[Reprinted from BULLETIN OF THE AMERICAN METEOROLOGICAL SOCIETY, Vol. 73, No. 4, April 1992]
Printed in U. S. A.

An Examination of Four-Dimensional Data-Assimilation Techniques for Numerical Weather Prediction

Dewey E. Harms*, Sethu Raman*, and Rangarao V. Madala+

Abstract

Four-dimensional data-assimilation methods, along with the most commonly used objective analysis and initialization techniques, are examined from a historical perspective. Operational techniques, including intermittent data assimilation and Newtonian nudging, and next-generation methods (Kalman–Bucy filtering and the adjoint method) are briefly described. Several methods are compared, with primary emphasis being placed on recent papers dealing with the operational assimilation techniques. Ongoing and future research is outlined, and some important implications of this research are discussed.

1. Introduction

Two major motivations for using data assimilation exist: as an analysis/diagnostic/research tool and for operational weather forecasting. Data assimilation has been applied not only in meteorology (air pollution and planetary boundary-layer studies, forecast case studies, quantitative assessment of new observing systems, among others), but also in oceanography for describing ocean currents (Ghil 1989; Robinson 1986). To summarize all relevant work pertaining to data assimilation is indeed a difficult task, because the contents are spread so widely. The purpose of the present paper is only to provide an overview of four-dimensional data assimilation with primary emphasis on assimilation methods currently useful for operational weather forecasting. To make this review more complete, attention has also been given to the state-of-the-art or "next-generation" techniques, but to a lesser extent. Excellent reviews covering the early days of data-assimilation research in both simulation and real-data studies are available in the literature. The interested reader is referred to Bengtsson (1975a), McPherson (1975), Hollingsworth (1986), and Bourke et al. (1985). Plus, a comprehensive history of data analysis and assimilation is given by Daley (1991).

Numerical weather prediction (NWP) has classically been viewed as an initial-value problem where the governing equations of geophysical fluid dynamics are integrated forward in time from a set of initial values. The quality of NWP is strongly dependent on the accuracy of specifying these initial conditions and on the ability to model mathematically the dynamics and physical processes of the atmosphere. In a pioneering paper, Charney et al. (1969) suggested combining past and current data in a numerical model such that the model's equations provide time continuity and dynamic coupling among the atmospheric fields. This concept, which has merged objective numerical analysis and numerical weather prediction, has become known as four-dimensional data assimilation (FDDA) and has proven to be a major advance in NWP during the past 20 years.

What inspired this major advance, or in other words, what spurred this development of data assimilation? The advent of meteorological satellites in the 1960s raised the possibility that nearly continuous atmospheric temperature observations would be available on a global basis. However, these space-based observing systems measure only radiance (or temperature) distributed in space and time, rather than at fixed locations and times. In order to fully utilize this new source of data, the numerical weather prediction techniques had to be adjusted. Beginning with Charney et al. (1969), Smagorinsky et al. (1970), Rutherford (1972), and Morel and Talagrand (1974), research progressed in reconstructing unobserved variables from the observed variables through the numerical model's dynamical coupling between those variables. By combining information about the state of the atmosphere, earlier observations are carried forward to provide an independent source of information to be added to the newly acquired observations.

Morel (1981) further illustrated why data assimilation is essential in NWP by listing five key reasons: 1) the inadequate distribution (spatial gaps) of the twice-daily upper-air sounding data, 2) the discrepancy between the conventional observations as point measurements and the true volume averages required by numerical models, 3) the inherently asynoptic character of remote observations obtained from sunsynch-

*Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, Raleigh, NC 27695-8208
+Naval Research Laboratory, Washington, D.C. 20375

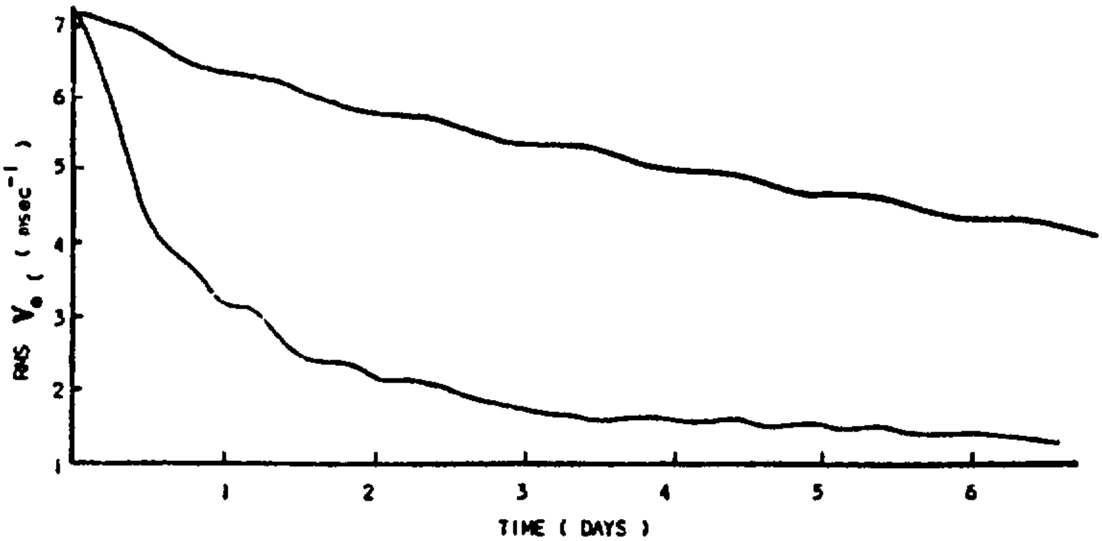

FIG. 1. Successive reduction of wind error using a direct insertion (upper curve) of height data in a barotropic model. Corresponding results using height as well as from the derived geostrophic winds (lower curve) (McPherson 1975).

ronous orbiting satellites, 4) the inadequate vertical resolution of remote observations of cloud motions from geostationary satellites, and 5) the significant random and systematic errors involved in the data processing required for reconstructing atmospheric fields from remotely measured physical quantities. He concluded that any weather prediction model must be initialized by merging the new observations with the currently estimated meteorological fields, computed on the basis of earlier observations, while taking into account the dynamical constraints between successive model states, specified by the governing dynamical equations.

This process of merging new observational data with the ongoing integration of a numerical forecasting model is known as "data assimilation" or, equivalently, "four-dimensional data assimilation," in consideration of the time–space distribution of the database. Then, as stated by Warner (1987), "the overall objective of the assimilation process is to provide the best possible initial state from which to begin a forecast, where the term 'best' implies an appropriate balance and reflects optimal use of four-dimensional data (synoptic as well as asynoptic) to define the structures on all scales at the initial time."

In the past 20 years, many techniques have been developed to insert data into dynamic models. However, the first data-assimilation experiments were simplistic and of limited success. The observation simply replaced the model forecast value at the model grid point nearest the observation location (Jastrow and Halem 1970). This technique, known as direct insertion, is inadequate from the initialization point of view (Bengtsson 1975b). Direct insertion gives a shock to the system and generates high-amplitude waves, or gravitational oscillations. For example, if observations of the mass field are inserted into a primitive equation model, an imbalance is created between the mass and wind fields. When the model integration is resumed, this imbalance is manifested as high-amplitude gravitational oscillations; this is the model's attempt to restore the dynamic balance, which was disrupted by the data insertion. Techniques such as damping time-integration schemes and time filters were developed to dampen these nonmeteorological waves. However, this damping must be rather strong and can be harmful to the meteorological modes. Bengtsson (1975b) stated that the shock effect will be reduced if locally analyzed data (interpolating the observation to several nearby grid points) are inserted into the model instead of observations only (indirect insertion). For example, applying a local multivariate analysis, or, correspondingly, relating wind and geopotential by the geostrophic relation and inserting both height and winds simultaneously proved successful by speeding the updating considerably (Fig. 1). The multivariate procedure mentioned here produces a simultaneous weighting of mass and motion observations, subject to the constraint of geostrophicity (Petersen 1968; Eddy 1973; Daley and Puri 1980).

Of the several FDDA methods that have been investigated over the last two decades, some have been implemented operationally while others have not yet been used or are in the developmental stage. In the 1970s, regional/mesoscale weather forecasting models were developed, normal mode initialization was introduced, and global data assimilation became operational. With the advent of mesoscale modeling, FDDA applications are also geared toward the mesoscale, with the emphasis shifting from applications on the global scale.

In operational numerical prediction systems, FDDA can be categorized into two broad areas: continuous method and intermittent (analysis–forecast cycle) method. The former refers to the insertion of data into a model as it is received, in a temporally continuous fashion. The analysis–forecast cycle clearly illustrates the four components of data assimilation: quality control, objective analysis, initialization, and an initial guess from a short-range forecast. With this method, the data are assimilated intermittently at specified intervals. So-called next-generation data-assimilation methods are being researched today, with the two most prevalent being the adjoint method based on the variations of calculus and the Kalman–Bucy filter technique.

Since quality control, objective analysis, and initialization are intricate parts of data assimilation, it is appropriate to consider each of these components in the present paper. Quality control methods are outlined in section 2. The most commonly used objective analysis and initialization techniques are overviewed in sections 3 and 4, respectively, followed by a discussion of various data-assimilation methods in section 5. Then, in section 6, two studies, comparing operational data-assimilation techniques, are reviewed. Section 7 outlines some implications of FDDA and highlights ongoing and future research in data assimilation. Finally, a summary is given in section 8.

