Theta}\left(C(Y) - Z_{\text{obs}}\right)\big\rangle_{\Theta',\Theta} \end{aligned} \tag{42}$$

where Λ_{Θ} is the canonical isomorphism between Θ and its dual space Θ', and $\langle \cdot, \cdot \rangle$ denotes the duality between Hilbert spaces.

ii) Let R be a linear operator from $\mathcal{Y}$ to $\mathcal{U}$, we define its dual operator to be the operator R^* from $\mathcal{U}'$ to $\mathcal{Y}'$ defined by

$$\big\langle R \cdot \mathcal{Y}, U' \big\rangle_{\mathcal{U}} = \big\langle Y, R^* \cdot U' \big\rangle_{\mathcal{Y}} \tag{43}$$

Using the dual operator of C' in (42) gives

$$\nabla J(U) \cdot V = \langle V, C'(Y)^* \Lambda_0 \left(C(Y) - Z_{\text{obs}}\right) \rangle_{\mathcal{U},\mathcal{U}'} \tag{44}$$

iii) Let us now define the adjoint system by

$$\left(\frac{\partial F}{\partial Y}\right)^* P = -C'(Y)^* \Lambda_H \left(CY(U) - Z_{\text{obs}}\right) \tag{45}$$

Then

$$\nabla J(U) \cdot V = \left\langle V, \left(\frac{\partial F}{\partial Y}\right)^* \cdot P \right\rangle_{\mathcal{U},\mathcal{U}'} \tag{46}$$

$$\nabla J(Y) \cdot V = \left\langle \frac{\partial F}{\partial Y} \cdot V, P \right\rangle_{\mathcal{Y},\mathcal{Y}'} \tag{47}$$

J is a functional defined on the space $\mathcal{U}$, so its gradient belongs to the dual space $\mathcal{U}'$. Theoretically, it is always possible to identify a Hilbert space to its dual. However, in practical problems there exist inclusion relations between the spaces used here, and when a space has been identified to its dual, it is no longer possible to identify subspaces with their duals.

In the practical phase of optimal control methods we were always operating in finite-dimensional spaces where no such problems exist.

Therefore Eq. (47) permits us to compute the gradient of J, applied to the direction V by:

1) determining P, the adjoint variable, as the solution of the adjoint system (46).

2) applying Eq. (47).

From this abstract situation let us extract two more practical examples enabling us to see how the gradient is computed. For an initial condition problem we will consider the case where the control variable is the initial condition, while for a boundary value problem we will see how to compute the gradient when the control variable is the value on the boundary.

5.2. Control of the Initial Condition

After a spatial discretization, we will assume that the state of the atmosphere, modelled by a vector Θ is verifying for the time interval $[0, T]$ the equation:

$$\frac{d\Theta(t)}{dt} = H(\Theta(t)) \tag{48}$$

where $\Theta(t)$ belongs to a finite dimensional space.

With an initial condition $\Theta(0) = \mu$, Eq. (48) has a unique solution $\Theta(\mu, t)$.

For the sake of simplicity, we will assume that a continuous observation $\widetilde{\Theta}$, in time, is given on the time interval $[0, T]$. The distance between a solution of (48) and the observation is defined by

$$J(\mu) = \frac{1}{2} \int_0^T \left\| \Theta(\mu, t) - \widetilde{\Theta}(t) \right\|^2 dt \tag{49}$$

where $\| \cdot \|$ is the Euclidian norm in finite dimensional space. With respect to the general theory developed above the space of the state variable is the same as the space of the observations. In practice, the observations are pointwise in both space and in time, therefore, Dirac's measures have to be introduced in the definition of J.

The derivation of the gradient of J with respect to μ is obtained as follows:

Let ν be some element belonging to the space of the initial conditions. The directional derivative of Θ in direction ν is defined by

$$\widehat{\Theta}(\mu, \nu) = \lim_{\alpha \to 0} \frac{\Theta[(\mu + \alpha), t] - \Theta(\mu, t)}{\alpha} \tag{50}$$

where $\widehat{\Theta}(\mu, \nu)$ is the solution of the differential system:

$$\begin{gathered} \frac{d\widehat{\Theta}(\mu, \nu)}{dt} = \frac{\partial H}{\partial \Theta} \left[\Theta(\mu, t) \right] \cdot \widehat{\Theta}(\mu, \nu) \\ \widehat{\Theta}(0) = \nu \end{gathered} \tag{51}$$

obtained by writing (48) with initial condition μ, then with initial condition $\mu_\alpha \nu$ and by letting the scalar α tend to zero. In (51) the expression $\frac{\partial H}{\partial \Theta}$ denotes the Jacobian of H.

the directional derivative of J in direction ν is obtained by taking the derivative of (49) leading to:

$$J'(\mu, \nu) = \int_0^T \left(\widehat{\Theta}(\mu, \nu, t), \Theta(\mu, t) - \widetilde{\Theta}(t) \right) dt \tag{52}$$

Let ψ be the dual variable to Θ, ψ is defined as the solution of the adjoint system to (48) given by

$$\begin{gathered} \frac{d\psi}{dt}(\mu, t) + \left[\frac{\partial H}{\partial \Theta} \Theta(\mu, t) \right]^T \cdot \psi(\mu, t) = \left(\Theta(\mu, t) - \widetilde{\Theta}(t) \right) \\ \psi(T) = 0 \end{gathered} \tag{53}$$

Let us write the scalar product of (52) with $\widehat{\Theta}$, then by integrating from 0 to T, we obtain:

$$J'(\mu,\nu) = \int_0^T \left(\frac{d\psi}{dt} + \left[\frac{\partial H}{\partial \Theta} \Theta(\mu,t) \right]^T \cdot \psi(\mu,t), \widehat{\Theta}(\mu,\nu,t) \right) dt \tag{54}$$

The time derivative in (53) is integrated by parts and then by using (51) we obtain:

$$J'(\mu,\nu) = \nabla J(\mu) \cdot \nu = \psi(\mu,0) \cdot \nu \tag{55}$$

Therefore, the gradient of J is obtained as the value at time zero of the dual variable. The backward integration of the adjoint system from T to 0 permits us to estimate the gradient of the cost functional and to perform a descent-type method.

An important remark for potential applications of control methods is the fact that with a different cost function only the right hand side of (53) has to be changed. The main difficulty encountered for programming optimal control methods is to write the left hand side of (53). This one is independent of the cost function and is intrinsic for a given model. Once it has been written and derived it can be used for other purposes such as data assimilation, initialization, sensitivity analysis, etc.

5.3. Control of the Boundary

For the sake of simplicity, we will suppose that on a domain Ω, of boundary Γ, some field is verifying the Laplace equation

$$\Delta U = f \tag{56}$$

Together with a boundary condition $U/\Gamma = V$, (56) has a unique solution, $U(V)$.

Let $\mathcal{T}$ be a set of points belonging to Ω, where some observations $\widetilde{U}$ of U are performed.

$$\mathcal{T} = \{Z_1, Z_2, \ldots, Z_N\} \tag{57}$$

The cost function is defined by

$$J(V) = \frac{1}{2} \sum_{i=1}^{N} \left(U\left(V, Z_i\right) - \widetilde{U}\left(Z_i\right) \right)^2 \tag{58}$$

the directional derivative $\overline{U}$ of U in a direction H is the solution of

$$\begin{aligned} \Delta\overline{U}(H) &= 0 \\ \overline{U}(H)/\Gamma &= H \end{aligned} \tag{59}$$

and the directional derivative of J verifies

$$J'(V,H) = \sum_{i=1}^{N} \left(\overline{U}\left(Z_i\right), U\left(V, Z_i\right) - \widetilde{U}\left(Z_i\right)\right)^2. \tag{60}$$

The adjoint system to (58) is introduced with P the dual variable to U.

$$\begin{aligned} \Delta P &= \sum_{i=1}^{N} U\left(V, Z_i\right) - \widetilde{U}\left(Z_i\right) \\ P/\Gamma &= 0 \end{aligned} \tag{61}$$

As above (61) is multiplied by $\overline{U}\left(H, Z_i\right)$ integrated on Ω, and after an integration by parts we find

$$\nabla J(V) = \frac{\partial P}{\partial n}/\Gamma \tag{62}$$

$\frac{\partial P}{\partial n}$ is the normal derivative of P on the boundary Γ. The estimation of the gradient for carrying out a descent-method requires the estimation of the gradient of J, which is obtained by solving the adjoint system (61).

Let us point out that this case is especially simple due to the fact that the Laplacian operator is self-adjoint. Therefore, a Laplace's equation solver may be used to solve both the direct and the adjoint problem.

This problem could have been solved using a classical variational formalism, for instance with a weak constraint formalism we would have to minimize the functional

$$J(U) = \frac{1}{2} \sum \left(U\left(Z_i\right) - \widetilde{U}\left(Z_i\right)\right)^2 + \frac{1}{C} \int_{\Omega} (\Delta U - f)^2 \, dy. \tag{63}$$

The Euler-Lagrange equation for (63) is a fourth order partial differential equation with complicated boundary conditions. From a numerical viewpoint the size of the discrete problem associated with (63) is equal to the number of grid points in the discrete point of view domain Ω.

By comparison, for the optimal control approach the dimension of the problem to be solved is only equal to the number of points on the discrete boundary. In this way we have obtained a significant reduction of the size of the problem.

5.4. Optimal Control Methods in Meteorology

Optimal control methods using the initial condition as control variables have been used by Lewis and Derber (1985) employing a forecast model in the form

$$\frac{\partial q}{\partial t} + J(\psi, q) + \beta \frac{\partial \psi}{\partial x} = 0 \tag{64}$$

where q is the partial differential operator

$$q = \left(\nabla^2 + \frac{\partial}{\partial p} \cdot \frac{f_0^2}{\sigma} \cdot \frac{\partial}{\partial p} \right) \psi, \tag{65}$$

ψ is the geostrophic stream function, β the meridional variation of the Coriolis parameter, and J is the Jacobian operator.

The cost function chosen is in the form

$$J = \sum_{p=1}^{P} \sum_{n=1}^{N} \left(\psi\left(t_p\right) - \widetilde{\psi}\left(t_p\right) \right)^2 \tag{66}$$

with $\widetilde{\psi}$ representing the analyses created from a primitive equation model.

The model is discretized in five levels on a 23×28 Lambert conformal grid with a resolution of 135.2 km at the standard latitudes. The analyses of the primitive equation model were inserted over two intervals, the first, a six hour interval and the second a complete twelve hour interval of the analysis period. The numerical results show (Derber (1987)) that the convergence rate is a function of the length of the assimilation period rather than a function of the density of data.

Courtier (1986) used the shallow water equations to test data assimilation with optimal control. The shallow water equations in this application were written with vorticity ξ and divergence η variables, assuming the form

$$\frac{\partial \xi}{\partial t} = \mathcal{J}\left(\xi + f, \Delta^{-1}\xi\right) - \nabla \cdot \left(\left(\xi + f\right) \nabla \Delta^{-1} \xi\right) - \Delta\Phi - \Delta K \tag{67}$$

$$\frac{\partial \eta}{\partial t} = \mathcal{J}\left(\xi + f, \delta^{-1}\eta\right) + \nabla \cdot \left(\left(\xi + f\right) \nabla \Delta^{-1} \xi\right) - \delta\Phi - \Delta K \tag{68}$$

$$\frac{\partial \Phi}{\partial t} = \mathcal{J}\left(\Phi, \Delta^{-1}\xi\right) - \Delta \cdot \left(\Phi \nabla \Delta^{-1} \eta\right) \tag{69}$$

$$\text{and } K = \frac{1}{2}\left(\nabla \Delta^{-1}\xi \cdot \nabla \Delta^{-1}\xi + \nabla \Delta^{-1}\eta + 2J, \Delta^{-1}\eta\right)$$

where $\mathcal{J}$ is the Jacobian operator.

The cost function J used is in the form

$$J = a \cdot J_h + J_v \tag{70}$$

J_h and J_v are the sums of the squares of the difference between the observed values and the model values. The discretization in space is performed using spherical harmonics with triangular truncation at order 21. The time integration of the discrete scheme has been carried out with a semi-implicit leapfrog scheme. The assimilation experiments were performed with wind and geopotential at 500 mb for a period of 24 hours. The numerical results show a good ability of these methods to retrieve dynamical fields from observations.

For a small scale model, methods of optimal control have been used by Le Dimet and Nouailler (1985). The model which is used corresponds to the study of a squall line on a 60km $\times$ 60km area. With u, v, w being the components of the wind speed and p being the atmospheric pressure, the model is written

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial w}{\partial z} - fv + \rho_0^{-1}\frac{\partial p}{\partial y} + \underline{C}_D|U|U = 0 \tag{71}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} + fu + \rho_0^{-1}\frac{\partial p}{\partial y} + C_D|U|v = 0 \tag{72}$$

$$\frac{\partial p}{\partial t} + C_x\frac{\partial p}{\partial x} + C_y\frac{\partial p}{\partial y} + k \cdot \mathrm{div}U = 0 \tag{73}$$

where C_D is a drag coefficient, $|U| = \left(u^2 + v^2\right)^{1/2}$, C_x and C_y are advective velocities of the squall line estimated from radar observations. Terms of vertical transport have been evaluated from the observations of a network of 18 stations measuring at the ground the wind and the atmospheric pressure stations every 30 seconds for some stations and every 2minutes 30 seconds for the others. Spatial discretization has been performed using a finite difference scheme with a 3 km gridsize in both directions. The time integration scheme was a leapfrog scheme. The numerical results were obtained by carrying out a method of optimization without constraint, intermediate between the conjugate gradient and quasi-Newton method (Lemarechal (1980), Lemarechal and Servigne (1984)), based on the Buckley-Lenir (1983,1985) method. The descent procedure exhibits a fast decrease of the cost function in the first few iterations, then becoming slower for the subsequent iterations. This is a common feature to all optimization methods which have been applied to minimize the gradient of the cost functional with respect to the initial conditions.

A main advantage of optimal control methods is to retrieve meteorological fields in conformity with the dynamics of the atmosphere modelled by a system of partial differential equations. Of course if the model used is not filtered (*i.e.*a primitive equations model), then the optimal solution may include gravity waves, especially if the data are noisy or contain some undue oscillation.

There are several ways to prevent the development of gravity waves in the optimal solution for the 4-D data assimilation problem.

The first method (Courtier (1985)) is to add to the cost function a penalty term.

For instance if the model is written as

$$\frac{dX}{dt} = F(X) \tag{74}$$

with initial condition $X(0) = U$, where X represents the meteorological variables. Then the cost function may be written as:

$$J(U) = \int_0^T \|X(U,t) - X_{\text{obs}}(t)\|^2 \, dt + \frac{1}{\epsilon} \int_0^T \left\| \frac{dX(U,t)}{dt} \right\|^2 dt \tag{75}$$

This transformation of the cost function may add only slightly to the computational cost, but nevertheless it has two main inconveniences:

a) using a L_2 norm does not prevent very fast and timewise localized variations of the term.

b) the optimal solution depends upon the value chosen for ϵ.

The questions to be answered are:

1) based on which physical considerations should ϵ be chosen?

2) to what extend is the optimal trajectory sensitive to the solution?

