NASA Contractor Report 3981

A Variational Assimilation Method for Satellite and Conventional Data: Development of Basic Model for Diagnosis of Cyclone Systems

G. L. Achtemeier, H. T. Ochs III,
S. Q. Kidder, R. W. Scott,
J. Chen, D. Isard, and B. Chance

CONTRACT NAS8-34902
MAY 1986

NASA

NASA Contractor Report 3981

A Variational Assimilation Method for Satellite and Conventional Data: Development of Basic Model for Diagnosis of Cyclone Systems

G. L. Achtemeier, H. T. Ochs III,
S. Q. Kidder, R. W. Scott,
J. Chen, D. Isard, and B. Chance
University of Illinois
Champaign, Illinois

Prepared for
George C. Marshall Space Flight Center
under Contract NAS8-34902

National Aeronautics
and Space Administration

**Scientific and Technical
Information Branch**

1986

Table of Contents

Chapter I

Summary of Research

1. Introduction

This chapter presents in Sections 2 - 7 the abstracts of the papers that appear in full in the remaining six chapters of this report. Each paper is a manuscript written for publication in a technical journal and is authored by one or more members of our research group. These manuscripts are in various stages of the review process and their final forms will be different from those presented here.

Sections 8 and 9 present work that was underway during the contract but which has not progressed sufficiently for the preparation of manuscripts. One study entitled "Application of Satellite Data to the Variational Analysis of the Three-Dimensional Wind Field" is the MS thesis topic for Ms. Barbara Chance. These studies will be finalized at a later time.

2. A Variational Assimilation Method for the Diagnosis of Cyclone Systems. Part I: Development of the Basic Model.

This paper outlines a theory for a variational objective analysis for the diagnosis of cyclone systems. Gridded fields of data from different type, quality, location and measurement source are weighted according to measurement accuracy and merged using a least squares criteria so that the two nonlinear horizontal momentum equations, the hydrostatic equation, and an integrated continuity equation are satisfied. We use the variational method of undetermined multipliers to derive the Euler-Lagrange equations necessary to create a dynamically consistent hybrid data set. A quasi-geostrophic solution sequence for these equations is described.

Other features of the variational diagnostic model include a hybrid nonlinear terrain-following vertical coordinate that eliminates truncation error in the pressure gradient terms of the horizontal momentum equations and easily accommodates TIROS-N mean layer temperatures in the middle and upper troposphere. A projection of the pressure gradient onto equivalent pressure surfaces removes most of the impacts of the lower coordinate surface on the variational adjustment. In addition, the local tendencies of the horizontal velocity components are reformulated to better diagnose these hypersensitive quantities.

An application of the variational diagnostic model to the study of a dissipating short wave appears in the following companion paper.

3. A Variational Assimilation Method for the Diagnosis of Cyclone Systems. Part II: Case Study Results with and without Satellite Data.

This paper presents the evaluation of a diagnostic multivariate data assimilation method described in a companion paper by Achtemeier et al.. Ground-based and space-based meteorological data are weighted according to the respective "measurement" errors and blended into a hybrid data set that is required to satisfy the two nonlinear horizontal momentum equations, the hydrostatic equation, and an integrated continuity equation for a dry atmosphere as dynamical constraints. Multivariate variational objective analyses with and without satellite data are compared with initial analyses and the observations to determine the accuracy and sensitivity of the assimilation to different data sets. Three evaluation criteria are developed that measure a) the extent to which the assimilated fields satisfy the dynamical constraints, b) the extent to which the assimilated fields depart from the observations, and c) the extent to which the assimilated fields are realistic as determined by pattern recognition. The last criterion requires that the signs, magnitudes, and patterns of the hypersensitive vertical velocity and local tendencies of the horizontal velocity components be physically consistent with respect to the larger scale weather systems.

The percent reduction of the initial RMS error is used to determine the extent to which the SAT and NOSAT blended data sets converge to the solution of the four dynamical constraints. There was approximately 90-95 percent error reduction for the two horizontal momentum equations when applied to the case of 1200 GMT 10 April 1979. The RMS error reductions for the integrated continuity and hydrostatic equations ranged from 90-100 percent except for the errors at levels 2 and 3 of the integrated continuity equation which were reduced to approximately 70 percent.

The pattern recognition analysis for the basic fields, height, temperature, and vector wind, revealed that the SAT and NOSAT analyses were similar with the following two exceptions. First. there were larger numerical differences between the SAT height analysis and the initial objective analysis than were found between the NOSAT analyses and the initial objective analysis. Second, large areas of the network were void of satellite data which caused the loss of important local details of the temperature field. One result was the introduction of a large (-40 m) height anomaly in the middle troposphere over the western U.S. Both NOSAT and SAT analyses corrected a rather poor univariate wind analysis and placed jet streaks over California, western Texas, and the Great Lakes.

In the analysis of hypersensitive variables, the variational method removed or reduced the magnitudes of several large vertical velocity centers (magnitudes greater than 10 cm sec^{-1}) which were placed between rawinsonde sites by conventional methods and replaced them with a zone of positive vertical velocity roughly parallel to the axis of an area of precipitation that was used as a check on the accuracy of the final fields. The variational analysis also concentrated an area of strong subsidence (-14 cm/sec) along the axis of a jet streak over the Northeastern States. It also placed a narrow band of ascending motion near the entrance region along the anticyclonic shear side of the jet streak, an area long noted by those experienced in the motion fields surrounding jet streams as favorable for upward vertical velocities.

The SAT and NOSAT tendency patterns were of approximately the same magnitude and scale as the observed tendency patterns that were obtained from NASA-AVE high frequency rawinsonde data. With the exception of a negative tendency center in the lower troposphere, the agreement among the tendency patterns was very good considering that the observed patterns were subject to interference by mesoscale phenomena and that the observed patterns were valid at 1330 GMT rather than at 1200 GMT. The relative accuracy of the variational tendencies was made more apparent upon comparison of the initial field tendencies with the observed patterns. The initial field tendencies consisted of relatively large amplitude centers of scale roughly equal to the

average separation between observing sites. The magnitudes of these centers became unrealistically large in the upper troposphere within high wind velocity areas.

These variational tendencies are the first relatively accurate diagnostic fields of local tendencies of the velocity components apart from initialization schemes for numerical prediction models.

4. Hybrid Vertical Coordinate and Pressure Gradient Formulations for a Numerical Variational Analysis Model for the Diagnosis of Cyclone Systems.

A hybrid nonlinear sigma vertical coordinate that is suitable for a diagnostic variational objective analysis model is presented and used for an analysis of the pressure gradient terms of the horizontal momentum equations. This vertical coordinate grades from the sigma coordinate to a pressure coordinate at some reference pressure level in the middle troposphere and thus eliminates hydrostatic truncation error from this level upward. For the lower troposphere, the nonlinear vertical coordinate is used to show that the truncation error for a horizontally homogeneous hydrostatic atmosphere with variable vertical temperature structure arises because of a faulty assumption in the transformation from pressure coordinates to the sigma coordinates. This error is eliminated through a "nonlocal formulation" for the pressure gradient terms that replaces the temperature with its lapse rate in the hypsometric equation. However, this solution is not incorporated into the variational constraints because of greatly increased complexity that would result in the Euler-Lagrange equations. We instead reduce the magnitudes of the individual terms of the pressure gradient terms approximately 30-fold by projecting the pressure gradient onto "equivalent pressure surfaces". This solution leaves the hydrostatic residual unchanged from the direct two-term calculation.

5. Day-Night Variation in Operationally-Retrieved TOVS Temperature Biases.

The variational assimilation model offers a means for blending satellite and conventional soundings in a way which preserves the information content of both data sources. However, the model requires input data which are as bias-free as possible and about which the error characteristics are known. Because previous studies of TOVS biases were insufficient for our purposes, we recalculated the biases. Tiros-N soundings and rawinsonde data for the period 26 March through 11 April 1979 were acquired. Layer mean virtual temperatures, derived from rawinsonde thicknesses each 12 hr were objectively analyzed on a 21 x 21 grid (260 km grid spacing at $45^{o}N$) covering most of North America. Biases were estimated by calculating the difference between satellite-estimated layer mean virtual temperatures and rawinsonde values interpolated in both time and space from the analyses to the satellite data. Biases were comparable to previous studies, however, biases for day and night soundings were found to be statistically different (95% confidence) at most levels for clear and partly cloudy soundings, and at several levels for cloudy soundings. Day-night differences are particularly large for clear soundings. In the mid-troposphere, nighttime soundings have little bias, while daytime soundings have a large cold bias. This day-night difference in TOVS biases has not previously been reported in the literature.

6. The Impact of Data Boundaries upon a Successive Corrections Objective Analysis of Limited-Area Datasets.

Successive corrections objective analysis techniques frequently are used to array data from limited areas without consideration of how the absence of data beyond the boundaries of the network impacts the analysis in the interior of the grid. This problem of data boundaries is studied theoretically by extending the response theory for the Barnes objective analysis method to include boundary effects. The results from the theoretical studies are verified with objective analysis of analytical data. Several important points regarding the objective analysis of limited-area datasets are revealed through this study.

(i) Data boundaries impact the objective analysis by reducing the amplitudes of long waves and shifting the phases of short waves. Further, in comparison with the infinite plane response, it is found that truncation of the influence area by limited-area datasets and/or the phase shift of the original wave during the first pass amplified some of the resolvable short waves upon successive corrections to that first pass analysis.

(ii) The distance that boundary effects intrude into the interior of the grid is inversely related to the weight function shape parameter. Attempts to reduce boundary impacts by producing a smooth analysis actually draw boundary effects farther into the interior of the network.

(iii) When analytical tests were performed with realistic values for the weight function shape parameters, such as the GEMPAK default criteria, it was found that boundary effects intruded into the interior of the analysis domain a distance equal to the average separation between observations. This does not pose a problem for the analysis of large datasets because several rows and columns of the grid can be discarded after the analysis. However, this option may not be possible for the analysis of limited-area datasets because there may not be enough observations.

The results show that, in the analysis of limited-area datasets, the analyst should be prepared to accept that most (probably all) analyses will suffer from the impacts of the boundaries of the data field.

7. On the Notion of Varying Influence Radii for a Successive Corrections Objective Analysis.

This study examines the NOTION that the best successive corrections objective analysis is obtained by first analyzing for the long wavelengths and then building in short wavelengths by successively reducing the influence radius for each correction pass. It is shown that the best objective analysis, as measured by filter fidelity (how well the objective analysis restores desired wavelengths and removes undesired wavelengths), is realized for the Barnes method if the effective influence area used for the correction pass is equal to the effective influence radius used for the first pass. The improvements are relatively small, ranging from a few percent for long wavelengths to about ten percent for short but resolvable waves. However, increased simplicity and potentially great reductions in computer time needed to analyze large masses of meteorological data advance these modest gains. Therefore, rather than attempt to build desired detail into an analysis, the analyst should determine the detail permitted by the data quality and distribution and analyze directly for these motion scales.

8. Application of Satellite Data to the Variational Analysis of the Three-Dimensional Wind Field

Two of the Euler-Lagrange equations, the integrated continuity equation and the velocity adjustment potential equation have been extracted from the general variational assimilation model and adapted for calculating mesoscale vertical velocities. This technique consists of a variational blending of vertical velocities obtained from the kinematic and adiabatic methods. The relative weights assigned to these methods are deduced from GOES infrared digital cloud top temperature data and from GOES visible brightness data. The kinematic method receives greatest weight near the surface and in deep cloudy areas. The adiabatic method is assigned the greatest weight in layers not strongly influenced by diabatic heating: clear regions and levels above the tropopause.

This algorithm reduces errors introduced by the kinematic method used alone. These errors arise from the wind measurements, data interpolation, and the finite difference approximation to the differential equation. Combining the kinematic method with the adiabatic approach provides an independent estimate of vertical velocity and removes sole reliance upon the kinematic method.

9. Estimating the Accuracies of Kinematic Vertical Velocity Methods by Pattern Comparisons with Precipitation.

The variational assimilation model requires as input the observations of the vertical velocity. Since the vertical velocity is not directly observed, it may be assumed to be zero initially or may be calculated from some equation that relates the vertical velocity to other variables that are observed. We desire that the vertical velocity be determined as accurately as possible and that it be determined from some algorithm that is independent from the dynamic constraints so as to assure some degree of independence among the initial variables.

There are a number of short-cut methods for calculating the vertical velocity. None of these methods give the true vertical velocity, however one may give more accurate estimates than the others. In this study we examine the relative merits of two kinematic methods for estimating vertical velocity. One makes use of the integrated continuity equation to estimate vertical velocity by the O'Brien (1970) method and the other is a simplification of the Petterssen (1956) development equation.

The first method (divergence method) is well known to be sensitive to small errors in the wind observations. It also has been found to locate centers of vertical velocity approximately midway between observation sites when the wind field is obtained by univariate objective analysis. The second method (vorticity method) requires specification of the vertical velocity profile.

Simplifying assumptions also restrict this method to the analysis of wind fields surrounding weakly baroclinic cyclones some of which are accompanied by widespread stratiform and embedded convective precipitation. These systems are common during summer and occur occasionally during fall and winter.

We are calculating vertical velocity fields with the two methods. The results will be compared with precipitation patterns available from hourly radar summary data. The results of these comparisons will determine whether one or the other or some combination of the two methods produces the best distribution of vertical velocity.

References

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

Petterssen, Sverre, 1956: Weather Analysis and Forecasting, Vol. 1, McGraw-Hill, New York, p. 320.

Chapter II

A Variational Assimilation Method for the Diagnosis of Cyclone Systems. Part I: Development of the Basic Model

Gary L. Achtemeier, Harry T. Ochs III,

and Julia Chen

Climate and Meteorology Section

Illinois State Water Survey

Champaign, IL 61820

A Variational Assimilation Method for the Diagnosis of Cyclone Systems. Part I: Development of the Basic Model

by

Gary L. Achtemeier, Harry T. Ochs III, and Julia Chen

Illinois State Water Survey

Champaign, IL 61820

ABSTRACT

This paper outlines a theory for a variational objective analysis for the diagnosis of cyclone systems. Gridded fields of data from different type, quality, location and measurement source are weighted according to measurement accuracy and merged using a least squares criteria so that the two nonlinear horizontal momentum equations, the hydrostatic equation, and an integrated continuity equation are satisfied. We use the variational method of undetermined multipliers to derive the Euler-Lagrange equations necessary to create a dynamically consistent hybrid data set. A quasi-geostrophic solution sequence for these equations is described.

Other features of the variational diagnostic model include a hybrid nonlinear terrain-following vertical coordinate that eliminates truncation error in the pressure gradient terms of the horizontal momentum equations and easily accommodates TIROS-N mean layer temperatures in the middle and upper troposphere. A projection of the pressure gradient onto equivalent pressure surfaces removes most of the impacts of the lower coordinate surface on the variational

adjustment. In addition, the local tendencies of the horizontal velocity components are reformulated to better diagnose these hypersensitive quantities.

An application of the variational diagnostic model to the study of a dissipating short wave appears in the following companion paper.

1. Introduction

The proliferation of new methods to measure the state of the atmosphere through remote sensing and advanced immersion techniques has led to the need to merge data collected from these new measurement systems with data collected routinely by traditional methods. These data include a number of different variables that are diverse in measurement accuracy and in observation location. Therefore, it is desirable to merge the data so that the hybrid product will contain the best approximation to the significant meteorological variables. This paper reports on the development of a diagnostic method to merge data of different source, type, quality and location in a dynamically consistent manner. We further develop and improve upon a diagnostic variational objective analysis technique developed by Achtemeier (1975) for the study of cyclone scale weather systems. A companion paper (Achtemeier, et al., 1986) deals with an evaluation of the method using a case study.

In most cases, the creation of a dynamically consistent hybrid set of data is accomplished through some form of data assimilation coupled with an initialization for a numerical model. Observational

data is blended with model forecast fields through an interpolation technique (Cressman. 1959; Gandin, 1963; Schlatter; 1975) in which the latter, used as a first guess, is updated with the observations. These methods are multivariate; wind observations are used in the interpolative analysis of the height and temperature fields and vice versa. Then some initialization procedure such as dynamic initialization (e.g., Miyakoda and Moyer, 1968; Nitta and Hovermale, 1969) or normal mode initialization (e.g., Baer and Tribbia, 1977; Machenhaur. 1977) brings the hybrid data set into consistency with a numerical model. Highly sophisticated dynamically consistent data assimilation schemes such as those described by McPherson, et al., (1979), Bengtsson, et al., (1982), Ghil, et al. (1979), Temperton (1984), and many others produce accurate representations of the state of the synoptic scale atmosphere. These hybrid data sets, though modified to be consistent with the scales of motion permitted by the models, may be useful for diagnostic studies as well as for initial states for forecast models.

The approach taken here is to develop a purely diagnostic method for the dynamic merger of diverse data. This is not to say that the variational method is superior to the data assimilation methods used with prognostic models, only that it is, by design, independent of numerical models and is therefore different from the other methods. It serves a different purpose, namely the diagnosis of the morphology and energetics of cyclone systems, whereas many studies of numerical forecasts with mixed data sets are focused upon improved forecast skill. However. the variational model may eventually be of value in

comparisons with existing data assimilations and may provide insights that could lead to improvements in both methods.

The goal of our research is a variational data assimilation method that incorporates as dynamical constraints, the primitive equations for a moist, convectively unstable atmosphere and the radiative transfer equation. Variables to be adjusted include the three-dimensional vector wind, height, temperature, and moisture from rawinsonde data, and cloud-wind vectors, moisture, and radiance from satellite data. This presents a formidable mathematical problem. In order to facilitate thorough analysis of each of the model components, we defined four variational models that divide the problem naturally according to increasing complexity. The first variational model (MODEL I) contains the two nonlinear horizontal momentum equations, the integrated continuity equation, and the hydrostatic equation. Problems associated with an internally consistent finite difference method, a nonlinear hybrid terrain-following vertical coordinate, formulations for the pressure gradient terms, formulations for the velocity tendency terms and the development of a convergent solution sequence are addressed with MODEL I and are the subject of this paper.

MODEL II contains MODEL I plus the energy equation for a dry adiabatic atmosphere. The introduction of this additional constraint violates the requirement that the number of subsidiary conditions (dynamic constraints) must be at least one less than the number of dependent variables (Courant, 1936). Inclusion of the same number of constraints as dependent variables overdetermines the problem and a

solution is not guaranteed. Therefore, we must develop a scheme to circumvent this problem or else the dynamically adjusted meteorological variables will not satisfy the closed set of primitive equations. MODEL III contains MODEL II plus an additional moisture variable and equation to describe moist adiabatic processes. MODEL IV includes MODEL III plus radiance as a dependent variable and the radiative transfer equation as a constraint.

The next section presents the philosophy of a variational diagnostic data assimilation method. Section 3 presents the dynamic equations in the forms they enter the variational formalism as constraints. The variational equations are derived in Section 4. Details concerning the grid mesh, boundary conditions, and convergence of the equations are found in Section 5. Section 6 summarizes the model.

2. A Variational Approach to Diagnostic Data Assimilation

A good diagnostic analysis includes appropriate mathematical algorithms applied to accurately gridded fields of meteorological variables. This diagnostic objective analysis is an adaptation of Sasaki's (1958) method of variational analysis. Data from different measurement systems are weighted according to measurement accuracies and are blended using a least squares method into a hybrid data set that satisfies a set of subsidiary conditions. Sasaki (1970a) has presented two variational formulations for the solution of the data assimilation problem. His "weak constraint" formalism requires only a

partial satisfaction of the subsidiary conditions through coefficients determined by the analyst. The subsidiary conditions are satisfied exactly through the "strong constraint" method. Ikawa (1984) has shown that the weak constraint algorithm converges to the strong constraint formalism as the coefficients become large.

This study makes use of the method of undetermined multipliers (strong constraint formalism). The constraints are the nonlinear horizontal momentum equations (products of a variational principle (Wang, 1984)), the hydrostatic equation and an integrated form of the continuity equation. The adjustments are carried out on fields of meteorological variables obtained through univariate objective interpolation. This kind of variational formulation has been criticized by Williamson and Daley (1983) on the grounds that the adjustments to the dynamic state are carried out from gridded fields rather than from the observations. Alternatively. observation statistics for different measurements of the same variable can be carried in the analyzed fields, perhaps as proposed by Baker (1983). Another implication of the method of undetermined multipliers is the extreme complexity of the variational equations which stimulates a need for simplier methods to create hybrid, dynamically balanced data sets. Wahba and Wendelberger (1980) have shown that multivariate statistical objective analysis and variational analysis are interchangeable for linear constraints. Our variational method permits nonlinear constraints, allows for the physical interpretation of the adjustments and provides mutual adjustment between the mass and wind fields (Temperton, 1984).

Accurately gridded meteorological variables are a requirement for any good diagnostic analysis. There are also quantities which, because of poor instrument accuracy or insufficient sampling frequency, cannot be measured directly and must be inferred through functions of other measured variables; in our case, they are products of the variational blending process. Among these are hypersensitive variables that are sensitive to small changes in the other variables, such as vertical velocity and the local tendencies of the horizontal velocity components which appear explicitly in the variational formulation. The variational diagnostic model must also produce accurate fields of these hypersensitive variables.

Krishnamurti (1968) calculated diagnostic vertical velocities through a 12-forcing function balance omega equation. More recently, Smith and Lin (1978) preferred vertical velocities diagnosed from the O'Brien (1970) variational method. Our variational model calculates vertical velocity from a generalized form of the kinematic method for which the O'Brien method can be shown to be a special case.

The local tendency terms of the horizontal velocity components are particularly difficult to determine with any accuracy because of the coarse sampling frequency of operational data collection networks. Local tendencies can be incorporated into the variational analysis by fixing them and assuming that the generated error will not appreciably contaminate the solution. But this ignores the fact that the tendency terms are of the same order of magnitude as the advection terms and that generated error undoubtedly will contaminate the solution.

especially the error sensitive divergence calculations. Sasaki (1970b), Sasaki and Lewis (1970), and Lewis and Grayson (1972) have used a "time-wise localized" method which physically is not a time adjustment, but rather a space filter designed to adjust variables in space at a particular time such that the local tendency is minimized with partial constraint satisfaction. Achtemeier (1975) included local rates of change in a primitive equation variational model through a subsidiary variational formulation based upon O'Brien's (1970) divergence adjustment method. This method was considered a failure after an extensive analysis (Achtemeier, 1978) found unrealistically large velocity component tendencies where actual velocity changes over a 12-hr period were small.

More recently, Lewis (1980, 1982) has examined the problem of time consistency from a Lagrangian approach through the application of Thompson's (1969) variational method. By requiring conservation of quasi-geostrophic potential vorticity, Lewis et al. (1983) combined rawinsonde data with VAS height data taken 2.5 hr later and found vertical velocity fields that compared favorably with space-observed cloud fields and surface weather reports. These studies and the results from Bloom's (1983) mesoscale analysis imply that variational methods can be used with some success in the direct determination of tendency variables, at least for observation frequencies on the order of 3-hr.

3. The Formulations for the Dynamic Constraints

The dynamic constraints, M_i, for MODEL I are the two nonlinear horizontal momentum equations, the hydrostatic equation, and an integrated continuity equation. These constraints (see equations (1)-(4)) have been transformed into the Lambert conformal map image projection and into a hybrid nonlinear sigma vertical coordinate. The equations have been nondimensionalized and presented in powers of the Rossby number. An additional transformation removes most of the impacts of unlevel terrain. Furthermore, thermodynamic variables are partitioned into mean and perturbation variables. The variational adjustments are carried out only on the scale of the meteorological perturbations. The equations are listed below. Discussions of the various transformations that render the equations into the forms shown and definitions of nonconventional symbols follow.

The four dynamic constraints as they appear in the diagnostic variational model are:

$$m_1 = R_o \left[\xi_u + m\,(u-c_x)\frac{\partial u}{\partial x} + m\,(v-c_y)\frac{\partial u}{\partial y} + R_o\,\dot{\sigma}\,\frac{\partial u}{\partial \sigma}\right] - (1 + R_1 C)v + (1 + R_1 K)\left[\frac{\partial \phi}{\partial x} + \eta_x\right] + f_u \tag{1}$$

$$m_2 = R_o \left[\xi_v + m\,(u-c_x)\frac{\partial v}{\partial x} + m\,(v-c_y)\frac{\partial v}{\partial y} + R_o\,\dot{\sigma}\,\frac{\partial v}{\partial \sigma}\right] + (1 + R_1 C)\,u + (1 + R_1 K)\left[\frac{\partial \phi}{\partial y} + \eta_y\right] + f_v \tag{2}$$

$$m_3 = \int \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) d\sigma + (\dot{\sigma} - \dot{\sigma}_o) + \int F\, d\sigma \tag{3}$$

where $$F = \frac{Lh}{H\ell}\, q_2\, \omega_s + R_1 K \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) - R_1 \left(u\,\frac{\partial K}{\partial x} + v\,\frac{\partial K}{\partial y}\right)$$

$$m_4 = \frac{\partial \phi}{\partial \sigma} + \overline{T}\,\frac{\partial\, \ell n\, P}{\partial \sigma} + \beta = 0\ . \tag{4}$$

a) A Hybrid Nonlinear Sigma Vertical Coordinate

The vertical coordinate used for this analysis model is an extension of the terrain-following coordinate system of Phillips (1957). Although this coordinate eliminates problems with the lower boundary encountered with other vertical coordinates, considerable error can be introduced into the pressure gradient terms of the momentum equations upon transformation into the Phillips coordinate. The pressure gradient terms transform into two large and compensating terms where there is steep sloping terrain. Pressure derivatives taken along the sloping sigma surface can contain a hydrostatic component that does not cancel among the two terms. Furthermore, the variational formalism will separate the pressure gradient terms and combine the large uncompensated terms with terms from the other equations. The large nonmeteorological contribution by these terms can also cause significant truncation error in the final solution unless methods are developed to remove them.

We have eliminated the above problems from the middle troposphere upward and have reduced them in the lower troposphere through the introduction of a hybrid nonlinear sigma vertical coordinate that blends from a terrain-following coordinate to a pressure coordinate at a reference pressure level p^* in the middle troposphere. For a complete description of this vertical coordinate, refer to Achtemeier

and Ochs (1986). All horizontal variations caused by the lower coordinate surface are confined to levels below p^*. The smooth transition from the sigma to the pressure coordinate is accomplished by fitting two curves which are piecewise continuous through the second derivatives. The curve from the top of the domain p_u to p^* is linear in pressure. The relationship between sigma and pressure is cubic between p^* and the surface pressure p_s. The equation for the hybrid vertical coordinate is

$$\sigma = \beta\,(p-p^*)^3 + \sigma^* \frac{(p-p_u)}{(p^*-p_u)} \tag{5}$$

$$\beta = \left[1- \sigma^* \left(\frac{p_s-p_u}{p^*-p_u}\right)\right] (p_s-p^*)^{-3} \tag{6}$$

The first term of (5) is zero where $p \leq p^*$.

The hybrid nonlinear vertical coordinate permits the dynamical equations to appear in their simplest forms on the pressure surfaces at and above the reference pressure level. Coding to omit terms that are zero for coordinate surfaces that are surfaces of constant pressure can result in a substantial reduction of computational overhead for this variational model. The tradeoff is that the complexity of the equations below p^* is increased over the complexity of the equations written for the linear sigma coordinate. However, the magnitudes of these additional terms become small in the sigma levels above the lower coordinate surface.

b) Expansion of the Local Tendencies

Local changes in the horizontal velocity components are caused by a combination of translation of existing disturbances and development. In partitioning the tendencies, we note that, for example, the local change in the u-component of the wind caused by a moving weather system is

$$\frac{\partial u}{\partial t} = - c \cdot \nabla u + \frac{du}{dt} \tag{7}$$

where c is the velocity of an advective or steering current (Fjortoft,1952), usually a smoothed middle tropospheric wind. Let $u=u_0+u'$ where u_0 is the u-component of the steady state part of the circulaton and u' is the u-component arising from development. Then,

$$\frac{\partial u}{\partial t} = - c \cdot \nabla u_o + (\frac{du'}{dt} - c \cdot \nabla u') \tag{8}$$

The first term of (8) is the local change in u caused by translation of the steady state part of the weather disturbance. The second term contains the local change of u from development. Note that the vertical advection of u is considered part of development.

The use of the advective current throughout the troposphere is valid because most synoptic systems tend to maintain vertical structure. Any changes in vertical structure are assumed to be the result of development. The variational formalism requires that the adjustments be carried out on the total velocity components. Therefore, we represent the local tendency of u by (7). The total derivative, an approximate developmental component, is defined as a new dependent variable. ξ_u=du/dt (ξ_v=dv/dt). With these definitions

substituted into (7), the local tendencies of u and v are replaced by the forms that appear in the constraints (1) and (2).

c) Scale Analysis

The equations of constraint are nondimensionalized following the methodology of Charney (1948), Haltiner (1971) and others. Our use of quasi-geostrophic scale theory revealed the need to transform the pressure gradient terms of the horizontal momentum equations in order to decrease the impacts of steeply sloping terrain upon several terms. Partitioning the thermodynamic variables into reference and perturbation variables is necessary for the variational adjustments to be with respect to meteorological scales of motion. Further, the expression of the terms of the equations in powers of the Rossby number permits a solution sequence in which the higher order nonlinear terms are gathered into forcing functions. As part of the expansion, the mapscale factor m and the Coriolis parameter f expand into $m=1+R_1K$ and $f=1+R_1C$ where the arrays K and C are of order one and $R_1=0.1$. There remains a set of linear algebraic and partial differential equations that can be solved easily by conventional techniques. However this method is not expected to yield a solution for the tropics nor for the mesoscale or wherever the Rossby number is greater than one. Finally, the scale analysis for our extratropical domain confirmed that map projection terms in the horizontal momentum equations may be dropped.

d) Transformation of the Integrated Continuity Equation

The mass continuity equation in generalized coordinates (Shuman and Hovermale, 1968) is

$$\frac{\partial}{\partial t}\left(\frac{\partial p}{\partial \sigma}\right)+\frac{\partial}{\partial \sigma}\left(\dot{\sigma}\frac{\partial p}{\partial \sigma}\right)+m\left[\frac{\partial}{\partial x}\left(u\frac{\partial p}{\partial \sigma}\right)+\frac{\partial}{\partial y}\left(v\frac{\partial p}{\partial \sigma}\right)\right]-\frac{\partial p}{\partial \sigma}\left(u\frac{\partial m}{\partial x}+v\frac{\partial m}{\partial y}\right)=0 \quad (9)$$

The material derivative in the Lambert map projection is

$$\frac{d}{dt}=\frac{\partial}{\partial t}+mu\frac{\partial}{\partial x}+mv\frac{\partial}{\partial y}+\dot{\sigma}\frac{\partial}{\partial \sigma} \quad (10)$$

Upon expanding the map scale factor m, (9) becomes

$$\nabla_3\cdot\vec{V}+R_1K^2\left(\frac{\partial u/K}{\partial x}+\frac{\partial v/K}{\partial y}\right)+\frac{d}{dt}\left(\ell n\frac{\partial p}{\partial \sigma}\right)=0 \quad (11)$$

The last term of (11) is determined from the equation for the hybrid vertical coordinate (5). Further, given (6) and the following definitions,

$$\alpha = \sigma^*/(p^*-p_u), \quad (12)$$

$$J = 3\beta\,(p-p^*)^3+\alpha\,(p-p^*), \quad (13)$$

it can be shown that

$$\frac{d}{dt}\left(\ell n\frac{\partial p}{\partial \sigma}\right)=q_1\dot{\sigma}+q_2\omega_s \quad (14)$$

where $q_1=\dfrac{-2[J-\alpha\,(p-p^*)]}{J^2}$

$$q_2=\frac{(p-p^*)^3}{(p_s-p^*)^4}\,\frac{J_s\,[J\;2\alpha\,(p-p^*)]}{J^2}$$

J_s is obtained by substituting p_s for p in (13). Including these modifications leads to the following form for the continuity equation.