2. Quality control

Quality control is an integral part of a data-assimilation system. Quality-control algorithms are designed to modify or reject erroneous meteorological data. Following Daley (1991), observational data errors can be classified into two groups: natural error and gross (or rough) error. The natural error includes instrument error and error of representativeness. The data describe the behavior of the instrument itself, not the behavior of the meteorological parameter it is intended to measure. Every instrument is approximate by its very nature. Errors of representativeness are deviations caused by small-scale perturbations and are also referred to as micrometeorological errors. Gross errors originate from improperly calibrated or malfunctioning instruments, incorrect registration of observations, incorrect coding of observations, and telecommunication errors. These errors (natural and gross) can be either random or spatially or temporally correlated, and there can be systematic biases. The systematic errors can result from improper calibration of an instrument or from the influence of some persistent factor that is not accounted for or is accounted for inaccurately.

Several quality-control techniques are used routinely to screen bad data. These techniques can be divided into four major categories (following Gandin 1988). The first two, plausibility check and check for contradictions, are used to identify gross errors based on the physical reasonableness of the data. Plausibility checks are the most widely used quality-control methods. These checks analyze each datum independently of other data. A simple plausibility check is one that rejects data values that can never occur in reality —for example, positive temperatures (°C) at 300 mb. Other versions compare the datum with the climatological mean or with a background field (numerical forecast). If the deviation is too large, the datum may be rejected. The check for contradictions is based on an analysis of two or more parameters at the same point. An example is the occurrence of rain in the absence of clouds.

Quality control is an integral part of a data-assimilation system. Quality-control algorithms are designed to modify or reject erroneous meteorological data.

The remaining quality-control procedures include checks that rely on some common information and redundancy between observations. The spatial continuity (or consistency) check compares a datum with data at adjacent locations. This method, also known as the buddy check, demands spatial consistency among the data. A temporal continuity check can also be made in which temporal continuity is required with past observations. Finally, checks using dynamic relations such as the geostrophic or hydrostatic relation can be used to check geopotential height with temperature and mass with wind. These checks require the data to obey the dynamic relation at least approximately, otherwise the data are suspected or rejected.

The necessity for including automated quality control in operational NWP was recognized during the early era of NWP. Significant advances have been made in quality control, including the development of methods using optimum interpolation (Rutherford 1976; Lorenc et al. 1977; Lorenc 1981). Recent advances include the complex quality-control method of Gandin (1988) and the Bayesian approach of Lorenc and Hammon (1988). Reviews on various quality control methods have been authored by Belousov et al. (1968), Gustavsson (1981), and Lorenc (1985).

3. Objective analysis techniques

Objective analysis (the second major component of FDDA) and quality control have become intertwined and, as a result, the data-assimilation process is more internally consistent. Excellent reviews on objective analysis have been presented by McPherson (1976) and Gustavsson (1981). The present discussion contains a synopsis of selected topics from these two reviews. According to McPherson (1976), objective analysis is a process in which meteorological observations distributed in space and time are combined with forecasts from previous analyses and perhaps with climatology to form a numerical representation of the state of the atmosphere. This representation takes the form of values of the meteorological variables at regularly spaced grid points subject to various mathematical and physical constraints.

The objective analysis process consists of three subprocesses that are essential for the overall success or effectiveness of the FDDA system: 1) filtering of small-amplitude random and systematic errors; 2) interpolation to a regular network of grid points or, in the case of spectral analysis, integration of the representing mathematical space functions over the irregularly spaced observations; and 3) forced adjustment of the meteorological variables using dynamic relationships among these variables.

Within a relatively short period of time, objective (numerical) analysis schemes were independently developed by several meteorologists. Panofsky (1949) devised the first objective analysis method—polynomial interpolation, or the so-called surface-fitting type. An extension of this method was developed by Gilchrist and Cressman (1954) and became the first operational objective analysis. In this technique, mathematical (polynomial) functions are adjusted to observed data in the close vicinity of the grid point, for which analyzed values are required. The adjustment or fit is obtained by a least squares technique. The polynomial method is nonlinear since nonlinear functions are used to approximate the variation of the analyzed variable. However, the resulting analyzed value at the grid point is a linear function of the observed data and, in this respect, this method is similar to the other analysis methods that will be described.

Bergthorsson and Döös (1955) introduced the successive correction method, and a similar method was devised by Cressman (1959). Cressman's successive correction technique essentially replaced the polynomial interpolation method because the latter produced unreasonable analyses at the edge of data-rich and data-sparse areas. In the successive corrective technique, a forecast model provides the preliminary estimate (first guess) of the field to be analyzed. The basic idea of the method is to correct this preliminary field iteratively during several analysis "scans"; the results of one scan become the first guess for the next scan. The estimate is modified by a combination of corrections computed for each grid point. The corrections, which are proportional to the difference between the observed and first-guess values, are weighted empirically, with observations nearest the grid point weighted the most heavily.

By using a short-range forecast based on a previous analysis as the first guess, the effect of past data is allowed to influence the analysis and thereby contributes to the temporal continuity and spatial coherence of the analysis. For many years, successive correction methods were most widely used in operational forecast centers.

By using a short-range forecast based on a previous analysis as the first guess, the effect of past data is allowed to influence the analysis and thereby contributes to the temporal continuity and spatial coherence of the analysis. For many years, successive correction methods were most widely used in operational forecast centers.

This early dominance of the successive correction methods was taken over during the last decade by statistical (optimal) interpolation schemes, which were originally introduced by Eliassen (1954) and Gandin (1963). Statistical interpolation was also studied extensively by Alaka and Elvander (1972) and Phillips (1976). This analysis technique is based on statistical linear regression and provides a systematic framework for blending observations of differing error characteristics with recent predictions or climatology. More accurate data receive more weight in the analysis. As in the successive correction method, the analyzed value at the grid point is the sum of the first guess and a linear combination of corrections, which are proportional to the difference between observational and first-guess values. The weighting coefficients are determined from the condition that the mean-square-error of the analyzed values be minimum, and they depend on the spatial covariances among the analyzed variables. This method is, in principle, spatially coherent and, like its counterpart of successive correction, incorporates temporal continuity through the use of a short-range forecast from the preceding analysis as a first guess.

The polynomial interpolation method mentioned earlier reappeared in the British Meteorological Office analysis (Dixon et al. 1972) and in the spectral analysis method devised by Flattery (1971). In the spectral analysis technique, mathematical functions are globally adjusted to fit observed data. One additional analysis method that has been utilized in the FDDA context (see discussion in section 5) is the variational technique introduced by Sasaki (1958). This method is a post-analysis adjustment technique based on the calculus of variations and is very effective in making the analyzed fields compatible with a forecast model. A functional is used, which minimizes the analyzed-minus-observed difference, filters undesirable high-frequency and high-wavenumber features, and employs dynamical constraints. These constraints may be strong (satisfied exactly) or weak (satisfied approximately).

Over the past 30 years, numerous variations of these different objective-analysis methods have been developed. Some versions are hybrids belonging to more than one of the aforementioned techniques.

According to McPherson (1976), objective-analysis methods used in operational meteorology can be divided into two basic categories. The first represents an analyzed field as a series expansion (spectral analysis):

$$Z_g = a_1 f_1 + a_2 f_2 + \ldots + a_m f_m, \qquad (3.1)$$

where Z_g represents the departure of the field from its mean value. The f_i represents a set of orthogonal functions—for example, a cosine series. The analysis procedure involves the determination of the time-dependent coefficients a_i, which make the series expansion best fit the observed data by, for example, a least-squares technique. Analysis methods of this type are in operational use at the National Meteorological Center (NMC) (Flattery 1971; Hayden 1976) and the United Kingdom Meteorological Office (UKMO) (Dixon 1976).

The second basic method, called the "gridpoint" method, includes statistical interpolation and successive correction techniques. Here, the analyzed value Z_g^a at a discrete point g is given by a linear combination of observations that are nearby in time and space:

$$Z_g^a = a_1 Z_1^o + a_2 Z_2^o + \ldots + a_n Z_n^o, \qquad (3.2)$$

where Z^o represents observed values at the several stations within some predetermined influence radius of point g, and the coefficients a_i determine the influence of each observation on the analyzed value. In this case, the analysis involves determining the coefficients a_i in the linear combination for each point of the analysis grid.

Two forms of representation are associated with these two basic methods of analysis: the discrete form and the spectral form. In the former, the analysis is a set of values at discrete points in space and time; in the latter, the analysis is represented by a series expansion such as (3.1). The spectral form has been used primarily for global and hemispheric applications. The discrete form has been extensively used for limited-area applications in mesoscale meteorology.

4. Initialization techniques

The objective-analysis methods described in the previous section generally do not provide balanced mass and wind fields to initiate a forecast. Uncompensated errors in wind and pressure–temperature observations, interpolation of observations to model grids, and the numerical model's inability to exactly describe the atmosphere are the primary sources of this dynamical inconsistency. The dynamical imbalances in the initial data lead to the generation of spurious inertia–gravity-wave oscillations, or "meteorological noise." Primitive equation models, unlike geostrophic models, admit these higher-frequency gravity-wave solutions that can have amplitudes much greater than their counterpart in the real atmosphere. These gravity-wave oscillations can obscure the lower-frequency Rossby-mode component of the model, which constitutes the meteorological signal. Early numerical modelers called attention to the need to eliminate the spurious high-frequency oscillations, which can compromise the forecast procedure. First of all, these fast-moving gravity waves require short computational time steps; second, they can seriously interfere with very short-range forecasts (<12 h); and third, they can impair vertical motion and, hence, precipitation forecasts (Daley 1981). Therefore, a long-standing approach has been to eliminate or effectively reduce these fast-moving inertia–gravity waves at initial forecast time. This process is known as model initialization. Charney (from unpublished letter to Phillip Thompson, 12 February 1947) provided insight for rectifying this initialization problem by suggesting that one should modify the initial state or modify the governing equations; that is, use filtered models.