An alternative method is to use a regularization-penalization method (Le Dimet, Sasaki, and White (1982)). the amplitude of the fast movement is supposed to be limited by some given constant H. In the formalization of the control problem, we introduce the following constraint on the state, requiring it to verify the following inequality

$$\left\| \frac{dX}{dt}(U,t) \right\|^2 \leq H \tag{76}$$

Therefore, (76) constitutes a pointwise constraint on the trajectory. The numerical solution of this problem of control with constraints is obtained by solving a sequence of unconstrained control problems with the cost function given by

$$J(U) = \frac{1}{2}\int_0^T \|X(U,t) - X_{\text{obs}}\|^2 \, dt + \frac{1}{\epsilon}\int_0^T g\left(\left\|\frac{dX(U)}{dt}\right\|\right) dt \tag{77}$$

where g is a function defined by

$$\begin{aligned} g: \quad & \mathbb{R} \to \mathbb{R} \\ g(Y) = \quad & 0 \qquad\qquad \text{if } |Y| \le H \\ = \quad & \frac{1}{2}(Y - H)^2 \quad \text{if } |Y| \ge H. \end{aligned} \tag{78}$$

Using such methods may prevent the need to perform a more sophisticated initialization, but more numerical experiments have to be carried out in order to evaluate the performance of these techniques for real data.

From the experience already accumulated using optimal control methods applied to meteorological problems it can be concluded:

i) In few iterations the retrieved fields are coherent with respect to the data.

ii) The numerical procedures are very sensitive to the quality of the gradient, therefore, the adjoint system has to be very carefully written and derived. For every operation done on the direct system, the respective adjoint must be carried out on the adjoint system.

Writing the adjoint of a system is a costly operation but it can be made more profitable if it is used in conjunction with other studies such as that of sensitivity analysis or evaluation of unknown coefficients *i.e.*parameter estimation.

5.5. Application of Sensitivity Analysis

In the above section the cost function has been presented as a measure of the difference between the solution of a model and the observations.

This interpretation can be extended to any cost function which would be the scalar response of the model to an external forcing, modelled in the cost function. The only restriction imposed, bears on the derivability of the cost function. The adjoint model is the same, and only the right hand side has to be changed for performing these studies.

This method has been applied by Hall, Cacuci and Schlesinger (1982) for the shallow water equations.

In a similar way, the adjoint of an atmospheric model is extensively used by Marchuck (1982) for environmental studies and especially for estimating the impact of industrial plants on the environment.

Some optimal control methods have also been used in oceanography (see for instance Reinhart (1985)). In his application, the boundary condition is the control variable and the method is applied for determining the optimal location of sensors.

5.6. Application for Parameter Identification

Many meteorological models contain numerical parameters which cannot be directly measured and are empirically estimated such as turbulent diffusion coefficients, drag-coefficients, etc. In the large majority of cases they represent subgrid effects and are, therefore, estimated and parameterized based on numerical considerations rather than on physical ones.

A way to properly estimate these parameters is to use them as control variables in a procedure of analysis. If the model may be written as

$$\begin{aligned} \frac{dX}{dt}(U,K) &= F(X,K) \\ X(0) &= 0 \end{aligned} \tag{79}$$

where K is some unknown and steady state coefficient. The cost function may then be defined by

$$J(U,K) = \left\| X(U,K) - \widetilde{X} \right\|^2 \tag{80}$$

where $\widetilde{X}$ is the observation.

As above, the gradient of J with respect to U and K will be computed by using the adjoint system to (80). Such a method has been used by Lamb, Chen, and Seinfeld (1975) for estimating coefficients of diffusion and by Le Dimet (1981) for the computation of a drag-coefficient in a two-dimensional model.

6. Variational Methods Applied for Numerical Purposes

Variational methods have been used mainly for static and dynamic situations in meteorology. Another use for variational methods resulting in a global approach was the '*a posteriori*' enforcement of integral invariants in numerical models.

Such methods were first proposed by Sasaki (1975, 1976, 1977) and Bayliss and Isaacson (1975), Isaacson (1977), and Isaacson *et al.* (1979). Independently Sasaki (1976) proposed a functional of the form

$$\begin{aligned} J = \sum \left[\overline{\alpha}\,(u-\tilde{u})^2 + \overline{\alpha}\,(v-\tilde{v})^2 + \beta \left(h-\tilde{h}\right)^2 \right] \\ + \lambda_E \left\{ \sum \left[\left(\frac{h}{2}\right) \cdot \left(u^2+v^2\right) + \left(\frac{g}{2}\right) h^2 \right] - T^0 \right\} \end{aligned} \tag{81}$$

where u and v are the x and y components of the velocity, h is the elevation of free water surface measured from the mean height, α and β are weights while T^0 is the total energy at time $t = 0$, λ_E is a Lagrange multiplier, constant with respect to time, but possibly variable in time. $\tilde{u}$, $\tilde{v}$, and $\tilde{h}$ are the values predicted for the $(N+1)^{\text{th}}$ time-step using a numerical weather prediction finite-difference discrete model.

Sasaki (1976) applied his method for the non-linear shallow-water equations on a rotating plane by solving iteratively the resulting Euler-Lagrange equations and obtained satisfactory numerical results. We will describe his method in ample detail in another subsection. Bayliss and Isaacson (1975) proposed independently a method making it possible to modify any given finite-difference scheme so as to ensure exact conservation of integral invariants. In their approach, Bayliss and Isaacson (1975) linearized the constraints about the predicted values.

The essence of their theoretical framework can be described as follows:

Assume we have an initial boundary value partial differential equation problem for the vector $\underline{u}$

$$u_t = B(u) \tag{82}$$

and that the solution u to (82) satisfies K integral invariants

$$g_k(u) = 0 \quad k = 1, 2, \ldots, K. \tag{83}$$

If we discretize the integral invariant constraints we obtain

$$G_K\left[U_{ij}^n\right] = 0 \quad k = 1, 2, \ldots, K \tag{84}$$

where U_{ij}^n is a net function defined at the grid points (x_i, y_j, t_n) and $U_{ij}^n = U(x_i, y_j, t_n)$ approximates $U(x_i, y_j, t_n)$.

At time t_{n+1}, the difference operator solving for the vector $\underline{u}$ (for instance $\underline{u} = (u, v, \phi)^T$ for the shallow-water equations) has the form

$$W(n+1) - C\left[W(n), W(n-1), \ldots, W_n - s)\right] = CW(n) \tag{85}$$

where $W(n)$ is a net function at time t_n.

We wish to modify the finite-difference scheme (85) in such a way as to produce a grid function $U(n+1)$ which will satisfy (84) – the discrete approximation of the integral invariants (83).

In other words, a corrective net function $V(n+1)$ is to be found such that

$$\begin{aligned} U(n+1) &= CU(n) + V(n+1) \\ G_k\left[U(n+1)\right] &= 0 \quad k = 1, 2, \ldots, K \\ \min\ \|V(n+1)\| & \end{aligned} \tag{86}$$

and such that the norm of the perturbation $V(n+1)$ is as small as possible $i.e.$min $\|V(n+1)\|$. The determination of $V(n+1)$ is a calculus problem of finding a net function that satisfied K simultaneous non-linear equations (86) and is of minimum norm (Isaacson (1977)).

Bayliss and Isaacson (1975) proposed to solve (86) by linearizing the discrete invariants $G_k\left[U(n+1)\right]$ about the predicted value $CU(n)$ which can be written as

$$\begin{aligned} G_K\left[U(n+1)\right] &= G_K\left[CU(n) + V(n+1)\right] \\ &\approx G_K\left[CU(n) + \mathrm{grad} G_K \cdot V(n+1)\right] \\ &= G_K\left[CU(n)\right] + \left.\frac{\partial G_K}{\partial U(n+1)}\right|_{U(n+1)=CU(n)\equiv L_K(V(n+1))} \cdot V(n+1) \end{aligned} \tag{87}$$

For a full implementation of the method, see Kalnay *et al.* (1977) and Navon (1987).

6.1. The Constraint-Restoration Method

Miele *et al.* (1968, 1969) proposed a constraint restoration method based on a least-square change of the coordinates in the state vector.

Their method assumes at the start that the vector $\underline{x}$

$$\underline{x}\left(\tilde{u}_{11}^{n}\ldots\tilde{u}_{N_xN_y}^{n},\tilde{v}_{11}^{n}\ldots\tilde{v}_{N_xN_y}^{n},\tilde{h}_{11}^{n},\ldots\tilde{h}_{N_xN_y}^{n}\right)^{T} \tag{88}$$

at the time $n\Delta t$ is in the vicinity of the optimal point $\underline{x}^*$ which satisfies exactly K discrete equality constraints given by

$$\phi\left(\underline{x}^*\right)=0 \tag{89}$$

where

$$\phi(x)=\begin{bmatrix}\phi_1(x)\\ \vdots\\ \phi_K(x)\end{bmatrix}\qquad K\leq 3N_xN_y=N \tag{90}$$

where N is the number of components of the vector $\underline{x}$.

If $\tilde{x}$ is a varied point related to the minimal point $\underline{x}$ by

$$\underline{\tilde{x}}=\underline{x}+\delta x \tag{91}$$

where δx is a perturbation of x. By using quasi-linearization, Eq. (89) is approximated by

$$\phi(x)+A^T(x)\delta x=0 \tag{92}$$

where A is the $(N\times K)$ matrix

$$A(x)=\begin{bmatrix}\frac{\partial\phi_1}{\partial x_1} & \cdots & \frac{\partial\phi_K}{\partial x_1}\\ \vdots & \ddots & \vdots\\ \frac{\partial\phi_1}{\partial x_N} & \cdots & \frac{\partial\phi_K}{\partial x_N}\end{bmatrix} \tag{93}$$

where the j-th column is the gradient of the integral constraint ϕ_j with respect to the vector $\underline{x}$.

If the vector $\underline{x}$ is an approximation to the desired solution, we wish to restore the K constraints (89) while causing the least change in the vector $\underline{x}$ components.

This means we wish to minimize the function

$$J=\frac{1}{2}\delta x^T\delta x \tag{94}$$

subject to the linearized constraint (92).

Using standard methods of theory of maxima and minima, the fundamental solution of this problem is given by

$$F - \frac{1}{2}\delta^T x\delta x + \underline{\lambda}^T \left[\phi(\underline{x} + A^T(\underline{x})\delta x\right] \tag{95}$$

where λ is a K component Lagrange multiplier vector to be determined.

The optimum change δx is obtained when the gradient of F with respect to δx vanishes, *i.e.*

$$\delta x = -A(\underline{x})\lambda. \tag{96}$$

Using Eqs. (95) and (96) we obtain an explicit expression for the Lagrange multiplier vector

$$\lambda = B^{-1}(\underline{x})\phi(x) \tag{97}$$

where

$$B(\underline{x}) = A^T(x)\phi(x) \quad (\text{a } K \times K \text{ matrix}). \tag{98}$$

so that

$$\delta x_{\text{opt}} = A(x)B^{-1}(x)\phi(x) \tag{99}$$

For a practical implementation of the method see Navon (1987), Miele *et al.* (1968, 1969, 1971).

The Bayliss-Isaacson algorithm and the constraint restoration method have been proven to be equivalent (Navon (1987b)), however, these methods of '*a posteriori*' enforcing of integral constraints do not exactly replicate the Arakawa (1966), and Arakawa and Lamb (1977,1981) '*a priori*' methods as shown in a study by Takacs (1988). While successfully conserving total energy and potential enstrophy, these methods seem to require the formulation of an additional constraint of mean wave number conservation (yet to be formulated) without which they introduce distortion in the energy spectra transfers as evidenced in the experiments with NASA/GLA shallow-water equations model (Takacs (1986, 1988)).

6.2. Other Approaches for Enforcing '*a posteriori*' Conservation of Integral Invariants

Sasaki (1975, 1976, 1977) proposed a variational approach for enforcing '*a posteriori*' integral invariants in a finite-difference model for an initial-value problem which consisted of a model of the shallow-water gravity waves on a rotating plane.

For the shallow-water equations an energy conservation law was written in a finite-difference analog as

$$TE = \sum \left(\frac{h}{2}\right) \cdot \left(u^2 + v^2\right) + \left(\frac{g}{2}\right) \cdot h^2 = T^0$$

where u and v are the 2 components of velocity, h is the depth of the fluid, T^0 is the value of the total energy TE at time $t = 0$.

If $\tilde{h}$, $\tilde{u}$, and $\tilde{v}$ are the values predicted for the $(n+1)^{\text{th}}$ time level by using a set of finite-difference equations discretizing the shallow-water equations, the variational problem can be formulated in terms of a cost functional as

$$\begin{aligned} J = &\sum \left[\tilde{\alpha}\,(u-\tilde{u})^2 + \tilde{\alpha}\,(v-\tilde{v})^2 + \tilde{\beta}\left(h-\tilde{h}\right)^2\right] \\ &+ \lambda_E \left\{\sum \left[\left(\frac{h}{2}\right) \cdot \left(u^2+v^2\right) + \left(\frac{g}{2}\right) \cdot h^2\right] - T^0\right\} \end{aligned} \tag{100}$$

where the relative weights $\tilde{\alpha}$ and $\tilde{\beta}$ are chosen so as to make the fractional adjustment of variables proportional to the fractional magnitude of the truncation errors in the predicted variables.

The stationary value of the functional results from setting its first variation to zero. The resulting Euler-Lagrange equations are

$$\begin{aligned} 2\tilde{\alpha}\,(u-\tilde{u}) + \lambda_E \cdot hu &= 0 \\ 2\tilde{\alpha}\,(v-\tilde{v}) + \lambda_E \cdot hv &= 0 \\ 2\tilde{\beta}\left(h-\tilde{h}\right) + \lambda_E \left[\frac{\left(u^2+v^2\right)}{2}\right] + \lambda_E \cdot gh &= 0 \\ \text{and } \sum \left[\left(\frac{h}{2}\right) \cdot \left(u^2+v^2\right) + \left(\frac{g}{2}\right) \cdot h^2 - T^0\right] &= 0. \end{aligned} \tag{101}$$

The numerical solutions of u, v, h, and λ_E are obtained using an iterative technique. Navon (1981) used an extension of Sasaki's approach to enforce conservation of potential enstrophy and mass in a long-term integration of two ADI finite-difference approximations of the non-linear shallow-water equations on a limited-area domain on a rotating β-plane. The Sasaki method was compared to the Bayliss-Isaacson method and the Bayliss-Isaacson method was found to be more robust and less demanding of CPU time. The filtering technique of Kalnay-Rivas *et al.* (1977, 1979) using GLAS fourth order global atmospheric model was also considered.

6.3. The Augmented-Lagrangian Method

Another novel approach was proposed by Navon and de Villiers (1983) consisting of applying an Augmented-Lagrangian method for enforcing conservation of integral invariants.