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial \dot{\sigma}}{\partial \sigma}+q_1\dot{\sigma}+F=0 \qquad \text{where } p>p^* \quad (15)$$

where $F=q_2\omega_s+R_1K^2\left(\dfrac{\partial u/K}{\partial x}+\dfrac{\partial v/K}{\partial y}\right)$

Both q_1 and q_2 are zero where $p \triangleq p^*$. When solved for $\dot{\sigma}$, (15) becomes the following non-homogeneous linear partial differential equation.

$$\dot{\sigma} = e^{-\int q_1 \delta\sigma}\, \dot{\sigma}_o - e^{-\int q_1 \delta\sigma} \int [(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}) + F]\, e^{\int q_1 \delta\sigma} d\sigma \qquad (16)$$

The domain of integration is arbitrary. For MODEL I, it is the depth of the atmosphere contained within the grid. In the more complex versions, the domain of integration will be sigma layers. In order to simplify the solution of the Euler-Lagrange equations we set q_1=0. This assumption removes the dependence of the integrated divergence on the variable pressure thickness of the sigma levels. Therefore, divergences in the levels near the surface over elevated terrain receive proportionally greater weight in the vertical velocity adjustment. We make this assumption because vertical velocities over elevated terrain are not important to this phase of the development of MODEL I. With q_1=0, (16) simplifies to the form in (4).

e) Pressure Gradient Force for the Diagnostic Variational Model.

The problem of hydrostatic inconsistency in the pressure gradient force has already been eliminated in the middle and upper troposphere through the hybrid nonlinear vertical coordinate. We reformulate the pressure gradient force at levels below p^* to reduce the magnitude

of the individual terms because the variational formalisms will separate the pressure gradient terms and combine the large uncompensated terms with terms from other equations. The large nonmeteorological contribution by these terms can cause significant errors in the final solution unless methods are developed to remove them. Achtemeier and Ochs (1986) present a thorough analysis of the hydrostatic equation and the pressure gradient force in the hybrid nonlinear vertical coordinate. Those results are summarized here.

We reduce the magnitude of the individual terms by projecting the pressure gradient onto equivalent pressure surfaces. The terminology "equivalent pressure surfaces" is used to avoid confusion with methods that calculate the pressure gradient on surfaces of constant pressure and then interpolate the results to sigma coordinate surfaces (Kurihara, 1968).

We first remove a hydrostatic component from both terms by partitioning the pressure gradient to cancel most of the orographic part. The separation is not complete because the mean layer temperature is not partitioned. The geopotential and pressure are expressed as an orograhic part plus a remainder, $\phi_w = \phi_T + \phi$ and $p_w = p_T + p$. Here the subscript w implies the whole or unpartitioned variable. Substitution into the hydrostatic equation gives

$$\frac{\partial \phi}{\partial \sigma} + \gamma R \overline{T} + \beta = 0 \qquad (17)$$

where

$$\gamma = \frac{\partial \ln (p)}{\partial \sigma} \qquad (18)$$

and

$$\beta = \left(\frac{\overline{p}_w}{\overline{p}} - 1\right) \frac{\partial \phi_w}{\partial \sigma} + \frac{\partial \phi_T}{\partial \sigma} + \frac{\overline{RT}_w}{\overline{p}} \frac{\partial p_T}{\partial \sigma} \qquad (19)$$

This equation describes the hydrostatic relationship between meteorological perturbations. The perturbations are subject to the variational adjustments. Most of the orographic component is located in β. Eq. 17 can be solved accurately if the layer average pressures are equal to the average of the arithmetic mean plus twice the geometric mean. The orographic variables are found by setting β=0 and defining p as equivalent pressure surfaces.

Having derived the relevant partitioned variables, the pressure gradient terms are easily transformed, e.g.,

$$PGX = \frac{\partial \phi}{\partial x} + \eta_x$$

where (20)

$$\eta_x = \frac{\partial \phi_T}{\partial x} + \overline{RT}^x \frac{\partial \ln (p_w)}{\partial x}$$

Fig. 1 shows the height of the lower coordinate surface for a grid to be used for the diagnostic variational analysis of data collected at 1200 GMT 10 April 1979. The heights on the unpartitioned terrain-following coordinate vary from 0-1800 m approximately (Fig. 1a) and show the steep gradients that surround a smoothed high elevation area over the western U.S. The heights remaining after the removal of the hydrostatic component that arises from variations in the elevation of the lower coordinate surface are shown in Fig. 1b. Calculations show that the projection onto equivalent pressure surfaces reduces the magnitudes of the these variations by about

30-fold. The equivalent 1000 mb heights resemble the actual 1000 mb heights (Fig. 1c) with the exception that the heights of the low center over the West are approximately 60 m higher in Fig. 1b. This residual orographic effect is retained through the unpartitioned mean layer temperatures.

Finally, we partitioned the thermodynamic variables into reference and meteorological perturbation atmospheres. Once determined, the reference atmosphere was not altered. However, this required that it be in hydrostatic balance initially. Removal of the reference atmosphere does not alter the form of the constraints from that given in (1)-(4).

4. Euler-Lagrange Equations for the Diagnostic Model

The previous section has presented the equations of constraint in the forms that they will appear in the diagnostic variational model. These equations are written in finite differences according to the grid structure of this model and the Euler-Lagrange equations derived from them. We derived two finite difference variational models, one with the dynamic equations written in uncentered differences on a nonstaggered grid and the other formulated with centered differences on a staggered grid. We sought a final difference formulation for the Euler-Lagrange equations that is symmetric about the central grid point. The centered difference formulation on a staggered grid proved most suitable from this standpoint.

Following Shuman and Hovermale (1968) and Anthes and Warner

(1978), we have defined the horizontal finite difference operators and the finite averaging operators as

$$\alpha_x \equiv (\alpha_{i+1/2,j} - \alpha_{i-1/2,j}) / \Delta x$$

$$\alpha_y \equiv (\alpha_{i,j+1/2} - \alpha_{i,j-1/2}) / \Delta y$$

$$\overline{\alpha}^x \equiv (\alpha_{i+1/2,j} + \alpha_{i-1/2,j}) / 2 \qquad (21)$$

$$\overline{\alpha}^y \equiv (\alpha_{i,j+1/2} + \alpha_{i,j-1/2}) / 2$$

The i is the east-west index, the j is the north-south index as measured at the grid origin which is located at the lower left corner of the grid. In addition, the vertical differences and averages are defined by

$$\alpha_\sigma \equiv (\alpha_{k+1/2} - \alpha_{k-1/2}) / \Delta\sigma \qquad (22)$$

$$\overline{\alpha}^\sigma \equiv (\alpha_{k+1/2} + \alpha_{k-1/2}) / 2$$

Figure 2 shows the staggered grid developed for this model. The geopotential ϕ is defined at the grid intersections, v is located at the top and bottom and u is located at the sides of the grid square. The divergence D is found at the center of the grid. The layer mean temperatures T are defined at one half grid length above and below the grid intersections and the vertical velocity $\dot{\sigma}$ is located one half grid space above and below the divergence. Mesinger and Arakawa (1976) have shown that phase speed and dispersion properties of this

staggered grid make it inferior relative to other grid configurations for numerical prediction. However, the grid with v located on the top and bottom and u located on the sides of the grid box is well suited for the solution sequence developed for the Euler-Lagrange equations later in this section. Other variables used in the variational analysis are collocated with the variables in Fig. 2 as follows: $\mathcal{E}_v$ and λ_1 at v. $\mathcal{E}_u$ and λ_2 at u, λ_3 at D, and λ_4 at $\bar{T}$.

The finite difference equations for the horizontal momentum equations written for the staggered grid are

$$M_1 = R_o[\bar{\varepsilon}_u^{xy} + \bar{m}^x(\overline{u - c_x})^{xy}\,\bar{u}_x^y + \bar{m}^x(v - c_y)\,\bar{u}_y^x + R_o\,\bar{\dot{\sigma}}^{y\sigma}\,\bar{u}^{xy\sigma}] - (1+R_1\bar{C}^x)v + (1+R_1\bar{K}^x)\,[\phi_x + \eta_x] + f_u = 0 \tag{23}$$

$$M_2 = R_o\,[\bar{\varepsilon}_v^{xy} + \bar{m}^y(u - c_x)\,\bar{v}_x^y + \bar{m}^y(\overline{v - c_y})^{xy}\,\bar{v}_y^x + R_o\,\bar{\dot{\sigma}}^{x\sigma}\,\bar{v}_\sigma^{xy\sigma}] + (1+R_1\bar{C}^y)u + (1+R_1\bar{K}^y)\,[\phi_y + \eta_y] + f_v = 0 \tag{24}$$

The analogs for the continuity and hydrostatic equations are

$$M_3 = \int (u_x + v_y)\,d\sigma + (\dot{\sigma} - \dot{\sigma}_o) + \int [\frac{Lh}{H\ell}\,q_2\omega_s + R_1\bar{K}^{xy}\,(u_x + v_y) - R_1(\bar{u}^x\,\bar{K}_x^y + \bar{v}^y\,\bar{K}_y^x)]\,d\sigma = 0 \tag{25}$$

$$M_4 = \phi_\sigma + \bar{T}^\sigma\,(\ell n\ p)_\sigma + \beta = 0 \tag{26}$$

The four dynamic constraints are referenced, respectively, to the following locations: M_1 at v, M_2 at u, M_3 at D, and M_4 at $\bar{T}$ on Fig. 2.

The variational analysis melds satellite data with conventional data at the second stage of a two-stage objective analysis. All data are gridded independently in the first stage and are combined in the second stage. The gridded observations to be modified are meshed with the dynamic constraints through Sasaki's (1970a) variational formulation. The finite difference analog of the adjustment functional is

$$\tilde{F} = \Delta x \Delta y \sum_i \sum_j a_i b_j I_{ij} \tag{27}$$

The integrand $I_{i,j}$ is

$$\begin{aligned} I = {} & \pi_1 (u-u^o)^2 + \pi_1 (v-v^o)^2 + \pi_2 (\dot{\sigma}-\dot{\sigma}^o)^2 + \pi_3 (\phi-\phi^o)^2 \\ & + \pi_4 (\overline{T}-\overline{T}^o)^2 + \pi_5 (\phi_x - \phi_x{}^o)^2 + \pi_5 (\phi_y - \phi_y{}^o)^2 + \pi_6 (\phi_\sigma - \phi_\sigma{}^o)^2 \\ & + \pi_7 (\xi_u - \xi_u{}^o)^2 + \pi_7 (\xi_v - \xi_v{}^o)^2 + 2\lambda_1 M_1 + 2 \lambda_2 M_2 + 2 \lambda_3 M_3 \\ & + 2 \lambda_4 M_4 \end{aligned} \tag{28}$$

The weights π_i, i=1,7 are Gauss' precision moduli (Whittaker and Robinson, 1926). The gridded observations (u^o, v^o, $\dot{\sigma}^o$, ϕ^o, T^o, $\mathcal{E}_u^o$, $\mathcal{E}_v^o$) to be adjusted enter in a least squares formulation and receive precision modulus weights according to their relative observation accuracies. The strong constraints to be satisfied exactly are introduced through the Lagrangian multipliers λ_i, i=1,4.

Objectively modified meteorological variables are determined by requiring the first variation on F to vanish. A necessary condition for the existence of a stationary set is that the functions are

determined from the domain of admissible functions as solutions of the Euler-Lagrange equations. The variation is to be carried out at every point (r,s) within the grid. Thus, setting the weights $a_i=b_j=1$ and differentiating the integrand (28) with respect to the arbitrary variable ($\alpha_{r,s}$), the Euler-Lagrange operator in finite differences is

$$\frac{\partial I_{i,j}}{\partial \alpha_{r,s}} = I_{\alpha_{i,j}} \frac{\partial \alpha_{i,j}}{\partial \alpha_{r,s}} = I_{\alpha_{i,j}} \delta_r^i \delta_s^j = 0 \tag{29}$$

The Kronecker delta functions δ_r^i, δ_s^j, equal 1 where r=i or s=j and are zero elsewhere. Each term in $I_{i,j}$ that contains an overbar term, e.g. $\bar{\alpha}_{r,s}^{x}$, produces a corresponding overbar term in the Euler-Lagrange equations when subjected to the operations specified by (29). It is convenient that the multiplicate overbar terms such as $\bar{\alpha}_{r,s}^{x,y}$ that appear in the nonlinear terms of the constraints be replaced by $\alpha_{r,s}$ so that fewer gridpoints are required to express these terms in the Euler-Lagrange equations.

The Euler-Lagrange equations for u, v, and $\dot{\sigma}$ result from the operations specified by (29). The equations are

$$\begin{aligned} &\Pi_1 (u-u^o) - (\int d\sigma)\, \lambda_{3x} + \lambda_2 (1+R_1\bar{C}^y) + R_o \{\bar{m}^y \bar{\lambda}_1^{xy} \bar{u}_x^x + \bar{m}^y \lambda_2 \bar{v}_x^y \\ &- [\bar{m}^{xy}\bar{\lambda}_1^y (\overline{u-c_x})^x]_x - [m\, \overline{\lambda_1 (v-c_y)^x}]_y - R_o (\dot{\bar{\sigma}}\, \bar{\lambda}_1^{xy\sigma})_\sigma\} \\ &- R_1 [(\lambda_3 \bar{K}^{xy})_x + \lambda_3 \bar{K}_x^{xy}] = 0 \end{aligned} \tag{30}$$

$$\begin{aligned} &\Pi_1 (v-v^o) - (\int d\sigma)\, \lambda_{3y} - \lambda_1 (1+R_1\bar{C}^x) + R_o \{\bar{m}^x \lambda_1 \bar{u}_y^x + \bar{m}^x \bar{\lambda}_2^{xy} \bar{v}_y^y \\ &- [m\, \overline{\lambda_2 (u-c_x)^y}]_x - [\bar{m}^{xy}\bar{\lambda}_2^x (\overline{v-c_y})^y]_y - R_o (\dot{\bar{\sigma}}\, \bar{\lambda}_2^{xy\sigma})_\sigma\} \\ &- R_1 [(\lambda_3 \bar{K}^{xy})_y + \lambda_3 \bar{K}_y^{xy}] = 0 \end{aligned} \tag{31}$$

$$\Pi_2 \left(\dot{\sigma} - \dot{\sigma}^o\right) + \bar{\lambda}_3^{\sigma} + R_o^2 \left[\bar{\lambda}_1^{y\sigma} \bar{u}_{\sigma}^{x} + \bar{\lambda}_2^{x\sigma} \bar{v}_{\sigma}^{y}\right] = 0 \tag{32}$$

The Euler-Lagrange equations for the thermodynamic variables ϕ and T are

$$\begin{aligned}\bar{\Pi}_5^{x} \left(\phi-\phi^o\right)_{xx} + \bar{\Pi}_5^{y} \left(\phi-\phi^o\right)_{yy} + \bar{\Pi}_6^{\sigma} \left(\phi-\phi^o\right)_{\sigma\sigma} + \Pi_{5x} \left(\bar{\phi}_x^{x}-\bar{\phi}_x^{o}\right) + \Pi_{5y} \left(\bar{\phi}_y^{y}-\bar{\phi}_y^{o}\right) \\ + \Pi_{6\sigma} \left(\bar{\phi}_{\sigma}^{\sigma}-\bar{\phi}_{\sigma}^{o}\right) - \Pi_3 \left(\phi-\phi^o\right) + \lambda_{1x} \left(1+R_1 K\right) + \lambda_{2y} \left(1+R_1 K\right) + \lambda_{4\sigma} = 0\end{aligned} \tag{33}$$

$$\Pi_4 \left(\bar{T}^{\sigma} - \bar{T}^{o}\right) + \gamma\lambda_4 = 0 \tag{34}$$

Similarly, the operations performed for $\mathcal{E}_u$ and $\mathcal{E}_v$ yield

$$\Pi_7 \left(\varepsilon_u-\varepsilon_u^o\right) + R_o \bar{\lambda}_1^{xy} = 0 \tag{35}$$

$$\Pi_7 \left(\varepsilon_v-\varepsilon_v^o\right) + R_o \bar{\lambda}_2^{xy} = 0 \tag{36}$$

Variation on the Lagrange multipliers restores the four original dynamic constraints (23)-(26).

Some of these Euler-Lagrange equations are complicated nonlinear partial differential equations for which solutions are difficult to obtain by direct conventional methods. An iterative method is proposed for the solution so that at the first cycle level, terms multiplied by R_o or R_1 are expressed with observed variables and are expressed by previously adjusted variables at subsequent cycles. At any particular solution cycle, these terms and the terms that are

determined by observed variables are specified and can be treated as forcing functions. Following this argument, the Euler-Lagrange equations for u, v, $\dot{\sigma}$, and ϕ become

$$\Pi_1 u - (\Delta\sigma)\,\lambda_{3x} + \lambda_2 + F_1 = 0 \tag{37}$$

$$\Pi_1 v - (\Delta\sigma)\,\lambda_{3y} - \lambda_1 + F_2 = 0 \tag{38}$$

$$\Pi_2\dot{\sigma} + \overline{\lambda_3}^{\sigma} + F_3 = 0 \tag{39}$$

$$\begin{aligned}\overline{\Pi_5}^{x}\,\phi_{xx} + \overline{\Pi_5}^{y}\,\phi_{yy} + \overline{\Pi_6}^{\sigma}\,\phi_{\sigma\sigma} + \Pi_{5x}\overline{\phi_x}^{x} + \Pi_{5y}\overline{\phi_y}^{y} + \Pi_{6\sigma}\overline{\phi_\sigma}^{\sigma} \\ - \Pi_3\phi + \lambda_{1x} + \lambda_{2y} + \lambda_{4\sigma} + F_4 = 0\end{aligned} \tag{40}$$

Similarly, the four dynamic constraints become

$$\phi_x - v + F_5 = 0 \tag{41}$$

$$\phi_y + u + F_6 = 0 \tag{42}$$

$$\int (u_x + v_y)\,d\sigma + \dot{\sigma} + \Delta\sigma\,F_7 = 0 \tag{43}$$

$$\phi_\sigma + \gamma\overline{T}^{\sigma} + \beta = 0 \tag{44}$$

Now these equations and (35) and (36) complete a set of eleven simple algebraic or linear partial differential equations. Variables may be easily eliminated to reduce the number of equations. Equations (37), (38), (41), and (42) formulated as vorticity expressions are combined to eliminate u and v.

$$\lambda_{1x} + \lambda_{2y} = (\Pi_1\phi_x)_x + (\Pi_1\phi_y)_y - F_{1y} + F_{2x} + (\Pi_1 F_5)_x + (\Pi_1 F_6)_x \tag{45}$$

Equation (43) is combined with (34) to eliminate T,

$$\lambda_{4\sigma} = \left(\frac{\Pi_4}{\gamma^2}\,\phi_\sigma\right)_\sigma + F_8 \tag{46}$$

where

$$F_8 = \left(\frac{\Pi_4}{\gamma}\,\overline{T^o}^{\sigma}\right)_\sigma + \left(\frac{\Pi_4 \beta}{\gamma^2}\right)_\sigma \tag{47}$$

Note that both (45) and (46) contain terms that obey the identity

$$(AB_z)_z = \overline{A}^z B_{zz} + A_z \overline{B}^z_z \tag{48}$$

Now λ_1, λ_2, and λ_4 can be eliminated in (45)-(48). This leaves a three-dimensional second-order partial differential equation with non-constant coefficients in ϕ ;

$$(\overline{\Pi}_1^x + \overline{\Pi}_5^x)\,\phi_{xx} + (\overline{\Pi}_1^y + \overline{\Pi}_5^y)\,\phi_{yy} + \left(\overline{\Pi}_6^\sigma + \left(\frac{\overline{\Pi}_4^\sigma}{\gamma^2}\right)\right)\phi_{\sigma\sigma} + (\Pi_1 + \Pi_5)_x\,\overline{\phi}_x^x$$

$$+ (\Pi_1 + \Pi_5)_y\,\overline{\phi}_y^y + \left(\Pi_6 + \frac{\Pi_4}{\gamma^2}\right)_\sigma \overline{\phi}_\sigma^\sigma - \Pi_3\phi + F_9 = 0 \tag{49}$$

where

$$F_9 = -F_{1y} + F_{2x} + (\Pi_1 F_5)_x + (\Pi_1 F_6)_y + F_4 + F_8 \tag{50}$$

All of the coefficients of the geopotential and its derivatives are functions of precision modulus weights and their derivatives. Note that the coefficients of the first three terms of (49) are sums of precision moduli. These are always positive, the three coefficients are always positive, and (49) is always elliptic over the analysis domain.

Variables are also eliminated to produce a diagnostic equation in λ_3. Dividing (37) and (38) by π_1 and reformulating as components of the divergence gives

$$u_x - \left(\frac{\int d\sigma}{\Pi_1} \lambda_{3x}\right)_x + \left[\frac{1}{\Pi_1} (\lambda_2 + F_1)\right]_x = 0 \tag{51}$$

$$v_y - \left(\frac{\int d\sigma}{\Pi_1} \lambda_{3y}\right)_y + \left[\frac{1}{\Pi_1} (-\lambda_1 + F_2)\right]_y = 0 \tag{52}$$

Then (51) and (52) combine into the divergence and after integration over some interval $d\sigma$, become

$$\begin{aligned} &- \int (u_x + v_y)\, d\sigma + \left(\lambda_{3x} \int \frac{\int d\sigma}{\Pi_1} d\sigma\right)_x + \left(\lambda_{3y} \int \frac{\int d\sigma}{\Pi_1} d\sigma\right)_y \\ &- \int \left\{\left[\frac{1}{\Pi_1} (\lambda_2 + F_1)\right]_x - \left[\frac{1}{\Pi_1} (-\lambda_1 + F_2)\right]_y\right\} d\sigma = 0 \end{aligned} \tag{53}$$

Now $\dot{\sigma}$ is eliminated from (47) and (51) so that

$$\int (u_x + v_y)\, d\sigma - \frac{\overline{\lambda_3}^{\sigma}}{\Pi_2} - \frac{F_3}{\Pi_2} + F_7 = 0 \tag{54}$$

A diagnostic equation in λ_3 is obtained upon elimination of the integrated divergence through combining (53) and (54):

$$\left(\lambda_{3x} \int \frac{\int d\sigma}{\Pi_1} d\sigma\right)_x + \left(\lambda_{3y} \int \frac{\int d\sigma}{\Pi_1} d\sigma\right)_y - \frac{\overline{\lambda_3}^{\sigma}}{\Pi_2} + F_{10} = 0 \tag{55}$$

$$F_{10} = - \int \left\{\left[\frac{1}{\Pi_1} (\lambda_2 + F_1)\right]_x + \left[\frac{1}{\Pi_1} (-\lambda_1 + F_2)\right]_y\right\} d\sigma - \frac{F_3}{\Pi_2} + F_7 \tag{56}$$

Since $\lambda_3 = \lambda_3(x,y)$, it follows that $\overline{\lambda_3}^{\sigma} = \lambda_3$. We also note that

several terms of (55) obey the identity (48). Therefore (55) transforms into the two-dimensional second-order elliptic partial differential equation with non-constant coefficients given by

$$\int \left(\frac{\overline{\int d\sigma}^{x}}{\Pi_1}\right) d\sigma \; \lambda_{3xx} + \int \left(\frac{\overline{\int d\sigma}^{y}}{\Pi_1}\right) d\sigma \; \lambda_{3yy} + \overline{\lambda}^{x}_{3x} \int \left(\frac{\int d\sigma}{\Pi_1}\right)_x d\sigma$$

$$+ \overline{\lambda}^{y}_{3y} \int \left(\frac{\int d\sigma}{\Pi_1}\right)_y d\sigma - \frac{\lambda_3}{\Pi_2} + F_{10} = 0 \qquad (57)$$

The relationship between u, v, and in (37) and (38) shows that (57) is an equation for an adjustment velocity potential. With the exception of a few small terms that contain the divergent part of the wind, (49) is a diagnostic equation for the rotational part of the wind. The solution sequence dictates that we adjust first for the rotational part of the wind and then for the divergent part of the wind. Our MODEL I is, in actuality, a variational adjustment within a variational adjustment. This formulation relaxes the requirement that the number of subsidiary conditions must be at least one less than the number of dependent variables (Courant, 1936) and therefore the energy equation may be included as the fifth constraint (MODEL II) without overdetermining the problem.

5. Computational Details

a) Grid Domain

The ten level hybrid vertical coordinate model has the state variables staggered in both the horizontal and vertical dimensions.

See Fig. 2 for the horizontal grid template. The variables u, v, ε_u, ε_v, and ϕ are located at 100 mb intervals from the top of the domain (100 mb) to 700 mb (p^*). The constraints M_1, M_2, and M_3 are referenced to these surfaces. T, $\dot{\sigma}$, and M_4 appear at 150-, 250-, 350-, 450-, 550-, and 650-mb surfaces. Further, the upper boundary on $\dot{\sigma}$ is at 50 mb ($\dot{\sigma}$ =0). The hybrid nonlinear vertical coordinate requires that the coordinate surfaces be standard levels above p^*. This choice allows the incorporation of TIROS-N mean layer temperatures directly without initial vertical interpolation and eliminates the need for any vertical interpolation in order to interpret the final analyses.

Below p^*, u, v and the developmental components appear on sigma surfaces and $\dot{\sigma}$ and M_4 are located at the half levels. The first three dynamic constraints and ? are referenced to the equivalent pressure levels and the mean layer temperature is located at the half levels. The lower boundary for $\dot{\sigma}$ ($\dot{\sigma}$ =0) is the ground. We have also chosen the surface observations to be representative of the average conditions of the lowest sigma layer. This means that the boundary layer divergence is representative of the mean divergence of this lowest layer.

b) Boundary Conditions

The correct number of boundary conditions are furnished by the variational formulation such that a unique solution is provided when natural and/or imposed boundary equations are satisfied (Forsythe and

Wasow, 1960). Natural boundary conditions are derived from the constraints as numerical expressions to be solved. However, the complexity of these expressons for the MODEL I dynamic constraints dictates the use of imposed boundary conditions whereby the dependent variables are specified on the boundaries. The solution sequence designed for MODEL I requires that boundary conditions be specified for the geopotential and adjustment velocity potential diagnostic equations, the remaining equations in the adjustment cycle, and the vertical velocity. In addition, special boundary conditions are imposed by the cyclic solution sequence. Details of the various boundary conditions follow.

Boundary conditions on the geopotential adjustment. Imposed boundary conditions for the geopotential adjustment equation (49) are supplied by the gridded fields of the observed meteorological component of the partitioned height field. The top boundary is provided by the analysis at 100 mb. The lower boundary is the meteorological height component transformed onto equivalent pressure surfaces (see Section 3e).

Boundary conditions on the velocity adjustment potential. Lateral boundary conditions are required for the adjustment velocity potential equation (57). They are imposed and may be either Dirichlet or Newmann boundary conditions. It is well known that the specification of boundary conditions on the velocity potential determines the structural details of the recovered wind field to some degree (Hawkins and Rosenthal, 1965). Furthermore, there appears to

be no method of uniquely specifying the boundary conditions (Shukla and Saha, 1974; Eskridge, 1977; Liu, 1977. Stephens and Johnson, 1978). It follows that the specification of boundary conditions for the adjustment velocity potential will determine the structural details of the recovered adjusted divergent part of the wind and therefore will determine the structural details of the total adjusted wind to some extent.

In order to determine the impacts of the boundary conditions upon the adjustment velocity potential, we make the simplifications; $\pi_1 = \pi_1(\sigma)$ and $\lambda_1 = \lambda_2 = 0$. Equation (53) becomes

$$-\int (\delta u_x + \delta v_y)\, d\sigma + \int \left(\frac{\Delta\sigma}{\pi_1}\right) d\sigma\, (\lambda_{3xx} + \lambda_{3yy}) = 0 \qquad (58)$$

where $\delta u_x = (u-u^o)_x$ is the adjustment required to satisfy the variational equations. Further, we separate the x-derivative from the y-derivative terms (see (51) and (52)) to obtain separate expressions for the x-boundaries and the y-boundaries:

$$\begin{aligned} \lambda_{3xx} &= \varepsilon\, \delta\bar{u}_x = \varepsilon\, (\delta\bar{D} - \delta\bar{v}_y) \\ \lambda_{3yy} &= \varepsilon\, \delta\bar{v}_y = \varepsilon\, (\delta\bar{D} - \delta\bar{u}_x) \end{aligned} \qquad (59)$$

Here ε is the inverse of the integrated coefficient of λ_3 in (58). The overbar on the divergence and velocity components indicates that these are averages over the vertical domain of integration. These equations show that Dirichlet boundary conditions force all of the divergence adjustment into the v-component along the x-boundaries and into the u-component along the y-boundaries. Neumann boundaries accomplish the opposite.

It appears that the ideal boundary conditions for the adjustment velocity potential are some combination of Dirichlet and Neumann boundary conditions. We define a variable r such that $r = \delta v_y / \delta u_x$. Then λ_3 can be expressed as functions of the adjustment divergence.

$$\lambda_{3xx} = \frac{\varepsilon}{1+r} \delta\bar{D}$$
$$\lambda_{3yy} = \frac{\varepsilon r}{1+r} \delta\bar{D} \qquad (60)$$

If, for example, r=1, (60) forces one half of the divergence adjustment into the boundaries. Then (60) can be solved as a line integral for λ_3 at the boundaries. Further, r may take on other values such as functions of the observed wind fields, however the structures of these functions are beyond the scope of this development.

Boundary conditions on the remaining variables. Horizontal boundary conditions are not required to determine the remaining variables in the interior of the analysis domain. However, boundary values of these variables are required in order to calculate horizontal derivatives for the forcing functions in subsequent solution cycles. Interior fields are extrapolated across the boundaries by using an approximation that is the sum of a locally averaged curvature with one half of a locally averaged gradient. This method provides boundaries that are compatible with the adjusted fields; they are generated, however, to eliminate boundary discontinuities, and do not satisfy the dynamic equations.

Boundary conditions on the vertical velocity. The boundary

conditions on $\dot{\sigma}$ are $\dot{\sigma}=0$ at the ground and at 50 mb. Because the lower coordinate surface slopes with the underlying terrain, there may exist finite actual vertical velocity near the ground. Given as $\omega_s = dp_s/dt$, the surface vertical velocity is a combination of flow over elevated terrain and through evolving meteorological pressure fields. Upon partitioning into terrain and meteorological components, the surface vertical velocity is

$$\omega_s = - \left(V \cdot \nabla p_T + \frac{\partial p_m}{\partial t} + V \cdot \nabla p_m\right) \tag{61}$$

Scale analysis of the terms of (61) (Achtemeier, 1972) showed the first term in parentheses to be at least an order of magnitude larger than the meteorological terms. This combined with our inability to accurately determine the local tendency of p_m on the synoptic scale, prompted us to neglect the meteorological components and approximate the surface vertical velocity with the first term.