The first and simplest approach was to exclude any possibility of the high-frequency oscillations by using a "filtered system" such as the balance equations, which simply reduce the model dynamics to the quasigeostrophic response. However, this approach severely restricts the model dynamics, which results in very poor forecasts beyond 24 hours. The so-called primitive equations account for more atmospheric *dynamics and can yield much better forecasts, so they* are generally used. However, the primitive equations do allow the amplification of fast-moving gravity waves, which requires some modification of the initial conditions to achieve the desired dynamical balance.

Over the years, many initialization methods have been developed. A summary of the more widely used techniques is presented here. For a more detailed review, the reader is referred to Daley (1991). In static initialization, the data are adjusted at a single time level to conform to some dynamical constraints. That is, certain time derivatives are identified as vanishing, in order to eliminate or reduce the generation of inertia–gravity-wave noise. In the conventional static initialization, a standard practice is to first analyze the geopotential field using pressure-height data and use wind observations to estimate the gradient of the geopotential using the geostrophic relation. The analyzed geopotential fields on pressure surfaces are

The most common approach of initialization in intermittent FDDA is normal mode initialization, which achieves dynamical balance using the normal modes of the linearized dynamical equations.

then used in the mass balance equation to obtain the streamfunction of the nondivergent wind, from which the rotational wind component can be computed. A major limitation in using the balance equation to determine a rotational wind for initializing a primitive equation model is that the lack of a divergent wind component insures the presence of gravitational modes (Haltiner and Williams 1980).

The most common approach of initialization in intermittent FDDA is normal mode initialization, which achieves dynamical balance using the normal modes of the linearized dynamical equations. The direct use of normal modes was introduced by Dickinson and Williamson (1972). They proposed that the amplitudes of the unwanted, fast-moving modes be set to zero. Their method was effective in suppressing the spurious noise in linear models, but failed in the nonlinear case. A nonlinear normal mode scheme was independently developed by Machenhauer (1977) and Baer (1977). In nonlinear normal mode initialization, the tendency of the unwanted modes are set to zero, versus setting the amplitude of these modes to zero. The solution of the nonlinear equation requires an iterative process. Unfortunately, this nonlinear scheme, without the inclusion of diabatic effects, suppressed the meridional circulation in the tropics. Puri and Bourke (1982) used the idea that the tropical divergent circulations driven by convection mainly influence the low-frequency gravity modes. Therefore, they excluded these low-frequency modes from the initialization using a frequency cutoff. Wergen (1982) introduced another method in which average diabatic heating is obtained by integrating the model for a few time steps prior to initialization. This model-produced diabatic heating is then included in the nonlinear forcing in the iterative process.

In global models (Andersen 1977; Daley 1979; Temperton and Williamson 1981; for example), after the normal modes of the model are computed, the high-frequency inertia–gravity waves can be removed by projecting the inertial wind and mass fields onto these normal modes. However, in limited-area models it is not possible to define the horizontal structure of the normal modes.

Bourke and McGregor (1983) introduced a method of initializing a limited area model without explicitly computing horizontal normal modes. In this technique, termed *vertical mode initialization,* the free modes of oscillation of the prediction model are identified by linearizing the equations about a basic state of rest. This linearization permits a simple decomposing of the three-dimensional eigenvalue problem into a series of two-dimensional problems. The vertical decomposing leads to a number of characteristic vertical modes, one corresponding to each discrete level in the model. Balance conditions on the horizontal structure equations are then derived for each vertical mode. Filtering conditions, in which the tendencies of divergence and ageostrophic vorticity are set to zero, are applied to derive linear diagnostic equations for the mass and divergency fields. In Fig. 2, from Bourke and McGregor (1983), graphs of surface pressure before and after initialization using the Australian regional primitive equation model are shown. It is obvious that the initialization procedure successfully removed noise from the integrations.

Temperton (1988) devised a method of applying Machenhauer's criterion without requiring the computation of the coefficients of the individual modes. Figure 3 (from Temperton 1988) shows graphs of 500- mb geopotential before and after the application of Temperton's implicit normal-mode initialization for the Canadian finite-element model. This method is also successful in removing the high-frequency noise.

Although appropriate for intermittent FDDA, the normal-mode initialization is a distinct, separate step from the objective analysis and usually leads to changes in model parameters. As a result, the initialized analysis may no longer fit the observations as closely as desired. An alternative, known as dynamic initializa-

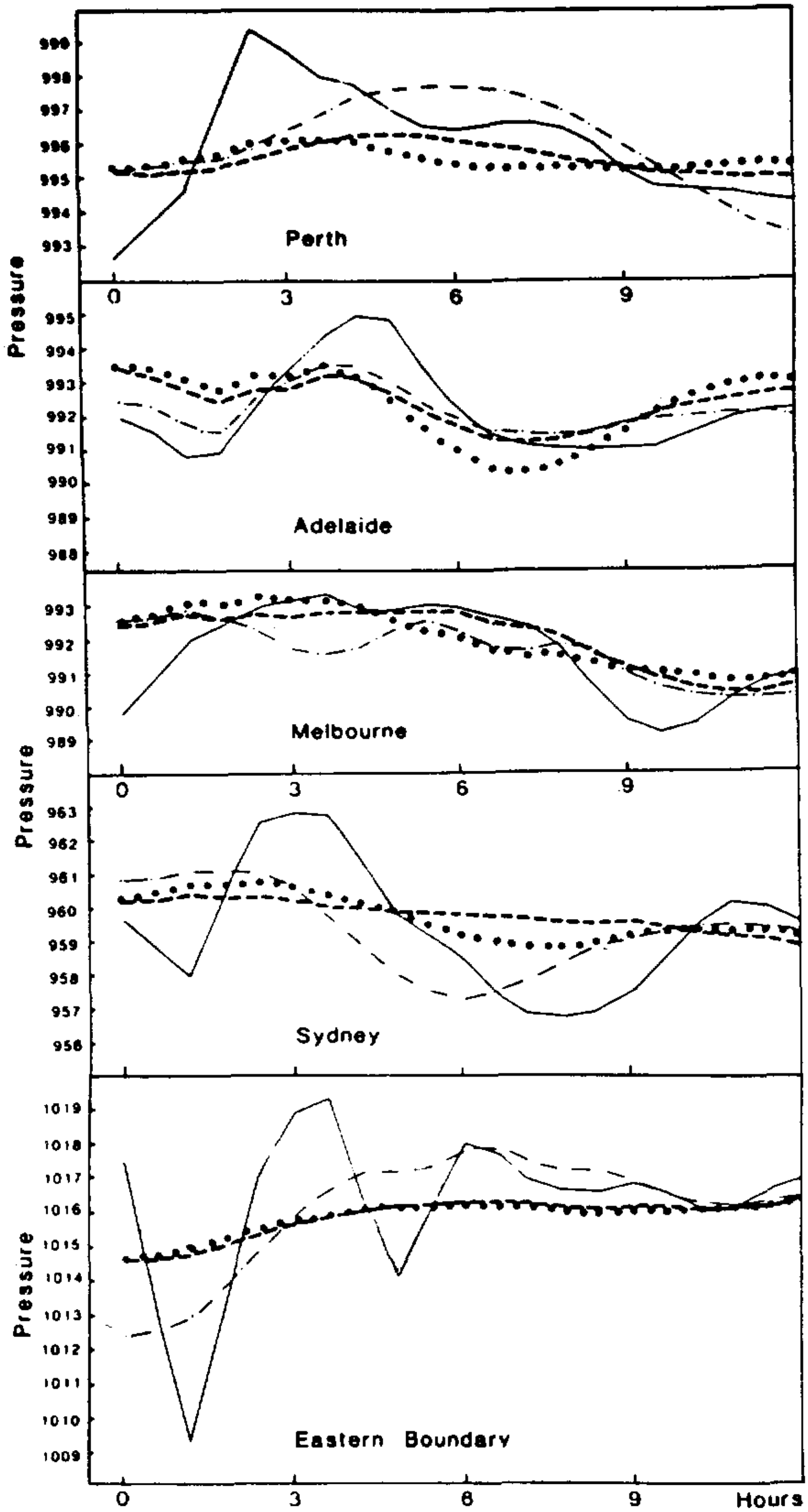

FIG. 2. Surface pressures in millibars for individual grid points for the first 12 h of prognosis using an uninitialized analysis (solid curves), after initializing two vertical modes (dot–dash), and after initializing two vertical modes (dashes—using Dirichlet boundary conditions; and dots—using Neumann boundary conditions) (Bourke and McGregor 1983).

tion, has the advantage of avoiding the complications of computing normal modes. In this method, which was introduced by Miyakoda and Moyer (1968), observations are inserted (intermittently or continuously) over a period of time. The method has the added advantage of simplicity and can balance physical processes as well as the mass and wind fields. In dynamic initialization, the model equations were integrated forward and backward through time under controls that encourage time derivatives to become small compared to spatial derivatives, which in turn selectively dampened the high-frequency components of the solutions. Early versions of this technique required several repetitions of the integration cycle to successfully reduce the gravity-wave oscillations and, as a result, these schemes were computationally expensive. Another disadvantage was that the slow modes of the model were dampened.