Using a similar functional as Sasaki (1976)

$$\begin{aligned} f &= \sum_{j=1}^{N_x} \sum_{k=1}^{N_y} \left[\tilde{\alpha} \left(u - \tilde{u} \right)^2 + \tilde{\alpha} \left(v - \tilde{v} \right)^2 + \tilde{\beta} \left(h - \tilde{h} \right)^2 \right]_{ij} \\ N_x \delta x &= L \\ N_y \Delta y &= D \end{aligned} \tag{102}$$

where L and D are respective dimensions of the rectangular domain, we define an Augmented-Lagrangian function L by

$$L \left(\underline{x}, u, r \right) = f(x) + u^T e(x) + \frac{1}{2r} \left| e(\underline{x}) \right|^2 \tag{103}$$

where

$$\underline{x} = \left(\tilde{u}_{11}^n \ldots \tilde{u}_{N_x N_y}^n, \tilde{v}_{11}^n \ldots \tilde{v}_{N_x N_y}^n, \tilde{h}_{11}^n \ldots \tilde{h}_{N_x N_y}^n \right)^T \tag{104}$$

subject to equality constraints

$$\underline{e} \left(\underline{x} \right) = 0 \tag{105}$$

where $e\left(X\right)$ is a vector of three non-linear quantities given by

$$e(\underline{x}) = \begin{cases} E^n - E^0 \\ Z^n - Z^0 \\ H^n - H^0 \end{cases} \tag{106}$$

where

$$\begin{aligned} E^n &= \frac{1}{2} \sum_{j=1}^{N_x} \sum_{k=1}^{N_y} \left[\tilde{h} \left(u^2 + v^2 \right) + g \tilde{h}^2 \right]_{jk}^n \Delta x \, \Delta y \\ Z^n &= \frac{1}{2} \sum_{j=1}^{N_x} \sum_{k=1}^{N_y} \left[\frac{\frac{\partial \tilde{v}^n}{\partial x} - \frac{\partial \tilde{u}^n}{\partial y} + f}{\tilde{h}} \right]_{jk}^2 \\ H^n &= \sum_{j=1}^{N_x} \sum_{k=1}^{N_y} \tilde{h}_{jk} \Delta x \, \Delta y \end{aligned} \tag{107}$$

where D^n, Z^n, and H^n are the values of the discrete integral invariants of total energy, potential enstrophy, and mass at time $t_n = n\Delta t$ while E^0, Z^0, and H^0 are corresponding values of the same integral invariants at time $t = 0$, and $u = (u_1 \ldots u_m)$ is an m-component Lagrange multiplier vector, while r is a penalty parameter. The basic idea of the Augmented-Lagrangian method is to solve the constrained minimization problem by transforming this problem into a sequence of unconstrained minimizations of the following Augmented-Lagrangian

$$\min L_{r_k}(\underline{x}, \underline{u}_k) = f(\underline{x}) + \sum_{i=1}^{n} u_k^i e_i(\underline{x}) + \frac{1}{2r_k} |\underline{e}(\underline{x})|^2 \tag{108}$$

The theory is explained in Bertsekas (1975, 1982) and is expressed in the following proposition:

Proposition (Bertsekas (1975))

For $k = 0, 1, \ldots,$ let x_k be a global minimum of the problem

$$\min L_{r_k}\left(\underline{x}, \underline{u_k}\right) \tag{109}$$

subject to

$$x \in \mathbb{R}^n$$

where $|u_k|$ is bounded and $0 < r_{k+1} < r_k$ for all k and $r_k \to 0$.

Then every limit point of the sequence $\{x_k\}$ is a global minimum of f subject to the equality constraints $\underline{e}(\underline{x}) = 0$. The method consists in a sequence of unconstrained minimizations of the augmented-Lagrangians $L_{r_k}(\underline{x}, u_k)$. Given a multiplier vector u_k and a penalty parameter r_k we minimize $L_{r_k}(x, u_k)$ over $\mathbb{R}^n$ and obtain a vector $\underline{x}_k$. The variable $\underline{u}_k$, the vector of Lagrange multipliers and the penalty parameters are held fixed during the minimization and then updated prior to the next unconstrained minimization for which powerful conjugate-gradient methods are used (see Navon and Legler (1987)).

The algorithm is typically terminated at a point x_r where

$$|\nabla x L_{r_k}(x_k, u_k)| \leq \epsilon_k \tag{110}$$

or

$$\epsilon_i(\underline{x}_k) < \epsilon_k' \quad i = 1, \ldots, m \tag{111}$$

where ϵ_k and ϵ_k' are small positive scalars.

One can use an inexact minimization by demanding only a moderate accuracy in the first unconstrained minimizations of the Augmented-Lagrangian and increasing the accuracy at later iterations by using a preselected decreasing sequence $\{\eta_k\}$, tending to zero. In practice a schematic Augmented-Lagrangian algorithm proceeds as follows:

a) Select initial vector of Lagrange multipliers $\underline{u}_0$ based on either prior knowledge or start with a null vector in absence of such knowledge.

Select penalty parameters $r_0^i > 0$ and a decreasing sequence $\{\eta_k\}$ with $\eta_0 \geq 0$.

Step 1: Given a multiplier vector u_k, penalty parameter r_k^i and η_k, find a vector x_k satisfying

$$\|\nabla k L_{r_k}(x_k, u_k)\| \leq \{\eta_k\} \left\|e\left(x_{\underline{k}}\right)\right\| \tag{112}$$

by using a conjugate-gradient method to solve the inexact unconstrained minimization job.

Step 2:

$$\text{If } |\epsilon_i(x_k)| < \epsilon_i \quad i = 1, \ldots, m \text{ Stop.}$$

Otherwise, proceed to Step 3.

Step 3: update the multiplier vector using

$$\eta_{k+1} = \eta_k + r_k^{-1} \cdot \underline{\epsilon}\left(x_{\underline{k}}\right). \tag{114}$$

Update and select penalty parameters $r_{k+1}^i \in \left(0, r_k^1\right)$ (see Navon and de Villiers (1983)).

Select $\eta_{k+1} \geq 0$ following a formula of the type

$$\eta_k = (\ell)^k \quad 0 < \ell < 1 \tag{115}$$

and return to Step 1.

Three to four cycles were generally required to obtain satisfactory results.

6.4. Other Variational Methods

Similar approaches were used by Schneider (1984) to answer to problem of the effect of horizontal eddy momentum fluxes on the equilibrium zonal mean motions. Specifically, one minimizes various globally integrated quantities such as ZKE, the zonal kinetic energy, or $ZKE + ZAPE$, where $ZAPE$ is the zonal available potential energy for any distribution of horizontal eddy momentum fluxes, ($i.e.$all possible distributions of $\overline{u'v'}$) and specified thermal forcing.

A two-level model of the zonally averaged steady state response to the heat and momentum sources was used. The variational problem was to find the minimum (or minima) of I where

$$I = \int_{y_0}^{y_1} F\left(y, v, \frac{\partial v}{\partial y}, \frac{\partial^2 v}{\partial y_2}\right) dy \tag{116}$$

over all functions $v(y)$ that satisfy the boundary conditions

$$v\left(y_0\right) = v\left(y_1\right) \neq 0 \tag{117}$$

subject to an integral constraint

$$\int_{y_0}^{y_1} G\left(y, v, \frac{\partial v}{\partial y}, \frac{\partial^2 v}{\partial y^2}\right) dy = 0. \tag{118}$$

Using calculus of variations and defining

$$\mathcal{H} = F + \gamma G \tag{119}$$

where γ is a constant, the extrema of I subject to the constraints results as a solution of the Euler-Lagrange equation

$$\frac{d^2}{dy^2} \cdot \frac{\partial \mathcal{H}}{\partial \left(v''\right)} - \frac{d}{dy} \cdot \frac{\partial \mathcal{H}}{\partial \left(v'\right)} = 0 \tag{120}$$

where primes denote differentiation with respect to y and $\mathcal{H}$ satisfies boundary conditions

$$\frac{\partial \mathcal{H}}{\partial \left(v''\right)} = 0 \tag{121}$$

at y_0 and y_1.

γ is chosen so that the solution to (116) and (117) satisfies the integral constraint.

Reddy (1982) proposed a penalty function method for finite-element models of fluid flow.

He considers the general variational problem of finding a minimum of the functional

$$I(u) = \int_{\Omega} F\left(x, y, u, u_x, u_y\right) dx\, dy \tag{122}$$

in a Hilbert space $\mathcal{H}$, subject to the constraint

$$G(u) = 0 \tag{123}$$

where G in general is a non-linear operator from $\mathcal{H}$, into some Hilbert space $\mathcal{H}_2$. The solution u belongs to a subspace of $\mathcal{H}_1$.

Usually the problem is solved by the Lagrange multiplier method which seeks stationary values (u, λ) of the modified functional

$$L(u, \lambda) = I(u) + \int_\Omega \lambda G(u) dx\, dy \tag{124}$$

on the product space $\mathcal{H}_L = \mathcal{H}_1 \times \mathcal{H}_2$, where λ is the Lagrange multiplier.

The penalty function method reduces problems of conditional or constrained minimization to problems without constraints by introducing a penalty for the infringement of the constraints. Instead of solving the original problem, one minimizes the augmented functional

$$J_u(u) = I(u) + \frac{1}{2}\alpha_\eta \left\| G(u) \right\|^2_{\mathcal{H}_2} \tag{125}$$

on the whole of the space $\mathcal{H}$ for some penalty parameter $\alpha_n > 0$. $\|\cdot\|_{\mathcal{H}_2}$ is the norm in $\mathcal{H}_2$.

A theorem due to Polyak (1971) guarantees the existence of the solution to the penalty problem. Reddy (1981) goes to show that in finite-element models of the Navier-Stokes equations, the type of numerical quadrature is crucial (for the penalty terms) for the success of the penalty method for incompressible fluid flow (see also Reddy (1981)).

Sasaki and Reddy (1980) used a variant of Sasaki's (1976) variational adjustment to compare stability and accuracy of some numerical models of two-dimensional circulation as well as to study the conservation of the mean-kinetic energy and enstrophy for long term integrations.

They found out that the variational adjustment has not improved the RMS error and in some cases made it worse, but enables the scheme to avoid computational instability.

6.5. Constrained Adjustment to Control Lamb External Gravity Waves

In many meteorological applications one is often interested in suppressing external gravity waves by modifying the observed wind field in such a way that the vertical motions at the lowest level of a three-dimensional baroclinic model vanish.

An alternative way is to regard this adjustment as a variational adjustment of the horizontal wind field in a pressure coordinate system (x, y, p) so that the pressure tendency $\frac{dp_s}{dt}$ is zero everywhere, where ps is the pressure surface.

The continuity equation in pressure coordinates is given by

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial p} = 0 \tag{126}$$

Integrating this equation from the top to the bottom of the atmosphere and assuming the vertical velocity $w = 0$ at both end points, we obtain (see Ramamurthy and Carr (1987))

$$\int_0^{p_s} \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) dp = 0 \tag{127}$$

The use of this equation as a strong constraint will ensure that

$$\frac{dp_s}{dt} = 0 \tag{128}$$

i.e., using the continuity equation as a strong constraint will enable us to suppress the Lamb waves which can be viewed as noise in a meteorological model and which moreover impose very stringent computational stability constraints on the allowable time-step Δt. The Augmented-Lagrangian functional, $\mathcal{L}$, for which the stationary value is to be found for this problem is:

$$\begin{aligned}
\mathcal{L} = & \int_x \int_y \int_p \left[(u - \tilde{u})^2 + (v - \tilde{v})^2 \right] dx\, dy\, dp \\
& \int_x \int_y \left[\lambda \int_0^{p_s} \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) dp \right] dx\, dy \\
& + \frac{1}{2} \int_x \int_y \left[C \left(\int_0^{p_s} \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial Y} \right) \right) \right]^2 dx\, dy
\end{aligned} \tag{129}$$

where C is a penalty term and λ is the vector of Lagrange multiplier. The same inexact minimization of the Augmented-Lagrangian of Bertsekas (1982) is applied using a conjugate gradient method of Shanno and Phua (1986) for the unconstrained minimization. For computational details see Navon, Phua, and Ramamurthy (1987).

6.6. Direct Minimization Techniques

Application of direct minimization techniques to objectively analyze meteorological fields was used by Hoffman (1984), Legler, Navon, and O'Brien (1988), and Navon and Legler (1987) which applied the method for objective analysis of wind stress over the Indian Ocean. Ramamurthy and Navon (1988) applied a direct minimization technique to variational blending of GFFE level II-b fields obtained from a high-resolution objective analysis scheme over the Indian Ocean basin with the ECMWF level III-b gridded analyses.

The underlying idea was to enhance ECMWF global analyses with enriched regional analyses in such a way as to retain the large-scale information from a global data assimilation system and, at the same time, add detailed information on small scale waves in the limited area of interest.

The functional F to be minimized is expressed as:

$$
\begin{aligned}
F = & \frac{1}{L^2}\rho \sum_x \sum_y \sum_p \left[(u - u_{\mathrm{FR}})^2 + (v - v_{\mathrm{FR}})^2 \right] \\
& + \frac{1}{L^2}\gamma \sum_x \sum_y \sum_p \left[(u - u_{\mathrm{ECMWF}})^2 + (v - v_{\mathrm{ECMWF}})^2 \right] \\
& + L^2 \Gamma \sum_x \sum_y \sum_p \left| \left[\nabla^2 (u - u_{\mathrm{ECMWF}}) \right]^2 + \left[\nabla^2 (v - v_{\mathrm{ECMWF}}) \right]^2 \right| \\
& + \beta \sum_x \sum_y \sum_p \left[\nabla \cdot (\underline{v} - \underline{v}_{\mathrm{FR}}) \right]^2 \\
& + \alpha \sum_x \sum_y \sum_p \left[\underline{k} \cdot \nabla X (\underline{v} - \underline{v}_{\mathrm{FR}}) \right]^2
\end{aligned}
\tag{130}
$$

where the subscripts *FR* and *ECMWF* stand for FGGE level II-b and ECMWF level III-b analyses respectively, w and v are the eastward and northward components of the wind respectively, while the coefficients ρ, γ, π, β, and α are weights which control how closely the direct minimization fits each constraint, while L is a convenient length scale allowing the bracketed expressions in the direct minimization functional to be of the same order of magnitude, thus facilitating the unconstrained minimization procedure.

Thacker, Eppel, and Hauser (1986) used a finite-element advective transport computational scheme where they enforced constraints of non-negativity and conservation using a variational multiplier method.

Using the conservation constraint the upwind scheme employed allowed Courant numbers larger than unity. This approach enabled the authors to derive a method of minimizing truncation error and connect it to a finite-difference scheme. The results might, however, suffer from the same problems as exposed in Navon (1987b) and Takacs (1988).

7. Connections Between the Variational and Optimization Method with Other Analysis Methods for Numerical Weather Prediction

The numerical weather prediction (NWP) analysis problem is underdeterminate when one uses observational data alone and in order to resolve the indeterminnancy one has to resort to four-dimensional data assimilation and use prior information to resolve the indeterminancy. One can formulate it by saying that the order of the NWP model N_x versus the observations alone is $N_x >> N_y$. Various methods of analysis related to the problem of determining the most adequate initial conditions for a numerical weather prediction model have been put forward by different researchers for the objective analysis of meteorological data.

When expressed in terms of multi-dimensional probability distribution functions (Kimeldorf and Wahba (1971), Wahba (1978, 1982), Ikawa (1984a, 1984b, 1984c), Pedder (1986), Lorenc (1986), Hollingsworth (1986), Schlatter (1988)) most of the analysis methods for NWP can be shown to be related to the variational approach as well as related to each other.

While each method requires the design of different computationally efficient algorithms, the fact that the methods are related through a general matrix expression whose minimum is sought in order to maximize a probability density function gives us a better insight into the nature of the analysis problem. In the following subsections we will briefly survey some of the connections between variational techniques, optimal interpolation, generalized cross-validation and smoothing splines, the Kalman-Bucy filter and universal Krieging and adjoint model data assimilation.

7.1. The Probability Distribution Function Formulation

A new view of statistical objective analysis using Bayesian probabilities, stimulated by the work of Phillips (1982), Lorenc (1981), and Wahba (1982) was proposed independently by Ikawa (1984a, 1984b, 1984c) and Purser (1984).