Boundary conditions required by the cyclic solution sequence.

Initial tests with MODEL I with the case study described in the following article revealed features near the lateral boundaries that gave reason to suspect local violations of linear stability. These features amplified and grew into the interior of the grid during successive cycles. The adjustment of the geopotential height field (49) is forced to take on the gridded values of the observed geopotential at the boundaries regardless of the relative weights ascribed to the other variables. Small perturbations where the heights are obtained by extrapolation are frozen into the geopotential

adjustment near the boundaries and cause the growth of the erroneous waves. We were unable to totally eliminate the perturbations but were able to eliminate the buildup of the undesired waves by requiring variational analysis to satisfy the geostrophic, hydrostatic equations near the boundaries. These solutions grade into the solutions for the full nonlinear dynamic equations at five grid spaces into the grid interior.

c) Convergence Criteria

The convergence criteria for the general second-order partial differential equation with nonconstant coefficients,

$$a\lambda_{xx} + b\lambda_{yy} + c\lambda_{\sigma\sigma} + d\lambda_x + e\lambda_y + f\lambda_\sigma - g\lambda + h = 0 \qquad (62)$$

obtained by the partial wave technique is

$$\left(\frac{a}{\Delta x^2} + \frac{b}{\Delta y^2} + \frac{c}{\Delta \sigma^2} + \frac{g}{2} \right)^2 > \left(\frac{d}{\Delta x} + \frac{e}{\Delta y} + \frac{f}{\Delta \sigma} \right)^2 \qquad (63)$$

Convergence of (49) is virtually assured because the coefficients a, b, c, and g are always positive. Further, the coefficient d is just the horizontal derivative of a, e is the horizontal derivative of b, and f is the vertical derivative of c. The most stringent convergence requirement is that the absolute magnitude of the derivative of a coefficient not exceed the value of the coefficient. This requirement can be easily satisfied through the definitions of the precision modulus weights. The coefficients of (57) are similarly related except that c=f=0.

Convergence of the cyclic solution sequence for MODEL I is assured for quasi-geostrophic motions; the Rossby number is much less than one. When the Rossby number approaches one, the adjustment terms in the forcing functions approach or exceed the magnitudes of the variables being solved for, a condition that favors the development of linear instability. A determination of the range of scales for which MODEL I will return a convergent solution is the subject of continuing research.

6. Some Concluding Remarks

We have presented an outline of the first of four models that will yield a general variational model for the diagnosis of cyclone systems. The method will meld data collected from rawinsonde (wind, temperature, height, moisture) with data collected from space-based platforms (cloud wind vectors, moisture, mean-layer temperatures). This method is, by design, independent of numerical prediction models. MODEL I incorporates as dynamical constraints, the two nonliner horizontal momentum equations, the hydrostatic equation, and an integrated continuity equation. The vertical coordinate minimizes the interpolation from pressure to terrain-following coordinates, easily accomodates TIROS-N mean-layer temperature data in the middle and upper troposphere, and decreases truncation error associated with the pressure gradient force in the horizontal momentum equations. Reformulations for the horizontal tendencies of u and v are designed to increase the accuracy of the variational analysis for these hypersensitive quantities.

We designed a cyclical solution sequence that is based upon a quasi-geostrophic linearization of the eleven Euler-Lagrange equations that comprise MODEL I. This solution form does not preclude ageostrophic motion. A solution is not guaranteed for scales of motion for which the Rossby number approaches one, however. MODEL I is actually a variational model within a variational adjustment model. Rotational and divergent parts of the wind are adjusted separately

This formulation does not include frictional effects as the lower boundary is held fixed. Moisture will be included later after problems associated with the incorporation of the thermodynamic equation are solved. Comparisons of fields of meteorological data obtained via the variational method with standard analyses are the subject of the following paper, Achtemeier et al., (1986).

Acknowledgements

This research was supported by the National Aeronautics and Space Administration (NASA) under contract NAS8-34902.

REFERENCES

Achtemeier, G. L., 1972: Variational initialization of atmospheric fields - a quasi-geostrophic diagnostic model. Ph. D. Disseration, Tallahassee, Florida State University.

_____, 1975: On the Initialization problem: A variational adjustment method. Mon. Wea. Rev., 103, 1090-1103.

_____, 1979: Evaluation of a variational initialization method. Preprints, 4th Conf. Num. Wea. Pred., Amer. Meteor. Soc. 1979, 1-8.

_____, S. Q. Kidder and R. W. Scott, 1986: A variational assimilation method for the diagnosis of cyclone systems. Part II: Case study results with and without satellite data. To be submitted.

_____, and H. T. Ochs III, 1986: A hybrid vertical coordinate and pressure gradient formulations for a numerical variational analysis model for the diagnosis of cyclone systems. To be submitted.

Anthes, R. A., and T. T. Warner, 1978: Development of hydrodynamic models suitable for air pollution and other mesometeorological studies. Mon. Wea. Rev., 106, 1045-1078.

Baer, F., and J. Tribbia, 1977: On complete filtering of gravity modes through nonlinear initialization. Mon. Wea. Rev., 105, 1536-1539.

Baker, W. E., 1983: Objective analysis and assimilation of observational

data from FGGE. Mon. Wea. Rev., 111, 328-342.

Bengtsson, L., M. Kanamitsu, P. Kallberg and S. Uppala, 1982: FGGE 4-dimensional data assimilation at ECMWF. Bull. Amer. Meteor. Soc., 63, 29-43.

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

Charney, J. G., 1948: On the scale of atmospheric motion. Geofys. Publikas., 17, 1-17.

Courant, R., 1936: Differential and Integral Calculus, Vol. 2, (E. J. McShane, translator), Wiley - Interscience, p198.

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

Eskridge, R. E., 1977: Comments on "An iterative algorithm for objective wind field analysis." Mon. Wea. Rev., 105, 1066.

Fjortoft, R., 1952: On a numerical method of integrating the barotropic vorticity equation. Tellus, 4, 179-194.

Forsythe. G. E., and W. R. Wasow, 1960: Finite Difference Methods for Partial Differential Equations. New York, John Wiley and Sons.

Gandin, L. S., 1963: Objective analysis of meteorological fields. Isdat., Leningrad. [Israel Program for Scientific Translations, Jerusalem, 1965, 242 pp.]

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

Haltiner, G. J., 1971: Numerical Weather Prediction, Wiley and Sons, USA. 46-61.

Hawkins, H. F., and S. L. Rosenthal, 1965: On the computation of stream functions from the wind field. Mon. Wea. Rev., 93, 245-252.

Ikawa, M., 1984: An alternative method of solving weak constraint problem and a unified expression of weak and strong constraints in variational objective analysis. Pap. Meteor. Geophys., 35, 71-80.

Krishnamurti, T. N., 1968: A diagnostic balance model for studies of weather systems of low and high latitudes. Rossby number less than one. Mon. Wea. Rev., 96, 197-207..bl .i-5

Kurihara, Y., 1968: Note on finite difference expressions for the hydrostatic relation and pressure gradient force. Mon. Wea. Rev., 96, 654-656.

Lewis, J. M., 1980: Dynamical adjustment of 500 mb vorticity using P.D. Thompson's scheme - a case study. Tellus, 32, 511-514.

_____, 1982: Adaptation of P.D. Thompson's scheme to the constraint of potential vorticity conservation. Mon. Wea. Rev., 110, 1618-1634.

_____, and T.H. Grayson, 1972: The adjustment of surface wind and pressure by Sasaki's variational matching technique.

J. Appl. Meteor., 11, 586-597.

_____, 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.

Liu, C. Y., 1977: Reply. Mon. Wea. Rev., 105, 1067.

Machenhaur, 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., K. H. Bergman, R. E. Kistler, G. E. Rasch, and D. S. Gorden, 1979: The NMC operational global data assimilation system. Mon. Wea. Rev., 107, 1445-1461.

Mesinger, F., and A. Arakawa, 1976: Numerical methods used in atmospheric models. Vol. 1, GARP Publications Series No. 17, p47.

Miyakoda, K., and R. M. Moyer, 1968: A method of inialization for dynamical weather forecasting. Tellus, 20, 115-130.

Nitta, T., and J. B. Hovermale, 1969: A technique of objective analysis and initialization for the primitive forecast equations. Mon. Wea. Rev., 97, 652-658.

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

Phillips, N. A., 1957: A coordinate system haveing some special advantages for numerical forecasting. J. Meteor., 14, 184-195.

Sasaki, Y., 1958: An objective analysis based upon the variational method. J. Meteor. Soc. Japan, 36, 77-88.

_____, 1970a: Some basic formalisms in numerical variational analysis. Mon. Wea. Rev., 98, 875-883.

_____, 1970b: Numerical variational analysis formulated under the constraints as determined by longwave equations and a low-pass filter. Mon. Wea. Rev., 98, 884-898.

_____, and Lewis, J., 1970: Numerical variational analysis of the planetary boundary layer in conjunction with squall line formation. J. Meteor. Soc. Japan, 48, 381-398.

Schlatter, T. W., 1975: Some experiments with a multivariate statistical objective analysis scheme. Mon. Wea. Rev., 103, 246-257.

Shukla, J., and K. R. Saha, 1974: Computation of non-divergence streamfunction and irrotational velocity potential from the observed winds. Mon. Wea. Rev., 102, 419-425.

Shuman, F., and J. B. Hovermale, 1968: An operational six-layer primitive equation model. J. Appl. Meteor., 7, 525-547.

Smith P. J., and C. P. Lin, 1978: A comparison of synoptic scale vertical motions computed by the kinematic method and two forms of the omega equation. Mon. Wea. Rev., 106, 1687-1694.

Stephens, J. J., and K. W. Johnson, 1978: Rotational and divergent wind potentials. Mon. Wea. Rev., 106, 1452-1457.

Temperton, C., 1984: Variational normal mode initialization for a multilevel model. Mon. Wea. Rev., 112, 2303-2316.

Thompson, P. D., 1969:Reduction of analysis error through constraints of dynamical consistency. J. Appl. Meteor., 9, 738-742.

Wahba, G., and J. Wendelberger, 1980: Some new mathematical methods for variational objective analysis using splines and cross validation. Mon. Wea. Rev., 108, 1122-1143.

Wang, P. K., 1984: A brief review of the Eulerian variational principle for atmospheric motions in rotating coordinates. Atmos.-Ocean, 22, 387-392.

Whittaker, E., and G. Robinson, 1926: The Calculus of Observations (2nd Edition). London, Blackie and Son, LTD., p176.

Williamson, D. L., and R. Daley, 1983: An iterative analysis initialization technique. Mon. Wea. Rev., 111, 1517-1536.

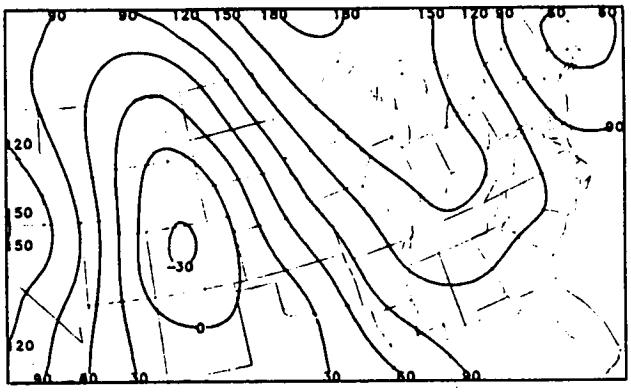
Figure Captions

Figure 1. Heights at the lower coordinate surface for a) unpartitioned terrain-following coordinate, b) the equivalent pressure surface, and c) the 1000 mb pressure surface.

Figure 2. A portion of the staggered grid used for the variational diagnostic model.

a)

b)



c)

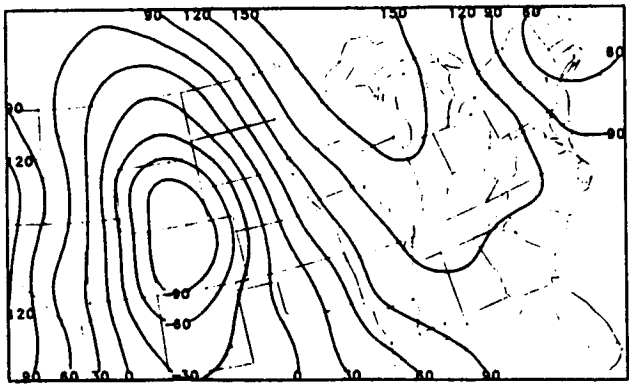


Figure 1

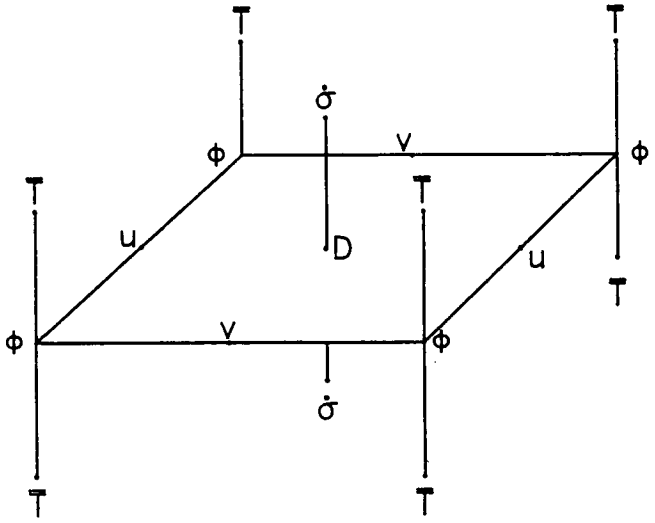


Figure 2

Chapter III

A Variational Assimilation Method for the Diagnosis of Cyclone systems. Part II: Case Study Results with and without Satellite Data

Gary L. Achtemeier, Stanley Q. Kidder,

and Robert W. Scott

Climate and Meteorology Section

Illinois State Water Survey

Champaign, IL 61820

A Variational Assimilation Method for the Diagnosis of Cyclone Systems. Part II: Case Study Results with and without Satellite Data.

Gary L. Achtemeier, Stanley Q. Kidder,

and Robert W. Scott

Illinois State Water Survey

Champaign, IL 61820

ABSTRACT

This paper presents the evaluation of a diagnostic multivariate data assimilation method described in a companion paper by Achtemeier et al.. Ground-based and space-based meteorological data are weighted according to the respective "measurement" errors and blended into a hybrid data set that is required to satisfy the two nonlinear horizontal momentum equations, the hydrostatic equation, and an integrated continuity equation for a dry atmosphere as dynamical constraints. Multivariate variational objective analyses with and without satellite data are compared with initial analyses and the observations to determine the accuracy and sensitivity of the assimilation to different data sets. Three evaluation criteria are developed that measure a) the extent to which the assimilated fields satisfy the dynamical constraints, b) the extent to which the assimilated fields depart from the observations, and c) the extent to which the assimilated fields are realistic as determined by pattern recognition. The

last criterion requires that the signs, magnitudes, and patterns of the hypersensitive vertical velocity and local tendencies of the horizontal velocity components be physically consistent with respect to the larger scale weather systems.

The percent reduction of the initial RMS error is used to determine the extent to which the SAT and NOSAT blended data sets converge to the solution of the four dynamical constraints. There was approximately 90-95 percent error reduction for the two horizontal momentum equations when applied to the case of 1200 GMT 10 April 1979. The RMS error reductions for the integrated continuity and hydrostatic equations ranged from 90-100 percent except for the errors at levels 2 and 3 of the integrated continuity equation which were reduced to approximately 70 percent.

The pattern recognition analysis for the basic fields, height, temperature, and vector wind, revealed that the SAT and NOSAT analyses were similar with the following two exceptions. First, there were larger numerical differences between the SAT height analysis and the initial objective analysis than were found between the NOSAT analyses and the initial objective analysis. Second, large areas of the network were void of satellite data which caused the loss of important local details of the temperature field. One result was the introduction of a large (-40 m) height anomaly in the middle troposphere over the western U.S. Both NOSAT and SAT analyses corrected a rather poor univariate wind analysis and placed jet streaks over California, western Texas, and the Great Lakes.

In the analysis of hypersensitive variables, the variational method removed or reduced the magnitudes of several large vertical velocity centers (magnitudes greater than 10 cm sec^{-1}) which were placed between rawinsonde sites by conventional methods and replaced them with a zone of positive vertical velocity roughly parallel to the axis of an area of precipitation that was used as a check on the accuracy of the final fields. The variational analysis also concentrated an area of strong subsidence (-14 cm/sec) along the axis of a jet streak over the Northeastern States. It also placed a narrow band of ascending motion near the entrance region along the anticyclonic shear side of the jet streak, an area long noted by those experienced in the motion fields surrounding jet streams as favorable for upward vertical velocities.

The SAT and NOSAT tendency patterns were of approximately the same magnitude and scale as the observed tendency patterns that were obtained from NASA-AVE high frequency rawinsonde data. With the exception of a negative tendency center in the lower troposphere, the agreement among the tendency patterns was very good considering that the observed patterns were subject to interference by mesoscale phenomena and that the observed patterns were valid at 1330 GMT rather than at 1200 GMT. The relative accuracy of the variational tendencies was made more apparent upon comparison of the initial field tendencies with the observed patterns. The initial field tendencies consisted of relatively large amplitude centers of scale roughly equal to the average separation between observing sites. The magnitudes of these centers became unrealistically large in the upper troposphere within high wind velocity areas.

These variational tendencies are the first relatively accurate diagnostic fields of local tendencies of the velocity components apart from initialization schemes for numerical prediction models.

1. Introduction

In a companion paper, Achtemeier et al., (1986) presented the description of a multivariate data assimilation method based upon Sasaki's (1958, 1970) method of variational objective analysis. It is the first of several variational numerical models designed to produce dynamically consistent fields of meteorological variables for the diagnosis of cyclone scale weather systems. Special emphasis is placed upon incorporating data from diverse sources and, in particular, upon meshing observations from space-based platforms with those from more traditional immersion techniques.

The variational diagnostic model (MODEL I) requires the two nonlinear horizontal momentum equations, the hydrostatic equation, and an integrated continuity equation for a dry atmosphere to be satisfied as dynamical constraints . (Later versions will include the energy equation for moist processes.) A hybrid nonlinear vertical coordinate allows for the easy incorporation of TIROS-N mean layer temperatures. Coordinate surfaces are pressure surfaces above 700 mb. The nonlinear vertical coordinate also makes possible the removal of much of the local variations with unlevel terrain at levels below 700 mb in the hydrostatic equation and the pressure gradient terms of the momentum equations. A complete development of the vertical coordinate and an analysis of the pressure gradient terms are

given by Achtemeier and Ochs (1986).

In addition, there are some quantities which, because of poor instrument accuracy or insufficient sampling frequency, cannot be measured directly and must be inferred through functions of other measured variables; in our case, they are determined as part of the variational blending processes. Among these are hypersensitive variables that are sensitive to small changes in the other variables, such as vertical velocity and the local tendencies of the horizontal velocity components. The local tendencies of u and v appear explicitly in the dynamic constraints and therefore must be solved for in the variational formulation. Various methods used in the past for accommodating the local tendencies are discussed in the companion paper.

The intent of this paper is to compare multivariate variational objective analyses with and without satellite data with initial analyses and the observations to determine the accuracy and sensitivity of MODEL I to different data sets. Because this assimilation is not an initialization for a numerical prediction model, the often used procedure of determining the best initial analysis by finding the best forecast does not apply. We instead use three diagnostic criteria which, although they may be somewhat more subjective than measures of forecast skill, have found use in the verification of diagnostic analyses (Krishnamurti, 1968; Achtemeier, 1975; Otto-Bliesner *et al*, 1977). These criteria are measures of a) the extent to which the assimilated fields satisfy the dynamical constraints, b) the extent to which the assimilated fields depart from the observations, and c) the extent to which the assimilated fields are realistic as determined by

pattern recognition. The last criterion requires that the signs, magnitudes, and patterns of the hypersensitive vertical velocity and local tendencies of the horizontal velocity components be physically consistent with respect to the larger scale weather systems.

The case study used for the verification of MODEL I was a short wave over the Central Plains on 1200 GMT 10 April 1979. Shown in an objective analysis of the 500 mb heights (Fig. 1), this disturbance was one of a progression of short waves that were embedded within southerly flow between a synoptic scale trough over the Great Basin and a high pressure ridge over the eastern United States. It was not accompanied by severe mesoscale convective systems as were the short waves that followed it through the Central Plains. Light precipitation (shaded patches) at 1235 GMT was mostly from relatively shallow convective elements embedded within middle tropospheric clouds (6 km) located along the upwind side of a large cold cloud mass (Fig. 2) along and ahead of the wave. This case was selected because TIROS-N temperature soundings coexist with NASA-AVE 3-hr rawinsonde data over a large area of the central United States. The 3-hr rawinsonde data are required to provide verification for the diagnosed 3-hr local tendencies of the horizontal velocity components. Furthermore, this SESAME case data has been the subject of several synoptic and mesoscale diagnostic analyses. These studies provide additional verification fields for MODEL I.

The time, 1200 GMT 10 April 1979, was selected because intense mesoscale convective systems were not significantly impacting the large scale dynamics. However, our preliminary analyses with the SESAME I

regional scale rawinsonde network revealed the presence of mesoscale features in the wind field for this relatively quiet period. Since the resolution of mesoscale systems is not part of the design for this study, only rawinsonde data for the NWS synoptic network were used (Fig. 3a). The distribution of TIROS-N temperature soundings is shown in Fig. 3b. Most notable is a large data void area over roughly the northwest quarter of the analysis domain. Data is also sparse over Texas and New England.

The methods used to prepare the data for insertion into MODEL I are described in the next Section. Methods for debiasing the TIROS-N temperature soundings are developed in Section 3. Section 4 contains the development of the precision moduli that weight the data in the assimilation. Comparisons of assimilations with and without satellite data are presented in Section 5 and the results of this study are summarized in Section 6.

2. Preparation of Data for the Variational Analysis

The ten level hybrid vertical coordinate model has the state variables staggered in both the horizontal and vertical dimensions. See Fig. 4 for the grid template. The variables u, v, ϕ, ε_u, and ε_v are located at 100 mb intervals from the top of the domain (100 mb) to 700 mb (p^*). $\bar{T}$ and $\dot{\sigma}$ appear at 150-, 250-, 350-, 450-, 550-, and 650-mb surfaces. Further, the upper boundary on $\dot{\sigma}$ is at 50 mb ($\dot{\sigma}$ =0). The hybrid nonlinear vertical coordinate is designed so that the coordinate surfaces are standard pressure levels above p^*. This choice allows the incorporation of TIROS-N mean layer temperatures directly at these levels without any

need for vertical interpolation and also eliminates the need for vertical interpolation at the end of the variational analysis in order to interpret the final analyses.

Below p^*, u, v and the developmental components appear on sigma surfaces and $\bar{T}$ and $\dot{\sigma}$ are located at the half levels. The lower boundary for $\dot{\sigma}$ ($\dot{\sigma}$ =0) is the ground. We also have chosen the surface observations to be representative of the average conditions of the lowest sigma layer. This means that the boundary layer divergence is representative of the mean divergence of this lowest layer.

Interpolation of observed quantities onto the 100 km by 100 km horizontal mesh at and above the p^* level is easily accomplished with a modification of the Barnes (1964) successive corrections technique (Achtemeier, 1986b). Impacts by this method upon the analyses near the boundaries of the data field can be significant for large amplitude, short wavelength patterns (Achtemeier, 1986a).

The procedure for obtaining the state variables for the sigma levels which vary nonlinearly with pressure below p^* first requires that all variables be gridded to constant pressure surfaces save for the surface winds and temperatures which are gridded directly to the lower coordinate surface. Vertical interpolation onto sigma surfaces awaits the definition of the sigma surfaces as functions of surface pressure. We first determine the mean temperature between the height of the p^* surface and a smoothed terrain that forms the lower coordinate surface by forming the weighted sum of the mean layer temperatures for the pressure layers from p^* to the first pressure surface above the lower coordinate surface and the mean

temperature from this level to the surface. We define a new surface temperature subject to the requirement that the mean temperature for the layer below p^* is equal to the average of the temperature at p^* and the modified surface temperature. This approach creates a mean lapse rate for the layer and eliminates cold thicknesses that would result if the mean layer temperatures were calculated from cold surface temperatures at the bases of shallow nocturnal inversions.

Once the mean layer temperatures are determined, the pressure at the lower coordinate surface may be found from

$$p_s = p^* \left[\frac{T^* + \Gamma\,(\phi^* - \phi_s)}{T^*}\right]^{\frac{1}{R\Gamma}} \qquad (1)$$

where $\Gamma = -\partial T/\partial \phi$. The pressures corresponding to the remaining sigma surfaces can be found from the definition of the nonlinear hybrid sigma coordinate (Achtemeier and Ochs, 1986). Once the locations of the sigma surfaces are known, the temperatures can be found by linear interpolation from the mean lapse rate of temperature between p^* and the surface. Then the corresponding geopotential heights are obtained with

$$\phi = \phi^* + R\,[(T^* + \gamma p^*)\ \ell n\ p^*/p - \gamma(p^* - p)] \qquad (2)$$

where $\gamma = -\partial T/\partial p$. The horizontal wind components are interpolated directly to the sigma surfaces from the pressure surfaces.

We nondimensionalize all variables. As shown by Achtemeier and Ochs (1986), the nonlinear vertical coordinate is used to remove a hydrostatic component that includes much of the vertical variations of the lower coordinate surface due to variable topography. This procedure transforms

the lowest three sigma coordinate surfaces onto "equivalent pressure surfaces". We arbitrarily let p_e = 800, 900, 1000 mb respectively for these surfaces. Figure 5 shows the untransformed heights (Fig. 5a) for the lowest coordinate surface and the heights remaining after the hydrostatic component was removed (Fig. 5b). Virtually all of the orographic contribution to the sigma coordinate system is removed by this method. However, the remaining heights at 1000 mb equivalent pressure do not equal the heights obtained directly from an objective analysis of the 1000 mb heights because the mean temperature of the layer between the surface and the reference pressure is used to calculate the thickness of the layer between the equivalent pressure level and the reference pressure.

Finally, a hydrostatic reference atmosphere is removed and the residual fields are multiplied by the ratio of the Rossby number to the Froud number to bring them into compatibility with the nondimensionalized dynamic equations. The variational adjustments are carried out on these residual variables.

3. Preparation of TIROS-N Temperature Data for the Variational Analysis

There are four steps in the process of preparing the satellite soundings for insertion in the variational analysis model: (1) determining and removing biases, (2) determining standard errors of the satellite temperatures to assist in the assigning of weights in the model, (3) converting layer mean temperatures (the form in which the Tiros-N soundings are supplied on tape) to level temperatures (which are required in the variational analysis model), and (4) determining what action to take to

correct for the non-synoptic nature of the satellite data.

a) Biases and Weights

The accuracy of satellite soundings has been the subject of a number of investigations (Phillips et al.,1979; Schlatter, 1981; Gruber and Watkins, 1982). However, the comparisons differ because the retrieval algorithms are occasionally updated and because the retrieval coefficients are regularly updated. In addition, the comparisons are made against different reference (truth) data sets. To make certain that the retrieval algorithms and coefficients for our calibration study are the same as for the case study used to test the variational analysis model (10-11 April 1979), we acquired Tiros-N soundings and RAOB data for the period 26 March through 11 April 1979. This is the same period (plus a few days) analyzed by Schlatter.

Schlatter compared the Tiros-N soundings with NMC Final Analyses; others have compared the satellite soundings with "co-located" RAOBS. Because Tiros-N is an afternoon satellite while the RAOBS are taken at 6 AM and 6 PM in the central U.S., comparisons with co-located rawinsonde data were not possible. Therefore, we decided to compare the satellite soundings with time- and space-interpolated, objectively-analyzed fields of RAOB data.

Mean layer virtual temperatures, derived from RAOB thicknesses, were objectively analyzed on a 21 x 21 grid covering most of North America (Fig. 6) for each 12 hr for the period 0000 GMT 26 March through 1200 GMT 11 April 1979. Biases were estimated by calculating the difference between

satellite-estimated mean virtual temperatures and RAOB values interpolated in both time and space from the analyses to the satellite data.

Figure 7 shows the 12 hr average biases as a function of time for each layer. The dashed lines represent the mean biases for the periods 26 March through 8 April and 10-11 April. The three sounding types (clear, partly cloudy, cloudy) have been kept separate. The error bars represent 95% confidence intervals assuming that the biases are normally distributed about the 12 hr mean, which proved to be a good assumption upon examination. As we studied these plots, two aspects were disturbing: (1) For a large number of points the error bars did not include the mean represented by the dashed line. Only one in 20 points should not include the mean if the long term average is representative. (2) There seems to be a 24-hr oscillation, which indicates that daytime and nighttime biases may be different.

Kidder and Achtemeier (1986) stratified the TIROS-N soundings by day and night and calculated the mean biases for the period 26 March through 8 April. It was found that biases for day and night soundings are statistically different for clear and partly cloudy soundings, and at most levels for cloudy soundings. It was also found if the biases are not properly removed by time and by sounding type, objectively analyzed fields of satellite temperatures can have variations of several degrees due solely to biases.

A second kind of bias associated with elevated terrain was also investigated. The statistical retrieval algorithm used by NESDIS is based upon the assumption that the base of the sounding is at (or near) 1000 mb. Over elevated terrain this assumption fails. There have been attempts to correct this problem in the operational scheme, but the process can generate a bias in the low levels. Failure of the assumption for elevated terrain should result in satellite temperatures that are too cold. To determine the magnitude of the cold bias, we correlated the deviations of the satellite temperatures from the rawinsonde analyses with terrain height. No significant correlation was found. Therefore, the cold bias over elevated terrain seems to be a negligible error, at least for the satellite data used for this study.

The TIROS-N temperature data inserted into the variational analysis model have been debiased with the mean biases tabulated in Table 1. Table 1 also includes the standard deviations of the biases. If biases are properly removed from the temperature data, these standard deviations are equal to the RMS errors. These RMS errors are included in the calculation of the precision modulus weights that determine the relative importance the satellite data receive in the variational meshing with the other observations.

c) Conversion from Layer Temperatures to Level Temperatures

Operationally retrieved Tiros-N soundings are supplied on tape as layer mean virtual temperatures at standard synoptic levels. The nonlinear sigma vertical coordinate was designed so that the coordinate surfaces for

the variational model are coincident with standard synoptic levels in the middle and upper troposphere. However, the coordinate surfaces for the lower troposphere are not coincident with standard synoptic levels and the layer mean temperatures must be decomposed into level temperatures at specified levels. The conversion can be accomplished easily if the temperature at one level is known. However, when we assumed that the 400 mb temperature is simply the average of the 300-400 and the 400-500 mb mean layer temperatures, the rawinsonde data revealed that the estimated 400 mb temperatures are too cold on the average by 0.65 K. A check using the U.S. Standard Atmosphere (1962), also showed that the 400 mb temperature estimated in this way should be 0.65 K too cold.