More recently, researchers including Bratseth (1982), Sugi (1986), and Satomura (1988) have developed schemes that are more computationally efficient and have more selective damping properties. Figure 4 (from Sugi 1986) depicts graphs of the gravity-wave activity (before and after initialization) for five vertical modes of a baroclinic model. The gravity-wave noise is dramatically reduced, particularly for the modes of large equivalent depths. Dynamic initialization, as well as the Laplace transform and bounded derivative methods (described below), is well suited for initializing data on a limited domain.

Lynch (1985a) developed an effective method of initialization based on a filtering scheme that uses a modified inverse Laplace transform. This technique is equivalent to the nonlinear normal-mode initialization method, but it has the advantage of not requiring a transformation of the model equations into normal mode space. Therefore, the Laplace trans-

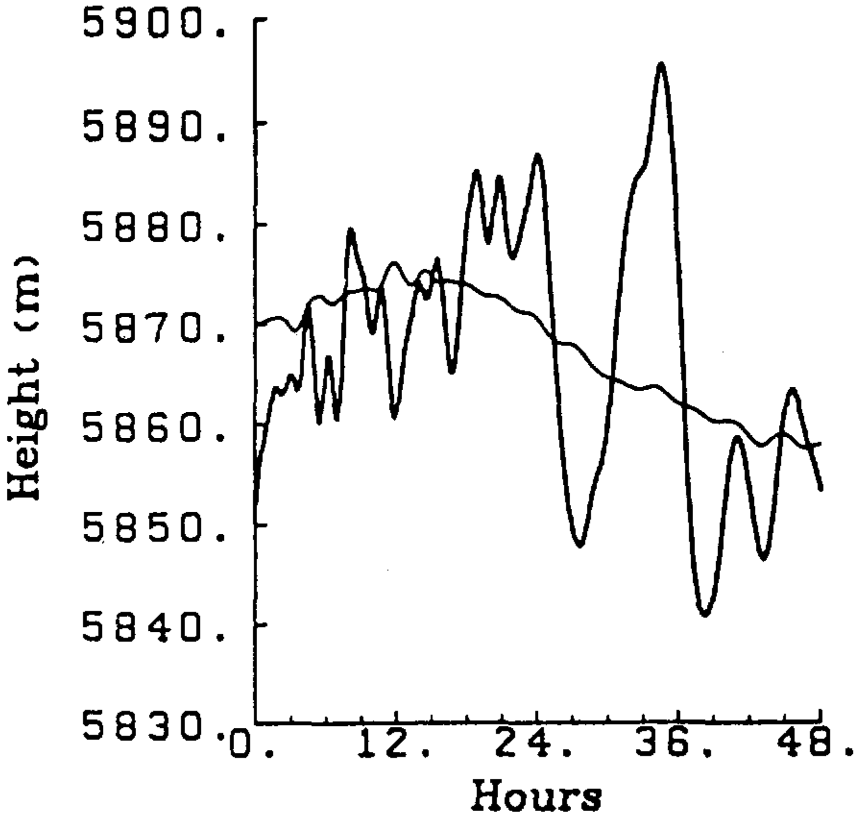

FIG. 3. Time trace of height field. Heavy line: no initialization. Light line: after two iterations of the implicit nonlinear normal-mode initialization scheme (Temperton 1988).

form method is well suited for initializing limited-area models with complex boundary conditions. Lynch (1985b) used his method to initialize data for a barotrophic limited-area model, successfully removing high-frequency gravity- wave oscillations during the model integration. This technique has also been applied in a filtering integration scheme for continuous data assimilation.

Another technique, which has been used to initialize models of limited domain, is the bounded derivative method. Kreiss (1979) developed the methodology of controlling the amplitudes of the high-frequency inertia–gravity waves by requiring the derivatives of the model's dependent variables with respect to time to be bounded, i.e., of order unity, as the initial time. First, the equations of motion are nondimensionalized so that certain terms are multiplied by a small parameter, ε. Then, if the first derivatives are bounded, the equations can only be satisfied if the model atmospheric flow is geostrophic and nondivergent to the order ε. If the second-order time derivatives are bounded, the resulting diagnostic relationships are the quasi-geostrophic omega equation and the nonlinear balance equation.

5. Various data-assimilation methods

An assimilation method extensively used in operational meteorology is the analysis–forecast cycle technique, commonly referred to simply as intermittent data assimilation (see Fig. 5). This process consists of four steps, which are repeated at each assimilation cycle (typically every 3–12 h). After the data have been checked (quality controlled), a static three-dimensional objective analysis (typically successive correction or optimal interpolation) is performed using observations and a background field. The background or "first guess" is usually a prior model forecast valid at the analysis time, or it can simply be climatology or a combination of both. Then, the analyzed fields are adjusted, or initialized, to conform to some dynamical constraint(s) in order to reduce or eliminate inertia–gravity-wave noise. The final step consists of a short-range numerical forecast to obtain first-guess fields for the analysis at the next assimilation cycle. Thus, the new estimates (analyzed fields) are clearly based on the past observations, being carried forward in time by the model forecast, and on the current observations. The intermittent updating process is entirely appropriate as long as most available data are taken at the same time, for example, at synoptic times. This technique is currently used at most of the world's major operational forecasting centers, including the NMC, (DiMego 1988); the Norwegian Meteorological Institute (Gronas and Midtbo 1986); and the European Centre for Medium-Range Weather Forecasts (ECMWF) (Hollingsworth 1986).

The extensive operational use of intermittent data assimilation is primarily due to its computational efficiency. In addition, this method normally includes a normal-mode initialization scheme that produces a balanced mass/wind initial state. A disadvantage of this method is that it is not totally suited for asynoptic data types; that is, it can not assimilate data continuously. However, updates on the order of every 2 to 3 h can be made, allowing some asynoptic data into the assimilation.

Two other mathematically elegant methods,

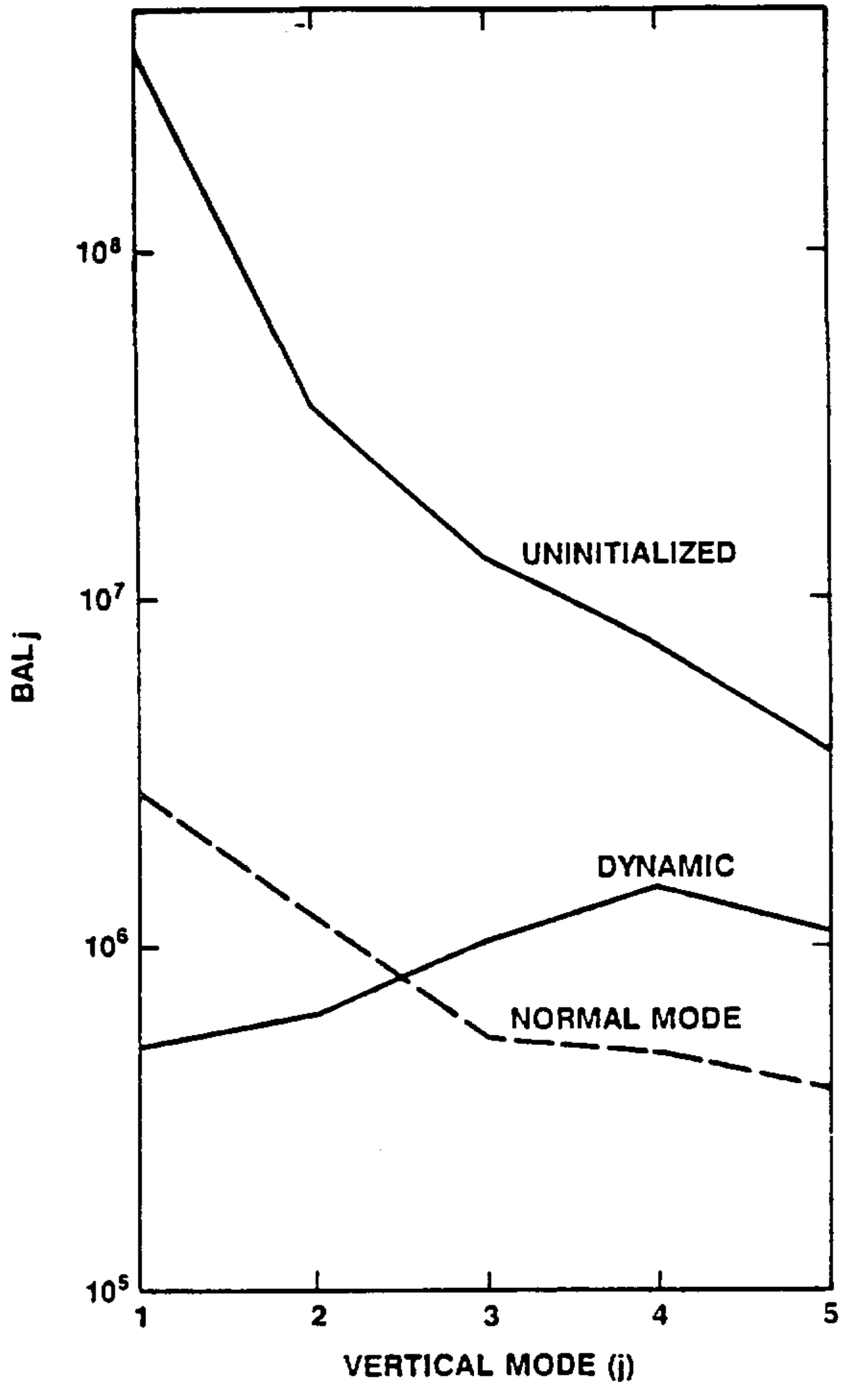

FIG. 4. Gravity-wave activity (before and after initialization) for five vertical modes of a baroclinic model (Sugi 1986).