Lorenc (1986) synthesized their views in a theoretical review paper which tied together the major meteorological analysis methods and constitutes an up to date reference. Hollingsworth (1986) and Schlatter (1988) have also provided reviews in which statistical objective analysis figures prominently.

The very general formulas for analysis which allow intercomparison between the different analysis methods all start from the basic question: What is the multi-dimensional minimum variance (maximum likelihood) of a particular atmospheric state x_p defined as a grid of a numerical weather prediction model given the numerical forecast data y_p and a set of observations y_0. The state estimator is in balance with a linear constraint and the NWP forecast data is also assumed to be in exact balance with the linear constraint.

Using Lorenc's (1986) formulation and the Bayes theorem which states the conditional probability of an event H occurring, given that event B is known to have occurred

$$\begin{aligned} P(A/B) &\sim P(B/A)P(A) && (131) \\ P_z(x) &\sim \left\{ \int P_0 \left(y_1 - y_0\right) P_f \left(y_1 - K_n(x)\right) dy_1 \right\} P_b \left(x - x_b\right) && (132) \end{aligned}$$

(Using the common assumption that PDF's are multi-dimensional Gaussian functions.)

For the minimum variance Lorenc (1986) shows that we minimize the expression

$$\mathcal{J} = \left\{y_0 - K_n(x)\right\}^T (O+F)^{-1} \left\{y_0 - K_n(x)\right\} + \left(x + x_b\right)^T B^{-1} \left(x - x_b\right) \tag{133}$$

where **B**, **O**, and **F** are covariance matrices for

$$x_b - x_t, \quad y_0 - y_t$$

the background, observation, and forecast errors

$$y = K_n(t)$$

given by

$$\begin{aligned} B &= \left\langle \left(x_b - x_t\right)\left(x_b - x_t\right)^T \right\rangle \\ O &= \left\langle \left(y_0 - y_t\right)\left(y_0 - y_t\right)^T \right\rangle \\ F &= \left\langle \left(y_t - K_n\left(x_t\right)\right)\left(y_t - K_n\left(x_t\right)\right)^T \right\rangle \end{aligned} \tag{134}$$

If one denotes by K the matrix of partial derivatives of K_n with respect to the components on x then

$$K_n\left(x + dx\right) = K_n(x) + K dx \tag{135}$$

and that x_a which minimizes $\mathcal{J}$ in (133) is given by

$$O = K^T (O+F)^{-1} \left\{y_0 - K_n\left(x_a\right)\right\} + B^{-1}\left(x_b - x_a\right) \tag{136}$$

This system can be solved iteratively to allow for nonlinearity in K_n, or if linearization of K_n is valid on the entire range of x, one can obtain (Lorenc (1986))

$$x_a - x_b + \left\{ \left\{ BK^T (O+F)^{-1} (K+I)^{-1} BK^T (O+F)^{-1} \right\} y_0 - K_n\left(x_b\right) \right\} \tag{137}$$

7.2. Duality Between Optimum Interpolation and Variational Objective Analysis

The first work in this domain was the work of Kimeldorf and Wahba (1970, 1971), Wahba (1978).

In terms of a univariate variable on the sphere if $\mathcal{L}(P)$ at a point P of a sphere represents the height minus the global average height at a given reference level and we have observations $y_i, i = 1, \ldots, n$ with zero mean independent measurement errors, ϵ_i, with common variance $\sigma^2 = E\epsilon_i^2$, i.e.,

$$y_i = h\left(P_i\right) + \epsilon \tag{138}$$

Using results from multivariate analysis

$$E\left(h(P)|y_1, \ldots, y_n\right) = \left[R\left(P, P_1\right), \ldots, R\left(P, P_n\right)\right]\left(R_n + n\lambda I\right)^{-1}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = h_\lambda(P). \tag{139}$$

Using that

$$Eh(P)h(Q) = bR(P, Q) \tag{140}$$

where $bR(P, Q)$ is a prior covariance.

In (139) $\lambda = \frac{\sigma^2}{nb}$, while R_n is an $n \times n$ matrix with entries $R\left(P_i, P_j\right)$ in the ij^{th} location.

If $\lambda = 0$ then $h_0(P)$ interpolates to the observed data exactly, i.e.,

$$h_0\left(P_i\right) = y_i \tag{141}$$

while for $\lambda > 0$, $h_\lambda(P)$ implies a smoothing of the data.

Wahba (1982) points out that λ controls the amount of smoothing and $h_\lambda(P)$ evaluated at grid points can be viewed as the optimum interpolant of Gandin (1965) if all available observations are used simultaneously.

The Duality Theorem (Kimeldorf and Wahba, (1970, 1971), Wahba (1978)) then states that for every covariance $R(P, Q)$ satisfying

$$\int\int R^2(P, Q)dP\,dQ < \infty \tag{142}$$

there is a variational problem for which $h_\lambda(P)$ is the solution. It is: find h in a certain reproducing kernel Hilbert Space H_R so as to minimize

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_i - h\left(P_i\right)\right)^2 + \lambda\mathcal{J}(h) \tag{143}$$

where $\mathcal{J}(h)$ is the square norm of h in H_R.

Using the Mercer-Hilbert-Schmidt expansion of the covariance $R(P,Q)$, Craven and Wahba (1979), and Wahba and Wandelberger (1980) show that

$$\mathcal{J}(h) = \sum_{\ell,s} \frac{h^2_{\ell,s}}{\lambda_{\ell,s}} \tag{144}$$

where $R(P,Q) = \sum_{\ell,s} \lambda_{\ell,s} y_{\ell,s}(P) y_{\ell,s}(Q)$ and where $\int R(P,Q) y_{\ell,s}(Q) dQ = \lambda_{\ell,s} y_{\ell,s}(P)$, where $y_{\ell,s}$, the eigenfunctions of R are spherical harmonics and $\lambda_{\ell,s}$ are eigenvalues of R, and

$$h_{\ell,s} = \int h(P) y_{\ell,s}(P) dP. \tag{147}$$

A detailed investigation of the relationship between variational analysis (Sasaki (1970a, 1970b, 1970c)) and the multivariate O/I algorithm has been carried out by Ikawa (1984a, 1984b, 1984c).

Ikawa (1984b, 1984c) shows that the slow-mode covariance matrix used in O/I has the same filtering properties as the variational method, i.e., that a covariance matrix consistent with the linear constraint operates as a filter without the explicit imposition of linear constraints as done in the variational objective analysis.

The work of Ikawa confirms the Phillips (1982) analysis about the completeness of multivariate O/I, i.e., that no more useful information can be extracted from data by performing a variational analysis with balance constraints which constitute slow-mode constraints. This in short was an equivalence result between a slow-mode O/I and the variational analysis. Wahba (1982) also discussed the possibility of including O/I and normal mode initialization balance constraints into one step in the framework of an "optimal" variational formulation which will have a bandwidth parameter λ, an error balancing parameter w which will control the relative weight to be given to forecast data and observational data and one partitioning parameter δ, governing the relative energy in the "signal" assigned to fast and slow modes.

These parameters are supposed to be chosen by generalized cross validation (Wahba and Wandeberger (1980)).

Lorenc (1986) generalized further the analysis problem to pose it as the minimization of a penalty function comprising terms depending on the distance of analysis from data as well as the distance of the analysis from prior information weighted by the relative accuracy of each term. One can use the property of L_2 norms by which any number of weak constraints can be combined into a single L_2 penalty.

However, the most complete proof of the equivalence relies on the work of Wahba and Wandelberger (1980) and Wahba (1982c) and references therein. To illustrate the ideas of their work one can take a vector of variables of interest $\hat{x} = (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n)^T$. Assume

$$z_i = \hat{x}_i + \epsilon_i \tag{148}$$

is observed for $i = 1, \ldots, n$, where ϵ_i are independent Gaussian random variables with zero mean and variance σ^2.

If x_i have a prior Gaussian distribution with

$$E\hat{x}_i = 0 \tag{149}$$

and

$$E\hat{x}_i\hat{x}_j = \sigma_{ij} \tag{150}$$

with E being the mathematical expectation, and if one defines by Σ the matrix with entries σ_{ij}(the covariance matrix), the conditional (Bayesian) expectation $E\hat{x}$ of $\hat{x}$, given the observation data $z = (z_1, \ldots, z_n)^T$ is

$$E\hat{x} = \Sigma\left(\Sigma + \sigma^2 I\right)^{-1} z \tag{151}$$

but $E\hat{x}$ is also the solution of the minimization problem: Find x to minimize

$$\frac{1}{n}\sum_{i=1}^{n} (\hat{x}_i z_i)^2 + \lambda \mathcal{J}\left(\underline{\hat{x}}\right) \tag{152}$$

where

$$\mathcal{J}\left(\hat{x}\right) = \hat{\underline{x}}^T \Sigma^{-1}\hat{x} \quad \text{and} \quad \lambda = \underline{\sigma}_n^2. \tag{153}$$

In general $\mathcal{J}$ assumes the form

$$\mathcal{J} = \int\int \left[\left(\frac{\partial^2 \hat{x}}{\partial x^2}\right) + 2\left(\frac{\partial^2 \hat{x}}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 \hat{x}}{\partial y^2}\right)^2\right] dx\, dy. \tag{154}$$

As discussed previously for spherical harmonics

$$\mathcal{J}\left(\hat{x}\right) = \int\int \left(\Delta^m \hat{x}\right)^2 dP \tag{155}$$

For instance (Wahba(1982)), given observed wind data (U_i, V_i) at a point P_i, $i = 1, 2, \ldots, n$ one can estimate vorticity and divergence as follows

$$\Psi(P) = \sum_{\ell=1}^{L} \sum_{s=-\ell}^{\ell} a_{\ell,s} y_{\ell,s}(P), \quad \Phi(P) = \sum_{\ell=1}^{L} \sum_{s=-\ell}^{\ell} b_{\ell,s} y_{\ell,s}(P) \tag{156}$$

where $y_{\ell,s}$ are the spherical harmonics.

Then for given δ and λ one can find coefficients $a_{\ell,s}$, $b_{\ell,s}$ to minimize

$$\begin{aligned} &\frac{1}{n} \sum_{i=1}^{n} \left(-\frac{1}{a} \frac{\partial \Psi}{\partial \Phi_i'} (P_i) + \frac{1}{a \cos \Phi_i'} \frac{\partial \Phi}{\partial x} (P_i) - u_i \right)^2 \\ &\qquad + \frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{a \cos \Phi_i} \frac{\partial \Psi}{\partial x'} (P_i) + \frac{1}{a} \frac{\partial \Phi}{\partial \Phi_i'} (P_i) - v_i \right)^2 + \lambda \left(\mathcal{J}_1(\Psi) + \frac{1}{\delta} \mathcal{J}_2(\Phi) \right) \end{aligned} \tag{157}$$

where

$$\mathcal{J}_1(\Psi) = \sum_{\ell=1}^{L} a_{\ell,s}^2 / \lambda_{\ell,s}^{(1)}, \quad \mathcal{J}_2(\Psi) = \sum_{\ell=1}^{L} b_{\ell,s}^2 / \lambda_{\ell,s}^{(2)} \tag{158}$$

Here λ can be viewed as a bandwidth parameter, δ as a signal partitioning parameter and $\lambda_{\ell,s}^{(1)}$ and $\lambda_{\ell,s}^{(2)}$ are weights adapted from collected data sets.

7.3. Smoothing Splines, Generalized Cross-Validation and Variational Analysis

Reinsch (1967) considered the following estimation problem which led to the smoothing spline interpolation method.

Given n discrete observations

$$x_j = x\left(t_j\right) + z_j \tag{159}$$

where x_j can be considered as observation data at position t_j and z_j a random normal error with zero mean and variance σ_j^2, estimate $x(t)$ as a linear function of x_j, i.e.,

$$\hat{x}(t) = \sum_{j=1}^{n} w_j x_j \tag{160}$$

Reinsch (1967) avoids the problems of both stochastic and deterministic models which require specification of a parametric model by seeking instead a solution of $\hat{x}(t)$ satisfying smooth interpolation requirements, but which does not involve specifying a parametric model for either $x(t)$ or for a stationary covariance function $v(\tau)$.

This method involves finding the solution for $\hat{x}(t)$ which minimizes a functional $\mathcal{J}$ given by

$$\mathcal{J} = \frac{1}{N_y} \sum_{i=1}^{N_y} w_j \left(y\left(t_i\right) - g\left(t_i\right)\right)^2 + \lambda \int \mathcal{J}_s(g) dt \tag{161}$$

or in Reinsch (1967) notation

$$\mathcal{J} = \lambda \int_{\Gamma} \left[\frac{\partial^2 \hat{x}(t)}{\partial t^2} \right]^2 dt + \frac{1}{N_y} \Sigma w_j \left[\frac{(\hat{x} - x)}{\sigma} \right]^2 - \delta^2 + \gamma^2 \tag{162}$$

where the $\mathcal{J}_s$ term is a penalty on smoothness with reference to the integral of squared curvature over the domain Γ. If the smoothness is related to the m^{th} derivative of $\hat{x}(t)$, $\frac{d^m(\hat{x}(t))}{dt^m}$, this leads to the polynomial splines of order $2m - 1$.

For prescribed values of m and λ the variational problem has a unique analytic solution given by

$$\hat{x}(t) = \Phi'(t)\underline{\alpha} + \sum_{j=1}^{n} \beta_j K\left(\Delta S_{j,\ell}\right) \tag{163}$$

where Φ represents a vector of polynomials complete to order $(m-1)$, α is an associated vector of parameters, β_j are parameters associated with $K\left(\Delta S_{j,\ell}\right)$ where $\Delta S_{j,\ell}$ is the distance in Euclidean d-space between t_j and t_ℓ

$$\underline{t} = (t_1, \ldots, t_d) \tag{164}$$

with d the dimension of d-dimensional Euclidean space and

$$K\left(\Delta S_{j,\ell}\right) = \left(\Delta S_{k,\ell}\right)^2 \log\left(\Delta S_{k,\ell}\right) \tag{165}$$

(see Pedder (1986), Thiebeaux and Pedder (1987)).

Solution of (162) can be viewed as a generalized spline, as the piecewise continuous (solution) property of $\hat{x}(t)$ is similar to univariate splines. Wahba and Wandelberger (1980) generalized this notion to seek the solution of the problem: find Φ in a suitable space x to minimize

$$\frac{1}{N_y} \sum_{i=1}^{n} \left[\Phi\left(x_i, y_i\right) - \tilde{\Phi}_i \right]^2 + \lambda \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \sum_{\nu=1}^{m} \binom{m}{\nu} \left(\frac{\partial^m \Phi}{\partial x^\nu \partial^{m-\nu}} \right) dx\, dy \tag{166}$$

which was obtained by Duchon (1976) and studied by Meinguet (1979) and Wahba (1979a, 1979b).

In frequency space it can be shown that λ controls the half-power point of the filter and m the steepness of the roll-off.