To correct this problem, the atmosphere was assumed to have a constant lapse rate between 300 and 500 mb. Using the equations which describe such an atmosphere (Hess, 1959), we derived an algorithm which, when given the average and the difference between the mean layer temperatures for the 300-400 and the 400-500 mb layers, would calculate a corrected 400 mb temperature. This method was tested with the rawinsonde data from 26 March through 11 April (includes over 3200 observations) and it was found that the mean of the difference between the corrected and actual 400 mb temperatures was 0.00 to within 2.77 K (95% confidence level). Since this method proved to be accurate, it was used to decompose the Tiros-N mean layer temperatures into level temperatures before insertion into the variational analysis model.

d) Time to Space Conversion for Asynoptic Satellite Data

Satellite soundings are not taken simultaneously with the semi-daily collection of rawinsonde data from the National Weather Service upper air network. Therefore, we considered several simple time to space conversions to make the satellite soundings more representative at synoptic times. Two of these methods made use of the univariate objective analyses of temperature and horizontal wind to calculate the advective rate of change of temperature and potential temperature. The calculated changes were compared with the actual temperature changes measured by rawinsonde at 3-hr intervals over the AVE/SESAME network for 10-11 April 1979. Very small correlations were found (See Table 2). A third method that determines horizontal temperature advection through the vertical shear of the geostrophic wind yielded similar results. The errors were so large that it was concluded that the satellite data should be included directly into the analysis as long as the difference in measurement times is less than three hours.

4. Precision Modulus Weights for the Variational Assimilation

The observations to be assimilated by the variational method are meshed with the dynamic constraints through the formulation described in the companion paper. The gridded observations to be adjusted receive precision modulus weights in proportion to their relative observation and interpolation accuracies according to

$$\Pi_i = \Pi_i^* G_i \qquad (3)$$

where Π^* for the ith observation is defined by $\Pi_i^* = 0.5\,\sigma_i^{-2}$ (Whittaker and Robinson (1926). The σ_i is the root mean square error (RMS) for the ith

observation. The function G_i that multiplies π_i^* is an additional weight that accounts for the variable density of observations surrounding a grid point. It varies inversely with the distances of observations from a grid point. G_i has been assigned a value of one for all data for this study. Consequences of this restriction upon the TIROS-N temperatures will be discussed later in the next section.

This development is restricted to securing values for π. Furthermore, we require that the π are functions of sigma only. The Lagrangian density (Eqn. 28 in the companion paper),

$$I = \Pi_1(u-u^o)^2 + \Pi_1(v-v^o)^2 + \Pi_2(\dot{\sigma}-\dot{\sigma}^o)^2 + \Pi_3(\phi-\phi^o)^2 + \Pi_4(\bar{T}^\sigma-\bar{T}^{\sigma o})^2 + \Pi_5(\phi_x-\phi_x{}^o)^2$$

$$+ \Pi_5(\phi_y-\phi_y{}^o)^2 + \Pi_6(\phi_\sigma-\phi_\sigma{}^o)^2 + \Pi_7(\xi_u-\xi_u{}^o)^2 + \Pi_7(\xi_v-\xi_v{}^o)^2 + 2\sum_{k=1}^{k}\lambda_k M_k , \quad (4)$$

contains precision modulus weights for seven quantities that enter the variational assimilation. Three of these (π_1 , π_3 , and π_4) weight observed variables, two more (π_5 and π_6) weight gradients of observed geopotential, and the remaining two (π_2 and π_7) weight the vertical velocity and the developmental components of the local tendencies of the velocity components.

The precision moduli for variables that are directly observed are calculated from the RMS values for that variable. Gradients of some observed variables are also included in the assimilation and the RMS errors for the gradient must be calculated from the observed RMS values. Finally, the RMS errors for some variables that are not directly observed must be estimated from the RMS errors of other observed variables.

Table 3 shows the RMS errors of observation for the meteorological variables that are observed. The first two columns give estimates for the RMS errors for the scalar wind speed as functions of elevation angle of the balloon (Fuelberg, 1974). The values for the 20 degree elevation angle compare favorably with the results from Hovermale's (1962) spectral decomposition of meteorological data. RMS values for heights and rawinsonde temperatures are from a composite of methods for estimating measurement error (Achtemeier, 1972). Estimates of the measurement error for the TIROS-N clear and cloudy temperature soundings are provided by this study (see Section 3).

The RMS errors for the horizontal gradient of geopotential are estimated based on the assumption that the errors of measurement of geopotential are uncorrelated among observation sites. If the representative gradient is the average separation between observing sites, then

$$\sigma_{\phi_x} = \sigma_{\phi_y} = \frac{\sqrt{2}\,\sigma_\phi}{\Delta S} \qquad (5)$$

where ΔS is the average separation between observation sites. The error in the geopotential thickness is related to the measurement error in the mean temperature for the layer through the hydrostatic equation. The RMS error is

$$\sigma_{\phi_\sigma} = \sigma_T \frac{\partial \ell n(p)}{\partial \sigma} \qquad (6)$$

for the transformed version of the nonlinear sigma vertical coordinate presented in this study.

The RMS errors for the developmental parts of the horizontal velocity components tendencies are estimated from the temporal gradient of the velocity in addition to the errors encountered upon estimating the advective part of the tendencies. The error in ξ_u at any observation site is thus

$$e_{\xi_u} = \frac{1}{\Delta t}\,(e_u^{\tau+1} - e_u^{\tau}) + \frac{C}{\Delta S}\,(e_{u_{x+1}}^{\tau} - e_u^{\tau}). \qquad (7)$$

Upon assuming that the errors are uncorrelated, the RMS error for the developmental part of the tendencies are related to the RMS errors for the velocity through

$$\sigma_{\xi_u} = \sqrt{2}\,\sigma_u \left[\frac{1}{(\Delta t)^2} - \frac{C^2}{(\Delta S)^2} + \frac{C}{\Delta t\,\Delta S}\right]^{1/2}. \qquad (8)$$

Finally, the RMS errors for the vertical velocity must be estimated through the algorithm that calculates vertical velocity, in this study, the integrated divergence. The mean square error for the divergence is twice the mean square error for the gradient of the velocity. This error is integrated vertically to find the RMS error for the vertical velocity at any level k:

$$\sigma_{\dot{\sigma}_k} = \frac{2\Delta\sigma}{\Delta S}\left(\sum_{j=1}^{k} \sigma_{u_j}^2\right)^{1/2}. \qquad (9)$$

Table 4 shows the nondimensional root-mean-square errors of "observation" for the variables and their derivatives to be adjusted in MODEL I. Small values mean greater accuracy. In order of decreasing accuracy, Table 4 ranks the geopotential height as the most accurately observed variable and then the winds and the temperatures. The developmental components of the local velocity tendencies are estimated to

be the least accurately determined variables. Note that the RMS errors for the rawinsonde temperatures are larger than the corresponding satellite temperatures for the upper troposphere. These RMS values are for vertically averaged layer temperatures and the error in representing the these temperatures by the average of the level temperatures increases near the tropopause where large changes in stability can be found within a layer. The RMS values for the satellite temperatures are also largest near the tropopause but are not degraded to the extent as the mean layer temperatures for rawinsonde data. Increasing the vertical resolution of this model will lead to improvements in the rawinsonde RMS values. Finally, the RMS value for level 10, the top boundary of the vertical velocity, has been set to near zero to guarantee that the adjusted vertical velocity will vanish at the top of the domain.

Table 5 gives the nondimensional precision moduli weights calculated from the RMS values presented in Table 4. Larger values imply greater accuracy and therefore smaller adjustments in the quantity that receives the large weights. These values are of course estimates of the actual weights because they are developed from mean observational errors and do not necessarily represent the actual observational error for this case study. Further, the estimated errors for non-observed quantities obtained through approximations do not necessarily strictly obey the assumptions made in their derivations.

The precision modulus weights that are used in the development of the variational hybrid data sets presented in the following sections differ from the weights shown in Table 5 by the following amounts for the

following reasons. First the weights for the temperature have been increased by a factor of 10. We have made this modification because a major purpose in undertaking this study is to assess the impact of temperatures measured from space-based platforms upon a hybrid data set that includes other variables that may be poorly measured or not observed. In addition, we have reduced the precision modulus weights for the geopotential by a factor of 10 for the above stated reason and because preliminary results indicated that the observed geopotential necessary to satisfy boundary conditions tends to force the solution toward the geopotential in the interior of the domain. Finally, we have increased the weights for the vertical velocity and the developmental components of the velocity tendencies to require the solution not to transfer error residuals into these adjusted variables. A more complete analysis of the sensitivity of MODEL I to different precision moduli is planned.

5. Results of the Assimilations

a) Satisfaction of Dynamical Constraints

The variational assimilation method we have developed is a physical model. Four of the Navier-Stokes equations that govern flow in free atmosphere subject to the assumptions that apply to hydrostatic and synoptic conditions have been used in the model derivation. Therefore, it is expected that the three dimensional fields of meteorological variables should be solutions of the dynamic equations.

Adjusted variables at two successive cycles were averaged and reintroduced into the dynamic constraints. Residuals were computed as remainders of algebraic sums of individual terms of each constraint. The RMS error (Glahn and Lowry, 1972) for each level was then found. Residuals vanish (constraint satisfaction) when variables at two successive cycles are unchanged. A measure of the convergence of the variational method to constraint satisfaction is the difference between the initial RMS error of the residuals of the unadjusted variables substituted directly into the dynamic equations and the RMS values at each cycle. These differences are divided by the initial RMS errors, converted to percent and expressed in Tables 6 and 7 as percent reduction of the initial RMS error.

Table 6 shows how the reductions of the initial RMS error for the two horizontal momentum equations varies for each pass through the cyclical solution sequence for the eight adjustable levels of the Model. The error is approximately cut in half with each cycle through the fifth cycle. The solution stabilizes near 90-95 percent error reduction and there is no further significant improvement in the assimilation after the fifth cycle. In fact the assimilated fields do not satisfy the constraints quite as well out to the eighth cycle. The percent RMS error reductions behave similarly for the two equations except that the final error reductions for the v-component equation are one to four percent less than for the u-component equation.

The percent RMS error reductions for the integrated continuity and hydrostatic equations are shown in Table 7. The errors for the integrated continuity equation are reduced approximately by 70 percent by the fourth

cycle and remain relatively unchanged through the eighth cycle at levels 2 and 3. The solution for the six upper levels improves to approximately 90 percent by the eighth cycle. These improvements are, of course, dependent upon the magnitudes of the initial RMS errors. We first calculated the vertical velocity by the O'Brien (1970) method and then determined the RMS errors for the integrated continuity equation. Had we assumed that the initial vertical velocity was zero, the initial RMS errors would have been much larger than the values used in Table 7 and the error reductions would have been 100 percent by the fourth cycle.

The RMS errors for the hydrostatic equation are halved at each cycle and the percent error reduction increases monotonically to near 100 percent by the eighth cycle.

b) Pattern Recognition. Results for the Major Fields

Analyses of height, streamlines and isotachs, and temperature developed by the objective and the NOSAT and SAT variational assimilations are presented for levels 3 and 6 for 1200 GMT, 10 April 1979. The lowest level is approximately the 800 mb surface except over the highest terrain while the upper level is identical with 500 mb. Height contours are drawn for every 30 m MSL, isotachs at 5 m s^{-1} intervals, and temperature at 2C. In addition, the variational height analyses include height differences from the initial objective analysis contoured at 5 m intervals. The dashed rectangle located in the interior of each plot, 3 grid distances from the network border, delineates the part of the variational analysis unaffected by the boundaries of the domain.

Height maps developed by the objective analysis, the NOSAT analysis, and the SAT analysis are shown on level 3 in Fig. 8. In general, all methods produced smooth analyses of the large scale height field. An elongated region of relatively low heights was oriented along the eastern slopes of the Rocky Mountains and centered in northwestern Colorado. The area was rather broad, covering the western half of the network. To the east, higher heights ridged across the U. S. from the central Gulf coast north northwestward into Wisconsin, generally dominating conditions from the eastern Plains to the Appalachians. Further east, a second trough, centered just outside the network over New England, created cyclonic conditions over the Atlantic coastal states from the Virginias northward.

NOSAT variational height contours (Fig. 8b) closely resembled those developed by the objective technique. Nearly all NOSAT heights were higher (dashed lines), but differences between the two analyses were relatively minor. Largest departures were centered close to extremes in the height field, varying from about 5 m higher near areas of lowest heights to just over 10 m higher across the ridge located in eastern Missouri and southern Illinois. In general, however, the NOSAT analysis tended to filter out most small scale features in the initial objective analysis.

The SAT height analysis (Fig. 8c) also was smoother than the original analysis. However, numerical differences between the SAT analysis and the initial objective analysis were substantially larger than those developed by the NOSAT variational analysis and included some areas with large negative departures in contrast to mostly higher heights in Fig. 8b. SAT adjusted heights within the western and northern sections of the trough in

the Rockies were 5 - 10 m lower than the objective analysis. Positive height departures were generated throughout the ridge in a pattern roughly similar to that shown by the NOSAT analysis but with larger magnitudes. Compared with the objective analysis in this area, heights were as much as 15 m higher. A 15 m positive departure was located also in the base of the trough across the mid-Atlantic states. In both of these areas, the use of satellite temperatures created departures three times larger than the amount calculated by the NOSAT analysis.

Unfortunately, large areas of the network were void of satellite data. Temperature sounding information (satellite subpoints are shown superimposed on the analysis in Fig. 3b) was unavailable over New England, Texas, and across the northern Plains north of a line from southwestern Iowa to northeastern Oregon. Temperatures over roughly the northwestern quarter of the network were interpolated from surrounding data rich areas. Thus, important local details of the temperature field were lost. One result was that warm thicknesses (observed by rawinsonde) were replaced by cold thicknesses from data located south of the data void area. There resulted a large negative height anomaly in comparison with the initial objective analysis.

Analyses from the three methods on level 6 (500 mb) are shown in Fig. 9. The objective technique (Fig. 9a) produced a deep, narrow trough along the western boundary of the network, centered from northeastern Nevada into western Idaho. Other analyses of this trough (Moore and Fuelberg, 1981; Kaplan et al., 1982) produced heights up to 40 m lower over Arizona, Nevada, and Utah. The axis of the trough was located several hundred

kilometers west of its position on level 3: a substantial tilt in the system between these two levels. The ridge also tilted westward into the northern Plains. The trough in the northeast remained in that area and a second, weak short wave was produced over the western Great Lakes.

The short wave, which is the focus of this study, stretched eastward from the trough in Utah through northern Colorado and into northwestern Missouri. Although appearing quite substantial here, this feature was purely a mid to upper level phenomenon and was not observed in any of the analyses on level 3. Showers associated with this wave were located across central Kansas (Fig. 2).

The 500 mb NOSAT variational height analysis (Fig. 9b) in general tended to smooth the initial objective height field. The western part of the long wave trough over the western U. S. was filled about 5 - 15 m in response to variational melding with a very poor wind field analysis that failed to produce the intense jet streak in an area centered over east central Nevada. The short wave was filled by approximately 5 m with a concomitant 5 m lowering of heights within the crest of the ridge to the north. These changes produced a reduction in the intensity of both features.

Further east, the weak perturbation over the Great Lakes region was filtered from the NOSAT analysis. Finally, an approximate 5 m filling occurred within the New England trough.

The SAT height analysis for level 6 is shown in Fig. 9c. The major trough was filled about 10 m over Nevada in response to the mutual adjustment with the poorly resolved wind field there as discussed in reference to the NOSAT height analysis. The major departure is the large region of deepening within the trough over the northwest quarter of the grid. Departures in excess of 45 m from the objectively analyzed heights were found over eastern Montana.

The reason for these large departures was discussed in reference to the level 3 height analysis. Satellite temperature data did not exist for the area and relatively cold temperatures along the edge of the available data were interpolated into the data void region. Thus, through adjustments to satisfy the hydrostatic equation, the SAT analysis lowered the heights in response to cold thicknesses in the temperature field.

Examples of the objective and variational temperature analyses for level 3 are shown in Fig. 10. The objective analysis developed a baroclinic zone in the central Rockies in association with the major trough in the height field (Fig. 7a). A large mass of warmer air extended northward over the entire Great Plains ahead of the trough. This feature was interrupted by a small pocket of cooler temperatures over north central Kansas and south central Nebraska that was associated with the short wave at 500 mb (Fig. 9a).

A second baroclinic region was produced by the objective analysis over the southern Plains, and a third, stronger thermal gradient from Minnesota to North Carolina. This latter feature was most noticeable over the western Appalachians in the form of an intense protrusion of cold air over

northeastern Tennessee and eastern Kentucky. An anomalously cold temperature at the Huntington, West Virginia, rawinsonde site produced the pattern. The pattern existed at all levels below 700 mb (level 4) and thus could have been an error in the data or a real observation through the top of a shallow cold air mass that upwelled along the eastern mountains and pushed southward by northerly flow behind the New England trough. The feature was not supported in either the height or wind fields and was not retained in either the NOSAT or SAT analyses.

The large scale features produced by the NOSAT temperature analysis on this level (Fig. 10b) were similar to the objective analysis in both pattern and magnitude but was much smoother across the Midwest where NOSAT temperatures were 2 - 5C warmer. The intensity of the gradient was reduced by more than one-half and was shifted northward over the Great Lakes. This loss of horizontal detail along a shallow, sloping baroclinic zone is at least partly the result of decomposition of the mean layer temperature obtained through the hydrostatic equation. Elsewhere, temperature differences between the two analyses were substantially less; however, all baroclinic zones were smoothed from their objectively-generated appearance.

An objective analysis of temperatures retrieved from satellite is shown in Fig. 10c. Although large scale features were similar to the rawinsonde temperatures in Fig. 10a, substantially colder temperatures were found within the long wave trough. These cooler temperatures were well documented by satellite soundings that show the data void region was further north. The warmer air of the Plains was centered further to the southeast over eastern Nebraska and Kansas. However, the center of warm

temperatures apparent in the rawinsonde data was located entirely within the data void region in the satellite data. These differences shifted the primary baroclinic zone to the east and south, but its intensity remained unchanged. A weak thermal gradient in the satellite temperatures also apparent over the eastern part of the country did not support the strong gradients shown in Fig. 10a over the central Appalachians.

The SAT variational analysis (Fig. 10d) located the large scale baroclinic zones with roughly the same orientation and intensity over the southern Plains and the Ohio Valley as did the NOSAT analysis. In contrast, the SAT analysis was colder by approximately 2C over the Southwest and the primary baroclinic zone was established farther east over Colorado and New Mexico. SAT temperatures, 2 - 5C colder than the original analysis, were found over the data rich areas of the central Rockies and interpolated into the data void region to the north.

The objective temperature analysis at 500 mb again showed a strong baroclinic zone in the west associated with the long wave trough (Fig. 11a). Elsewhere, a relatively flat temperature field replaced the strong zonal temperature gradient found at level 3. A small area with cold temperatures was found over northeastern Kansas in association with the middle tropospheric trough.

Both the NOSAT and SAT analyses (Fig. 11b and 11d) weakened the temperature gradient preceding the major trough. Both analyses weakened the long wave ridge over the northern Plains and smoothed out the cold pocket associated with the central Plains trough although a flat, possibly weakly reversed thermal pattern remains. The SAT analysis increased the

intensity of the baroclinic zone over Oklahoma and Texas in comparison with the NOSAT analysis; however, both analyses produced weaker gradients than were observed in the initial objective analysis of rawinsonde temperatures. Both NOSAT and SAT analyses radically altered the temperature field near the left boundary of the domain in response to the fixed height field boundaries and concomitant poor wind field analyses there. Boundary effects are mostly contained in the area between the dashed rectangle and the boundaries, and these areas are not included in our evaluation of the variational models.

Objective analysis of the satellite temperatures on level 6 (Fig. 11c) was not substantially different from the rawinsonde analysis in Fig. 11a. Rawinsonde temperatures were slightly colder in the Rockies with a more intense baroclinic zone along the eastern Rockies. However, the failure of the satellite temperature analysis to develop a thermal gradient of this strength was related again to the satellite data void region in the vicinity of warmest temperatures. Across the Plains, including the area of the weak short wave, temperature differences were less than a degree. Further east, satellite temperatures were slightly warmer over the Southeast and colder to the north, generating a moderate thermal gradient over the central Appalachians that was not observed in Fig. 11a.

Streamline and isotach maps for the initial objective analysis and NOSAT analysis are presented in Fig. 12a and 12b for level 3. The SAT variational streamline and isotach analyses on this level closely resemble the NOSAT variational analysis and are not presented.

All of the analyses produced a large cyclonic circulation in the streamline pattern along the Rockies from north central Montana to western Colorado. Although the alignment of the trough coincided well with the low in the height field (Fig. 8), there were significant differences in the streamlines between the objective and NOSAT analysis. The objective analysis (Fig. 12a) developed a sharp trough line that stretched southward from a center of cyclonic flow over north central Montana to a secondary center in northwestern Colorado. Comparison with the height field (Fig. 8a) shows that there existed strong ageostrophic flow west of the trough axis. This anomalous condition was likely the result of the interruption of the wind field by high mountain ridges.

The NOSAT variational analyses produced a much broader, elongated area of cyclonic circulation that slowly converged into the center in Colorado (Fig. 12b). This circulation center was coincident with the position of the height center in Fig. 8b. The dynamic constraints in MODEL I permit frictionless flow at level 3 and the resulting dynamic assimilation forces the wind field toward geostrophic flow. As a result, the objectively-generated cross-contour flow on the western side of the long wave trough was eliminated by the NOSAT analysis.

Placement of southerly winds over nearly the entire Great Plains was very similar between the initial and variational analyses. However, wind speeds calculated by the variational method were 5 - 8 m s^{-1} higher than the univariate objective analysis on both the eastern and western sides of the Colorado cyclone. Isotachs and unit vectors of the vector differences between the NOSAT and initial objective analyses for level 3 are shown in

Fig. 13. Vectors oriented roughly along the streamlines of Fig. 12 identify where the variational analysis increased the wind speeds over the initial objective analysis. The cyclonic pattern of vectors surrounding the circulation center over Colorado shows that the variational analysis increased the strength of the entire circulation by up to 8 - 10 m s^{-1} over the mountainous areas.

Elsewhere, in the eastern part of the country, two areas show relatively large differences that are a result of large directional changes in the wind necessary to bring the initial wind field into variational adjustment with the height field.

The 500 mb (level 6) objective streamline and NOSAT variational analysis are shown for comparison in Fig. 14. The patterns over the interior of the western half of the grid are similar. The flow has been turned more northwesterly to bring the wind field into balance with the height field over the eastern interior of the grid. Rather large modifications are found near the grid boundaries where the variational adjustments for both NOSAT and SAT placed jet streaks over California, western Texas, and the Great Lakes. Anthes et al. (1982) showed a 50 m s^{-1} wind speed maximum in southern California in an analysis that used a site located within the jet maximum, an observation not available to our analysis.

The vector difference (Fig. 15) at 500 mb showed the appearance of the jet streaks north of the Great Lakes and Texas and the removal of the wind maximum south of California to the actual jet streak location further north. The southerly vectors over Utah on the west side of the trough show

where the NOSAT winds were reduced to accommodate the slight filling in this area. Virtually no vector difference was found over the Plains, including the area of the short wave in Kansas.

c) Pattern Recognition. Results for the Hypersensitive Variables

The variational assimilation has produced significant adjustments in height, temperature, and wind velocity in order that the values of these variables are solutions of the dynamic constraints. However, these modifications can cause large and physically unrealistic changes in other important variables such as vorticity, divergence, and vertical velocity and other quantities that involve derivatives of the basic variables. In addition, the local tendencies of the horizontal velocity components are sensitive to small errors in the basic variables when they are determined from the arithmetic sum of the other terms of the horizontal momentum equations. The patterns of these hypersensitive variables must be physically realistic when compared with other data sets such as cloud fields, precipitation, and independent measurements of the variable itself. Thus, the hypersensitive variables provide a critical test of the accuracy of the MODEL I dynamic assimilation.

Patterns of the relative vorticity of the observed wind and the vorticity of the NOSAT variational assimilated wind for level 6 (500 mb) are shown in Fig. 16. The SAT vorticity fields are similar to the NOSAT vorticity fields save for a slight increase in cyclonic vorticity over Wyoming, Montana, and the Dakotas as the wind field was modified slightly to balance the 40 m SAT height anomaly (Fig. 9c) and will therefore not be

included in the discussion. Comparison of the figures reveals that the variational analysis increased the detail in the relative vorticity patterns. The magnitude of the vorticity center over Kansas was reduced slightly while negative vorticity was increased across the ridge in response to increased curvature and shear. Vorticity centers and the associated gradients were strengthened over the Eastern States and over Nevada and California in response to intensification of the jet streaks near the major troughs.

Figure 17 shows the vertical velocities in cm sec^{-1} at level 6 (500 mb) for 1200 GMT 10 April 1979 as developed from the kinematic method of O'Brien (1970) applied to the objective analyses of the initial observations and from the more general kinematic method that is part of MODEL I. The O'Brien kinematic method, which uses gridded wind data only, produced a general area of rising motion over the Rocky Mountain States and the plains west of the Mississippi River within the area of southwesterly flow in advance of the upper tropospheric trough and behind the preceding upper level ridge (Fig. 17a). General subsidence is over the Eastern States in advance of the ridge. Upon closer inspection, it is seen that the field of rising motion breaks down into several large vertical velocity centers (magnitudes greater than 10 cm/sec) which are located between rawinsonde sites (see Fig. 3). These strong ascent centers are separated by areas of weak vertical motion of either sign. This type of pattern (vertical velocity centers located between data collection sites) is typical of the kinematic technique when applied to univariate objective analyses of the components of the horizontal wind field. Our wind data was gridded with a successive corrections method (Achtemeier, 1986b) that is an

adaptation of the method developed by Barnes (1964).

The relationship between this vertical velocity field and the precipitation areas as observed by radar (shaded areas) within the region of interest for this study, namely, the weak short wave trough over Kansas and Missouri, is not quite fortuitous. Precipitation is not located within the 10 cm sec^{-1} vertical velocity center located over southeastern Colorado nor within the 4 cm sec^{-1} center over northern Missouri. However, precipitation is located along the axis of upward motion that connects these two centers. The precipitation area in Missouri turns southward into Arkansas, a region of weak vertical velocity of both signs. It is also typical with the kinematic method to find areas of precipitation overlapping into regions with downward vertical motion. Several isolated showers are found within an area identified by the GOES visible imagery as containing fields of swelling cumulus, the area located over Oklahoma and northeast Texas. The kinematic analysis positions a center of downward motion over this area.

The vertical motion field developed by the SAT variational analysis (Fig. 17b) is representative of both SAT and NOSAT analyses. This pattern is similar to the kinematic analysis on the large scale in that the ascent areas are positioned from the Mississippi River westward into the Rocky Mountain States. However the smaller scale patterns are different. Gone are the large centers of upward vertical motion. The 10 cm sec^{-1} positive motion center over Montana in the kinematic analysis is not found in the variational analysis. This reduction was brought about by the decrease of convergence in the lower troposphere as the horizontal wind

field was made to more closely balance the height field. The 10 cm sec^{-1} ascent center over southeastern Colorado has been reduced and shifted into western Kansas in alignment with precipitation observed there. The variational analysis also places a 6 cm sec^{-1} center of upward motion over Oklahoma and northeastern Texas as part of a zone of positive vertical velocity that extends from western Kansas through eastern Oklahoma and into Louisiana. This pattern compares favorably with the convective showers over Oklahoma and Texas but is located behind the position of the short wave trough (Fig. 1). It is oriented along the main precipitation band but is generally located to far to the southwest of it. We note that the observations lag the precipitation patterns by approximately one hour and that, given the rapid northeastward movement of the short wave trough, could account for about 50 km of the displacement between the two patterns.

The most notable impact by the variational analysis upon the vertical velocity is the increased detail of the adjusted patterns associated with a vigorous, fast moving short wave and accompanying jet streak over the northeastern United States (see Fig. 14). The kinematic analysis produced a rather large region of subsidence over the Great Lakes States with the maximum sinking exceeding -8 cm sec^{-1} over western Pennsylvania. The variational analysis concentrated the area and increased subsidence to -14 cm sec^{-1} over West Virginia. The variational analysis also increased the magnitude of the upward motion area in the largely data void area along the East Coast and introduced a narrow zone of upward motion along an axis from northern Kentucky to southeast Georgia and beyond. This is a significant departure from the kinematic analysis which produced a broad area of weak subsidence over the same region.

A comparison with the features of the wind field in Fig. 14 reveals that the most prominent subsidence was located along the axis of the jet streak. The narrow band of ascending motion was found near the entrance region along the anticyclonic shear side of the jet streak, an area long noted by those experienced in the motion fields surrounding jet streams as favorable for upward vertical velocities (Riehl, 1952; Reiter, 1961). Thus these rather extreme vertical velocities are not without support from pattern analysis.

The solution sequence of the variational assimilation requires that the developmental components of the local tendencies of u and v be found from the arithmetic sum of the other terms of the horizontal momentum equations. The developmental components are recombined with the advective components, redimensionalized, and expressed as 3-hr changes. These tendencies are compared with the observed 3-hr tendencies of u and v calculated from the high frequency rawinsonde data collected over the central part of the U. S. as part of the NASA-AVE SESAME project and with the 3-hr tendencies calculated with values from the initial gridded fields substituted in place of the adjusted fields in the horizontal momentum equations. In making these comparisons, we assume that the observed 3-hr tendencies represent "ground truth" subject to the following qualifications. First, in keeping with the synoptic scale of the analysis, we have gridded only 3-hr tendencies taken from data collected at standard NWS observing sites. Second, the ground truth tendencies are calculated over the 3-hr interval from 1200-1500 GMT and are therefore centered at 1330 GMT. The tendencies found from MODEL I are centered at 1200 GMT. Therefore, some phase shift should be observed between the patterns.

Third, to the extent that the tendencies calculated from the SESAME data suffer from mesoscale "noise", the patterns will not accurately represent the true pattern of synoptic scale tendencies.

The following figures present the comparisons for the initial field, observed (ground truth), and SAT variational 3-hr v-tendencies for three representative levels within the troposphere. We have omitted discussion of the 3-hr u-tendencies without loss of content. The u-tendency patterns, though different from the v-tendency patterns, are comparatively similar with the poorest results found in the lower troposphere. In addition, the NOSAT results are similar to the SAT results and will not be included in the comparisons.

Figure 18 shows the fields of the three v-tendencies for level 4 (700 mb) for the central region of the U. S. roughly covered by the SESAME-AVE rawinsonde network. The initial field tendencies consist of relatively large amplitude centers of scale roughly equal to the average separation between observing sites. The pattern has no relationship with the position of the short wave trough over Kansas (Fig. 8) nor with the wind field at 700 mb. The pattern is more suggestive of the tendency field that would excite inertial-gravitational oscillations within a numerical model initialized with unadjusted initial data. The pattern also shows little correspondence with the observed 3-hr tendencies (Fig 18b) in either the locations, amplitudes or the scale of the features. The one plausible exception, the -8 m sec^{-1} $3h^{-1}$ center would have had to move 300 km in 1.5 hr in order to locate over Arkansas, a distance about twice as far as expected from the observed speed of movement of weather systems on 10

April 1979. The SAT tendency field (Fig. 18c), though of the same magnitude and scale as the observed tendency field, suffers from the same excessive displacement of the negative center over the Central Plains. Elsewhere, the SAT 3-h tendencies are in general agreement with the observed 3-h tendencies over the eastern and northern portions of the domain.