Kalman–Bucy filtering and the adjoint method using variational techniques, have emerged as state-of-the-art methodology in FDDA. The latter is briefly reviewed first. For a more rigorous and complete discussion of the general theory of these two techniques, the reader is referred to Daley (1991).

Variational assimilation, based on the calculus of variations, involves the incorporation of dynamical constraints in a variational treatment and has been pursued by Sasaki (1969), Stephens (1970), and others for several years. In variational calculus, stationary points (extrema) of integral expressions known as functionals are determined. J is a functional of the function $q(t)$ in the interval (t_m, t_n), if it depends on all the values $q(t)$ for $t_m \leq t \leq t_n$. In this approach, successive analyses are mutually adjusted to effectively increase the database at each time step by using information at other analysis times through the forecast equations. The objective is to produce initialized values of q subject to certain constraints, such as the hydrostatic relation, the continuity equation, the geostrophic relation, or the nonlinear balance equation. The approach is designed to keep the initialized fields close to the observations while satisfying the constraint (Daley 1991).

Lewis (1972) developed a variational scheme using a thermal wind relationship and the hydrostatic equation as constraints. A more recent variational approach, known as the adjoint method (Lewis and Derber 1985), uses a complete dynamic model as a strong constraint. This method fits a model to observational data distributed over a finite period by computing derivatives of model output. An iterative method minimizes the weighted squared difference between the original analyses at several times and the coincident solutions to the model (constraint) for a given output variable. The final analyses are constrained to satisfy the model forecast from a set of initial conditions. The functional, J, is minimized by finding the gradient of J with respect to the initial conditions.

The output of any model depends upon a set of input variables: initial conditions, boundary conditions, and even modeling parameters of physical processes. Because of this dependence, the "adjoint" can be used to determine the sensitivity of the model output to any input parameter, i.e., initial conditions. Neglecting forecast errors over the assimilation period

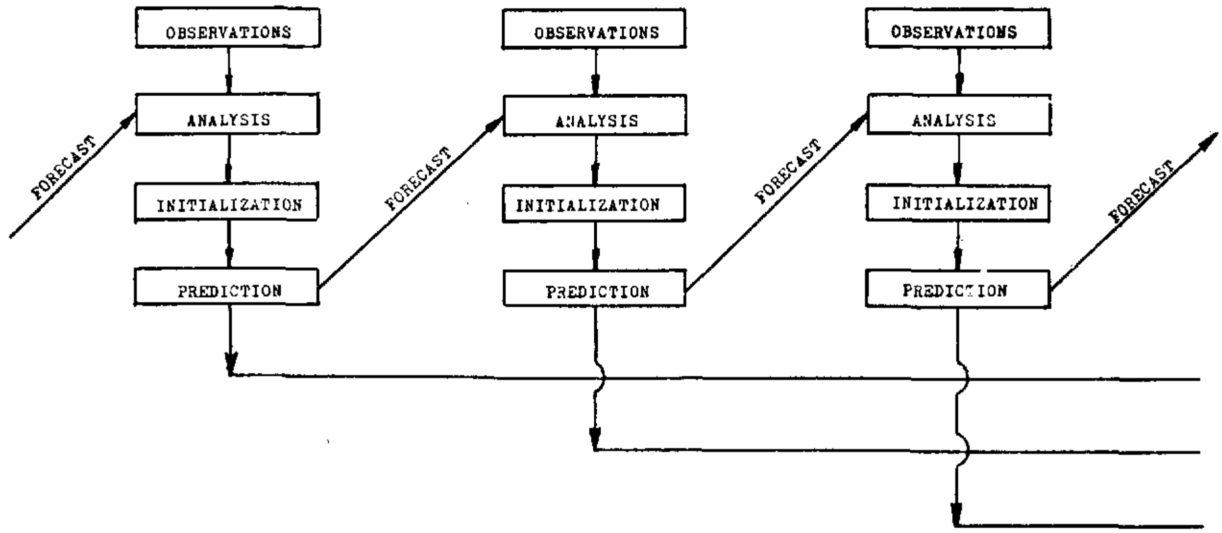

FIG. 5. Intermittent data assimilation using an analysis–forecast cycle (Bengtsson 1975).

(time period spanning the observations), this technique can produce the optimal initial state (end of the assimilation period) such that the model best fits the observations through the entire period.

The alternate "next-generation" FDDA method is called Kalman–Bucy (K–B) filtering. The K–B method can be thought of as a continuous dynamic FDDA where the weighting factors are optimally determined by explicitly calculating the error covariance of the analysis. In 1960, Kalman developed the basic theory for the linear, unbiased data-assimilation scheme known as the Kalman filter (Kalman 1960).

The K–B filter (Kalman and Bucy 1961; Ghil et al. 1979, 1981), which is the time-continuous counterpart to Kalman's original scheme, uses the forecast model itself to predict the background error statistics. The goal of this method is to obtain the most accurate analysis value for all time during the assimilation period using only present and past observations. The weight given to the current observations is inversely proportional to their variance, and the accuracy of the analysis is the sum of the accuracies of the forecast, based on the past observations, and of the current observations. The K–B filter minimizes the analysis error variance not only at every time step, but over the entire assimilation interval in which data are provided (Ghil 1989). Through an application of Bayesian ideas in a dynamical sense (Kalman 1960; Lorenc 1986), the filter is able to extract all useful information from the observational increment/residual at each time step, thus allowing observations to be discarded as soon as they are assimilated. As a result, the method is sequential.

The K–B filter and the adjoint method are very promising assimilation techniques. Both methods have been shown to produce improved assimilated states for model forecast integration as compared to those from the operational techniques. However, these "next-generation" data-assimilation methods are of limited practicality at the present because of their complexity and extensive computational requirements. Therefore, operational implementation is still several years away (Stauffer and Seaman 1990). However, Lorenc (1988) has pursued approximations to the adjoint technique in a quest to make it operational in numerical weather prediction.

Most research has been directed toward the less elegant but more practical method of dynamic assimilation, in which the numerical prediction model serves as an integrator of observations distributed in time and space. In this approach, which has become known simply as "nudging" or Newtonian relaxation, the model integration is interrupted periodically and the current model state is updated with the new observations. During the assimilation cycle, or preforecast integration period, the model variables are gradually driven, or nudged, toward the observations by extra forcing terms in the equations (Anthes 1974; Kistler 1974; Hoke and Anthes 1976; Davies and Turner 1977). As a result, the model fields are gradually corrected and no further dynamic balancing through initialization is required. The general form of the predictive equation of variable S is:

$$\frac{\partial S}{\partial t} = F(S, \tilde{x}, t) + G\, W\, (S_o - S). \qquad (5.1)$$

All of the model's physical forcing terms (Coriolis, advection, etc.) are represented by F, where S is a model-dependent variable, S_o represents observations of S, $\tilde{x}$ is the independent spatial variable, and t

is time. The second term on the right is the nudging term, where G is the nudging constant (generally 10^{-3} to 10^{-4}) and W represents a four-dimensional weighting function. The data to be nudged can be either derived or measured, analyzed to a grid for assimilation into the model, or inserted as individual observations.

This technique, which has been successful in bringing the data and the model in harmony and providing a relatively noise-free start for the forecast, has been widely used on the global scale (Lyne et al. 1982; Krishnamurti et al. 1988) and on the regional scale in limited-area models (Anthes 1974; Hoke and Anthes 1977). Of late, several researchers have developed a new interest in the technique (Stauffer et al. 1985; Bell 1986; Ramamurthy and Carr 1987; Kao and Yamada 1988; Wang and Warner 1988; Kuo and Guo 1989; Stauffer and Seaman 1987, 1990). It is currently used operationally at the UKMO for both global (Lyne et al. 1982) and regional (Bell 1986) data assimilation.

The nudging technique mentioned above does possess a few desirable attributes; these were summarized by N. L. Seaman (from his lecture notes presented at the 1990 Summer Colloquium on Mesoscale Data Assimilation, Boulder, Colorado): 1) The assimilating model is complete, so irreversible processes are included without difficulty. 2) Any data type that can be represented as a tendency of a prognostic variable can be assimilated. 3) Observation nudging can easily assimilate asynoptic and single-level data. 4) Analysis nudging requires that the analyses be performed only once prior to model integration; i.e., it is economical. 5) Nudging does not require a separate balancing/initialization step. 6) Nudging is conceptually and computationally simple.

On the other hand, this method has a few disadvantages: 1) The nudging constant is generally assigned in an application-dependent semiarbitrary manner. 2) Observation nudging is based on "continuous analysis" during model integration and can become computationally expensive. 3) Analysis nudging is not well suited for asynoptic data types. 4) Use of accurate data in observation nudging may cause assimilation of local or unrepresentative components (e.g., microscale observations spread over a large area).