7.4. Equivalence of Best-Fit Trajectory to Kalman Filtering

This method uses the representation of the cost functional proposed by Thacker (1986), where the new observations $\hat{x}_n$ and forecast $\hat{z}_n$ at time t_n have error-covariance matrices R_n and M_n, respectively, giving an estimate of the state

$$x_n = \hat{z}_n + \beta_n \left(\hat{x}_n - z_n\right) \tag{167}$$

$$\beta_n = M_n \left[M_n + R_n\right]^{-1} \tag{168}$$

which is a weighted average of the new data with the forecast with weights proportional to the inverse of the respective error covariance matrices

$$P_n^{-1} = M_n^{-1} + R_n^{-1}. \tag{169}$$

If we use the best fit dynamic trajectory, i.e., finding the best trajectory passing as near as possible to the asynoptic data, while minimizing a cost function measuring the misfit of the dynamics to the data,

$$F = \sum_{n=3}^{N} \left(x_n - \hat{x}_n\right)^T R_n^{-1} \left(x_n - x_n^T\right) + \sum_{n=1}^{N} \left(f_n - \hat{f}_n\right) Q_n^{-1} \left(f_n - \hat{f}_n\right) \tag{170}$$

where f_n represent forcing terms and $\hat{x}_n$ are the state observations with weights proportional to inverse error covariance matrices. Optimal x_n and f_n are obtained by minimizing cost function subject to the strong constraint of satisfying the model's dynamics. If the dynamics (forecast equations) are represented by a linear model

$$x_n = A_n x_{n-1} + f_n \tag{171}$$

$$\begin{aligned} c = {} & \textstyle\sum_{N=0}^{n} \left(x_n - \hat{x}_n\right)^T R_n^{-1} \left(x_n - \hat{x}_n\right) \\ & + \textstyle\sum_{N=1}^{n} \left(x_n - A_n x_{n-1} - \hat{f}_n\right)^T \hat{R}_n^{-1} \left(x_n - A_n x_{n-1} - f_n\right) \end{aligned} \tag{172}$$

If one wishes to pose the problem as an Augmented-Lagrange problem, i.e.,

$$L = C + \sum_{n=1}^{N} \Lambda_n^T \left[x_n - A_n x_{n-1} - f_n\right] \tag{173}$$

where L is the Augmented-Lagrangian, C is Eq. (171) and if one requires the gradient of L with respect to each Λ_n and x_n to vanish, one obtains the same equation as in the adjoint approach (Le Dimet and Talagrand (1986)), as the Lagrange multipliers can be shown to be identical to the adjoint variables.

To show that the Kalman filtering method produces the same result when the two methods use the same information for x_n as the best fit dynamical trajectory approach, Thacker (1986) designs a sequential algorithm to solve for x_1 as a function of Λ_2 and of the data at the first two time levels.

As mentioned previously the big disadvantage of the Kalman-Bucy filtering technique is the requirement for the calculation of the error-covariance matrix at each time-step. This in general requires a computational effort equivalent to that required by the full numerical weather prediction system multiplied by twice the number of degrees of freedom of the full system which for present day 3-dimensional models is computationally premature.

7.5. Splines and Universal Krieging

In mining practice, one problem is to find the best possible estimator of the mean grade of a mined block taking into account the assay values of the different samples available either inside or outside the block to be estimated. Krieging (following the basic regression procedure of D. G. Kriege (1951)) is a procedure of selecting within a given class of possible estimator, the estimator with a minimum estimation variance, *i.e.*, the estimator, leading to a minimum variance of the resulting estimation error. The method is amply described by Matheron (1963, 1970, 1981), Riccardo (1974), and Journel (1977). This minimum variance estimate of deviations from a trend field can be a low-order polynomial obtained by minimizing the variance best fit to the observational data.

For a random variable $Z(X)$ with covariance

$$\langle Z(X)Z(Y)\rangle = E\left\{Z(X)Z(Y)\right\} = \sigma_{xy} \tag{174}$$

The minimum distance

$$E\left\{\left[Z\left(X_0\right) - Z^*\right]^2\right\} \tag{175}$$

is the called minimum estimation variance.

The Krieging estimator is

$$Z^* = \lambda_0 + \sum_{\alpha} \lambda_\alpha Z\left(X_\alpha\right) \tag{176}$$

If the expectation is neither stationary not know, but of a known linear combination of L known functions $f_\ell(X)$, $\ell = 1, \ldots, L$

$$E\{Z(X)\} = m(X) = \sum_\ell A_\ell f_\ell(X) \tag{177}$$

where the L parameters A_ℓ are unknown, we define $\nu(X)$, the nonstationary expectation as a trend or drift.

In Universal Krieging (UK) the unbiasedness of the linear estimator $\lambda_\alpha Z(X_\alpha)$ restricts it to a linear manifold defined by the L conditions on the weights

$$\sum_\alpha \lambda_\alpha f_\ell(X_\alpha) = f_\ell(X_0) \quad \text{for } \ell = 1, \ldots, L \tag{178}$$

and the Krieging estimator Z^*_{KL} is the element of the linear manifold nearest to the unknown $Z(X_0)$ (Journel (1977)), and the weights λ_α of Z^*_{KL} must satisfy

$$\sum_\alpha \lambda_\alpha f_\ell(X_\alpha) = f_\ell(X_0) \tag{179}$$

and

$$\|Z(X_0) - \lambda_\alpha Z_\alpha\| = d^2 = \min. \tag{180}$$

The second requirement amounts to minimizing the expression

$$\mathcal{L} = d^2 - \sum_\ell 2\mu_\ell \sum \lambda_\alpha f_\ell(X_\alpha) \tag{181}$$

where μ_ℓ, $\ell = 1, \ldots, L$ are Lagrange multipliers.

If we denote by

$$\sigma_{\alpha\beta} = \left\langle Z(X_\alpha), Z\left(X_\beta\right)\right\rangle = E\left\{Z(X_\alpha) Z\left(X_\beta\right)\right\} \tag{182}$$

the non-centered variance, the minimum of $\mathcal{L}$ is obtained by setting to zero the partial derivatives of Q with respect to the unknown weights λ_α

$$\frac{\partial \mathcal{L}}{\partial \lambda_\alpha} = 0 \quad \alpha = 1, \ldots, u \tag{183}$$

with

$$d^2 = \sigma_{X_0 X_0} - 2\lambda_\alpha \sigma_{\alpha X_0} + \lambda_\alpha \lambda_\beta \sigma_{\alpha\beta}. \tag{184}$$

One can show (Matheron (1970, 1981)) that the universal Krieging solution minimizing estimation error variance $E\left\{\left[Z\left(X_0\right)-Z_{KL}^*\right]^2\right\}$ can be written as

$$\hat{Z}\left(X_0\right)=m(X)+\nu'(X)\left(\nu+\mathbf{D}\right)^{-1}\left(Z-m(X)\right) \tag{185}$$

where $m(X)$ is the estimated drift component, $\nu'(X)$ is a vector of "station-to-gridpoint" covariances of the form

$$\lambda_\alpha \sigma_{\alpha X_0} \tag{186}$$

(see Pedder (1986)) while $(\nu+\mathbf{D})$ is the observation covariance matrix where D is a diagonal matrix of observation error covariance of the form

$$\sigma_{X_0 X_0} \tag{187}$$

As shown by Matheron (1981) and Pedder (1986) the generalized spline approach solution takes a similar form to that of the universal Krieging.

Thus, the minimum variance estimate of deviations from a trend field is shown to be equivalent to the use of polynomial splines in the Reinsch (1971) smoothing spline approach using the "thin-plate" solution obtained by Duchon (1976).

8. Conclusion

In recent years, variational methods have experienced a large expansion in their use. This is due in our opinion to their ability to be very flexible, as well as to their ability to be adapted to various physical frameworks.

Furthermore, many significant developments have been carried out by the mathematical community in what concerns the development of efficient optimization methods. It is clearly evident by now that these methods may be extended without difficulty to meteorological cases of interest.

A large amount of research work has also been carried out especially for creating a link with stochastic techniques of the Kalman-Bucy filter type as well as to obtain well adapted optimization algorithms (Ghil, *et al.* (1981), Navon and Legler (1987)).

Variational methods are not merely a computational trick, but they constitute another way to conceive meteorological modelling by working on both the data and the model.

Acknowledgements

The authors are indebted to Professor Jesse Stephens of the Department of Meteorology, Florida State University, who has contributed valuable constructive criticism as well as allowing them access to his unique collection of variational papers. Donna Middleton and Mimi Burbank did a splendid job of professionally typing a difficult manuscript.

This work was supported by the Florida State University Supercomputer Computations Research Institute which is partially funded by the U.S. Department of Energy through cooperative agreement number DE-FC05-85ER250000.

References

Achtemeier, G. L., 1972: "*Variational initialization of atmospheric fields: a quasi-geostrophic diagnostic model*," Ph. D. thesis, Dept. of Meteorology, Florida State University, Tallahassee, Florida, 32306, 216 pages.

Achtemeier, G. L., 1975: "*On the initialization problem: a variational adjustment method*," *Mon. Wea. Rev.*, **103**, 1089–1103.

Achtemeier, G. L., H. T. Ochs, S. Q. Kidder and R. W. Scot, 1986: "*Evaluation of a multivariate variational assimilation of conventional and satellite data for the diagnosis of cyclone systems*," in Variational Methods in Geosciences, (Y. K. Sasaki, ed.) Elsevier p. 49–54.

Arakawa, A., 1966: "*Computational design for long-term numerical integration of the equations of fluid motion: Part I: Two-dimensional incompressible flow*," *Journal of Computational Physics*, **1**, 119–143.

Arakawa, A. and V. R. Lamb, 1977: "*Computational design of the basic dynamical processes of the UCLA general circulation model*," *Methods in Computational Physics*, **17**, J. Chang, Ed., Academic Press, 173–265.

Arakawa, A. and V. R. Lamb, 1981: "*A potential enstrophy conserving scheme for the shallow-water equations*," *Monthly Weather Review*, **109**, 18–36.

Barker, E., G. Haltiner, and Y. Sasaki, 1977: "*Three dimensional initialization using variational analysis*," in Proc. 3rd Conference on Numerical Weather Prediction of the AMS, Omaha, 169–181.

Bayliss, A. and E. Isaacson, 1975: "*How to make your algorithms conservative*," *Notices Amer. Math. Soc.*, Aug. A594–A595.

Bertsekas, D. P., 1975: "*Combined primal-dual and penalty methods for constrained optimization*," *SIAM Journal of Optimal Control*, **13**, 521–544.

Bertsekas, D. P., 1976: "*Multiplier methods: a survey*," *Automatica*, **12**, 133–145.

Bertsekas, D. P., 1982: "*Constrained optimization and Lagrange multiplier methods*," Addison-Wesley, 491 pages.

Bloom, S., 1983: "*The use of dynamical constraints in the analysis of mesoscale rawinsonde data*," *Tellus*, **35**, 363–378.

Bourke, W. and J. L. McGregor, 1983: "*A nonlinear vertical mode initialization scheme for a limited area model*," *Mon. Wea. Rev.*, **111**, 2285–2297.

Brody, R. L., 1974: "*Fine mesh sea level pressure analysis scheme incorporating effects of elevation and terrain*," U. S. Env. Pred. Research Facility NPS, Naval Postgraduate School, Monterey, CA 93940, Tech. Note 16, 31 pages.

Buckley, H. G., and A. Lenir, 1983: "*QN-like variable storage conjugate gradients*," *Math Program*, **27**, 155–175.

Buckley, H. G., and A. Lenir, 1985:"*Algorithm 630-BBVSCG – A variable storage algorithm for function minimization*," *ACM Trans. Math Software*, **11(2)**, 103–119.

Cacuci, D. G., 1981: "*Sensitivity theory for nonlinear systems I. Nonlinear functional analysis approach*," *J. Math. Phys.*, **22**, 2794–2802.

Cacuci, D. G., 1981: "*Sensitivity theory for nonlinear systems II. Extensions to additional classes of responses*," *J. Math. Phys.*, **22** (12), 2803–2812.

Cacuci, D. G. and M. C. G. Hall, 1984: "*Efficient estimation of feedback effects with application to climate models*," *J. Atmos. Sci.*, **41**, 2063–2068.

Charba, J. and Y. Sasaki, 1981: "*Structure and Movement of the Severe Thunderstorm of April 13, 1964, as Revealed from Radar and Surface Mesonetwork Data Analysis*," *Journal of Meteorology Japan*, **49**, 191–214.

Chen, W. H. and J. H. Seinfeld, 1972: "*Estimation of spatially varying parameters in partial differential equations*," *Int. J. Control*, **15**, 487–495.

Chong, M. and F. Roux, 1983: "*Three-dimensional wind field analysis from dual-Doppler radar date. Part II: Minimizing the error due to temporal variation*," *J. Climate Appl. Meteor*, **22**, 1216–1226.

Chong, M. and J. Testud, 1983: "*Three-dimensional wind field analysis from dual-Doppler radar data. Part III: The boundary condition. An optimum determination based on a variational concept*," *J. Climate Appl. Meteor*, **22**, 1227–1241.

Cohn, S. E., 1982: "*Methods of sequential estimation for determining initial data in numerical weather prediction*," Ph. D. thesis, Courant Institute for Mathematical Sciences, Report CI-6-82, New York University, New York, NY, 16012, 183 pages.

Courtier, P., 1985: "*Experiments in data assimilation using the adjoint model technique*," Workshop on High-Resolution Analysis ECMWF, UK, June 1985, 20 pages.

Courtier, P., 1986: "*Use of adjoint equations for assimilation of meteorological observations by barotropic models*," in Variational Methods in Geosciences (Y. K. Sasaki, ed.), Elsevier, P. 35–42.

Courtier, P. and O. Talagrand, 1987: "*Variational assimilation of meteorological observations with the adjoint vorticity equation. Part II. Numerical results*," *Quart. Jour. Roy. Met. Soc.*, **113**, 1329–1347.

Cram, J. M. and M. L. Kaplan, 1985: "*Variational Assimilation of VAS data into a mesoscale model: assimilation method and sensitivity experiments*," *Mon. Wea. Rev.*, **113**, 467–484.

Craven, P., and G. Wahba, 1979: "*Smoothing noisy data with spline functions*," *Numer. Math*, **31**, 377–403.

Daley, R., 1978: "*Variational non-linear normal mode initialization*," *Tellus*, **30**, 201–218.

Daley, R. and K. Puri, 1980: "*Four dimensional data assimilation and the slow manifold*," *Mon. Wea. Rev.*, **108**, 85–99.

Derber, J. C., 1985: "*The variational 4-D assimilation of analyses using filtered models as constraints*," Ph. D. Thesis, Dept. of Meteorology, Univ. of Wisconsin-Madison, Madison, WI 53706, 142 pages.

Derber, J. C., 1986: "*The variational four-dimensional assimilation of analysis using quasi-geostrophic models as constraints*," in Variational Methods in Geosciences (Y. K. Sasaki, ed.), Elsevier, p. 43–48.

Derber, J. C., 1987: "*Variational four-dimensional analysis using quasi-geostrophic constraints*," *Mon. Wea. Rev.*, **115**, 998–1008.

Dickerson, M. H., 1971: "*Objective analysis of lower stratospheric temperature and height data*," Ph. D., Dept. of Meteorology, Florida State University, Tallahassee, Florida, 32306–3034, 117 pages.

Dickerson, M. H., 1978: "*MASCON - a mass-consistent atmospheric flux model for regions with complex terrain*," *J. Appl. Meteor.*, **17**, 241–253.

Duchon, J., 1976: "*Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces*," *RAIRO Anal. Numer.*, **10**, 5–12.

Duewer, W. H., *et al.*, 1978: "*The Livermore regional air quality model*," *J. Appl. Meteor.*, **17**, 273–311.

Eddy, A., 1967: "*The statistical objective analysis of scalar data fields*," *J. Appl. Meteor.*, **6, No. 4**,597–609.