The initial field 3-h tendencies (Fig. 19a) at level 6 (500 mb) feature a center of large (12 m sec^{-1}) increases in v over Kansas. The observed 3-h tendency field replaces this pattern with a smaller center of opposite sign (Fig. 19b). Only along the eastern part of the domain near the position of the long wave ridge does the initial field analysis have some correspondence with the observed tendencies. The agreement is to the extent that both analyses show increases in v over three hours. The magnitudes and locations of the maxima are more poorly related. The SAT 3-h tendency field (Fig. 19c) reproduces the observed 3-h tendency pattern over the Mississippi Valley and over western Texas. The relative minimum over Kansas and Oklahoma consists of small increases in v and the negative center over northeastern Texas is too far south to relocate at the observed position by 1330 GMT.

Strongest jet stream winds were located at level 8 (300 mb). Large magnitudes and gradients of the velocity can combine to create large tendencies if the terms of the horizontal momentum equations do not compensate. The 3-h tendency field obtained from the initial data (Fig. 20a) is a pattern of large magnitude centers of alternating sign spaced at approximately the average observation separation. These centers imply

unrealistically large changes in v over three hours. With allowances for horizontal displacement of the pattern over 1.5 h, the only correspondence with the observed 3-h tendency field is the sign of the pattern along the eastern part of the domain. The SAT analysis (Fig. 20c) reproduces most of the features of the observed 3-h tendencies. The positive tendency center near the lower boundary of the grid (Texas-New Mexico border) in the SAT analysis appears over the Texas panhandle at 1330 GMT in the observed tendencies (Fig. 20b). Furthermore, the relative minimum over Oklahoma is moved into southeastern Kansas. These displacements are in accord with the rapid northeastward movement of the weather systems through the area. Relative horizontal displacements were smaller within the weaker flow near the long wave ridge over the eastern part of the domain. Here the SAT analysis preserved the area of larger positive v-tendencies but located the maximum over Illinois rather than over Mississippi as found in the observed tendencies.

6. Discussion

This paper has presented the evaluation of a diagnostic multivariate data assimilation method described in a companion paper (Achtemeier et al., 1986). Special emphasis is placed upon incorporating data from diverse sources and, in particular, upon meshing observations from space-based platforms with those from more traditional immersion techniques. Meteorological data are weighted according to the respective "measurement" errors and blended into a hybrid data set that is required to satisfy the two nonlinear horizontal momentum equations, the hydrostatic equation, and an integrated continuity equation for a dry atmosphere as dynamical

constraints. A hybrid nonlinear vertical coordinate allows for the easy incorporation of TIROS-N mean layer temperatures. Coordinate surfaces are pressure surfaces above 700 mb. The nonlinear vertical coordinate also makes possible the removal of much of the local variations with unlevel terrain at levels below 700 mb in the hydrostatic equation and the pressure gradient terms of the momentum equations.

The intent of this paper is to compare multivariate variational objective analyses with and without satellite data with initial analyses and the observations to determine the accuracy and sensitivity of MODEL I to different data sets. Because this assimilation is not an initialization for a numerical prediction model, the often used procedure of determining the best initial analysis by finding the best forecast does not apply. We instead use three diagnostic criteria which, although they may be somewhat more subjective than measures of forecast skill, have found use in the verification of diagnostic analyses. These criteria are measures of a) the extent to which the assimilated fields satisfy the dynamical constraints, b) the extent to which the assimilated fields depart from the observations, and c) the extent to which the assimilated fields are realistic as determined by pattern recognition. The last criterion requires that the signs, magnitudes, and patterns of the hypersensitive vertical velocity and local tendencies of the horizontal velocity components be physically consistent with respect to the larger scale weather systems.

The case study used for the verification of MODEL I was a short wave over the Central Plains on 1200 GMT 10 April 1979. This case was selected because TIROS-N temperature soundings coexist with NASA-AVE 3-hr rawinsonde

data over a large area of the central United States; the latter data set was required to provide verification for the diagnosed 3-hr local tendencies of the horizontal velocity components. In addition, the case was selected because intense mesoscale convective systems were not significantly impacting the large scale dynamics.

Methods to debias the TIROS-N temperatures and to insert the initial data into the variational assimilation model were described. These data received weights according to the respective precision moduli developed in Section 4.

We developed the percent reduction of the initial RMS error as a measure of the convergence of the SAT and NOSAT blended data sets to satisfaction of the four dynamical constraints. Upon application of this measure to the two horizontal momentum equations, we found that the error is approximately cut in half with each cycle through the fifth cycle for the eight adjustable levels of the Model. The solution stabilizes at approximately 90-95 percent error reduction and there is no further significant improvement in the assimilation after the fifth cycle. The RMS error reductions for the integrated continuity and hydrostatic equations ranged from 90-100 percent except for the errors at levels 2 and 3 of the integrated continuity equation which were reduced to approximately 70 percent by the fourth cycle and remained relatively unchanged through the eighth cycle. These improvements for the integrated continuity equation are, of course, dependent upon the magnitudes of the initial RMS errors. We determined the initial vertical velocity by the O'Brien kinematic method. Had we assumed that the initial vertical velocity was zero, the

initial RMS errors would have been much larger than the values used for this study and the RMS error reductions would have been 100 percent by the fourth cycle.

The pattern recognition analysis for the basic fields, height, temperature, and vector wind, revealed that the SAT and NOSAT analyses were similar with the following two exceptions. First, there were larger numerical differences between the SAT analysis and the initial objective analysis than were found between the NOSAT analyses and the initial objective analysis. In this respect, the comparisons are unfair because the initial height fields were in an approximate hydrostatic balance with the NOSAT temperatures. Adjustments between these two fields were only necessary to remove error introduced by the gridding of the data. The satellite temperatures are initially independent of the heights. Thus the adjustments must be made for both gridding error and nonhydrostatic residuals. Second, large areas of the network were void of satellite data. Temperature sounding information was unavailable over New England, Texas, and across the northern Plains to near the Pacific coast north of a line from southwestern Iowa to northeastern Oregon. Temperatures over roughly the northwestern quarter of the network were interpolated from surrounding data rich areas. Thus, important local details of the temperature field were lost. One result was that warm thicknesses (observed by rawinsonde) were replaced by cold thicknesses from data located south of the data void area. Because the G-function that weights the gridpoint values of the initial fields according to data density was set to one for this study, there resulted a large (-40 m) height anomaly in the middle troposphere in comparison with the initial objective analysis.

Streamline and isotach maps for the SAT analyses closely resembled the NOSAT analyses. The variational blending increased the horizontal detail of the wind field in comparison with the initial univariate analyses in the middle troposphere. Rather large modifications of the initial wind field were found near the grid boundaries where the variational adjustments for both NOSAT and SAT placed jet streaks over California, western Texas, and the Great Lakes. The jet streak over California was verified by a study by Anthes et al. (1982) who used data not available to this study to diagnose a 50 m s^{-1} wind speed maximum in southern California.

The most significant departure between the initial and variational analyses was in the lower troposphere where the variational analysis replaced an area of strong ageostrophic flow over the Western States with a broad area of cyclonic circulation that slowly converged into the center of a major low pressure system located in Colorado. The ageostrophic flow was likely the result of the interruption of the wind field by high mountain ridges. The dynamic constraints in MODEL I permit frictionless flow in the free atmosphere over the smoothed terrain that forms the lower coordinate surface and the dynamic assimilation forces the wind field into greater alignment with the height field.

In all other features of the height, temperature, and wind velocity fields, the pattern recognition technique is suggestive that the modifications to satisfy the dynamic constraints does not alter the physical realism of the final fields. However, small modifications can cause large and physically unrealistic changes in other important variables such as vertical velocity and the local tendencies of the horizontal

velocity components which are not observed directly and must be calculated from other variables through the dynamic constraints.

The variational analysis removed or reduced the magnitudes of several large vertical velocity centers (magnitudes greater than 10 cm sec^{-1}) which were placed between rawinsonde sites by the O'Brien kinematic technique applied to the univariate objective analyses of the components of the initial horizontal wind field and developed a zone of positive vertical velocity roughly parallel to the axis of an area of precipitation that was used as a check on the accuracy of the final fields. This pattern compared favorably with the convective showers over Oklahoma and Texas but was located behind the position of the short wave trough over the Central Plains. It was oriented along the main precipitation band but was generally located too far to the southwest of it. However, we note that the observations lagged the precipitation patterns by approximately one hour and that, given the rapid northeastward movement of the short wave trough, could have accounted for about 50 km of the displacement between the two patterns.

The most notable impact by the variational analysis upon the vertical velocity was the increased detail of the adjusted patterns associated with a vigorous, fast-moving short wave and accompanying jet streak over the northeastern United States. The variational analysis concentrated an area of strong subsidence (-14 cm/sec) along the axis of the jet streak. It also placed a narrow band of ascending motion near the entrance region along the anticyclonic shear side of the jet streak, an area long noted by those experienced in the motion fields surrounding jet streams as favorable

for upward vertical velocities. Thus these rather extreme vertical velocities are not without support from pattern analysis.

The solution sequence of the variational assimilation requires that the developmental components of the local tendencies of u and v be found as the summation of the other terms of the horizontal momentum equations. We recombined the developmental components with the advective components, redimensionalized, and expressed the results as 3-hr changes. These tendencies were compared with the observed 3-hr tendencies of u and v calculated from the high frequency rawinsonde data collected over the central part of the U. S. as part of the NASA-AVE SESAME project and with the 3-hr tendencies calculated with values from the initial gridded fields substituted in place of the adjusted fields in the horizontal momentum equations.

We found that the SAT and NOSAT tendency patterns were of approximately the same magnitude and scale as the observed tendency patterns. There was a significant disagreement between the two fields in the placement of a large negative tendency center in the lower troposphere, however. Elsewhere, particularly in the upper troposphere, the agreement among the tendency patterns was very good considering that the observed patterns were subject to interference by mesoscale phenomena and that the observed patterns were valid at 1330 GMT rather than at 1200 GMT. The relative accuracy of the variational tendencies was made more apparent upon comparison of the initial field tendencies with the observed patterns. The initial field tendencies consisted of relatively large amplitude centers of scale roughly equal to the average separation between observing sites. The

magnitudes of these centers became unrealistically large in the upper troposphere where large magnitudes and gradients of the velocity can combine to create large tendencies if the terms of the horizontal momentum equations do not compensate. These patterns have no relationship with the position of the short wave trough over Kansas nor with the wind field nor with the observed 3-h tendencies. The patterns are more suggestive of the tendency fields that would excite inertial-gravitational oscillations within a numerical model initialized with unadjusted initial data.

Now, does this variational assimilation method produce better hybrid data fields than other methods? Since no intercomparison studies have been performed, we cannot offer definitive answers to the question. However, we believe that the variational model should provide quality analyses if the following two criteria are satisfied. First, the variational assimilation method we have developed is a physical model. Four of the basic primitive equations that govern flow in free atmosphere subject to the assumptions that apply to hydrostatic and synoptic conditions have been used in the model derivation. Since the real atmosphere obeys these equations, it is expected that the three dimensional fields of meteorological variables should be reasonable approximations to the true atmosphere if they are solutions of the dynamic equations. Furthermore, advanced versions of this model that include the energy equation as a fifth constraint should provide analyses that are superior to the results presented here.

Second, the dynamical equations permit many solutions. Therefore, the error characteristics of the observations and the horizontal distributions of the G-function that make up the precision moduli should be known with

accuracy. The sensitivity of the variational model to the values given to these weights is currently not fully known and is the subject of investigation in the ongoing model development.

Finally, we note from the results of the pattern recognition that the variational analysis produced physically reasonable fields of the hypersensitive variables. Model I produced the first relatively accurate diagnostic fields of local tendencies of the velocity components apart from initialization schemes for numerical prediction models. Our continued model developments should improve upon these results.

Acknowledgments. This work was supported by the National Aeronautics and Space Administration (NASA) under Contract NAS8-34902. Ms. Donna Isard is acknowledged for the preparation of the figures.

REFERENCES

Achtemeier, G. L., 1972: Variational initialization of atmospheric fields - a quasi-geostrophic diagnostic model. Ph.D. Dissertation, Tallahassee, Florida State University, 101-103.

_____, 1975: On the Initialization problem: A variational adjustment method. Mon. Wea. Rev., 103, 1090-1103.

_____, 1986a: The impact of data boundaries upon a successive corrections objective analysis of limited-area data sets. Mon. Wea. Rev., 114, 40-49.

_____, 1986b: On the notion of varying influence radii for a successive corrections objective analysis. Mon. Wea. Rev., xxx, xxx.

_____, H. T. Ochs III, and J. Chen, 1986: A variational assimilation method for the diagnosis of cyclone systems. Part I: Development of the basic model. To be submitted.

_____, and H. T. Ochs III, 1986: A hybrid vertical coordinate and pressure gradient formulations for a numerical variational analysis model for the diagnosis of cyclone systems. To be submitted.

Anthes, R.A., Y.-H. Kuo, S.G. Benjamin, and Y.-F. Li, 1982: The evolution of the mesoscale environment of severe local storms: Preliminary modeling results. Mon. Wea. Rev., 110, 1187-1213.

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. J. Appl. Meteor. 3, 396-409.

Fuelberg, H. E., 1974: Reduction and error analysis of the AVE II pilot experiment data. NASA CR-120496, Marshall Space Flight Center, Alabama, p. 60

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. J. Appl. Meteor., 11., 1203-1211.

Gruber, A., and C.D. Watkins, 1982: Statistical assessment of the quality of TIROS-N and NOAA-6 satellite soundings. Mon. Wea. Rev., 110, 867-876.

Hess, S.L., 1959: Introduction to Theoretical Meteorology. Holt, Rinehart and Winston, New York, 362 pp.

Hovermale, J. B., 1962: A comparison of data accuracy in the troposphere and in the stratosphere. Dept. Meteorology, Penn. State Univ., Contract AF 19-(604)-6261, Rept. No. 2.

Kaplan, M.L., J.W. Zack, V.C. Wong, and J.J. Tuccillo, 1982: Initial results from a mesoscale atmospheric simulation system and comparisons with the AVE-SESAME I data set. Mon. Wea. Rev., 110, 1564-1590.

Kidder, S. Q., and G. L. Achtemeier, 1986: Day-night variation in operationally-retrieved TOVS temperature biases. Mon. Wea. Rev., xxx, xxx.

Krishnamurti, T. N., 1968: A diagnostic balance model for studies of weather systems of low and high latitudes. Rossby number less than one. Mon. Wea. Rev., 96, 197-207..bl .i-5

Moore, J.T., and H.E. Fuelberg, 1981: A synoptic analysis of the first AVE-SESAME '79 period. Bull. Amer. Meteor. Soc., 62, 1577-1590.

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

Otto-Bliesner, B., D. P. Baumhefner, T. W. Schlatter, and R. Bleck, 1977: A comparison of several data analysis schemes over a data-rich region. Mon. Wea. Rev., 105, 1083-1091.

Phillips, N., L. McMillin, A. Gruber, and D. Wark, 1979: An evaluation of early operational temperature soundings from TIROS-N. Bull. Amer. Meteor. Soc., 60, 1188-1197.

Reiter, E., 1961: Jet Stream Meteorology. Univ. Chi. Press, 515 pp.

Riehl, H., 1952: Forecasting in middle latitudes. Meteor Monog., 5, Boston, 80 pp.

Sasaki, Y., 1958: An objective analysis based upon the variational method. J. Meteor. Soc. Japan, 36, 77-88.

_____, 1970: Some basic formalisms in numerical variational analysis. Mon. Wea. Rev., 98, 875-883.

Schlatter, T.W., 1981: An assessment of operational TIROS-N temperature retrievals over the United States. Mon. Wea. Rev., 109, 110-119.

U.S. Standard Atmosphere, 1962: U.S. Government Printing Office, Washington, D.C.

Whittaker, E., and G. Robinson, 1926: The Calculus of Observations, (2nd Edition). London, Blackie and Son, Ltd., p.176.

Table 1

Biases and standard deviations of Tiros-N layer mean virtual temperatures (K).

Layer (mb)	Day			Night		
	Clear	Partly Cloudy	Cloudy	Clear	Partly Cloudy	Cloudy
			Biases			
200-100	-0.02	-0.05	0.54	0.19	0.18	0.51
300-200	0.98	0.65	1.20	0.61	1.49	1.62
400-300	-0.04	0.20	-0.41	0.41	0.26	-0.02
500-400	-0.87	0.33	-1.14	0.14	-0.25	-1.11
700-500	-0.90	0.29	-1.35	-0.10	-0.23	-1.37
850-700	-0.16	0.67	-0.65	0.09	1.08	-0.31
1000-850	0.32	1.03	1.87	-0.42	1.35	2.06
			Standard Deviations			
200-100	1.29	1.63	1.48	1.38	1.12	1.47
300-200	1.57	1.40	1.74	1.18	1.34	1.45
400-300	1.50	1.37	1.83	1.38	1.40	1.99
500-400	1.43	1.43	1.70	1.59	1.30	1.89
700-500	1.62	1.72	1.77	1.25	1.44	1.77
850-700	1.84	2.42	2.31	1.84	2.28	2.60
1000-850	2.11	2.52	2.60	2.64	3.12	3.39

Table 2

Correlation coefficients between advective temperature change and observed temperature change over a 3 hr period.

Level (mb)	Correlation Coefficient
850	.20
700	.27
500	.02
400	.04
300	.09
200	.11
100	.04

Table 3

RMS errors of observation for meteorological variables that are observed. Units are standard.

Model Level	Pressure (mb)	VARIABLE u^1	u^2	H	Rawin T	Satellite T(c1)	T(cy)
10	100	4.5	2.3	25.0	2.0	1.9	1.9
9	200	4.5	2.3	19.8	3.0	2.4	2.6
8	300	4.2	2.1	18.0	3.0	2.0	2.7
7	400	3.6	1.8	15.0	2.6	1.9	2.6
6	500	3.2	1.6	11.6	2.0	1.8	2.4
5	600	3.0	1.5	9.2	1.5	1.7	2.3
4	700_3	2.8	1.4	7.8	1.5	2.0	2.8
3	800_3	2.4	1.2	7.0	1.5	2.0	2.8
2	900_3	2.1	1.1	6.5	1.5	2.4	3.5
1	1000_3	2.0	1.0	6.0	1.5	3.0	4.0

^{1}
^{2}...at 20 degree elevation angle
^{3}...at 40 degree elevation angle
...approximate pressure level

Table 4

Nondimensional RMS errors for all variables to be adjusted.

Model Level	Pressure (mb)	VARIABLE u	u	H	H	H	Mean Temperature Rawin	T(cl)	T(cy)		
											0.00
10	100	0.45	0.23	0.25	0.71						
						3.68	0.59	0.57	0.57	2.13	6.98
9	200	0.45	0.23	0.20	0.56						
						3.21	0.88	0.70	0.76	1.88	6.98
8	300	0.42	0.21	0.18	0.51						
						2.28	0.88	0.59	0.79	1.64	6.51
7	400	0.36	0.18	0.15	0.42						
						1.53	0.76	0.56	0.76	1.43	5.58
6	500	0.32	0.16	0.12	0.33						
						0.97	0.59	0.53	0.70	1.24	4.65
5	600	0.30	0.15	0.09	0.26						
						0.61	0.44	0.50	0.67	1.04	4.34
4	700	0.28	0.14	0.08	0.22						
						0.53	0.44	0.53	0.70	0.84	3.72
3	800	0.24	0.12	0.07	0.20						
						0.47	0.44	0.59	0.82	0.64	3.26
2	900	0.21	0.11	0.06	0.18						
						0.42	0.44	0.70	1.03	0.44	3.10
1	1000	0.20	0.10	0.06	0.17						

Table 5

Nondimensional precision modulus weights for MODEL I variational assimilation.

Model Level	Pressure (mb)	VARIABLE u	H	H	H	Mean Temperature Rawin	T(c1)	T(cy)			
									10^6		
10	100	2.5	8.0	1.0							
					0.04	1.4	1.5	1.5	0.11	0.01	
9	200	2.5	12.5	1.6							
					0.05	0.6	1.0	0.9	0.14	0.01	
8	300	2.8	15.4	1.9							
					0.10	0.6	1.4	0.8	0.19	0.01	
7	400	3.9	22.2	2.8							
					0.21	0.9	1.6	0.9	0.24	0.02	
6	500	4.9	34.7	4.6							
					0.53	1.4	1.8	1.0	0.33	0.02	
5	600	5.6	61.7	7.4							
					1.34	2.6	2.0	1.1	0.46	0.03	
4	700	6.4	78.1	10.3							
					1.78	2.6	1.8	1.0	0.71	0.04	
3	800	8.7	102.0	12.5							
					2.26	2.6	1.4	0.7	1.22	0.05	
2	900	11.3	138.9	15.4							
					2.83	2.6	1.0	0.5	2.58	0.05	
1	1000	12.5	138.9	17.3							

Table 6

Percent NOSAT RMS error reduction with respect to initial RMS residuals for the u- and v-horizontal momentum equations.

CYCLE	LEVEL 2	3	4	5	6	7	8	9
				u-component				
0	0	0	0	0	0	0	0	0
1	52	52	52	50	50	49	50	49
2	78	77	76	73	73	73	75	75
3	90	88	87	84	85	84	86	86
4	95	93	91	90	90	90	91	90
5	96	95	93	92	93	93	92	91
6	95	95	93	93	94	94	93	91
7	94	94	93	93	94	94	92	91
8	92	94	92	92	93	93	92	90
				v-component				
0	0	0	0	0	0	0	0	0
1	52	52	51	52	50	50	50	49
2	75	78	76	78	76	75	75	73
3	84	87	86	89	87	86	86	83
4	88	90	89	91	90	90	90	86
5	92	92	90	91	90	91	90	87
6	94	93	90	91	90	91	90	87
7	94	93	90	90	90	90	90	87
8	90	93	90	90	90	90	89	86

Table 7

Percent NOSAT RMS error reduction with respect to initial RMS residuals for the integrated continuity and hydrostatic equations.

CYCLE	LEVEL 2	3	4	5	6	7	8	9
				Integrated Continuity				
0	0	0	0	0	0	0	0	0
1	50	50	50	50	50	50	50	50
2	65	60	39	51	57	57	63	72
3	70	65	43	56	62	60	67	78
4	70	68	55	65	71	69	75	82
5	70	69	66	74	79	78	82	84
6	71	70	75	81	85	85	86	87
7	70	69	82	87	90	90	90	89
8	70	68	87	90	92	93	92	91
				Hydrostatic				
0	0	0	0	0	0	0	0	0
1	50	50	50	50	50	50	50	50
2	75	75	75	75	75	75	75	75
3	87	88	88	87	87	87	87	87
4	93	94	94	94	94	94	94	94
5	96	96	97	97	97	97	97	97
6	98	98	98	98	98	98	98	98
7	98	98	99	99	99	99	99	99
8	98	98	100	100	100	100	100	100

FIGURE CAPTIONS

Fig. 1. The 500 mb height field at 1200 GMT, 10 April 1979 showing a weak short wave disturbance over the Central Plains.

Fig. 2. Middle and upper tropospheric cloud (bright white) and low cloud (light gray) as observed in IR from GOES-EAST and precipitation echoes from radar summary charts in association with the short wave over the Central Plains at 1200 GMT, 10 April 1979.

Fig. 3. The locations of the data used for this study; (a) rawinsonde data for the NWS synoptic network, and (b) TIROS-N temperature soundings.

Fig. 4. The grid template for the variational assimilation model.

Fig. 5. (a) The heights of the lowest coordinate surface before transformation and (b) the heights remaining after transformation of the hydrostatic equation into "equivalent pressure surfaces".

Fig. 6. Asterisks indicate the location of the 101 rawinsonde stations used to construct the objective analyses for comparison with satellite soundings. The dashed line encloses the satellite soundings. Note that the satellite soundings are well within the boundaries of the rawinsonde objective analysis area, thus edge effects should be minimal.

Fig. 7. 12 h average biases of satellite soundings: (a) clear soundings, (b) partly cloudy soundings, (c) cloudy soundings. The error bars represent the 95% confidence interval. Dashed lines represent the average of the biases. The temperature scale on the right is in Kelvin.

Fig. 8. Height (m) analyses on level 3 (approximately 800 mb) at 1200 GMT, 10 April 1979: (a) objective analysis, (b) NOSAT analysis, and (c) SAT analysis. Dashed lines on the NOSAT and SAT analyses represent the adjustment in height from the analysis.

Fig. 9. Same as Fig. 8, but for level 6 (500 mb).

Fig. 10. Temperature (C) on level 3 at 1200 GMT, 10 April 1979: (a) objective analysis of rawinsonde temperatures, (b) NOSAT analysis

Fig. 10 (continued). (c) objective analysis of satellite temperatures, and (d) SAT analysis. Satellite sounding locations from Fig. 3b are superimposed in (c).

Fig. 11. Same as Fig. 10, but for level 6.

Fig. 11 (continued).

Fig. 12. Streamline (solid lines) and isotach (dashed lines, m s^{-1}) analyses on level 3 at 1200 GMT, 10 April 1979: (a) objective analysis and (b) NOSAT analysis.

Fig. 13. Isotachs (m s^{-1}) and unit vectors of the vector difference between the NOSAT and objective analyses on level 3 at 1200 GMT, 10 April 1979.

Fig. 14. Same as Fig. 12, but for level 6.

Fig. 15. Same as Fig. 13, but for level 6.

Fig. 16. Patterns of a) the relative vorticity of the observed wind and b) the vorticity of the NOSAT variational assimilated wind for level 6 (500 mb) at 1200 GMT, 10 April 1979. Units are in 10^{-5} s^{-1}.

Fig. 17. Vertical velocities (cm s^{-1}) on level 6 (500 mb) at 1200 GMT, 10 April 1979 from a) the kinematic method (O'Brien, 1970) applied to the objective analyses of the initial wind field and b) from the MODEL I adjusted wind field. Hatching delineates areas of precipitation.

Fig. 18. Fields of a) initial, b) observed, and c) SAT v-tendencies for level 4 (700 mb) for the central region of the U. S. roughly covered by the SESAME-AVE rawinsonde network. Units are in m s^{-1} $3h^{-1}$.

Fig. 19. Same as Fig. 18, but for level 6 (500 mb).

Fig. 20. Same as Fig. 18, but for level 8 (300 mb).

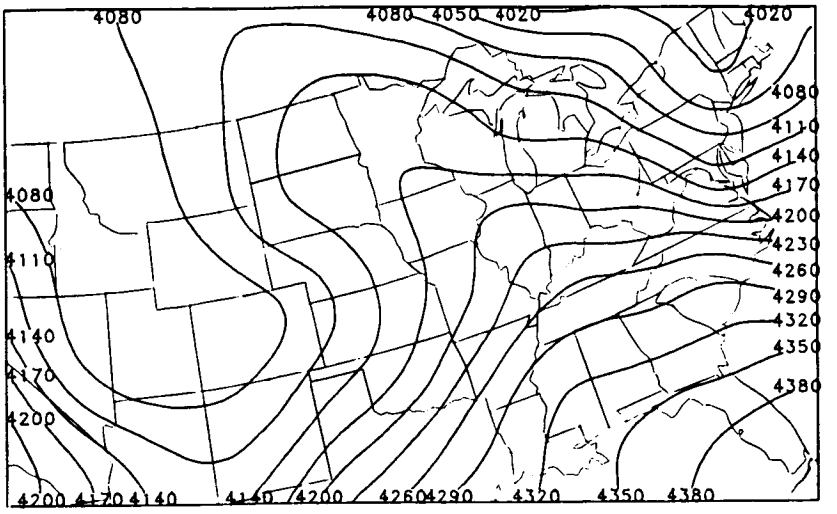


Fig. 1. The 500 mb height field at 1200 GMT, 10 April 1979 showing a weak short wave disturbance over the Central Plains.

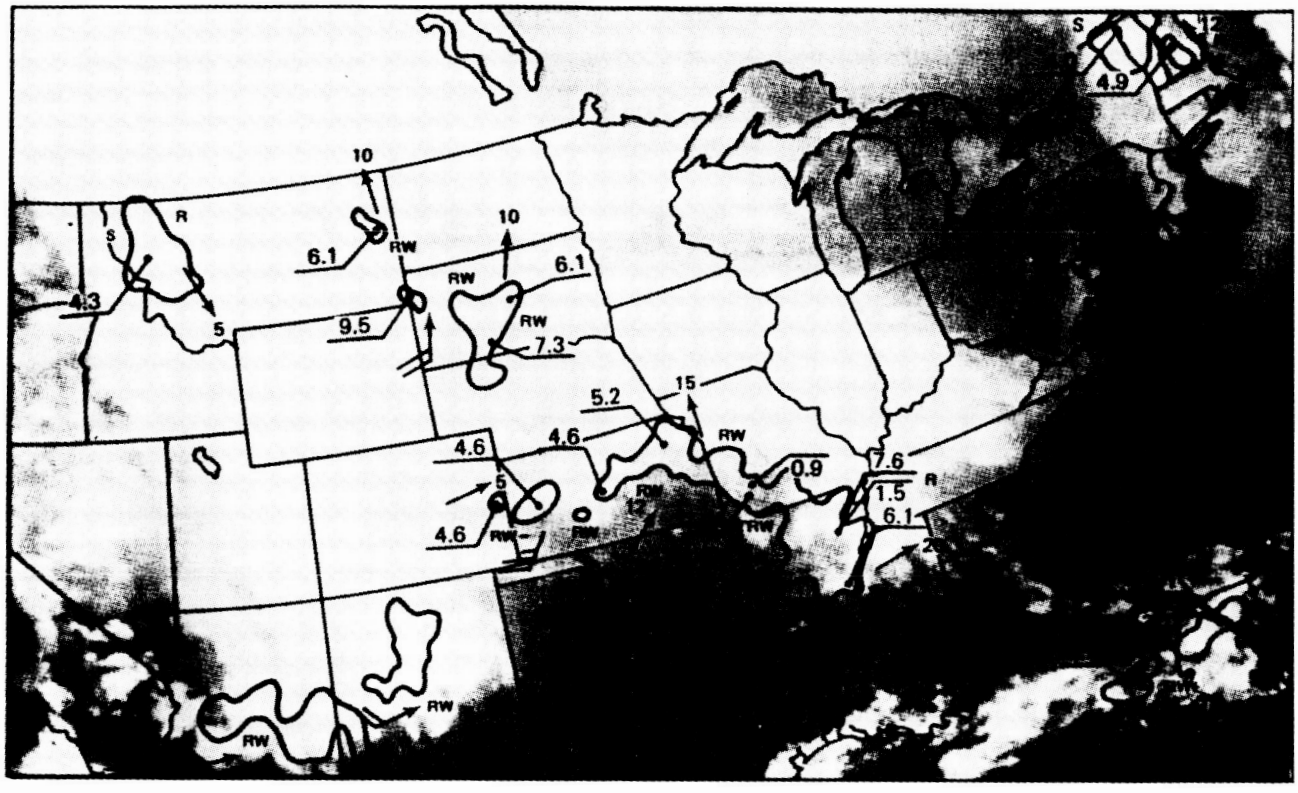
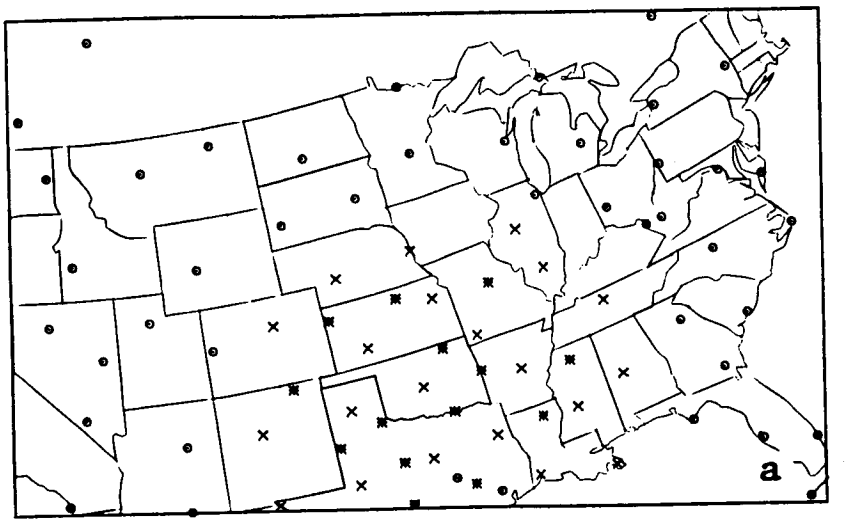
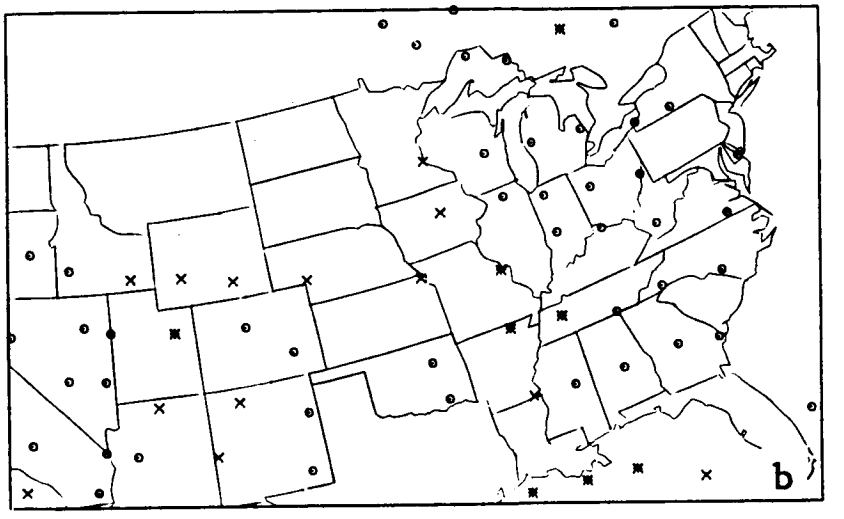


Fig. 2. Middle and upper tropospheric cloud (bright white) and low cloud (light gray) as observed in IR from GOES-EAST and precipitation echoes from radar summary charts in association with the short wave over the Central Plains at 1200 GMT, 10 April 1979.