6. Comparisons of operational assimilation techniques

Two papers from the recent literature (Ramamurthy and Carr 1987; Kuo and Guo 1989) compared different data-assimilation methods. Unlike the purpose of the present paper, both of the aforementioned articles advocated Newtonian nudging. The sole purpose here is to offer a short review of the available comparative studies.

Ramamurthy and Carr (1987) studied the applicability of several assimilation techniques currently being employed in operational models. A prime scientific objective was to determine the "best" way to assimilate asynoptic observations in limited-area models. A sequence of ten assimilation experiments were conducted using different update procedures. In each experiment, their limited-area model was initialized with ECMWF FGGE level III-b data, and then 12-h assimilations were performed using level II-b data from the 1979 Summer Monsoon Experiment (SMONEX). Forecasts were then made from these assimilated states. The first experiment served as the control run, since no assimilation was performed. Figure 6 is a schematic of the overall assimilation-forecast strategy; four types of data assimilation were compared: 1) In the static assimilation, the model was updated only once at the end of the preforecast (assimilation) period. An initialization step was then taken to suppress the noise associated with the external inertia–gravity mode; the internal modes were not initialized. 2) The intermittent data-assimilation experiment was similar to the static case, except the model was updated twice during the assimilation period. 3) Four experiments were conducted using continuous indirect data assimilation in which the model was updated whenever new observations became available. 4) Newtonian relaxation was used in the last three experiments; the model state was nudged toward analyses produced from the observations. The reader should see Ramamurthy and Carr (1987) for a detailed description of the experiments.

Comparisons among these differing techniques were made by examining assimilated states (analyses obtained at the end of the 12-h assimilation period). Newtonian nudging produced better assimilated states than did the continuous assimilations via indirect insertion. Also, the continuous assimilation experiments produced noisy assimilated fields due to insertion shocks.

From each of the assimilated states, 24-h forecasts were made and the results were compared against each other and with the observations. The continuous assimilation (via indirect insertion) continued to suffer from the ill effects of the insertion shock. However, the forecasts from the nudging experiments had a minimal amount of noise.

The degree of spinup was examined in terms of the development of precipitation. The excessive shocking associated with continuous insertion was detrimental to the spinup process and consequently to the rainfall predictions. The rotational nudging experiment (only the rotational component of the wind was nudged)

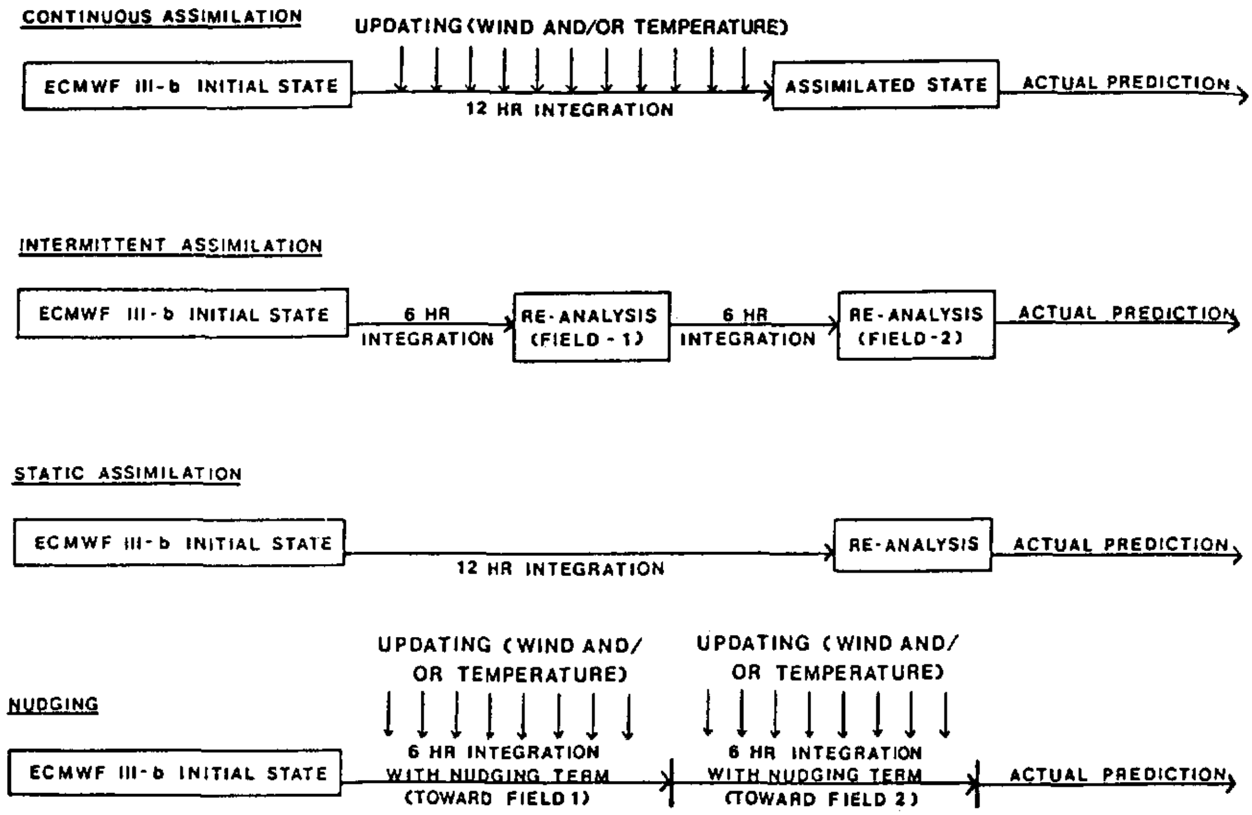

FIG. 6. A schematic of the assimilation-forecast cycle for various experiments (Ramamurthy and Carr 1987).

produced the most accurate rainfall predictions. The lack of divergence-related shock in this approach aided in the evolution of the physical processes.

The predicted tracks of the cyclone in the SMONEX case study were compared. During the first 12 h (assimilation period), the cyclone's movement was erratic in all of the experiments due to the rapid mutual adjustment between the mass and wind fields that occurred in the early stages of integration. In the static and intermittent assimilation experiments in which the fields were replaced completely at the end of the assimilation period, this problem extended into the forecast stage for another 12 h. Overall, the nudging experiments predicted more accurate tracks than the other experiments.

In another comparative study, Kuo and Guo (1989) examined different data-assimilation strategies. They conducted a series of observing-system simulation experiments (OSSEs) to test a Newtonian nudging technique for continuous assimilation of observations from a hypothetical network of profilers. Twenty experiments were described in their paper. Similar to the work of Ramamurthy and Carr, they investigated static initialization, intermittent assimilation, and Newtonian nudging.

Comparisons of the three assimilation techniques were made by obtaining time series of wind and temperature errors during the assimilation period. With nudging, the errors decreased gradually during the period. The intermittent assimilation produced a stepwise decrease in the errors and the results after four assimilation cycles were considerably better than in the static assimilation case. This reveals that intermittent data assimilation was also effective in producing an improved initial state for the model forecast.

When considering model noise generated during the preforecast period, a strong difference emerged between continuous nudging and the intermittent assimilation. In the nudging experiment, the model noise gradually decreased during the assimilation period. In contrast, the intermittent data assimilation produced model noise with large spikes immediately after each reanalysis of the wind field. However, this noise could have been reduced substantially if a normal-mode initialization procedure had been incorporated.

The two papers reviewed above suggest that Newtonian nudging is the best assimilation method among those compared. However, the investigators' findings may be scientifically inconclusive due to their incomplete tests (notably, their failure to include an effective normal-mode initialization procedure in the intermittent data-assimilation experiments). Further-more, their comparisons were limited to operational data-assimilation techniques. The next-generation methods will produce more optimal assimilated states for initiating a numerical weather prediction model. But as stated previously, these methods are still in the developmental research stage and computationally efficient versions are not currently available for operational use.

7. FDDA research: Implications and the future

FDDA has been developed tremendously during the past 15 years and is now an essential component of numerical analysis and prediction systems in both research and operations. The improving performance of medium-range global and hemispheric prediction and the rapid development of limited-area and mesoscale models at centers such as the NMC, ECMWF, and UKMO clearly illustrate the practical benefits of research in data assimilation.

The systematic development of data-assimilation methods has made possible the use of unconventional and asynoptic observations from satellites, aircraft, drifting buoys, and soon from wind profilers and doppler radars. As a result, the accuracy of short-range forecasts has significantly improved, although only minor changes have occurred in the global observing system since the Global Weather Experiment in 1979 (Bengtsson and Shukla 1988). According to Lange and Hellsten (1986), the 3-day rms forecast error for the Northern Hemisphere dropped by more than 35% between 1979 and 1986. During the same period, medium-range weather prediction has been usefully extended in the time scale from 3–4 days to about 7 days in the Northern Hemisphere and to 4 days in the Southern Hemisphere (Bengtsson 1985; Bourke et al. 1985). The largest improvement has occurred at middle and high latitudes of the Northern Hemisphere. Weather prediction in the tropics has not improved nearly as much, due to insufficient observations and deficiencies in the formulation of critical physical processes.

Lorenz (1982) addressed the limit of medium-range predictability and concluded that it is possible to predict instantaneous weather patterns with better accuracy than guesswork nearly two weeks in advance. Such extensions in predictability will depend on the ability to improve current data-assimilation systems as well as the numerical prediction models themselves.