Endlich, R. M. and R. L. Mancuso, 1968: "*Objective analysis of environmental conditions associated with severe thunderstorms and tornadoes*," *Mon. Wea. Rev.*, **96** (6), 342–350.

Euler, L., 1744: "*Methodus inveniendi lineas curvas maximi minimive propietate gandentes sive solutio problematis isoperimetrici latissimo sensu accepti*," Lausanne, Opera S. I., 24 (1952) 174–189, ed. Constantin Caratheodory.

Euler, L., 1764: "*Elementa calculi variationum novi comentarii academiae scietiarum petropolitane*," 10 (1764) 1766 pages. 51–93 Opera S. I., 25, Caratheodory, (1952), Basel, 141–176.

Fletcher, R., 1980a: "*Practical methods of optimization. Vol. 1: Unconstrained optimization*," Wiley and Sons, New York, 120 pages.

Fletcher, R., 1980b: "*Practical methods of optimization. Vol. 2: Constrained optimization*," Wiley and Sons, New York, 224 pages.

Fraser, C., 1985: "*J. L. Lagrange's changing approach to the foundations of the calculus of variations*," *Archive for History of Exact Sciences*, **32, No. 2**, 159–191.

Gal-Chen, T., 1982: "*Errors in fixed and moving frame of references: applications for conventional and Doppler radar analysis*," *Jour. Atmos. Sci.*, **39**, 2279-2300.

Gal-Chen, T., B. D. Schmidt, L. W. Lucellin, 1986: "*Simulation Experiments for Testing the Assimilation of Geostationary Satellite Temperature Retrievals IntoNumerical Prediction Models*," *Monthly Weather Review*,**114**, 1213–1230.

Gal-Chen, T., 1987: "*A theory for the retrievals of virtual temperature from remote measurements of horizontal winds and thermal radiation.*," *Monthly Weather Review*, **116**, 1302–1319.

Galerkin, B. G., 1915: "*Series occurring in various questions concerning the elastic equilibrium of rods and plates*," *Vestnik Inzhenerov Tech.*, **19**, 897–908.

Gandin, L. S., 1965: "*Objective analysis of meteorological fields*," Translated from Russian by Israel Programs for Scientific Translations, Jerusalem, Israel, 292 pages [NTIS–TT65–50007].

Ghil, M. S., and R. Mosebach, 1978: "*Asynoptic variational method for satellite-data assimilation*," in Halem *et al.* The GISS sounding temperature impact test, NASA Tech. Memo 78063, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, 332–349, 422 pages.

Ghil, M., S. Cohn, J. Tavantis, V. Bube and E.Isaacson, 1981: "*Applications of estimation theory to numerical weather prediction in dynamic meteorology: data assimilation methods*," (Bengtsson, Ghil, and Källén, eds.) Springer-Verlag, **36**, New York, p. 139–224.

Ghil, M., S. E. Cohn and A. Dalcher, 1982: "*Sequential estimation, data assimilation and initialization*," in Proc. of the Fourteenth Stanstead Seminar (D. Williamson, ed.), McGill University, Montreal, Quebec, Canada, p. 83–97.

Gill, P. E., W. Murray and M. H. Wright, 1981: "*Practical Optimization*," Academic Press, 401 pages.

Glowinski, R., 1984: "*Numerical methods for non-linear variational problems*," Springer-Series, *Computational Physics*, New York, 493 pages.

Halem, M. and E. Kalnay, 1983: "*Variational global analysis of satellite temperature soundings*," 6$^{\text{th}}$ Conference on Numerical Weather prediction of the AMS, June 6–9, 1983, Omaha, Nebraska, pp. 120–127.

Hall, M. C. G., D. G. Cacuci and M. E. Schlesinger, 1982: "*Sensitivity analysis of a radiative-convective model by the adjoint method*," *J. Atmos. Sci.*, **39**, 2038–2050.

Hall, M. C. G. and D. G. Cacuci, 1983: "*Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models*," *J. Atmos. Sci.*, **40**, 2537–2546.

Haltiner, G., Y. K. Sasaki and E. Barker, 1975: "*A variational procedure for obtaining global balanced winds*," in Proceedings J. O. C. Study Group Conf. on Four Dimensional Data Assimilation, 17–22 Nov., Paris, 198–223.

Haltiner, G. and E. H. Barker, 1976: "*Initial balancing with a variational method*," *Annalen der Meteorolgic, Neue Folge*, **111**, 119–121.

Hantel, M., 1986: "*Variational modification of the 3-d wind field*," in *Variational Methods in the Geosciences* (Y. K. Sasaki, ed.), 107–112.

Harlan, J. and J. O'Brien, 1986: "*Assimilation of scatterometer wind into surface pressure fields using a variational method*," *J. of Geophys. Res.*, **91**, D7, 7816–7836.

Heimbach, J. A. and Y. K. Sasaki, 1975: "*A variational technique for mesoscale objective analysis of air pollution*," *J. of Appl. Met.*, **14**, 194–203.

Hoffman, R. N., 1982: "*SASS wind ambiguity removal by direct minimization*," *Mon. Wea. Rev.*, **110**, 434–445.

Hoffman, R. N., 1984: "*SASS wind ambiguity removal by direct minimization: Part II: Use of smoothness and dynamic constraints*," *Mon. Wea. Rev.*, **112**, 1829–1852.

Hoffman, R. N., 1986: "*A four-dimensional analysis exactly satisfying equations of motion*," *Mon. Wea. Rev.*, **114**, 388–397.

Holl, M. and B. Mendenhall, 1972: "*Fields by information blending*," Tech. Note 72-2, Fleet Numerical Weather Central, Monterey, CA, 66 pages, WTIS-AD-A094-082.

Hollingsworth, A., 1986: "*Objective Analysis for Numerical Weather Prediction*," *Tech. Memo No. 128*, ECMWF, Shinfield Park, Berkshire, England (also paper presented at WMO Symposium on NWP, Tokyo, August 1986), 91 pages.

Ikawa, M., 1984a: "*An Alternative Method of Solving Weak Constraint Problems and a Unified Expression of Weak and Strong Constraints in Variational Objective Analysis*," *Papers in Meteorology and Geophysics*, **35, No. 3**, 71–79.

Ikawa, M., 1984b: "*Generalization of Multivariate Optimum Interpolation Method and the Roles of Linear Constraint and Covariance Matrix in the Method, Part I: In Discrete Form*," *Papers in Meteorology and Geophysics*, **35**, 81–102.

Ikawa, M., 1984c: "*Generalization of Multivariate Optimum Interpolation Method and the Roles of Linear Constraint and Covariance Matrix in the Method, Part II: In Continuous Form*," *Papers in Meteorology and Geophysics*, **35, No. 4**, 169–180.

Isaacson, E., 1977: "*Integration schemes for long-term calculations*," *Advances in Computer Methods for PDE's* (A. Vishnevetsky, ed.), AICA 251-255.

Isaacson, E., D. Marchesin and G. Zwas, 1979: "*Numerical methods for meteorology and climatology*," *Fourth NASA Weather and Climate Program Science Review* (Earl E. Kreins, ed.), NASA Conf. Pub. 1076, paper 31, 183–190.

Johnson, K. W., 1975: "*A dynamic programming approach to the variational adjustment of meteorological data on isotropic surfaces*," M. Sc. Thesis, Dept. of Meteorology, Florida State University, Tallahassee, FL, 32306-3034, 78 pages.

Jongdao, J., 1978: "*Application of the variational adjustment method in evaluation of the performance of various conventional objective analysis schemes*," in *Mesoscale diagnostic numerical variational analysis models*, J. J. Stephens and R. L. Inman, Final Report DAAD07-76-C-0037, Atmospheric Sciences Lab., White Sands Missile Range, New Mexico 88002, 99 pages.

Journel, A.G., 1977: "*Krieging in Terms of Projections*," *Math. Geology*, **9**, 563–586.

Kalman, R. E., 1960: "*A New Approach to Linear Filtering and Prediction Problems*," *J. Basic Eng (Trans. Amer. Soc. Mech. Eng.)*, **82, Series D, No. 1**, 35–45.

Kalnay-Rivas, E., A. Bayliss, and J. Storch, 1977: "*The 4th order GISS model of the global atmosphere*," *Contrib. Atmos. Phys.*, **50**, 306–311.

Kalnay-Rivas, E., 1979: "*The effect of accuracy, conservation and filtering on numerical weather forecasting*," in the Proceedings of the Fourth Conference on Numerical Weather Prediction, Silver Spring, MD, American Meteorological Society, 302–312.

Kasahara, 1982: "*Nonlinear normal mode initialization and the bounded derivative method*," *Rev. Geophys. Space Res.*, **20**, 385–397.

Kimeldorf, G. S. and G. Wahba, 1970: "*A correspondence between Bayesian estimation on stochastic processes and smoothing by splines*," *Ann. Math. Statist.*, **41**, 495–502.

Kimeldorf, G. and G. Wahba, 1971: "*Some Results on Tchebycheffian Spline Functions*," *J. Math. Anal. and Applic.*, **33**, 1.

Kontarev, G., 1980: "*The adjoint equation technique applied to meteorological problems*," ECMWF Tech. Rep. No. 21, Reading, U.K., 21 pages.

Kriege, D. J., 1951: "*A statistical approach to solve mine valuations and allied problems on the Witwatersrand*," unpublished Master Thesis, Univ. of Witwatersrand, Johannesburg, South Africa, 45 pages.

Lagrange, J. L., 1760a: "*Essai d'nne nouvelle methode pour determiner les maxima et les minima des formules indefinies*," *Miscellanea Taurinensia*, **2** (1762) Oeuvres, 1, 365–468.

Lagrange, J. L., 1760b: "*Application de la methods expose dans le memoire precedent a la solution de differentes problemes de dynamique*," *Miscellanea Taurinensia*, **2** (1762) Oeuvres, 1, 365–468.

Lamb, A., W. H. Chen, and J. H. Seinfeld, 1975: "*Numerico-empirical analyses of atmospheric diffusion theories*," *Jour. Atmos. Sciences*, **32**, 1794–1807.

Le-Dimet, F. X., 1981: "*Une application des methodes de côntrole, optimal à l'analyse variationnelle*," *Rapport Scientifique LAMP*, Université Blaise-Pascal, 63170, Aubiere Cedex, 42 pages.

Le-Dimet, F. X., 1982: "*A general formalism of variational analysis*," CIMMS Report No. 22, 34 pages, Norman, OK 73019.

Le-Dimet, F. X., Y. K. Sasaki, and L. White, 1982: "*Dynamic initialization with filtering of gravity waves*," CIMMS Report No. 40, Norman, OK 73019, 12 pages.

Le-Dimet, F. X. and J. Segot, 1985: "*Variational analysis of wind field and geopotential at 500 mb*," in Proc. Int. Symp. on Variation Meth. in Geosciences, Norman, OK, 73019, Oct. 15-17, 1985, pg. 13–15.

Le-Dimet, F. X. and A. Nouallier, 1986: "*Assimilation of dynamic data in a limited-area model in variational methods in geosciences*, (Y.K. Sasaki, editor) Elsevier, Amsterdam 181–198.

Le-Dimet, F. X. and O. Talagrand, 1986: "*Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects*," *Tellus*, **38A**, 97–110.

Legler, D., I. M. Navon, and J. J. O'Brien, 1988: "*A direct minimization approach to objective analysis of pseudo-stress over the Indian Ocean*," *to appear in the Monthly Weather Review*, also Supercomputer Computations Research Institute Technical Report FSU–SCRI–88–42, Florida State University, Tallahassee, FL, 32306–4052.

Lemarechal, C., 1980:"*Non-differentiable optimization in Non-linear Optimization: Theory and algorithms*," Dixon, Spedicato, and Szego (Editors), (Proc. International Summer Seminar, Univ. of Bergamo, Italy, 1979), Birkhauser, Cambridge, Boston, MA, 486 pages.

Lemarechal, C., and P. Servigne, 1984: "*Le module M1GC2 in Modulopt*," Institut National de Recherche en Informatique et Automatique, Roequencourt, B.P. 105, 78153 Le Chesnay, Cedex, France, 7 pages.

Lewis, J. M., 1971: "*Variational subsynoptic analysis with applications to severe local storms*," *Mon. Wea. Rev.*, **99**, 786–795.

Lewis, J. M., 1972: "*The operational upper-air analysis using the variational method*," *Tellus*, **24**, 514-530.

Lewis, J. M. and T. H. Greyson, 1972: "*The adjustment of surface wind and pressure by Sasaki's variational matching technique*," *J. Appl. Meteor.*, **11**, 586–597.

Lewis, J. M. and S. C. Bloom, 1978: "*Incorporation of time- continuity into subsynoptic analysis by using dynamic constraints*," *Tellus*, **30**, 496–515.

Lewis, J. M., 1980: "*Dynamic adjustment of 500 mb vorticity using P. D. Thompson's scheme:a case study*," *Tellus*, **32**, 511–524.

Lewis, J. M., 1982: "*Adaptation of P. D. Thompson's scheme to the constraint of potential vorticity conservation*," *Mon. Wea. Rev.*, **110**, 1618–1634.

Lewis, J. M. and L. Panetta, 1983: "*The extension of P. D. Thompson's scheme to multiple times*," *J. Clim. and Appl. Meteor.*, **22**, 1649–1653.

Lewis, J. M., C. M. Hayden, and A. J. Schreiner, 1983: "*Adjustment of VAS and RAOB geopotential analysis using quasi- geostrophic constraints*," *Mon. Wea. Rev.*, **111**, 2058–2067.

Lewis, J. M., C. Hayden, C. Veldon, T. Stewart, R. Lord, S. Goldenberg, and S. Aberson, 1985: "*The use of VAS winds and temperatures as input to barotropic hurricane track forecasting*," extended abstracts, 16th AMS Conf. on Hurricanes and Tropical Meteorology, May 14–17, 1985, Houston, TX, 77004, 40–41.

Lewis, J. M. and J. C. Derber, 1985: "*The use of adjoint equations to solve a variational adjustment problem with advective constraints*," *Tellus*, 37A, **4**, 309–332.

Lewis, J. M., 1986: "*The use of adjoint equations to solve variational adjustment problems subject to vorticity constraints*," in Variational Methods in Meteorology, Elsevier Pub. Co. (Y. K. Sasaki, Ed.), 101–106.

Liang, W. J. and R. L. Inman, 1978: "*Variational optimization of wind fields to improve estimations of vertical velocity*," in (J. J. Stephens and R. L. Inman, eds.) Final Report, DAAD07-76-C-0037, Atmospheric Sciences Lab, White Sands Missile Range, New Mexico 88002, 55 pages.

Lions, J. L., 1968: "*Sur le contrôle optimal des systemes gouvernés par des equations aux deriveés partielles*," Ganthier-Villars, Dunod (Paris) (English translation by S. K. Mitter, Springer-Verlag, **170**, Berlin Heidelberg, New York, 1971), 397 pages.

Liu, C. Y. and W. R. Goodin, 1976: "*An iterative scheme for objective wind field analysis*," *Mon. Wea. Rev.*, **104**, 784–792.

Lorenc, A. C., 1981: "*A global three-dimensional multivariate statistical interpolation scheme*," *Monthly Weather Review*, **109**, 701–721.

Lorenc, A. C., 1982: "*Analysis methods for meteorological observations*," Rivista di Meteorologia Aeronautica, **XLII, No. 2–3**, 155–176.

Lorenc, A. C., 1986: "*Analysis methods for numerical weather prediction*," *Quart. J. R. Met. Soc.*, **112**, 1177–1194.