○ = 12 HOUR SOUNDINGS ONLY
✳ = 3 HOUR SOUNDINGS ONLY
× = BOTH 12 AND 3 HOUR SOUNDINGS



⊙ = CLEAR
✳ = CLOUDY
× = PARTLY CLOUDY

Fig. 3. The locations of the data used for this study; (a) rawinsonde data for the NWS synoptic network, and (b) TIROS-N temperature soundings.

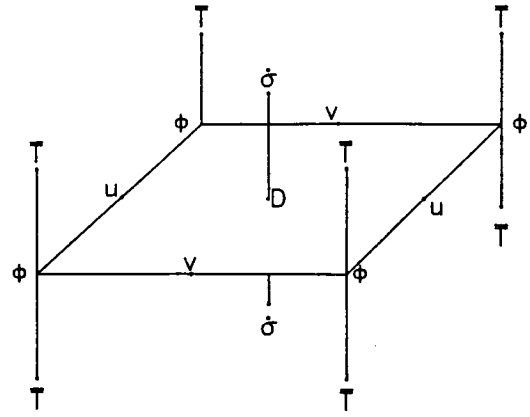


Fig. 4. The grid template for the variational assimilation model.

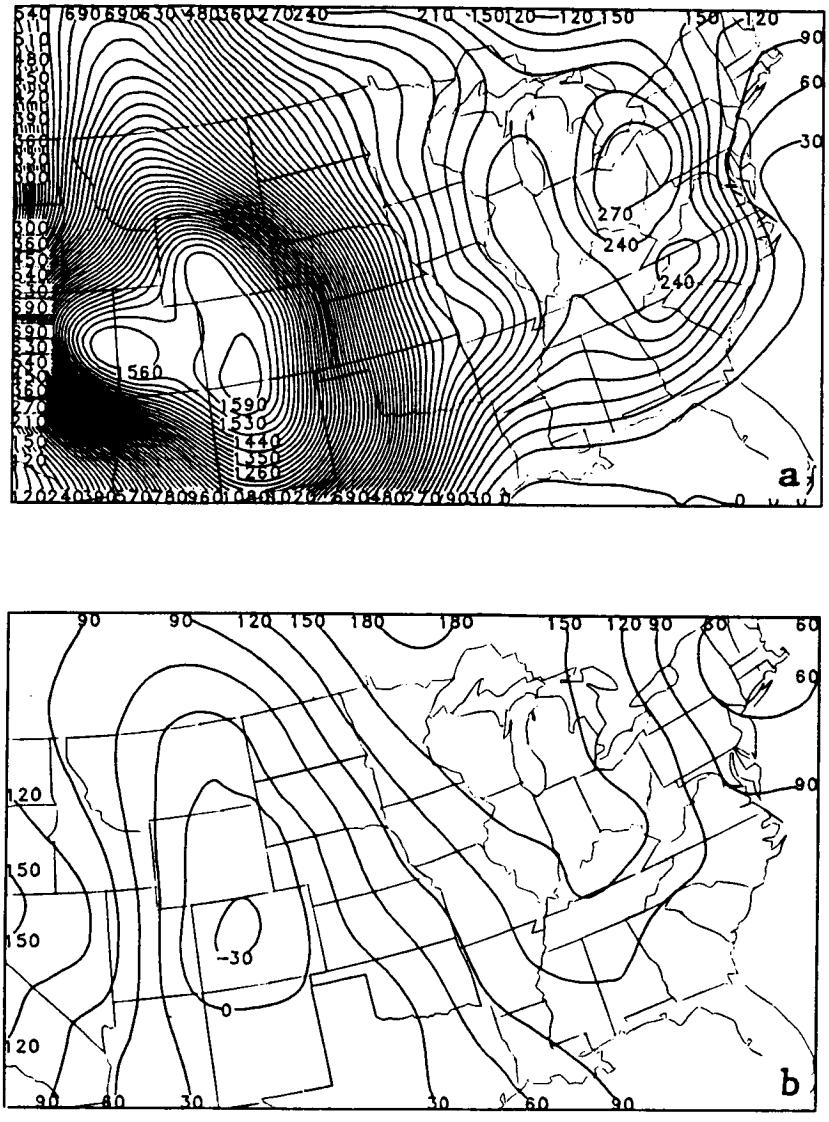


Fig. 5. (a) The heights of the lowest coordinate surface before transformation and (b) the heights remaining after transformation of the hydrostatic equation into "equivalent pressure surfaces".

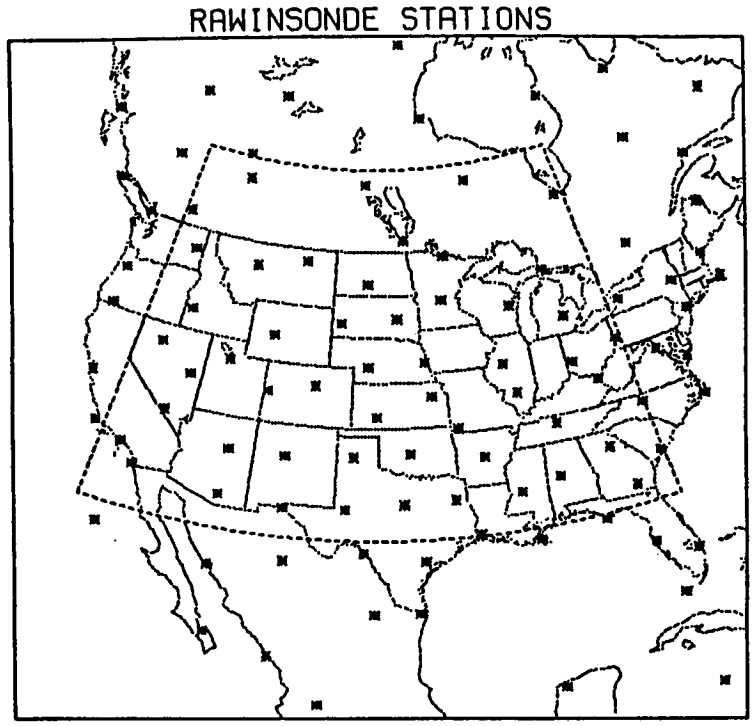


Fig. 6. Asterisks indicate the location of the 101 rawinsonde stations used to construct the objective analyses for comparison with satellite soundings. The dashed line encloses the satellite soundings. Note that the satellite soundings are well within the boundaries of the rawinsonde objective analysis area, thus edge effects should be minimal.

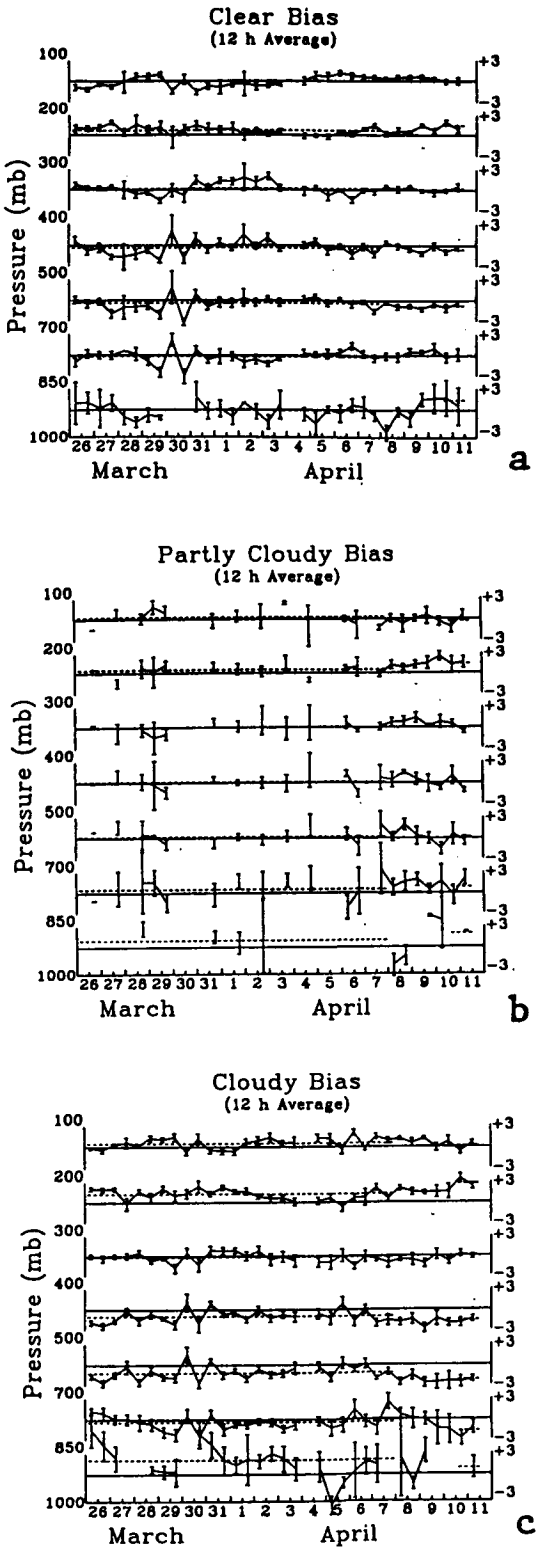


Fig. 7. 12 h average biases of satellite soundings: (a) clear soundings, (b) partly cloudy soundings, (c) cloudy soundings. The error bars represent the 95% confidence interval. Dashed lines represent the average of the biases. The temperature scale on the right is in Kelvin.

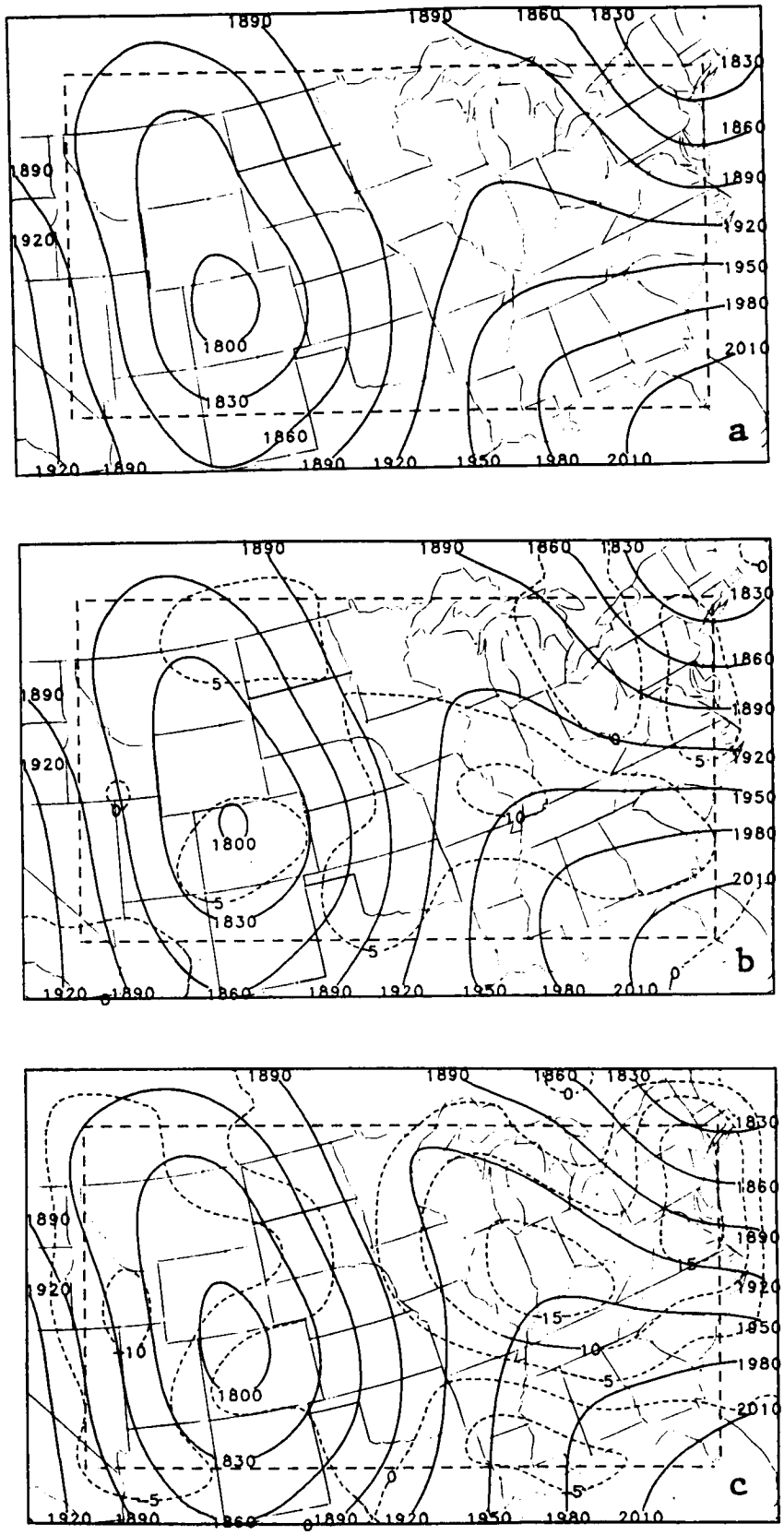


Fig. 8. Height (m) analyses on level 3 (approximately 800 mb) at 1200 GMT, 10 April 1979: (a) objective analysis, (b) NOSAT analysis, and (c) SAT analysis. Dashed lines on the NOSAT and SAT analyses represent the adjustment in height from the analysis.

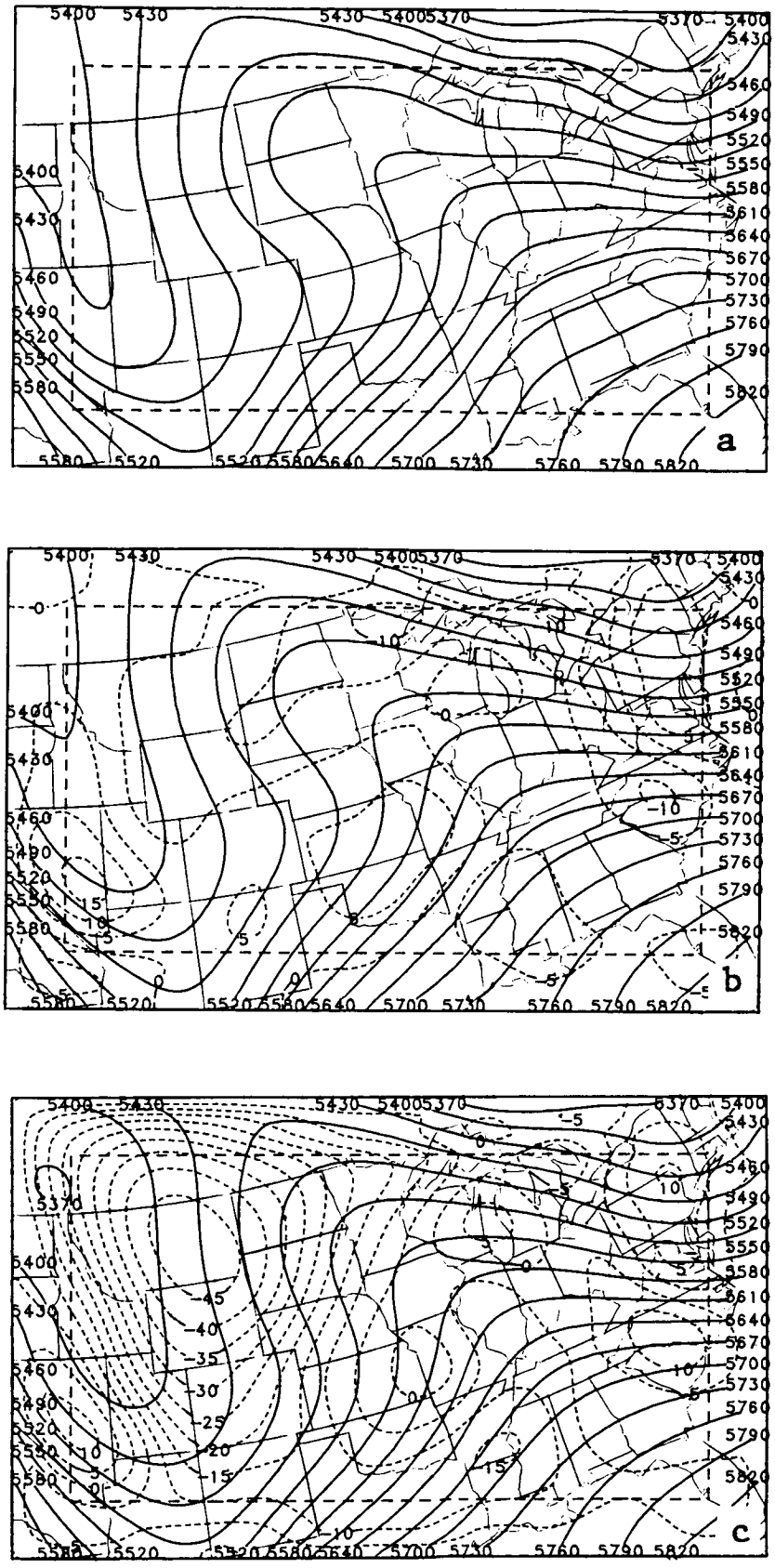


Fig. 9. Same as Fig. 8, but for level 6 (500 mb).

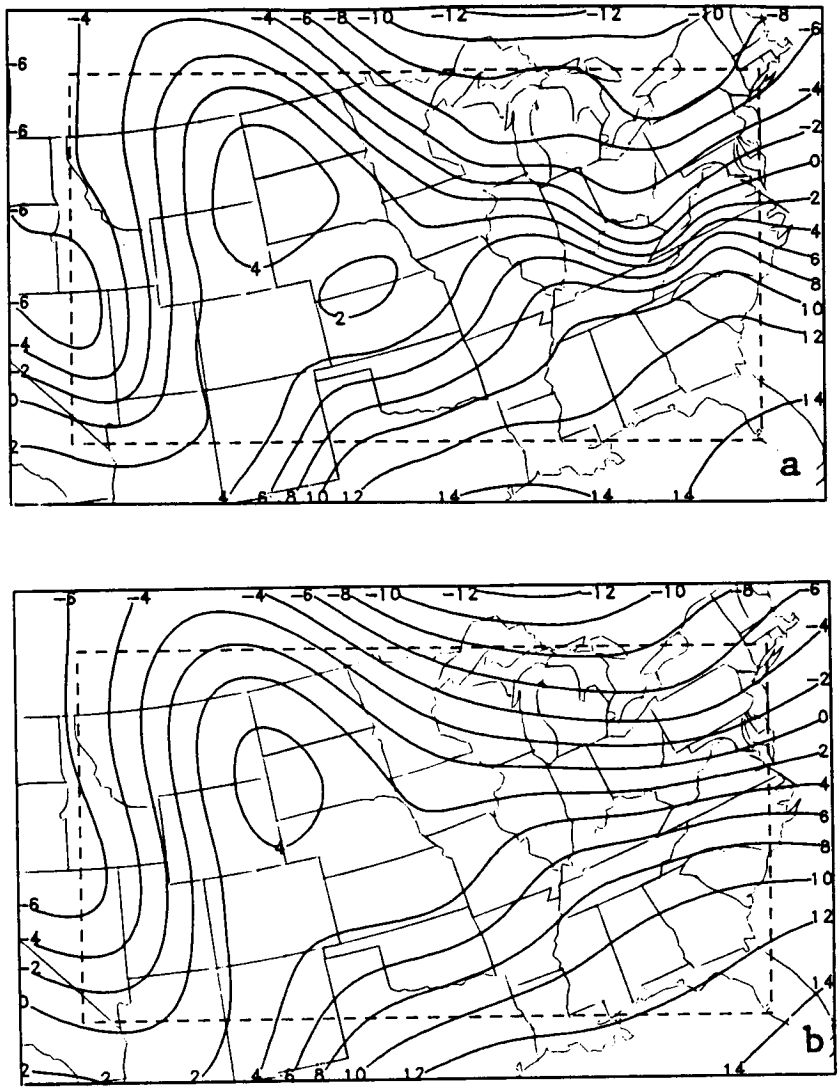


Fig. 10. Temperature (C) on level 3 at 1200 GMT, 10 April 1979: (a) objective analysis of rawinsonde temperatures. (b) NOSAT analysis

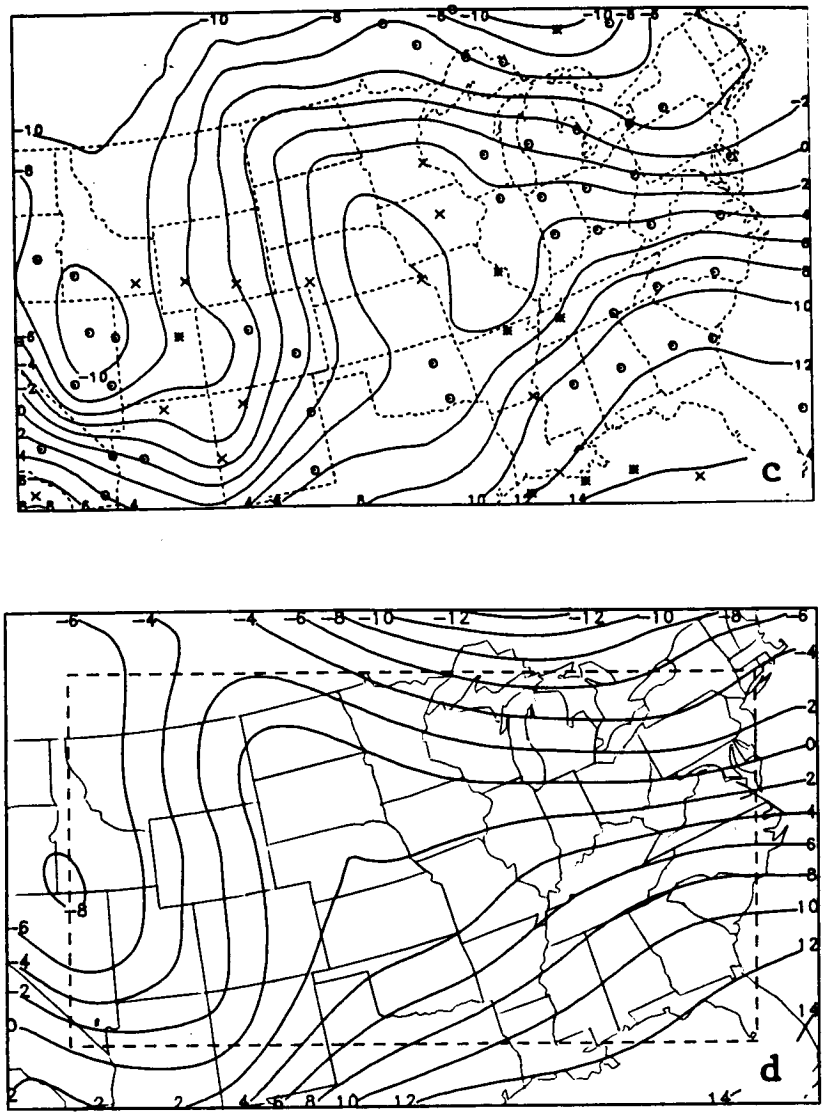


Fig. 10 (continued). (c) objective analysis of satellite temperatures, and (d) SAT analysis. Satellite sounding locations from Fig. 3b are superimposed in (c).

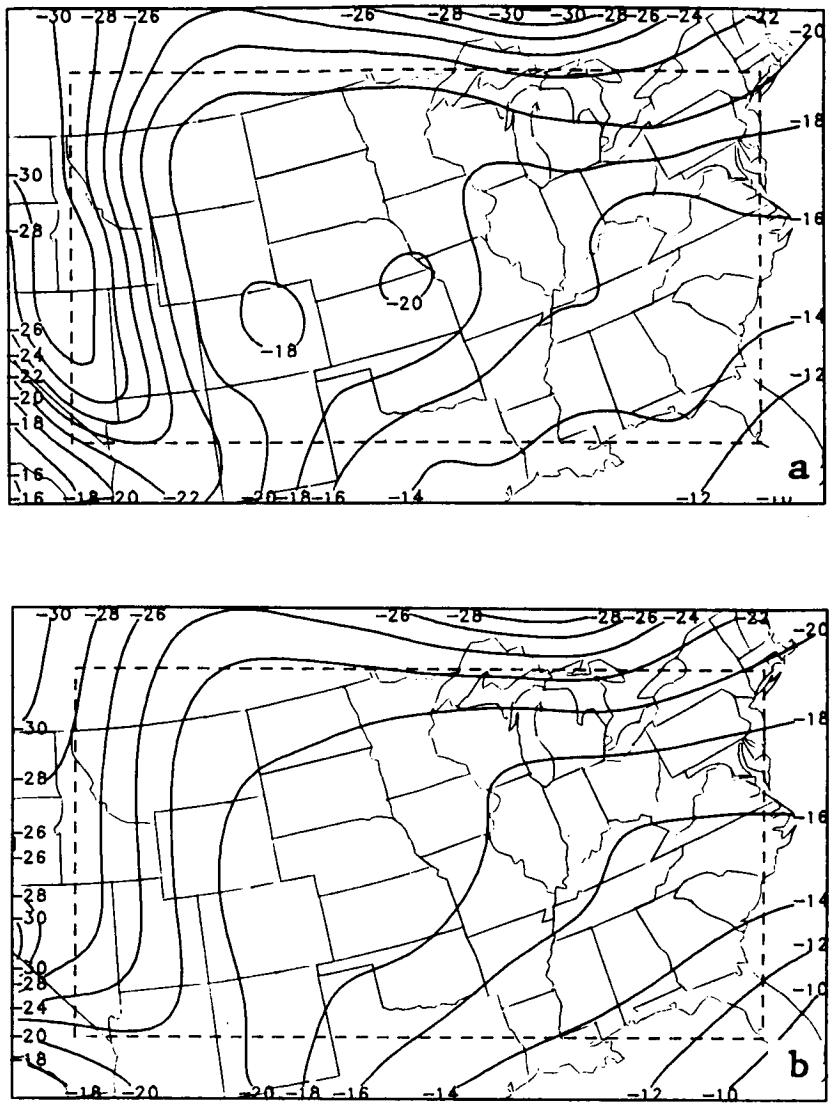


Fig. 11. Same as Fig. 10, but for level 6.

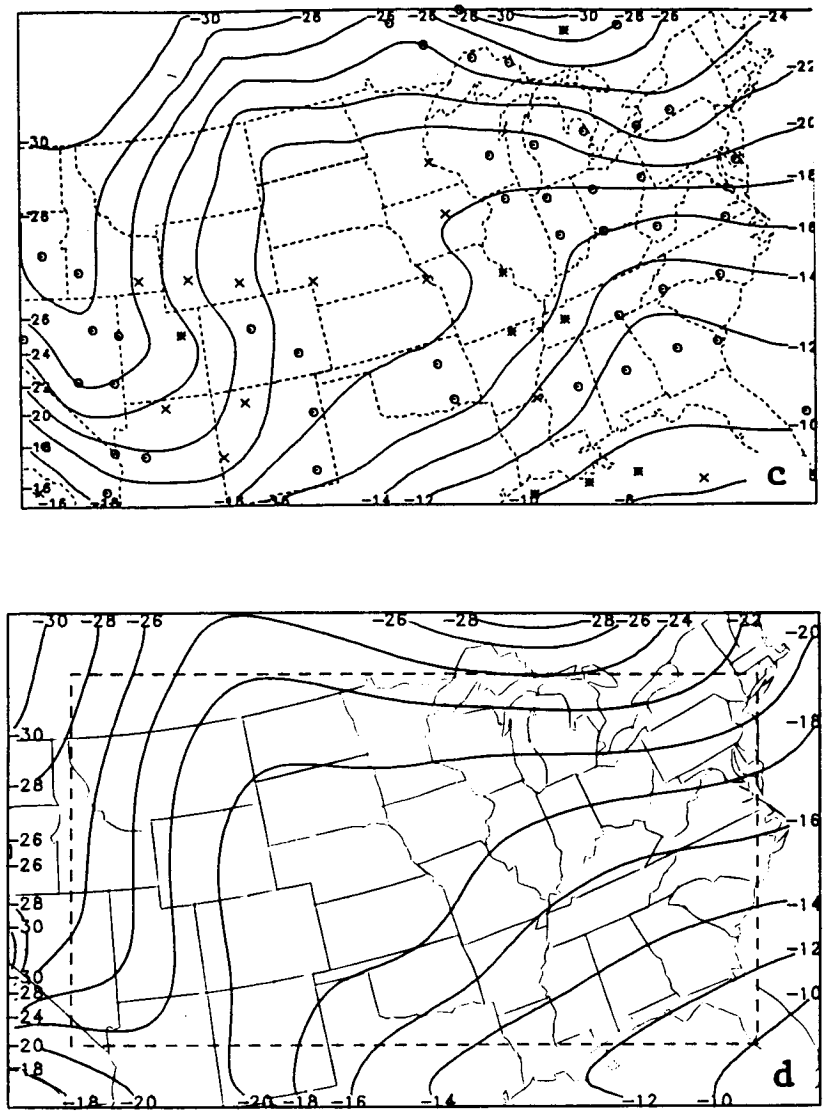


Fig. 11 (continued).