Research is ongoing to find ways to improve data-assimilation methods. A number of research centers are investigating higher-resolution models, better numerical techniques, improved physical parameterizations, and improved models of auto- and cross-correlation functions for prediction error. Another area of interest is in specifying the diabatic heating at the initial forecast time. A detailed specification is necessary in order to correctly analyze and forecast the divergent wind field over the tropics. A few universities are making major strides in developing the next-generation FDDA techniques and enhancing existing methods. For example, adjoint methods are being intensely investigated at the University of Oklahoma; K–B filtering at McGill University, Montreal, Canada; and Newtonian relaxation at the Pennsylvania State University.

Several other critical research topics need attention during the 1990s. They include: assimilation of moisture processes (rainfall data and integrated liquid water), lateral boundary conditions and mesoscale predictability (predictive skill of limited-area models is strongly controlled by the lateral boundary conditions), assimilation of surface conditions/characteristics, and

During the next decade, FDDA will play a role in solving or alleviating important environmental issues such as air pollution and climate control (increasing CO_2 and the ozone hole). Recent clean-air legislation requiring the use of low-sulfur fuels and extremely expensive equipment will make accurate numerical models of atmospheric-chemistry transport and removal more important than ever.

quality control, which will continue to be a priority as new data types are applied in FDDA. Data assimilation is relatively new on the meso-alpha and sub-alpha scales and needs considerable research with improved datasets. Existing first-generation assimilation systems must be improved to effectively forecast on these small scales.

During the next decade, FDDA will play a role in solving or alleviating important environmental issues such as air pollution and climate control (increasing CO_2 and the ozone hole). Recent clean-air legislation requiring the use of low-sulfur fuels and extremely expensive equipment will make accurate numerical models of atmospheric-chemistry transport and removal more important than ever. FDDA can be used to obtain accurate meteorological fields for input to complex air-chemistry models. Bengtsson and Shukla (1988) have suggested that a comprehensive analysis of global observations based on an FDDA system with a realistic physical model should be used to produce internally consistent, homogeneous datasets for the earth's climate system. These global observations will include many new environmental variables such as external forcing variables, concentrations of radiatively and chemically important rare species, and land surface and oceanic variables. New remote-sensing systems such as EOS (Earth Observing System) will monitor these variables.

With these environmental issues in mind, Daley (1991) elaborated on a possible vision of the future of data assimilation:

> Firstly, data assimilation will no longer be entirely or even primarily concerned with short (or medium) range weather forecasting. Secondly, the data base will become incredibly diverse in both variables measured and type of observing system. Thirdly, assimilating models will be much more comprehensive, involving ocean, land surface and stratospheric components. Finally, there will be considerably more emphasis on the long term stability of algorithms to facilitate climate change signal detection.

8. Summary

In this paper, an attempt has been made to offer a review of four-dimensional data assimilation from a historical perspective with emphasis on two of the major components of data assimilation: objective analysis and initialization. Major objective analysis methods were reviewed by McPherson (1976) and Gustavsson (1981) and have been summarized here. These include polynomial interpolation, successive correction, optimal interpolation, spectral analysis, and variational techniques. Several model-initialization methods have been highlighted; among these are static initialization, nonlinear normal-mode initialization, vertical-mode initialization, implicit normal-mode initialization, dynamic initialization, Laplace transform, bounded derivative technique, and the variational method.

Several data-assimilation methods have been reviewed, including the two leading next-generation FDDA techniques: the adjoint method and the Kalman–Bucy filtering. Newtonian relaxation and the analysis–forecast cycle (intermittent data assimilation) were identified as the two major types of FDDA used today in operational numerical weather prediction, and they will likely be needed to meet the challenges imposed by new observing systems at least into the mid-1990s. These two methods, along with continuous assimilation via indirect insertion and static assimilation, were compared.

Finally, some important implications of FDDA research have been given. FDDA has undoubtedly led to significant improvements in short- and medium-range weather forecasting and may provide the much-needed edge in coping with the pressing environmental issues (air pollution and climate control) of the nineties.

Acknowledgments. The authors extend their appreciation to the Naval Research Laboratory and the U.S. Air Force for supporting this work.

References

Alaka, M. A., and R. C. Elvander, 1972: Optimum interpolation from observations of mixed quality. *Mon. Wea. Rev.*, **100**, 612–624.

Andersen, J. H., 1977: A routine for normal mode initialization with non-linear correction for a multi-level spectral model with triangular truncation. ECMWF Internal Report No. 15, European Centre for Medium-Range Weather Forecasts, Bracknell, U.K., 41 pp.

Anthes, R. A., 1974: Data assimilation and initialization of hurricane prediction models. *J. Atmos. Sci.*, **31**, 702–719.

Baer, F., 1977: Adjustments of initial conditions required to suppress gravity oscillations in non-linear flows. *Beitr. Phys. Atmos.*, **50**, 350–366.

Bell, R. S., 1986: The Meteorological Office fine-mesh data assimilation scheme. *Meteor. Mag.*, **115**, 161–177.

Belousov, S. L., L. S. Gandin, and S. A. Mashkovich, 1968: Computer processing of current meteorological data. *Hydrometeoizdat* [Meteor. Trans. No. 18, 1972. Canada Dept. of the Environment, Atmospheric Environment Service, Downsview, Ontario], 227 pp. (in Russian).

Bengtsson, L., 1975a: Four-dimensional data assimilation of meteorological observations. GARP Publ. Ser., No. 15, WMO/ICSU, Geneva, 76 pp.

——, 1975b: Problems in four-dimensional data assimilation. *Seminars of Scientific Foundation of Medium Range Weather Forecasts, Part I*, Reading, 113–138.

——, 1985: Medium range forecasting at the ECMWF. *Adv. Geophys.*, **28B**, 3–56.

——, and J. Shukla, 1988: Integration of space and in situ observations to study global climate change. *Bull. Amer. Meteor. Soc.*, **69**, 1130–1143.

Bergthorsson, P., and B. R. Döös, 1955: Numerical weather map analysis. *Tellus*, **7**, 329–340.

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

——, R. Seaman, and K. Puri, 1985: Data assimilation. *Adv. Geophys.*, **28B**, 124–149.

Bratseth, A. M., 1982: A simple and efficient approach to the initialization of weather prediction models. *Tellus*, **34**, 352–357.

Charney, J., M. Halem, and R. Jastrow, 1969: Use of incomplete historical data to infer the present state of the atmosphere. *J. Atmos. Sci.*, **26**, 1160–1163.

Cressman, G., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374.

Daley, R., 1979: The application of non-linear normal mode initialization to an operational forecast model. *Atmos. Ocean*, **17**, 97–124.

——, 1981: *Predictability experiments with a baroclinic model. Atmos. Ocean*, **19**, 77–109.

——, 1991: *Atmospheric Data Analysis.* Cambridge University Press, 457 pp.

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

Davies, H. C., and R. E. Turner, 1977: Updating prediction models by dynamical relaxation: An examination of the technique. *Quart. J. Roy. Meteor. Soc.*, **103,** 225–245.

Dickinson, R. E., and D. L. Williamson, 1972: Free oscillations of a discrete stratified fluid with application to numerical weather prediction. *J. Atmos. Sci.*, **29,** 623–640.

DiMego, G. J., 1988: The National Meteorological Center regional analysis system. *Mon. Wea. Rev.*, **116,** 977–1000.

Dixon, R., 1976: An objective analysis system using orthogonal polynomials. GARP WGNE Report No. 11, 73–85. [Available from World Meteorological Organization, Geneva.]

——, E. Spackman, I. Jones, and A. Francis, 1972: The global analysis of meteorological data using orthogonal polynomial base functions. *J. Atmos. Sci.*, **29,** 609–622.

Eddy, A., 1973: The objective analysis of atmospheric structure. *J. Meteor. Soc. Jpn.*, **51,** 450–457.

Eliassen, A., 1954: Provisional report on calculation of spatial covariance and autocorrelation of the pressure field. Rept. No. 5, Institute of Weather and Climate Res., Academy of Science Oslo, 11 pp. (reprinted in *Dynamic Meteorology: Data Assimilation Methods,* L. Bengtsson, M. Ghil, and E. Kallen, Eds., Springer-Verlag, 319–330).

Flattery, T., 1971: Spectral models for global analysis and forecasting. *Proc. Sixth AWS Technical Exchange Conf.*, Air Weather Service Technical Report No. 242, 42–54.

Gandin, L. S., 1963: Objective analysis of meteorological fields. *Hydrometeoizdat* (in Russian), English translation by Israel Program for Scientific Translations, Jerusalem, 1965, 242 pp. (available from NTIS, as N66-18047).

——, 1988: Complex quality control of meteorological observations. *Mon. Wea. Rev.*, **116,** 1137–1156.

Ghil, M., 1989: Meteorological data assimilation for oceanographers, Part I: Description and theoretical framework. *Dyn. Atmos. and Oceans,* **13,** 171–218.

——, M. Halem, and R. Atlas, 1979: Time-continuous assimilation of remote-sounding data and its effect on weather forecasting. *Mon. Wea. Rev.,* **107,** 140–171.

——, S. Cohn, J. Tavantzis, K. Bube, and E. Issacson, 1981: Applications of estimation theory to numerical weather prediction. *Dynamic Meteorology: Data Assimilation Methods,* L. Bengtsson, M. Ghil, and E. Kallen, Eds., Springer-Verlag, 139–224.

Gilchrist, B., and G. P. Cressman, 1954: An experiment in objective analysis. *Tellus,* **6,** 97–101.