Lorenc, A. C., 1987: "*A practical approximation to optimal 4- dimensional objective analysis*," National Meteorological Center, Office Note 328, WWB/NMC, Washington, D. C. 20322, 45 pages.

Lorenc, A. C., 1988: "*Optimal nonlinear objective analysis*," *Quart. J. R. Met. Soc.*, **114**, 205–240.

Lugenberger, 1984: "*Linear and nonlinear programming*," Second edition, Addison-Wesley, 491 pages.

McCracken, M. C., *et al.*, 1978: "*The Livermore regional air quality model*," *J. Appl. Meteor.*, **17**, 254–272.

McEwen, J. N., M. B. Danard and G. A. Davidson, 1981: "*Variationally adjusted surface winds*," *Boundary Layer Met.*, **20**, 473–483.

McFarland, M. J., 1974: "*Variational optimization analysis of temperature and moisture advection in severe storm environment*," Ph. D. Thesis, Dept. of Meteor., Univ. of Oklahoma, Norman, OK, 73019, 78 pages.

McGinley, J. A., 1973: "*Environmental energy fields associated with severe storms*," M. S. Thesis, Dept. of Meteorology, Univ. of Oklahoma, Norman, OK, 73019, 143 pages.

McGinley, J. A., 1984: "*Analysis of flow blocking during ALPEX by imposition of mass continuity*," *Revista di Meteorologia Aeronautica*, Rome, Italy, **44**, 37–45.

Machenhauer, B., 1977: "*On the dynamics of gravity oscillations in a shallow-water model, with application to normal mode initialization*," *Contrib. Atmos. Phy.*, **50**, 253–271.

Marchuk, G. I., 1974: "*Numerical solution of the problems of the dynamics of the atmosphere and ocean*," (in Russian) Leningrad, Gidrometeoizdat, 303 pages.

Marchuk, G. I., 1975: "*Formulation of the theory of perturbations for complicated models*," *Appl. Math. and Optim.*, **2**, No. 1, 1–45.

Marchuk, G. I., 1982: "*Mathematical issues of industrial effluent optimization*," *Jour. Met. Soc.*, Japan, **60**, 481–485.

Mashkovich, S. A., 1972: "*Application of the methods of congruence of meteorological fields and optimum interpolation to 4-D analysis*," U.S.S.R. Gidrometeorologia N.I.T. SSSR Leningrad, Trudy, No. 103, p. 3–12.

Masuda, Y., 1984: "*On the use of surface - pressure changes in variational optimization analysis for horizontal wind and vertical velocity fields on "SIGMA" surface*," *J. Met. Soc.*, Japan, **62**, No. 4, 597–612.

Matheron, G., 1963: "*Traite de Géostatique appliquée, Vol. 1*," 333 pages, and "*Le Krigenge, Vol 2*," 173 pages, Memoires Bureau de Recherches Geologiques et Minières, 19, Rue Vangirand, 75737, Paris, Cedex 15, France.

Matheron, G., 1970: "*Random Functions and Their Applications in Geology*," in Geostatics (D. F. Merriam, editor), Plenum, New York, 79–87.

Matheron, G., 1981: "*Splines and Krieging: Their Formal Equivalence*," Syracuse University Geology Contribution No. 8, Dept. of Geology, Syracuse University, Syracuse, New York, 13210.

Maxwell, B. R., 1976: "*A generalized version of the fields by information blending (FIB) technique*," Technical Note 24, Environmental Prediction Research Facility, N.P.S., Monterey, CA 93940, 87 pages.

Mienguet, J., 1979: "*Multivariate interpolation at arbitrary points made simple*," *Zeit Ang. Math. Phys (ZAMP)*, **30**, 292–304.

Miele, A. and J. C. Heideman, 1968: "*The restoration of constraints in holonomic and non-holonomic problems*," Aero-Astronautics Report No. 39, Rice University, Houston, TX, 77251, 14 pages.

Miele, A., J. C. Heideman, and J. N. Damoulakis, 1969: "*The restoration of constraints in holonomic and non-holonomic problems*," *J. Optimiz. Theory and Appl.*, **3**, No. 5, 361–381.

Miele, A., A. V. Levy, and E. E. Cragg, 1971: "*Modifications and extensions of the conjugate-gradient restoration algorithm for mathematical programming problems*," *Journal Optimization Theory and Application*, **7, No. 6**, 450–472.

Mills, G. A. and J. L. McGregor, 1983: "*Vertical mode initialization in a limited area data assimilation experiment*," *Mon. Wea. Rev.*, **111**, 2110–2121.

Navon, I. M., 1981: "*Implementation of 'a posteriori' methods for enforcing conservation of potential enstrophy and mass in discretized shallow-water equation models*," *Mon. Wea. Rev.*, **109**, 946–958.

Navon, I. M., 1982a: "*A posteriori numerical techniques for enforcing simultaneous conservation of integral invariants upon finite- difference shallow-water equations models*," in Notes on Numerical Fluid Dynamics, **5**, Henri-Viviand Ed., Vieweg., 230–240.

Navon, I. M., 1983: "*A Numerov-Galerkin technique applied to a finite-element shallow-water equations model with enforced conservation of integral invariants and selective lumping*," *Jour. Comput. Phys.*, **52**, No. 2, 313–339.

Navon, I. M. and R. de Villiers, 1983: "*Combined penalty multiplier optimization methods to enforce integral invariants conservation*," *Mon. Wea. Rev.*, **11**, 1228–1243.

Navon, I. M., 1984: "*Application of Augmented Lagrangian methods in meteorology. Comparison of different conjugate-gradient algorithms for large scale minimization*," NASA Tech. Memo 86053 Research Review 1983, Laboratory for Atmospheres. NASA/GSFS, Greenbelt, Maryland 20771, 150–157.

Navon, I. M. and R. de Villiers, 1985: "*The application of the Turkel-Zwas explicit large-time step scheme to a hemispheric barotrophic model with constraint restoration*," *Mon. Wea. Rev.*, **115**, No. 5, 1036-1057.

Navon, I. M. and R. de Villiers, 1986: "*GUSTAF: A Quasi-Newton nonlinear ADI FORTRAN IV program for solving the shallow-water equations with Augmented-Lagrangians*," *Computers & Geosciences*, **12**, 2, 151–173.

Navon, I. M., 1986: "*A review of variational and optimization methods in meteorology*," in Variational Methods in Geosciences, Y. K. Sasaki ed., Elsevier, p. 29–34.

Navon, I. M., 1987a: "*FEUDX - a two stage high accuracy finite- element Fortran program for solving shallow-water equations*," *Computers & Geosciences*, **13**, No. 3, 255–285.

Navon, I. M., 1987b: "*The Bayliss-Isaacson algorithm and the constraint restoration method are equivalent*," *Journal of Meteorology and Atmospheric Physics*, **13**, 143–152.

Navon, I. M. and D. Legler, 1987: "*Conjugate-gradient methods for large-scale minimization in meteorology*," *Mon. Wea. Rev.*, **115**, No. 8, 1479–1502.

Navon, I. M., P. K. H. Phua, and M. K. Ramamurthy, 1988: "*Vectorization of conjugate-gradient methods for large-scale minimization*," *to appear in the Journal of Optimization Theory and Applications*, also in the proceedings of the Supercomputing '88 Conference, Orlando, FL, Nov. 14–18, 1988, pp. 410–418, 458 pages.

O'Brien, J. J., 1970: "*Alternative solutions to the classical vertical velocity problem*," *Jour. Appl. Meteor.*, **9**, 197–203.

O'Sullivan, F., 1983: "*On the Analysis of Some Penalized Likelihood Schemes*," Tech. Report 726, Statistics Dept., Univ. of Wisconsin-Madison, Madison, WI, 53706, 118 pages.

O'Sullivan, F. and Wahba, G., 1985: "*A Cross-validated Bayesian retrieval algorithm for non-linear remote sensing experiments*," *Journal of Computational Physics*, **59**, 441–455.

Pagava, T. S., 1973: "*Variation analysis of standard and asynchronous satellite information*," Meteorologiya I Gidrologiya, No. 4, 55–66.

Parrish, F. D. and S. E. Cohn, 1985: "*A Kalman filter for a two- dimensional shallow-water equations model: formulation and preliminary experiments*," Office Note 304, NMC, Washington, D. C. 20233, 64 pages.

Parrish, F. D. and S. E. Cohn, 1986: "*Assimilation of wind profiler data by a two dimensional Kalman filter*," in Proc. WMO/IUGG Intl. Symp. on Short- and Medium-Range WMP, Tokyo, 4-8 August 1986, p 31–34.

Pedder, M. A., 1986: "*Generalised spline surfaces*," in workshop on High-Resolution Analysis, June 24–26, 1985, ECMWF, Shinfield Park Reading, RG2-9AX, U.K., 259–275, 325 pages.

Penenko, V. V., 1974: "*On a method for objective analysis of meteorological element fields on the sphere*," *Meteorologiya I Gidrologiya*, Moskow, No. 5, May 1974, 11–23.

Penenko, V. V. and N. N. Obratsov, 1976: "*A variational initialization method for the fields of the meteorological elements*," English translations *Soviet Meteorology and Hydrology*, **11**, 1–11.

Petersen, D. P., 1968: "*On the Concept and Implementation of Sequential Analysis for Linear Random Fields*," *Tellus*, **20, No. 4**, 673–686.

Petersen, D. P. , 1969: "*Algorithms for Sequential and Random Observations*," *Meteorological Monographs, AMS*, **11, No. 33**, 100–109.

Petersen, D. P., 1973a: "*Static and dynamic constraints on estimation of space-time covariance and wave number frequency spectral fields*," *J. Atmos. Sci.*, **30**, 1252–1266.

Petersen, D. P., 1973b: "*Transient Suppression in Optimal Sequential Analysis*," *Journal of Applied Meteorology*, **12, No. 3**, 437–440.

Pfeffer, R. Z., D. Mardan, B. Sterbentz and W. Fowlis, 1965: "*A new concept of available potential energy*," in Proc. Intl. Symp. on Dynamics of Large-Scale Processes in the Atmosphere, Moscow, June, 1965, p. 179–189.

Phillips, N. A., 1981: "*Variational analysis and the slow manifold*," *Mon. Wea. Rev.*, **109**, 12, 2415–2426.

Phillips, N. A., 1982: "*On the Completeness of Multi-Variate Optimum Interpolation for Large-Scale Meteorological Analysis*," *Monthly Weather Review*, **110**, 1329–1334.

Polyak, B. T., 1971: "*The convergence rate of the penalty function method*," *Zh. Vychisl. Mat. math. fiz.*, **11**, 3–11 (English translation: U.S.S.R. Computational Mathematics and Mathematical Physics, Volume 11, 1–12).

Pontryagin, L. S., V. G. Botyanskii, R. V. Gamkrelidze, and E. E. Mishkenko, 1960: "*The theory of optimal processes I. The maximum principle*," IZV. Akad. Nauk, SSR, Ser Math 24, 3. Also English translation, Tuterscience, New York, 1962.

Powell, M. J. D., 1982: "*Nonlinear optimization 1981*," Academic Press, London, 298 pages.

Provost, C., 1986: "*The variational inverse method for the general circulation in the ocean*," in *Variational Methods in Geosciences*, (Y. K. Sasaki, ed.), Elsevier, p. 55–76.

Puri, K., 1982: "*Normal mode initialization in the ANMRC data assimilation scheme*," in *The Interaction between Objective Analysis and Initialization*, The Proc. Stanstead Seminar, Publication in Meteorology, No. 127, Dept. of Meteorology, McGill University, Montreal, Canada, 137–142.

Puri, K., 1983: "*Some experiments in variational normal mode initialization in data assimilation*," *Mon. Wea. Rev.*, **111**, 1208–1218.

Purser, R. J., 1984: "*A New Approach to the Optimal Assimilation of Meteorological Data by Iterative Bayesian Analysis*," 10^{th} Conference on Weather Forecasting and Analysis, Clearwater Beach, Florida, American Meteorological Society, Boston, pp. 102–105.

Purser, R. J., 1986: "*Bayesian Optimal Analysis for Meteorological Data*," in Variational Methods in Geosciences (Y. K. Sasaki, editor), Elsevier, Amsterdam, 167–172, 309 pages.

Purser, R. J., 1986: "*The application of adaptive methods in a Bayesian scheme of objective analysis*," in Proc. WMP/IUGG Sym. on Short- and Medium-Range NWP, Tokyo 4–8 August 1986, 59–63.

Racher, P. and R. Rosset, 1985: "*Formulation tridimensionnelle par elements finis d'une procedure variationelle d'ajustement du champ de vent sur une topographie complexe: tests preliminaires sur l'île d'Hawaii*," *Atmos. Ocean*, **23**, (1), 1–22.

Ramamurthy, M. and F. Carr, 1987: "*Four-dimensional data- assimilation in the Monsoon Region. Part I: Experiments with wind data*," *Monthly Weather Review*, **115**.

Ramamurthy, M. and I. M. Navon, 1988: "*Application of a conjugate-gradient method for variational blending of meteorological fields*," in the Proceedings of the Eighth Conference on NWP, Feb. 22–26, 1988, Baltimore, MD, 202–208, 860 pages, American Meteorological Society.

Ray, P. S., K. K. Wagner, K. W. Johnson and J. J. Stephens, W. C. Bumgardner and E. A. Mueller, 1978: "*Triple-Doppler radar observations of a convective storm*," *J. Appl. Meteor.*, **173**, 1201–1212.

Ray, P. S., C. Zeigler, W. Bumgardner, and R. Serafin, 1980: "*Single and multiple Doppler-radar observations of tornadic storms*," *Mon. Wea. Rev.*, **108**, 1607–1625.

Reddy, J. N., 1981: "*A finite-element analysis of steady incompressible flows in 3 dimensions by a penalty function formulation*," Res. Report No. VPI-E-81-4, Dept. of Eng. Sci. Mec., Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.

Reddy, J. N., 1982: "*On penalty function methods in the finite- element analysis of flow problems*," *Int. Jour. Numer. Meth. in Fluids*, **2**, 151–171.

Reinhart, 1985: "*Identification des conditions aux limites pour les equations de Sant-Venant*," *Rapport de Recherche, INRIA*, 458, Domaine de Volucean Roequencourt, 78153 Le Chesnay, Cedex, France, 72 pages.

Reinsch, C. H., 1967: "*Smoothing by Spline Functions*," *Numerical Math*, **10**, 177–183.

Riccardo, O. O., 1974: "*Optimal Contour Mapping Using Universal Krieging*," *J. Geophys. Res.*, **79**, 695–702.

Ritz, W., 1908: "*Uber line neue methods zur lösung gewisser variations probleme der mathematische physik*," *J. Reine Phys. Math.*, **135**, 1–61.

Rusk, D. J., M. Ramamurthy and F. H. Carr, 1986: "*Tests of a simple variational scheme for a limited area primitive equations model*," WMO/IUGG Intl. Symp. on Short and Medium Range Numerical Weather Prediction, Tokyo, August, 1986, p. 117–118.

Sadokov, V. P. and D. B. Shteinbok, 1977: "*Application of conjugate functions in analysis and forecasting of the temperature anomaly*," (English translation) *Soviet Meteorology and Hydrology*, **10**, 16–21.

Sasaki, Y. K., 1955: "*A fundamental study of the numerical prediction based on the variational principle*," *J. Meteor. Soc. Japan*, **33**, 262–275.