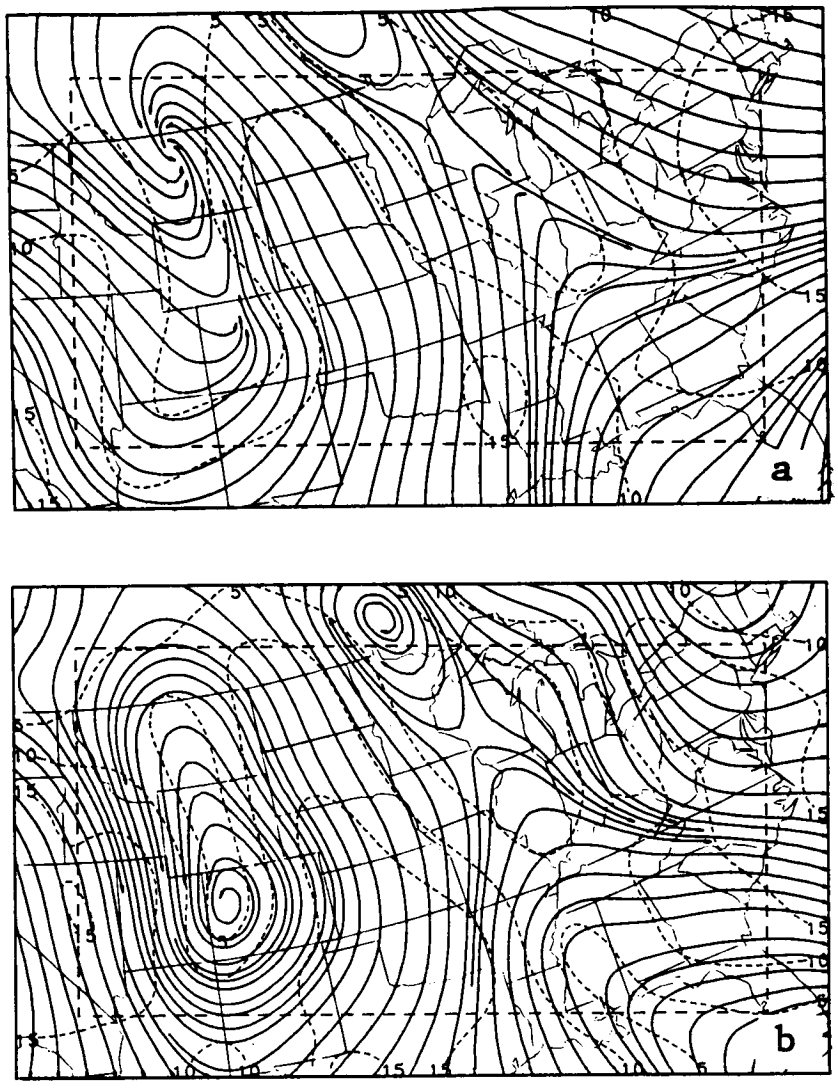


Fig. 12. Streamline (solid lines) and isotach (dashed lines, m s^{-1}) analyses on level 3 at 1200 GMT, 10 April 1979: (a) objective analysis and (b) NOSAT analysis.

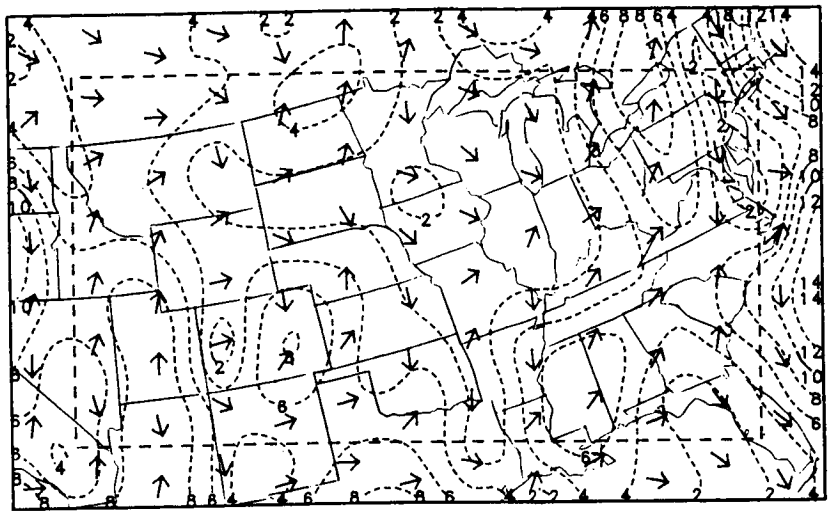


Fig. 13. Isotachs (m s^{-1}) and unit vectors of the vector difference between the NOSAT and objective analyses on level 3 at 1200 GMT, 10 April 1979.

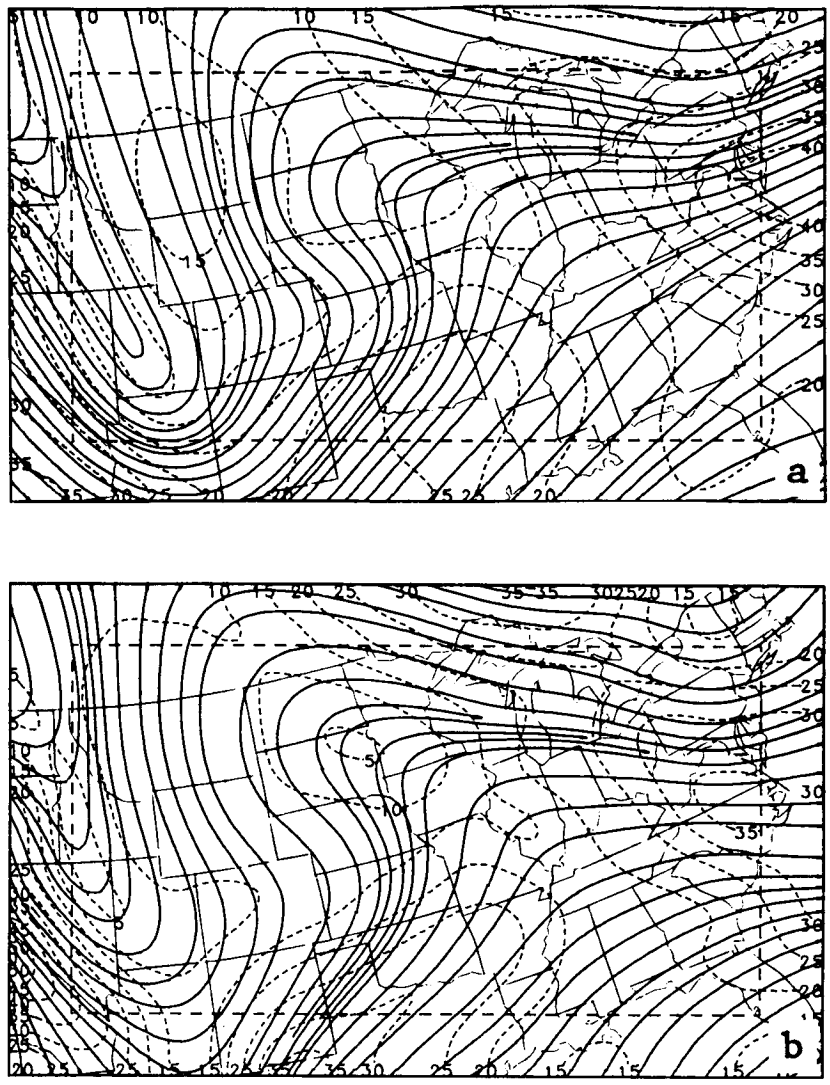


Fig. 14. Same as Fig. 12, but for level 6.

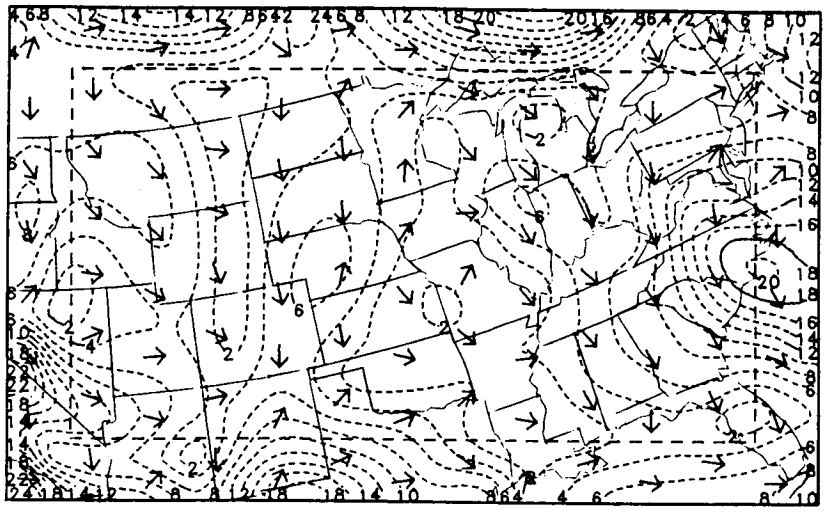


Fig. 15. Same as Fig. 13, but for level 6.

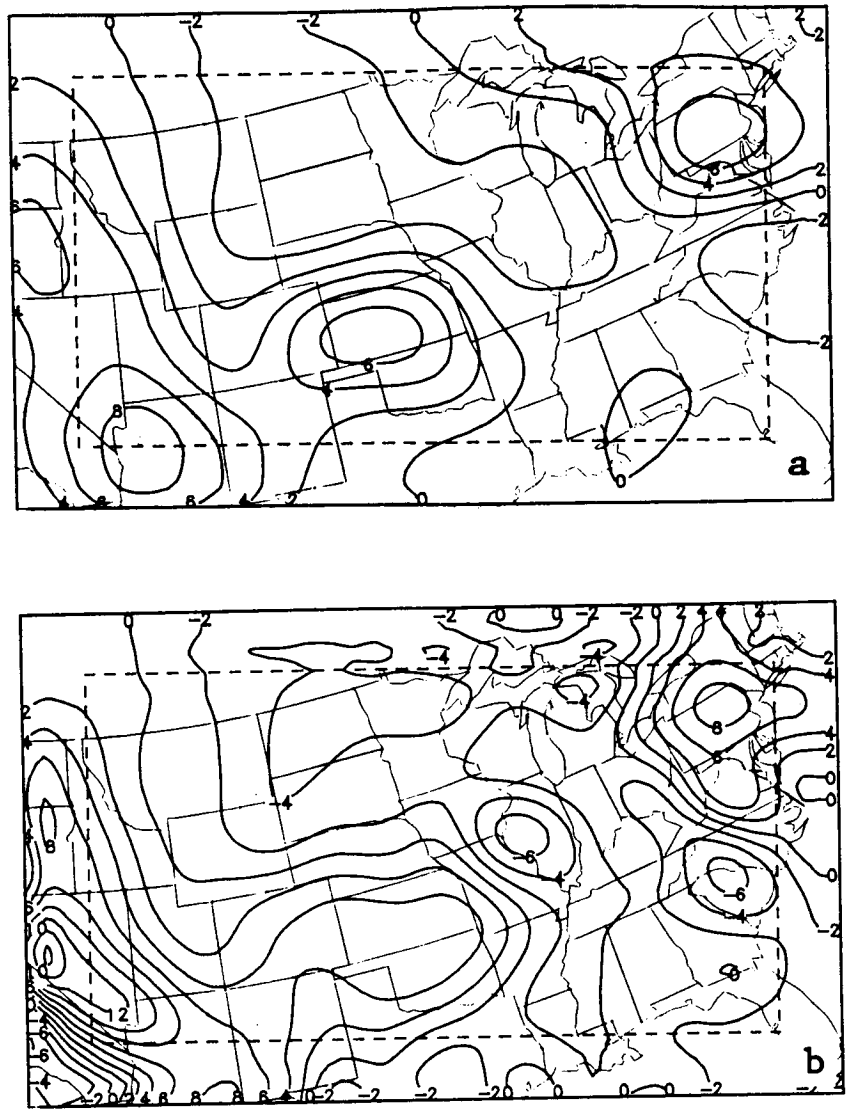


Fig. 16. Patterns of a) the relative vorticity of the observed wind and b) the vorticity of the NOSAT variational assimilated wind for level 6 (500 mb) at 1200 GMT, 10 April 1979. Units are in 10^{-5} s^{-1}.

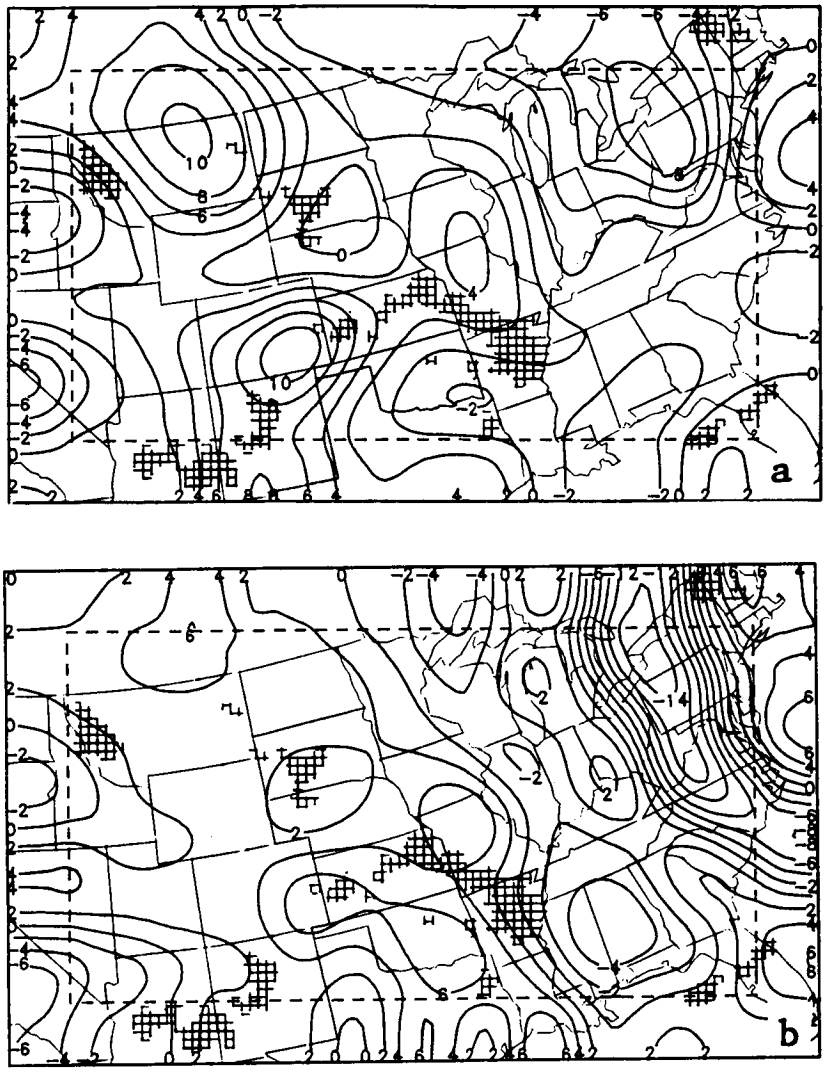


Fig. 17. Vertical velocities (cm s^{-1}) on level 6 (500 mb) at 1200 GMT, 10 April 1979 from a) the kinematic method (O'Brien, 1970) applied to the objective analyses of the initial wind field and b) from the MODEL I adjusted wind field. Hatching delineates areas of precipitation.

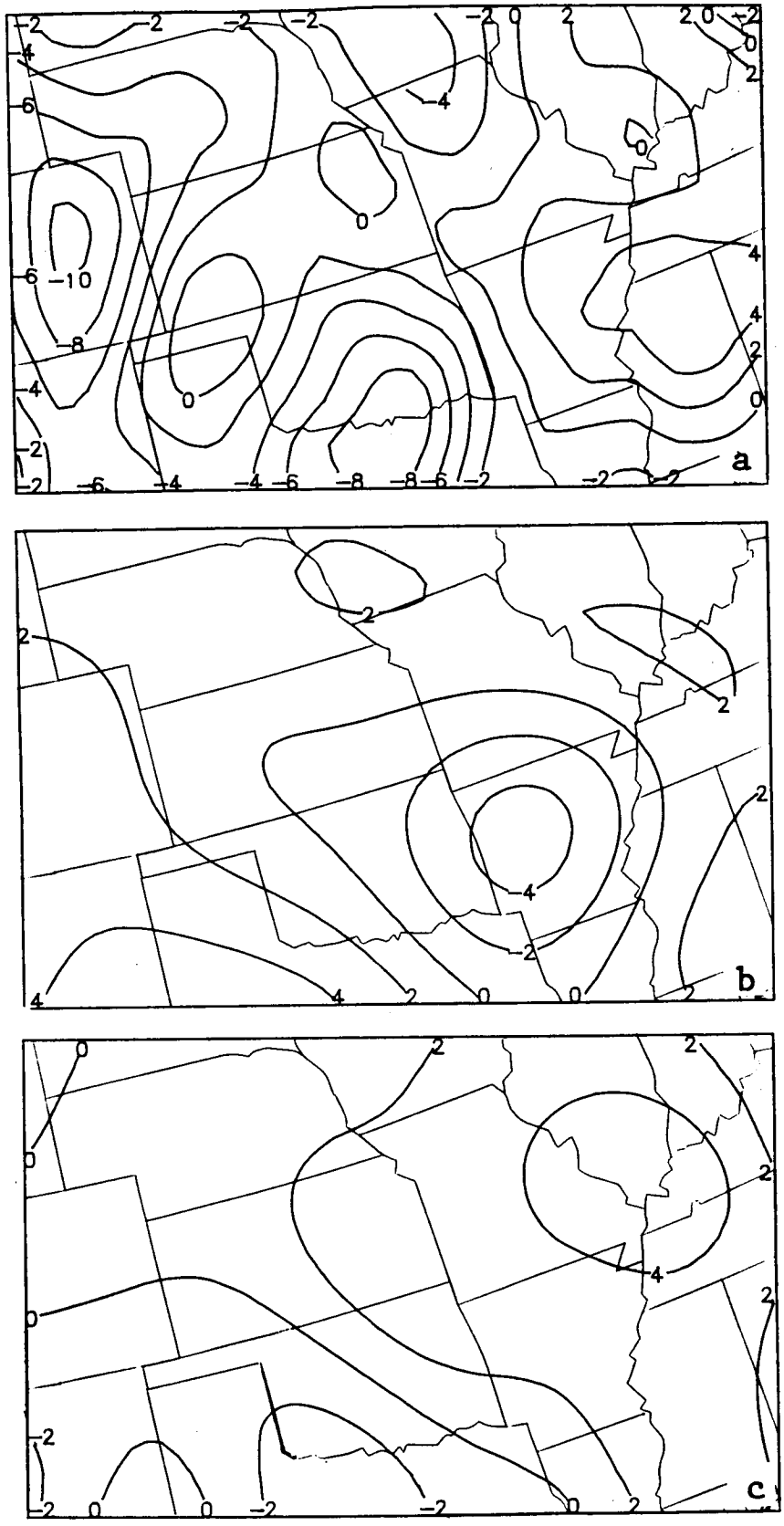


Fig. 18. Fields of a) initial, b) observed, and c) SAT v-tendencies for level 4 (700 mb) for the central region of the U. S. roughly covered by the SESAME-AVE rawinsonde network. Units are in m s^{-1} 3h^{-1}.

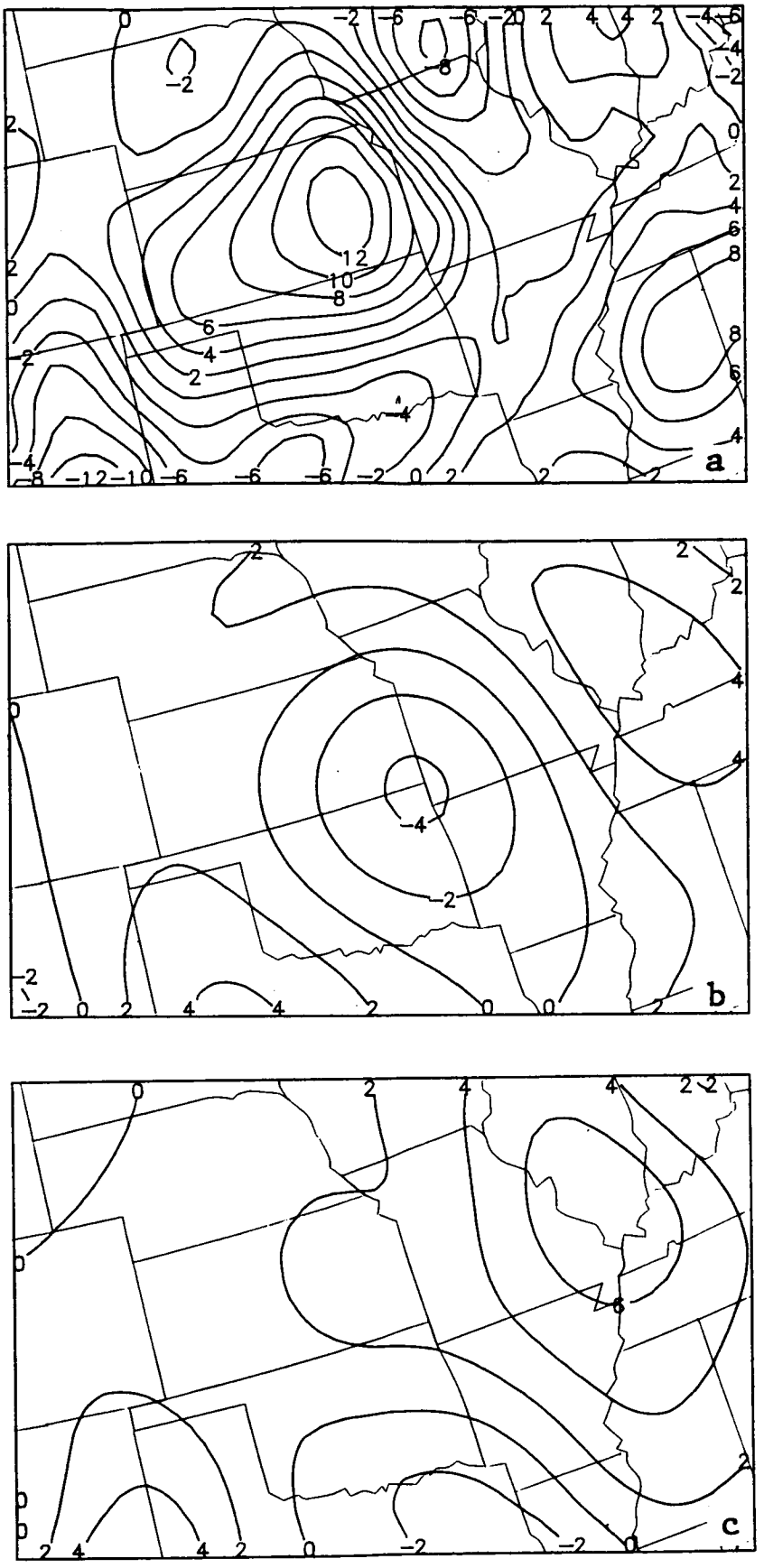


Fig. 19. Same as Fig. 18, but for level 6 (500 mb).

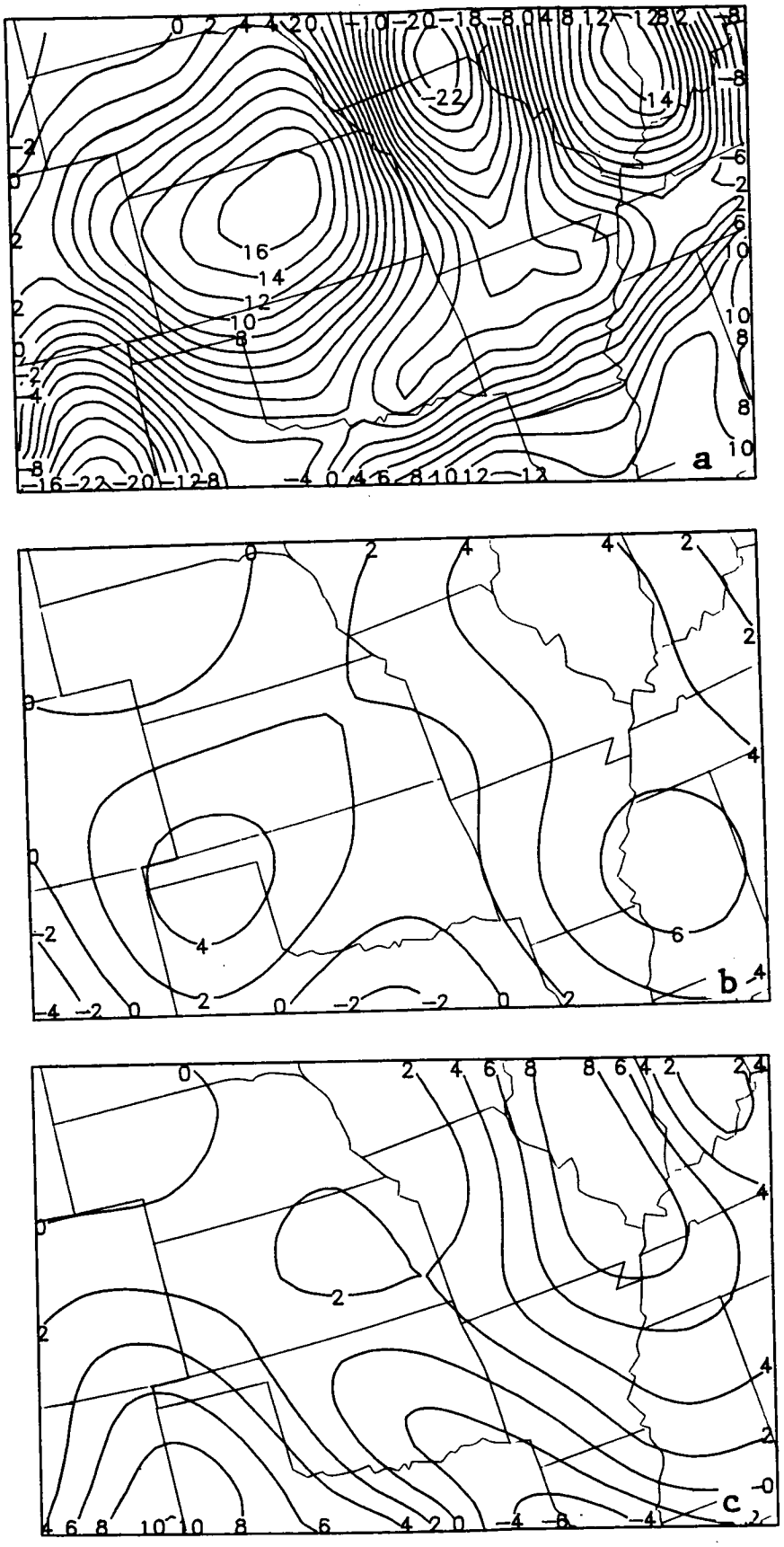


Fig. 20. Same as Fig. 18, but for level 8 (300 mb).

Chapter IV

Hybrid Vertical Coordinate and Pressure Gradient Formulations for a Numerical Variational Analysis Model for the Diagnosis of Cyclone Systems

Gary L. Achtemeier and Harry T. Ochs III

Climate and Meteorology Section

Illinois State Water Survey

Champaign, Il 61820

Hybrid Vertical Coordinate and Pressure Gradient Formulations for a Numerical Variational Analysis Model for the Diagnosis of Cyclone Systems

by

Gary L. Achtemeier and Harry T. Ochs

Illinois State Water Survey

Champaign, IL 61820

ABSTRACT

A hybrid nonlinear sigma vertical coordinate that is suitable for a diagnostic variational objective analysis model is presented and used for an analysis of the pressure gradient terms of the horizontal momentum equations. This vertical coordinate blends from the sigma coordinate to a pressure coordinate at a reference pressure level in the middle troposphere and thus eliminates hydrostatic truncation error above this level. For the lower troposphere, the nonlinear vertical coordinate is used to show that the truncation error for a horizontally homogeneous hydrostatic atmosphere with variable vertical temperature structure arises because of the representative temperature used for the transformation from pressure coordinates to the sigma coordinates. This error is eliminated through a "nonlocal formulation" for the pressure gradient terms that replaces the temperature with its lapse rate in the hypsometric equation. However, this solution is not incorporated into the variational constraints because of greatly increased complexity that would result in the Euler-Lagrange equations. We instead reduce the magnitudes of the

individual pressure gradient terms approximately 30-fold by projecting the pressure gradient onto "equivalent pressure surfaces". This solution leaves the hydrostatic residual unchanged from the direct two-term calculation.

1. Introduction

A variational assimilation model for diagnosis of cyclone systems under development will combine data sets obtained from space-based platforms and from other remote sources with more traditional data systems. Fields of observations will be weighted according to observed and diagnosed measurement accuracies and will be blended to satisfy four dynamical constraints; the nonlinear horizontal momentum equations, the hydrostatic equation, and an integrated form of the continuity equation. Fundamentals of the "strong constraint" variational formalisms as applied to meteorological problems have been presented by Sasaki (1958, 1970), Stephens (1965, 1970), and Wang (1984). Later versions of the assimilation model will include the thermodynamic equation for adiabatic motions and for moist processes.

The adjustment equations will approach or exceed the equations of many numerical prediction models in complexity and number of equations. Although we do not expect to encounter some of the problems posed by numerical prediction models such as the temporal buildup of high frequency waves, we do expect to encounter problems that are endemic to prognostic models and a diagnostic model of this type simply because the dynamics are the same and because the

difference equations are in many respects analogous to the difference formulations for time-dependent models. This paper deals with two such problems, the formulations for the vertical coordinate and for the pressure gradient terms of the horizontal momentum equations.

Numerous vertical coordinate systems have been developed for use in numerical weather prediction models (Kasahara, 1974). The advantages of using the pressure coordinate are that it is the vertical coordinate preferred by diagnosticians and much of the physical processes of the atmosphere are understood in relation to these surfaces. In addition, the physical equations appear in a simplified form. Some disadvantages of the pressure coordinate are that the coordinate surfaces intersect the ground surface, which leads to an irregular mesh and coding complications, and that values of meteorological variables must be extrapolated below ground surface. The alternative terrain-following or sigma coordinate system of Phillips (1957) eliminates these problems. However, considerable error can be introduced in the pressure gradient terms of the momentum equations. These transform into two large and compensating terms where there is steep sloping terrain. Pressure derivatives taken along the sloping sigma surface contain a hydrostatic component which must cancel in the two terms or else large forecast errors can occur. We are not concerned about forecast errors in our diagnostic model. However, it is necessary that the equations be formulated to produce a physically realistic dynamical balance.