Gronas, S., and K. H. Midtbo, 1986: Four dimensional data assimilation at the Norwegian Meteorological Institute. Tech. Report 66, Research Dept., The Norwegian Meteorological Institute, Oslo, Norway, 66 pp.

Gustavsson, N., 1981: A review of methods for objective analysis. *Dynamic Meteorology: Data Assimilation Methods,* L. Bengtsson, M. Ghil, and E. Kallen, Eds., Springer-Verlag, 17–76.

Haltiner, G. J., and R. T. Williams, 1980: *Numerical Prediction and Dynamic Meteorology.* Second ed., John Wiley and Sons, 477 pp.

Hayden, C. M., 1976: Satellite reference level experiments with VTRR and the NMC global spectral analyses. GARP WGNE Report No. 11, 57–72. [Available from World Meteorological Organization, Geneva.]

Hoke, J. E., and R. A. Anthes, 1976: The initialization of numerical models by a dynamical initialization technique. *Mon. Wea. Rev.*, **104,** 1551–1556.

——, and R. A. Anthes, 1977: Dynamic initialization of a three-dimensional primitive-equation model of Hurricane Alma of 1962. *Mon. Wea. Rev.*, **105,** 1266–1280.

Hollingsworth, A., 1986: Objective analysis for numerical weather prediction. *Short and Medium-Range Weather Prediction, Proc. WMO/IUGG NWP Symposium,* Tokyo, 11–59.

Jastrow, R., and M. Halem, 1970: Simulation studies related to GARP. *Bull. Amer. Meteor. Soc.*, **51,** 490–513.

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. Trans. ASME, *J. Basic Eng.*, **82D,** 35–45.

——, and R. S. Bucy, 1961: New results in linear filtering and prediction theory. Trans. ASME, *J. Basic Eng.*, **83D,** 95–108.

Kao, C.-Y. J., and T. Yamada, 1988: Use of the CAPTEX data for evaluations of a long-range transport numerical model with a four-dimensional data assimilation technique. *Mon. Wea. Rev.*, **116,** 293–306.

Kistler, R. E., 1974: A study of data assimilation techniques in an autobarotropic primitive equation channel model. M.S. thesis, The Pennsylvania State University, 84 pp.

Kreiss, H.-O., 1979: Problems with different time scales for ordinary differential equations. *SIAM J. Num. Anal.*, **16,** 980–998.

Krishnamurti, T. N., H. S. Bedi, W. Heckley, and K. Ingles, 1988: Reduction of the spinup time for evaporation and precipitation in a spectral model. *Mon. Wea. Rev.*, **116,** 907–920.

Kuo, Y.-H., and Y.-R. Guo, 1989: Dynamic initialization using observations from a network of profilers. *Mon. Wea. Rev.*, **117,** 1975–1998.

Lange, A., and E. Hellsten, 1986: Results of the WMO/CAS NWP data study and intercomparison project for forecasts for the Northern Hemisphere in 1981–82. WMO Short- and Medium-Range Weather Prediction Research, Publication Series No. 2, WMO, Geneva.

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

——, and J. C. Derber, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus,* **37A,** 309–322.

Lorenc, A., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109,** 701–721.

——, 1985: Analysis methods for the quality control of observations. *Workshop on the Use and Quality Control of Meteorological Observations,* Reading, U.K., European Centre for Medium-Range Weather Forecasts, 397–428.

——, 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112,** 1177–1194.

——, 1988: A practical approximation to optimal four-dimensional objective analysis. *Mon. Wea. Rev.*, **116,** 730–745.

——, and O. Hammon, 1988: Objective quality control of observations using Bayesian methods: Theory and a practical implementation. *Quart. J. Roy. Meteor. Soc.*, **114,** 515–543.

——, I. Rutherford, and G. Larsen, 1977: The ECMWF analysis and data assimilation scheme—Analysis of mass and wind fields. European Centre for Medium-Range Weather Forecasts, Reading, Berkshire, Tech. Rep. No. 6.

Lorenz, E. N., 1982: Atmospheric predictability experiments with a large numerical model. *Tellus,* **34,** 505–513.

Lynch, P., 1985a: Initialization using Laplace transforms. *Quart. J. Roy. Meteor. Soc.*, **111,** 243–258.

——, 1985b: Initialization of a barotropic limited-area model using the Laplace transform technique. *Mon. Wea. Rev.*, **113,** 1338–1344.

——, 1991: Filtering integration schemes based on the Laplace and Z transforms. *Mon. Wea. Rev.*, **119,** 653–666.

Lyne, W. H., R. Swinbank, and N. T. Birch, 1982: A data assimilation experiment and the global circulation during the FGGE special observing periods. *Quart. J. Roy. Meteor. Soc.*, **108,** 575–594.

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model with applications to normal mode initialization. *Beitr. Phys. Atmos.*, **50,** 253–271.

McPherson, R. D., 1975: Progress, problems, and prospects in meteorological data assimilation. *Bull. Amer. Meteor. Soc.,* **56,** 1154–1166.

——, 1976: Operational objective analysis techniques and potential applications for mesoscale meteorology. *Mesoscale Meteorology and Forecasting,* P. S. Ray, Ed., Amer. Meteor. Soc., 151–172.

Miyakoda, K., and R. Moyer, 1968: A method of initialization for dynamical weather forecasting. *Tellus,* **20,** 115–128.

Morel, P., 1981: An overview of meteorological data assimilation. *Dynamic Meteorology: Data Assimilation Methods,* L. Bengtsson, M. Ghil, and E. Kallen, Eds., Springer-Verlag, 5–16.

——, and O. Talagrand, 1974: Dynamic approach to meteorological data assimilation. *Tellus,* **26,** 334–344.

Panofsky, H., 1949: Objective weather map analysis. *J. Meteor.,* **6,** 386–392.

Petersen, D. P., 1968: On the concept and implementation of sequential analysis for linear random fields. *Tellus,* **20,** 673–686.

Phillips, N. A., 1976: The impact of synoptic observing and analysis systems on flow pattern forecasts. *Bull. Amer. Meteor. Soc.,* **57,** 1225–1250.

Puri, K., and W. Bourke, 1982: A scheme to retain the Hadley circulation during nonlinear normal mode initialization. *Mon. Wea. Rev.,* **110,** 327–335.

Ramamurthy, M. K., and F. H. Carr, 1987: Four-dimensional data assimilation in the monsoon region. Part I: Experiments with wind data. *Mon. Wea. Rev.,* **115,** 1678–1706.

Robinson, A. R., 1986: Data assimilation, mesoscale dynamics and dynamical forecasting. *Advanced Physical Oceanographic Numerical Modelling,* J. J. O'Brien, Ed., D. Reidel Pub. Co., 465–483.

Rutherford, I. D., 1972: Data assimilation by statistical interpolation of forecast error fields. *J. Atmos. Sci.,* **29,** 809–815.

——, 1976: An operational three-dimensional multivariate objective analysis scheme. *Proc. GARP Conf. on Four Dimensional Data Assimilation,* Paris, GARP WGNE Rep. No. 11, 98-121.

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

——, 1969: Proposed inclusion of time variation terms, observational in numerical variational objective analysis. *J. Meteor. Soc. Jpn.,* **47,** 115–124.

Satomura, I., 1988: Dynamic normal mode initialization of a limited-area model. *J. Meteor. Soc. Jpn.,* **66,** 261–276.

Smagorinsky, J., K. Miyakoda, and R. Strickler, 1970: The relative importance of variables in initial conditions for dynamical weather prediction. *Tellus,* **22,** 141–157.

Stauffer, D. R., and N. L. Seaman, 1987: A real-data numerical study and four-dimensional data assimilation application for mesobeta-scale flow in complex terrain. *Proc. Symp. Mesoscale Analysis and Forecasting,* Vancouver, ESA, 533-538.

——, and ——, 1990: Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Mon. Wea. Rev.,* **118,** 1250–1277.

——, T. T. Warner, and N. L. Seaman, 1985: A Newtonian "nudging" approach in four-dimensional data assimilation: Use of SESAME-IV data in a mesoscale model. Preprints, *Seventh Conf. on Numerical Weather Prediction,* Montreal, Amer. Meteor. Soc., 77–82.

Stephens, J., 1970: Variational initialization of the balance equation. *J. Appl. Meteor.,* **9,** 732–739.

Sugi, M., 1986: Dynamic normal mode initialization. *J. Meteor. Soc. Jpn.,* **64,** 623–632.

Temperton, C., 1988: Implicit normal mode initialization. *Mon. Wea. Rev.,* **116,** 1013–1031.

——, and D. L. Williamson, 1981: Normal mode initialization for a multilevel grid-point model, Part I: Linear aspects. *Mon. Wea. Rev.,* **109,** 729–743.

Wang, W., and T. T. Warner, 1988: Use of four-dimensional data assimilation by Newtonian relaxation and latent heat forcing to improve a mesoscale-model precipitation forecast, a case study. *Mon. Wea. Rev.,* **116,** 2593–2613.

Warner, T.T., 1987: Four-dimensional initialization of mesoscale models. In *Proc. Symp. Mesoscale Analysis and Forecasting,* Vancouver, Canada, ESA SP-282, 663–665.

Wergen, W., 1982: Incorporation of diabatic effects in non-linear normal mode initialization. *The GARP Programme on Numerical Experimentation,* Rep. No. 3, GARP/WCRP, WMO/ICSU, Geneva, 2.8–2.10.