Sasaki, Y. K., 1958: "*An objective analysis based on the variational method*," *J. Meteor. Soc. Japan*, **36**, 77–88.

Sasaki, Y. K., 1960: "*An objective analysis for determining initial conditions for the primitive equations*," Tech. Rep. Proj. 208, Texas A& M University, College-Station, TX, 77843, 22 pages.

Sasaki, Y. K., 1969: "*Proposed inclusion of time-variation terms, observational and theoretical in numerical variational objective analysis*," *J. Meteor. Soc. Japan*, **47**, 115–203.

Sasaki, Y. K., 1970a: "*Some basic formalisms in numerical variational analysis*," *Mon. Wea. Rev.*, **98**, 857–883.

Sasaki, Y. K., 1970b: "*Numerical variational analysis formulated under the constraints as determined by long-wave equations as a low-pass filter*," *Mon. Wea. Rev.*, **98**, 884–898.

Sasaki, Y. K., 1970c: "*Numerical variational analysis with weak constrain and application to the surface analysis of severe storm gust*," *Mon. Wea. Rev.*, **98**, 899–910.

Sasaki, Y. K., 1970d: "*A theoretical interpretation of anisotropically weighted smoothing on the basis of numerical variational analysis*," *Mon. Wea. Rev.*, **99**, ndeg 9, 698–707.

Sasaki, Y. and J. M. Lewis, 1970: "*Numerical variational objective analysis of the planetary boundary layer in conjunction with squall line formation*," *J. of Japan Meteor. Soc.*, **5**, 381–399.

Sasaki, Y. K., 1971: "*Low pass and band pass filters in numerical variational optimization*," *J. Met. Soc. Japan*, **49**, 766–773.

Sasaki, Y. K., 1973: "*Mechanism of squall line formation as suggested from variational analysis of hourly surface observation*," preprint volume of the Eighth Conference on Severe Local Storms, October 15–17, AMS, Boston, MA.

Sasaki, Y. K., 1975: "*Variational design of finite-difference schemes for initial value problems of conservative system*," Dept. of Meteorology, University of Oklahoma, Norman, OK, 73019.

Sasaki, Y. K., 1976: "*Variational design of finite-difference schemes for initial value problems with an integral invariant*," *Journal of Computational Physics*, **21**, 270–278.

Sasaki, Y. K., 1977: "*Variational design of finite-difference schemes for initial value problems with a global divergent barotropic mode*," *Beitr. Phys. Atmos.*, **50**, 284–289.

Sasaki, Y. K., T. Baxter and J. S. Goerss, 1979: "*Dynamic data assimilation by the noise freezing method*," NERPF Report CR-79-03, 80 pages, NEPRF, Monterey, CA 93940, 80 pages.

Sasaki, Y. K., P. S. Ray, J. S. Goerss and P. Soliz, 1979: "*Inconsistent finite differencing errors with the variational adjustment of horizontal wind components*," *J. Met. Soc. Japan*, **57**, 88–92.

Sasaki, Y. K. and J. N. Reddy, 1980: "*A comparison of stability and accuracy of some numerical models of two-dimensional circulation*," *Intl. J. Numer. Math. in Eng.*, **16**, 149–170.

Sasaki, Y. K. and J. McGinley, 1981: "*Application of the inequality constraints in adjustment of superadiabatic layers*," *Mon. Wea. Rev.*, **109**, 194–196.

Sasaki, Y. K., 1982a: "*A variational initialization assimilation scheme: basic concept and preliminary result*," CIMMS Rep. No. 7, 19 pages, Norman, Ok, 73019.

Sasaki, Y. K., 1982b: "*Variational methods in meteorology*," in Proc. Fourteenth Stanstead Seminar, Publication in Meteorology, McGill University, 127, 143–149.

Sasaki, Y. K. and J. Goerss, 1982: "*Satellite data assimilation using NASA Data Systems Test 6 observations*," *Mon. Wea. Rev.*, **110**, 1635–1644.

Schaefer, J. T. and C. A. Doswell, 1979: "*On the interpolation of a vector field*," *Mon. Wea. Rev.*, **107**, 458–476.

Schneider, E. K., 1984: "*Some minimization calculations concerning the effect of horizontal eddy momentum fluxes on the equilibrium zonal mean motions*," *Jour. Atmos. Sci.*, **41**, No. 9, 1640–1647.

Seaman, R. S., 1977: "*Absolute and differential accuracy of analysis achievable with specified observational network characteristics*," *Mon. Wea. Rev.*, **105**, 1211–1222.

Seaman, R. S., R. Falconer, and J. Brown, 1977: "*Application of a variational blending technique to numerical analysis in the Australian region*," *Austr. Meteor. Mag.*, March, 1–22.

Seaman, R. S., and M. F. Hutchinson, 1985: "*Comparative Real Data Tests of Some Objective Analysis Methods by Witholding Observations*," *Aust. Met. Mag.*, **33**, 37–46.

Semazzi, F. H. M. and I. M. Navon, 1986: "*A comparison of the bounded derivative and the normal mode initialization methods using real data*," *Mon. Wea. Rev.*, **114**, (11), 2106–2121.

Shanno, D. F. and P. K. H. Phua, 1980: "*Remark on algorithm 500*, *ACM Trans. Math. Software*, **6(4)**, 618–622.

Sheets, R. C., 1973a: "*Analysis of hurricane data using the variational optimization approach with a dynamic constraint*," *J. Appl. Meteor.*, **12**, 963–976.

Sheets, R. C., 1973b: "*Analysis of Hurricane Debbie modification results using the variational optimization approach*," *Mon. Wea. Rev.*, **101**, No. 9, 633–684.

Sherman, C., 1978: "*A mass-consistent model for wind fields over complex terrain*," *J. Appl. Met.*, **17**, 312–319.

Shian, J. J., G. Wahba, and D. R. Johnson, 1986: "*Partial Spline Models for the Inclusion of Tropopause an Frontal Boundary information in Otherwise Smooth Two and Three-Dimensional Objective Analysis*," *J. of Atmos. and Oceanic Technology*, **3, No. 4**, 714–725.

Soliz, P. and J. Fein, 1980: "*Kinematic analysis for a major winter storm during AMTEX 75*," *J. Met. Soc. Japan*, **58**, 16–32.

Stephens, J. J., 1965: "*A variational approach to numerical weather analysis and prediction*," Ph. D. Dissertation, Texas A& M University, College-Station, TX, 77863, Rep. 3, 243 pages.

Stephens, J. J., 1967: "*Filtering responses of selected distance dependent weight functions*," *Mon. Wea. Rev.*, **95** (1), 45–46.

Stephens, J. J., 1968: "*Variational resolution of wind components*," *Mon. Wea. Rev.*, **96**, 229–231.

Stephens, J. J., 1970: "*Variational initialization with the balance equation*," *J. Appl. Meteor.*, **9**, 732–739.

Stephens, J. J., K. W. Johnson and T. J. Whiteside, 1977: "*OBAN user's guide*," Dept. of Meteorology, Florida State University, Tallahassee, Florida, 32306-3034, 70 pages.

Stephens, J. J., 1978: "*Variational adjustment of atmospheric fields: a review*," Dept. of Meteorology, Florida State University, Tallahassee, FL, 32306-3034, 45 pages.

Stephens, J. J. and K. W. Johnson, 1978: "*Middle/Large-Scale variational adjustment of atmospheric fields in mesoscale diagnostic numerical variational analysis models*," Final Report, Dept. of Meteorology, Florida State University, Tallahassee, FL, 32306-3034, pp. 1–38.

Takacs, L. L., 1986: "*Documentation of the Goddard Laboratory for atmospheres fourth-order two layer shallow-water model*," NASA Tech. Memo. 86227, 80 pages, NASA/GSFC, Greenbelt, Maryland 20771.

Takacs, L. L., 1988: "*Effects of using 'a posteriori' methods for the conservation of integral invariants*, *Monthly Weather Review*, **116, No. 3**, 525–545.

Talagrand, O. and P. Courtier, 1985: "*Formalization of the adjoint method*," Meteorological applications, Working note 117, 10 pages (in French).

Talagrand, O., 1985: "*The adjoint model technique and meteorological applications*," in Proceedings Workshop on High Resolution Analysis, ECMWF, Reading, U. K., p. 325.

Talagrand, O., 1986: "*Application of optimal control to meteorological problems*," in Variational Methods in Geosciences, (Y. K. Sasaki, ed.), Elsevier, 13–29.

Talagrand, O. and P. Courtier, 1987: "*Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I: theory*," *Quart. J. Roy. Met. Soc.*, **113**, 1311–1328.

Temperton, C., 1982: "*Variational normal mode initialization for the ECMWF gridpoint model*," in Proc. Fourteen Stanstead Seminar, publication in *Meteorology*, McGill University, **127**, 160–164.

Temperton, C., 1984: "*Variational normal mode initialization for a multilevel model*," *Monthly Weather Review*, **112**, 2303–2316.

Temperton, C., 1985: "*Application of a new principle for normal mode initialization*," preprints Seventh conference on Numerical Weather Prediction, Montreal, American Meteorological Society, 105–107.

Temperton, C., 1988: "*Implicit normal mode initialization*," *Monthly Weather Review*, **116, No. 5**, 1013–1031.

Testud, J., and M. Chong, 1983: "*Three-dimensional wind field analysis from dual-Doppler radar data. Part I: Filtering interpolating and differentiating the raw-data*," *J. Climate Appl. Meteor.*, **22**, 1204–1215.

Thacker, W. C., 1986: "*Relationship between statistical and deterministic methods of data assimilation*," in *Variational Methods in Geosciences* (Y. K. Sasaki, ed.) Elsevier, 173–189.

Thacker, W. C., D. Eppel, and J. Hauser 1986: "*Advective Transport Via Error Minimization: Enforcing Constraints of Non-Negativity and Conservation*," *Appl. Math. Modelling*, **10**, 438–444.

Thacker, W. C. and R. B. Long, 1988: "*Fitting dynamics to data*," *Journal of Geophysical Research*, **93**, 1227–1240.

Thiebaux, H. J., and M. A. Pedder, 1987: "*Spatial Objective Analysis with Applications in Atmospheric Science*," Academic Press, New York, 299 pages.

Thompson, P. D., 1969: "*Reduction of analysis error through constraints of dynamical consistency*," *J. Appl. Meteor.*, **8**, 738–742.

Tribbia, J. J., 1982a: "*Experiments with variational normal mode initialization*," in Proc. of the Fourteenth Stanstead Seminar (D. Williamson, ed.), McGill University, Montreal, Quebec, Canada, 171–177.

Tribbia, J. J., 1982b: "*On variational normal mode initialization*," *Monthly Weather Review*, **110**, 6, 445–470.

Tseng, C. Y., 1985: "*Variational initialization and determination of weighting factors*," in *Variational Methods in Geosciences*, (Y. K. Sasaki, ed.) Elsevier, 119–125.

Tuerpe, D. R., P. M. Gresho and R. L. Sani, 1978: "*Variational wind field adjustment over complex terrain using finite-element techniques*," in Proceedings AMS Conf. on Sierra-Nevada Meteorology, Lake Taho, NV, June 19–21, 1978, pp. 68–70.

Wagner, K. K., 1971: "*Variational analysis using observational and low pass filtering constraints*," M. S. Thesis, Dept. of Meteorology, The University of Oklahoma, Norman, OK, 73019, 39 pages.

Wagner, K. K., 1977: "*Wind analysis within the planetary boundary layer*," in *Mesoscale Diagnostic Numerical Variational Analysis Models.* J. J. Stephens and R. L. Imman Report DAAD07-76-C-0037, White Sands Missile Range, New Mexico 88002, 150 pages.

Wahba, G., 1975: "*Smoothing noisy data with spline functions*," *Numerical Mathematics*, **24**, 383–393.

Wahba, G., 1978: "*Improper Priors, Spline Smoothing and the Problem of Guarding Against Model Errors in Regression*," *J. Roy. Stat. Soc. B*, **40**, 364–372.

Wahba, G., 1979a: "*Spline interpolation and Smoothing on the Sphere*," Tech. Report 584, Dept. of Statistics, University of Wisconsin, Madison, WI, 53706, 38 pages.

Wahba, G., 1979b: "*Convergence rates of thin plate smoothing splines*," in Smoothing Techniques for Curve Estimation Lecture Notes in Mathematics, **752**, Th. Gasser and M. Rosenblatt, Editors, Springer-Verlag, New York, pp. 233–245.

Wahba, G. and J. Wendelberger, 1980: "*Some new mathematical methods for variational objective analysis using splines and cross- validation*," *Monthly Weather Review*, **108**, 1122–1143.

Wahba, G., 1981a: "*Spline Interpolation and Smoothing on the Sphere*," *SIAM J. Sci. Stat. Comput.*, **2, No. 1**, 5–16.

Wahba, G., 1981b: "*Some new techniques for variational objective analysis on the sphere using splines, Hough-functions, and sample spectral data*," in Proc. 7$^{\text{th}}$ Conference on Probability and Statistics in the Atmospheric Sciences, Monterey, CA, Nov. 26, 1981, 213–216.

Wahba, G., 1982a: "*Vector Splines on the Sphere with Application to the Estimation of Vorticity and Divergence from Discrete, Noisy Data*," in Multivariate Approximation Theory II, Walter Schempp and Karl Zeller, editors, Birkhauser Verlag, Basel, PP. 407–418.

Wahba, G., 1982b: "*Constrained Regularization for Ill-Posed Linear Operator Equations with Applications in Meteorology and Medicine*," Third Purdue Symposium on Statistical Decision Theory, **2**, S. S. Gupta and J. O. Berger, editors, Academic Press, New York, pp. 384–418.

Wahba, G., 1982c: "*Variational methods in simultaneous optimum interpolation and initialization*," in *The Interaction between Objective Analysis and Initialization*, Proc. Stanstead Seminar, publication in *Meteorology*, No. 127, Dept. of Meteorology, McGill University, Montreal, 178–185.

Washington, W. M. and R. J. Duquet, 1963: "*An objective analysis of stratospheric data by Sasaki's method*," Dept. of Meteorology, Penn. State University, University Park, PA, 16802, 23 pages.

Washington, W. M., 1964: "*Initialization of primitive equations models*," for N.W.P. Final Report (1), Coint. AF 191604-7261, Dept. of Meteorology, Penn. State University, University Park, PA, 16802, 83 pages.

Washington, W. M. and D. P. Baumhefner, 1975: "*A method for removing Lamb waves from initial data for the primitive equation models*," *J. Appl. Meteor.*, **14**, 114–119.

Whiteside, T. J., 1977: "*The adjustment of spatial covariances subject to dynamical constraints*," Dept. of Meteorology, Florida State University, Tallahassee, FL, 32306-3034, 29 pages.

Wilkins, E. M., 1971: "*Variational principle applied to numerical objective analysis of urban air pollution distributions*," *Journal of Applied Meteorology*, **10**, 974–981.

Wilkins, E. M., 1972: "*Variationally optimized numerical analysis equations for urban air pollution monitoring networks*," *Journal of Applied Meteorology*, **11**, 1334–1341.

Wunsch, C. and J. F. Minster, 1982: "*Methods for box models and ocean circulation tracers: Mathematical programming and nonlinear inverse theory*," *Journal of Geophysical Research*, **87**, C8, 5647–5662.

Ziegler, C. C., 1978: "*A dual Doppler variational objective analysis as applied to studies of convective storms*," M. S. Thesis, Dept. of Meteorology, University of Oklahoma, Norman, OK, 73019, 115 pages.