Attempts to overcome the large truncation errors in the pressure

gradient terms have included interpolation back to pressure surfaces in order to calculate the terms (Kurihara, 1968; Sundqvist, 1976), the development of finite difference methods that satisfy conservation constraints (Johnson and Uccellini, 1983; Arakawa and Suarez, 1983; Corby et al., 1972; Simmons and Burridge, 1981), and the formulation of hybrid vertical coordinates that change from sigma coordinates near the ground to pressure coordinates at the top of the model atmosphere (Simmons and Burridge, 1981). Bleck (1978) also found that forecasts were more stable if the vertical coordinate blended from the sigma coordinate into, in his case, isentropic coordinates.

We have attempted to avoid complicated difference formulations for the pressure gradient terms in the development of the diagnostic variational model because the complexity is greatly increased in the adjustment equations. Instead we introduce a nonlinear vertical coordinate that changes from the sigma coordinate into a pressure coordinate at a pressure level p^* in the middle troposphere. This hybrid vertical coordinate has the following advantages for a diagnostic variational analysis model:

(1) All coordinate surfaces at and above p^* are pressure surfaces. The pressure gradient is expressed by one term and there is no truncation error.

(2) The pressure surfaces are the coordinate surfaces most preferred by diagnosticians. There is no need to interpolate from sigma coordinates back to pressure coordinates in order to interpret the variationally adjusted fields of meteorological variables.

(3) Vertical interpolation of the initial meteorological fields from pressure coordinates to sigma coordinates is required only for the lower troposphere.

(4) The dynamical equations are presented in their simplest form

on the pressure surfaces at and above p^*. Coding to omit terms that are zero for coordinate surfaces that are surfaces of constant pressure can result in a substantial reduction of computational overhead (Simmons and Strufing, 1983). The tradeoff is that the equations below p^* are more complex than the equations written for the linear sigma coordinate. However, the magnitudes of these additional terms become small in the sigma levels above the lower coordinate surface.

We describe the hybrid nonlinear vertical coordinate in Section 2. In Section 3, we use the nonlinear vertical coordinate to determine the origin of the truncation error for a horizontally homogeneous hydrostatic atmosphere with variable vertical temperature structure. Section 4 presents our method to partition the pressure gradient terms to reduce truncation.

2. A Hybrid Nonlinear Vertical Coordinate

The hybrid sigma coordinate blends from a terrain-following coordinate in the lower troposphere into a pressure coordinate in the middle troposphere. All horizontal variations with the lower coordinate surface are confined to levels below a reference pressure level p^*. The smooth transition from the sigma to the pressure coordinate is accomplished by fitting two curves which are piecewise continuous through the second derivatives. The curve for the upper layer bounded by p_u at the top and by p^* at the bottom is given by a straight line subject to the boundary conditions that $\sigma = 0$ at $p=p_u$ and that $\sigma = \sigma^*$ at $p=p^*$. This equation is

$$\sigma = \sigma^* \frac{p-p_u}{p^*-p_u} , \qquad (1)$$

The equation for the nonlinear part of the hybrid vertical coordinate

between p^* and the surface pressure p_s is found subject to the following four conditions:

$$\left.\begin{aligned} \sigma &= 1.0 && \text{at } p=p_s \\ \sigma &= \sigma^* \\ \frac{\partial\sigma}{\partial p} &= \sigma^*/(p^*-p_u) \\ \frac{\partial^2\sigma}{\partial p^2} &= 0\ . \end{aligned}\right\} \quad \text{at } p=p^*$$

These four conditions specify the equation as a cubic polynomial which takes the form

$$\sigma = \beta\,(p-p^*)^3 + \sigma^* \frac{(p-p_u)}{(p^*-p_u)}\ , \qquad (2)$$

$$\beta = \left[1- \sigma^* \left(\frac{p_s-p_u}{p^*-p_u}\right)\right] (p_s-p^*)^{-3}.$$

Fig. 1 shows the relationship between sigma and pressure for the levels below the elevation of the 600 mb pressure surface for the coordinate parameters that have been selected for the variational objective analysis. The reference pressure p^* is at 700 mb. A straight line from 700 mb to 1000 mb separates two sets of curves which describe the relationship between sigma and pressure for low surface pressure (high elevation) from those for high surface pressure. If the surface pressures were everywhere equal to 1000 mb, the hybrid sigma coordinate would be identical to a pressure coordinate system. The thicknesses between sigma coordinate surfaces are compacted over higher elevations wherever the slopes of the curves in Fig. 1 are less than the slope of the straight line. The greatest

packing of the coordinate surfaces is found for levels nearest the lower coordinate surface. These layer depths increase to approach the thicknesses of the pressure layers at levels above p^*.

The slopes of the curves in Fig. 1 exceed the slope of the straight line at locations where the surface pressure is greater than 1000 mb. The pressure thicknesses between the sigma coordinate surfaces here are greater than are the pressure thicknesses over linear part of the coordinate. Note how the curve from 1100 mb to 700 mb has approached the straight line by sigma equal to 0.88. Whenever the surface pressure is greater than 1000 mb, the nonlinear vertical coordinate will force most of the transition between terrain-following coordinate surfaces and pressure-following coordinate surfaces into the layer immediately above the ground. Thus the nonlinear sigma coordinate surfaces in the lower troposphere tend to behave as pressure surfaces that are punctuated by areas of higher elevation.

Fig. 2 shows the distribution of hybrid coordinate surfaces below 600 mb as the surface pressure varies from 800 to 1025 mb, the approximate range of surface pressures for the smoothed orography of the variational model. The greater compression of the coordinate surfaces over higher elevation nearest the surface is clearly evident. Notice how the nonlinear coordinate surfaces tend to become surfaces of constant pressure at locations away from the areas of high elevation. Note also the increased pressure depth of the lowest layer where the surface pressures exceed 1000 mb. Clearly this nonlinear vertical coordinate does not provide for a boundary layer of uniform

thickness. Use of this coordinate can increase the complexity of some numerical models in that the boundary layer will have to be parameterized as a function of layer thickness.

3. The Truncation Error in the Pressure Gradient

It is well known that, upon transformation from the pressure coordinate system into the sigma coordinate system, the pressure gradient becomes the sum of two terms the total of which may be an order of magnitude or more smaller than the individual terms in areas where coordinate surfaces pass over steeply sloping terrain. Gary (1973) noted that pressure derivatives taken on a sloping surface contain a hydrostatic component that must cancel in the two terms that make up the pressure gradient in the sigma system or else there will be a risk of considerable error. In this section, we use the hybrid nonlinear vertical sigma coordinate to isolate the hydrostatic component. Then we show how the truncation error in the presure gradient originates, and finally show how the error can be eliminated.

a) Origin of the Hydrostatic Component

Consider from Fig. 2 two sigma surfaces over steeply sloping terrain. Let the top surface be at p^* (700 mb), a surface of constant pressure. At any arbitrary point on the lower, sloping sigma surface p, T, and ϕ are known. The pressure gradient force per unit mass at this point is given by

$$PGF = RT \frac{\partial \ell n\, p}{\partial X} + \frac{\partial \phi}{\partial X} \qquad (3)$$

We proceed to express PGF as the sum of the pressure gradient force at the * level,

$$PGF^* = \frac{\partial \phi^*}{\partial X} \tag{4}$$

and the incremental pressure gradient force for the layer between the two sigma levels. In so doing, we will make use of the hypsometric equation,

$$p = p^* \exp\left[(\phi^* - \phi)/R\bar{T}\right] \tag{5}$$

where $\bar{T} = 0.5(T^* + T)$. Substitution of (5) into (3) gives

$$PGF = \frac{\partial \phi}{\partial X}\left(1 - \frac{T}{\bar{T}}\right) + \frac{T}{\bar{T}}\frac{\partial \phi^*}{\partial X} - \frac{T}{\bar{T}^2}(\phi^* - \phi)\frac{\partial \bar{T}}{\partial X} \tag{6}$$

Furthermore, we can rewrite the horizontal gradient of the mean temperature if

$$\bar{T} = T^* + 0.5\,\Gamma\,(\phi^* - \phi) \tag{7}$$

where the vertical temperature lapse rate Γ is defined by $\frac{\partial T}{\partial \phi}$. Using (7) to eliminate the mean temperature gradient in (6), the pressure gradient force becomes

$$PGF = \left(\frac{\Delta T}{2\bar{T}}\right)^2 \frac{\partial \phi}{\partial X} + \frac{TT^*}{\bar{T}^2}\frac{\partial \phi^*}{\partial X} - \frac{T}{\bar{T}^2}(\phi^* - \phi)\frac{\partial T^*}{\partial X} - \frac{T}{\bar{T}^2}\left(\frac{\phi^* - \phi}{2}\right)^2 \frac{\partial \Gamma}{\partial X} \tag{8}$$

where $T = T^* - T$.

If the atmosphere is horizontally invariant, $\partial\phi^*/\partial x = \partial T^*/\partial x = \partial\Gamma/\partial x = 0$ and there remains a residual,

$$PGF = \left(\frac{\Delta T}{2\bar{T}}\right)^2 \frac{\partial\phi}{\partial X} \qquad (9)$$

which as a function of the slope of the sigma coordinate surface and the temperature change between the two sigma coordinate surfaces. This uncancelled hydrostatic component is for a layer bounded by a pressure surface and a height surface and is therefore unique to this nonlinear vertical coordinate.

We calculated the hydrostatic residual for an adiabatic atmosphere between the surface (1000 mb) and the p* level (700 mb) with a surface temperature of 285K. The lower coordinate surface sloped from sea level to the top of a mountain (elevation 1800 m) in a horizontal distance of 200 km. We chose an adiabatic atmosphere to maximize T. The residual, expressed as a geostrophic wind error, was 1.16 m s^{-1}. Given the differences in modeling constraints, this result is considerably less than the errors found by Johnson and Uccellini (1983) for five methods for calculating the pressure gradient force.

It would appear that a nontrivial improvement can be gained if the pressure gradient force is calculated as a function of the sum of the pressure gradient force on the pressure coordinate p^* and the hydrostatic thickness of the intervening layer. However, a drawback of this approach is that the pressure gradient force is rendered as a "nonlocal" calculation which can seriously affect the local accuracy

of the hydrostatic equation (Arakawa and Suarez, 1983).

b) Origin of the Hydrostatic Residual in Nonlinear Sigma Coordinates

The transformation from the pressure coordinate to the sigma coordinate is given by

$$\frac{\partial \phi}{\partial X_p} = \frac{\partial \phi}{\partial X_\sigma} + \frac{\partial \phi}{\partial p}\frac{\partial p}{\partial X} \qquad (10)$$

where, substitution of the hydrostatic and state equations for the vertical gradient of geopotential height as a function of pressure $\partial\phi/\partial p$ yields the familiar form for the pressure gradient in sigma coordinates (3). If (3) is calculated by using the temperature located on either the sigma surface or the pressure surface, there will result a small error where the incremental separation between the two surfaces is large. As shown by Fig. 3, the vertical gradient of ϕ is better represented by T_m which is the mean temperature of the layer of interpolation between the sigma and pressure surfaces. T approximates T_m only at locations where the slopes of the sigma coordinate surfaces do not depart much from the slopes of the pressure surfaces. This is not the case over steep sloping terrain where the sigma coordinate surfaces depart appreciably from pressure surfaces. Here T_m differs significantly from T and a nontrivial hydrostatic residual will exist unless the lapse rate of temperature over the layer of the coordinate transformation is isothermal. Then (9) is equal to zero. Sundqvist (1975) also found that the hydrostatic residual vanishes for an isothermal atmosphere.

c) Elimination of the Hydrostatic Residual

If T, T_m, and the mean layer temperature are found within an atmospheric layer of constant lapse rate, we can retain (3) with T and eliminate the hydrostatic residual upon expressing the pressure gradient force in terms of the pressure gradient force at p^* and the lapse rate of the intervening layer. To accomplish this, we rewrite the hypsometric equation as

$$\int_p^{p^*} d \ln p = -\int_\phi^{\phi^*} \frac{d\phi}{RT} \tag{11}$$

If we substitute for T with

$$T = T^* + \Gamma (\phi^* - \phi) \tag{12}$$

and integrate (11) over the layer between the two sigma surfaces, we find that p can be related to p^* through

$$p = p^* \left[\frac{T^* + \Gamma (\phi^*-\phi)}{T^*}\right]^{\frac{1}{\Gamma R}} \tag{13}$$

Substitution of (13) into (3) gives the pressure gradient force as a function of the horizontal gradients of ϕ^*, T^*, and Γ;

$$PGX = \frac{\partial \phi^*}{\partial X} + \frac{1}{\Gamma}\left[1 - \frac{T}{T^*}\right]\frac{\partial T^*}{\partial X} + \frac{T}{\Gamma^2}\left[\ln \frac{T^*}{T} + \frac{T-T^*}{T}\right]\frac{\partial \Gamma}{\partial X} \tag{14}$$

This equation is an improvement over (8) in that all of its terms are functions of variables at the p^* level and the horizontal gradient

of the lapse rate. Therefore, if the atmosphere is horizontally invariant, these terms vanish and there exists no uncancelled hydrostatic component.

4. Pressure Gradient Force for the Diagnostic Variational Model.

We recognize the need to reduce the truncation error between the two terms of the pressure gradient in the lower troposphere in the development of the diagnostic variational model. However there are other important factors that must be taken into consideration. The "strong constraint" version of the variational formalism (Sasaki, 1970) produces a set of Euler-Lagrange adjustment equations that have complexity greater than the complexity of the original dynamical constraints. We therefore seek the simplest formulations for the constraints. The nonlocal formulations that eliminated the truncation error in the pressure gradient contain nonlinear terms that would greatly increase the difficulty in obtaining a solution for the Euler-Lagrange equations. However, if the hydrostatic terms are not reformulated, the variational algorithm will separate the pressure gradient terms and combine the large uncompensated terms with terms from other equations. The large nonmeteorological contribution by these terms can cause significant errors in the final solution unless methods are developed to remove them (Achtemeier, 1975).

We are therefore motivated to develop a method that compensates the two pressure gradient terms in the levels below p^* and also retains the simple formulation that is required for the variational

model. We remove a hydrostatic component from both terms by partitioning the pressure gradient to cancel most of the orographic part. The separation is not complete because the mean layer temperature is not partitioned. We restate the hydrostatic equation as

$$\frac{\partial \phi_w}{\partial \sigma} + R\overline{T}_w \frac{\partial \ln (p_w)}{\partial \sigma} = 0 \qquad (15)$$

where the subscript w implies the whole or unpartitioned variable. The geopotential and pressure are expressed as an orograhic part plus a remainder. $\phi_w = \phi_T + \phi_a$ and $p_w = p_T + p$. Substitution into the hydrostatic equation gives

$$\frac{\partial \phi_a}{\partial \sigma} + \gamma R\overline{T}_w + \beta = 0 \qquad (16)$$

where

$$\gamma = \frac{\partial \ln (p)}{\partial \sigma} \qquad (17)$$

and

$$\beta = \left(\frac{\overline{p}_w}{\overline{p}} - 1\right) \frac{\partial \phi_w}{\partial \sigma} + \frac{\partial \phi_T}{\partial \sigma} + \frac{R\overline{T}_w}{\overline{p}} \frac{\partial p_T}{\partial \sigma} \qquad (18)$$

We further remove a reference atmosphere by defining $\phi_a = \phi_R + \phi$ and $\overline{T}_w = \overline{T}_R + \overline{T}$ and requiring that

$$\frac{\partial \phi_R}{\partial \sigma} + \gamma R\overline{T}_R = 0 \qquad (19)$$

The remaining equation

$$\frac{\partial \phi}{\partial \sigma} + \gamma R\overline{T} + \beta = 0 \qquad (20)$$

describes the hydrostatic relationship between meteorological

perturbations. The perturbations are subject to the variational adjustments. Most of the orographic component is located in ϕ_T. Eq. 18 can be solved accurately if the layer average pressures are equal to the average of the arithmetic mean plus twice the geometric mean. The orographic variables are found by setting $\beta = 0$ and defining p as equivalent pressure surfaces. We use the terminology "equivalent pressure surfaces" to avoid confusion with methods that calculate the pressure gradient on surfaces of constant pressure and then interpolate the results to sigma coordinate surfaces (Kurihara, 1968). The equivalent pressure surfaces can be easily determined from the definition of the nonlinear vertical coordinate. Fig. 1 shows that the relationship between pressure and sigma is linear in the lower troposphere if the surface pressure is equal to 1000 mb. Choosing $p_s = p_{se} = 1000$ mb uniquely determines the remaining pressures through (2) and therefore p_T. Then ϕ_T is found by downward integration from the reference pressure level.

Having derived the relevant partitioned variables, the pressure gradient terms are easily transformed, e.g.,

$$PGX = \frac{\partial \phi}{\partial x} + \eta_x \tag{21}$$

where

$$\eta_x = \frac{\partial \phi_T}{\partial x} + R\overline{T}^x \frac{\partial \ln (p_w)}{\partial x} \tag{22}$$

Fig. 4 shows the height of the lower coordinate surface for a grid to be used for the diagnostic variational analysis of data collected at 1200 GMT 10 April 1979. The heights on the unpartitioned

terrain-following coordinate vary from 0-1800 m approximately (Fig. 4a) and show the steep gradients that surround a smoothed high elevation area over the western U.S. The heights remaining after the removal of the hydrostatic component that arises from variations in the elevation of the lower coordinate surface are shown in Fig. 4b. Calculations show that the projection onto equivalent pressure surfaces reduces the magnitudes of the these variations by about 30-fold. The equivalent 1000 mb heights resemble the actual 1000 mb heights (Fig. 4c) with the exception that the heights of the low center over the West are approximately 60 m higher in Fig. 4b. This residual orographic effect is retained through the unpartitioned mean layer temperatures.

5. Discussion

We have presented a hybrid vertical coordinate for a diagnostic variational objective analysis model. The coordinate blends from a terrain-following sigma coordinate in the lower troposphere to constant pressure surfaces in the middle troposphere. There are several advantages to this nonlinear vertical coordinate. All coordinate surfaces from the middle troposphere to the top of the analysis domain are pressure surfaces. The pressure gradient is expressed by one term and there is no hydrostatic truncation error. This is also true for the lower stratosphere for which the pressure gradient calculated on sigma coordinates is extremely sensitive to truncation.

Pressure surfaces are the coordinate surfaces most preferred by diagnosticians. There is no need to interpolate from fields presented on sigma coordinates to fields presented on pressure coordinates in order to interpret the patterns of variationally adjusted meteorological variables. Further, vertical interpolation of the initial meteorological data from pressure coordinates to sigma coordinates is required only for the lower troposphere.

The dynamical equations are presented in their simplest form on the pressure surfaces at and above p^*. Coding to omit terms that are zero for coordinate surfaces that are surfaces of constant pressure can result in a substantial reduction of computational overhead (Simmons and Strufing, 1983). The tradeoff is that the complexity of the equations below p^* is increased over the complexity of the equations written for the linear sigma coordinate. However, the magnitudes of these additional terms become small in the sigma levels above the lower coordinate surface.

We used the nonlinear vertical coordinate to derive an equation for the hydrostatic residual. We found that an uncancelled hydrostatic residual is present in the transformation of the dynamic equations from pressure to sigma coordinates. The residual results when the temperature on the sigma level is used in the transformation instead of the mean temperature of the incremental layer between the sigma and the pressure coordinate surfaces. If the temperature at the sigma level is used, the hydrostatic residual vanishes only if the coordinate surfaces are coincident or if the lapse rate of temperature

for the layer is isothermal. We were able to eliminate the hydrostatic residual in this nonlinear vertical coordinate by replacing the temperature with its lapse rate in the hypsometric equation. However. this "nonlocal solution" was not incorporated into the variational constraints because of greatly increased complexity that would result in the Euler-Lagrange equations.

We developed a method that removes most of the influence of unlevel terrain in the pressure gradient terms by projecting the pressure gradient onto equivalent pressure surfaces. This method reduced the magnitudes of the individual terms by approximately 30-fold. Some hydrostatic residual remains, however. the nonlinear formulation for this hybrid vertical coordinate requires the hydrostatic residual to decrease upward to smaller values in comparison with the linear sigma coordinate in order that it vanish at the lower-middle troposphere (700 mb in our model) and upward.

Acknowledgements

This research was supported by the National Aeronautics and Space Administration (NASA) under contract NAS8-34902. Ms. Donna Isard is acknowledged for the preparation of the figures.

References

Achtemeier, G. L., 1975: On the initialization problem: A variational adjustment method. Mon. Wea. Rev., 103, 1089-1103.

Arakawa, A., and M. J. Saurez, 1983: Vertical differencing of the primitive equations in sigma coordinates. Mon. Wea. Rev., 111, 34-45.

Bleck, R., 1978: On the use of hybrid vertical coordinates in numerical weather prediction models. Mon. Wea. Rev., 106, 1233-1244.

Corby, G. A., A. Gilchrist and R. L. Newson, 1972: A general circulation model of the atmosphere suitable for long period integrations. Quart. J. Roy. Meteor. Soc., 98, 809-832.

Gary, J. M., 1973: Estimate of truncation error in transformed coordinate primitive equation atmospheric models. J. Atmos. Sci. 30, 223-233.

Janjic, Z. I., 1977: Pressure gradient force and advection scheme used for forecasting with steep and small scale topography. Beit. Phys. Atmos.,50, 186-189.

Johnson, D. R., 1980: A generalized transport equation for use with meteorological coordinate systems. Mon. Wea. Rev., 108, 733-745.

_____, and L. W. Uccellini, 1983: A comparison of methods for computing the sigma-coordinate pressure gradient force for flow over sloped terrain in a hybrid theta-sigma model. Mon. Wea. Rev., 111, 870-886.

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

Kurihara, Y., 1968: Note on finite difference expressions for the hydrostatic relation and pressure gradient force. Mon. Wea. Rev., 96, 654-656.

Phillips, N. A., 1957: A coordinate system having some special advantages for numerical forecasting. J. Meteor., 14, 184-185.

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

_____, 1970: Some basic formalisms in numerical variational analysis. Mon. Wea. Rev., 98, 875-883.

Simmons, A. J., and D. M. Burridge, 1981: An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. Mon. Wea. Rev., 109, 758-766.

_____ and R. Strufing, 1983: Numerical forecasts of stratospheric warming events using a model with a hybrid vertical coordinate. Quar. Jour. Royal Met. Soc., 109, 81-112.

Stephens, J. J., 1965: A variational approach to numerical weather analysis and prediction. Ph.D Dissertation, Austin, University of Texas.

_____, 1970: Variational initialization with the balance equation. J. Appl. Meteor., 9, 732-739.

Sundqvist, H., 1975: On truncation errors in sigma-system models. Atmosphere, 13, 81-95.

_____, 1976: On vertical interpolation and truncation in connexion with use of sigma system models. Atmosphere, 14, 37-52.

Wang, P. K., 1984: A brief review of the Eularian variational principle for atmospheric motions in rotating coordinates. Atmosphere-Ocean, 22, 387-392.

Figure Captions

Figure 1. The relationship between sigma and pressure for the levels below the elevation of the 600 mb pressure surface for the coordinate parameters that have been selected for the variational objective analysis.

Figure 2. The distribution of hybrid coordinate surfaces below 600 mb as the surface pressure varies from 800 to 1025 mb.

Figure 3. The relationship between the geopotential expressed on pressure and sigma coordinate surfaces.

Figure 4. Heights at the lower coordinate surface for a) unpartitioned terrain-following coordinate, b) the equivalent pressure surface, and c) the 1000 mb pressure surface.

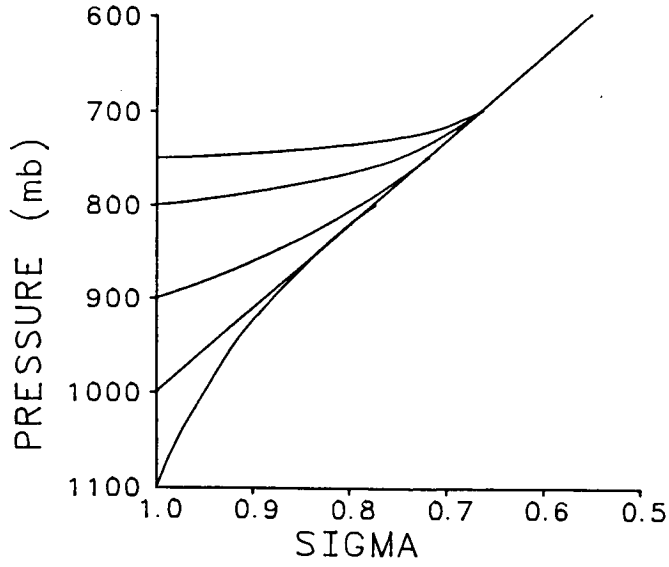


Figure 1

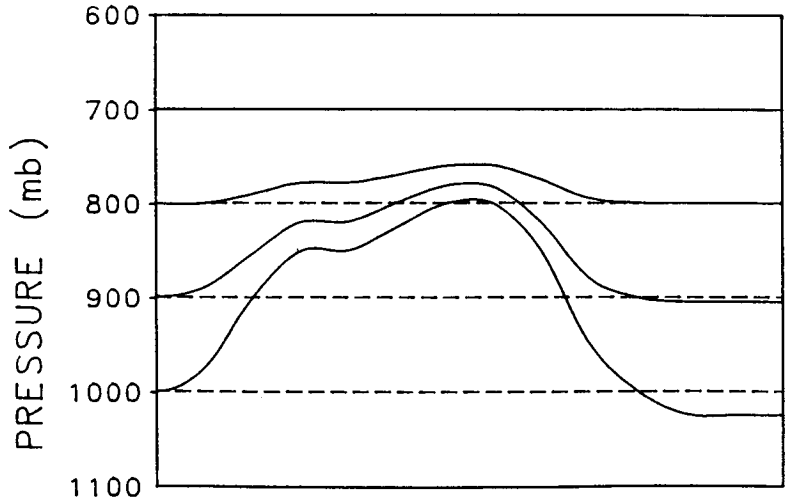


Figure 2

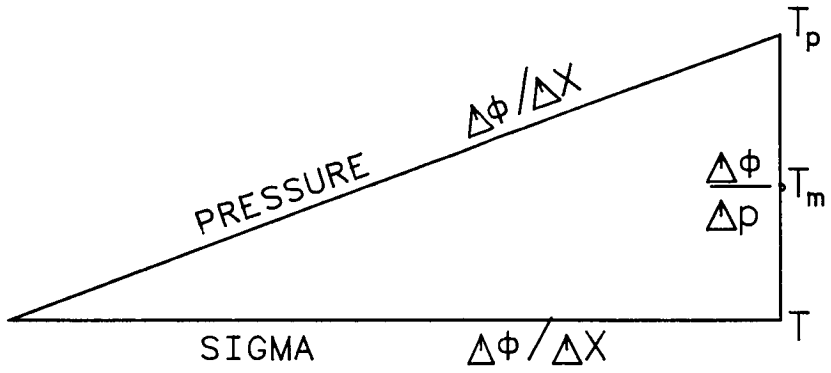
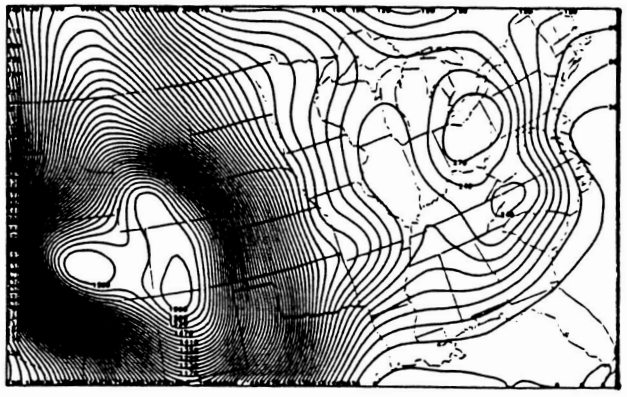
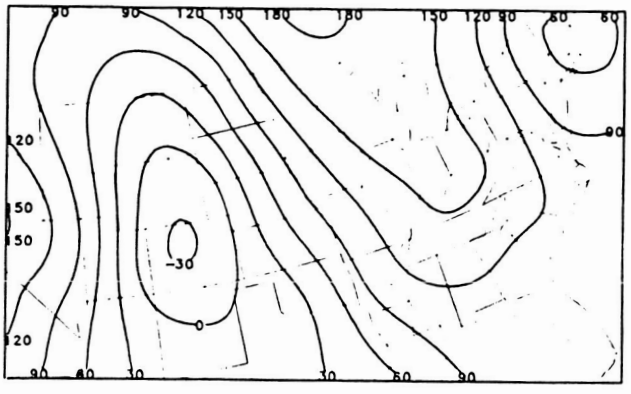

T_p
$\Delta\phi/\Delta X$
PRESSURE
$\frac{\Delta\phi}{\Delta p}$
T_m
T
SIGMA
$\Delta\phi/\Delta X$

Figure 3

a)



b)


90
90
120
150
180
180
150
120
90
60
60
120
150
150
-30
0
120
90
60
30
30
60
90

c)

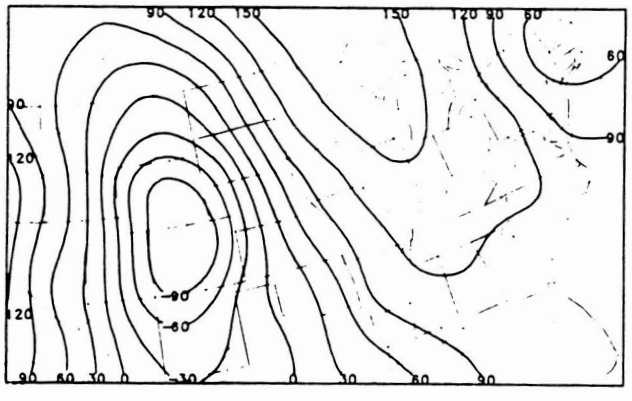

90
120
150
150
120
90
60
60
90
90
120
120
-90
-60
90
60
30
0
-30
0
30
60
90

Figure 4

Chapter V

Day-Night Variation in Operationally-Retrieved TOVS Temperature Biases

Stanley Q. Kidder and Gary L. Achtemeier

Climate and Meteorology Section

Illinois State Water Survey

Champaign, Il 61820

Day-Night Variation in Operationally-Retrieved TOVS Temperature Biases

by

Stanley Q. Kidder and Gary L. Achtemeier

Climate and Meteorology Section
Illinois State Water Survey
Champaign, IL 61820

1. Introduction

Several authors have reported that operationally-retrieved TOVS (Tiros Operational Vertical Sounder) temperatures are biased with respect to rawinsonde temperatures or temperature analyses (Phillips et al., 1979; Schlatter, 1981; Gruber and Watkins, 1982). Not appearing in the literature, however, is an indication of how these biases may vary diurnally. This note documents a significant day-night variation in the biases over the United States during one time period.

2. Background

Under development at the Illinois State Water Survey is a sophisticated variational analysis model (Achtemeier et al., 1986a) which offers a means for blending satellite and conventional soundings in a way which preserves the information content of both data sources. However, the model requires input data which are as bias-free as possible and about which the error characteristics are known. Because gridded data are required as input for the model, we needed to know the bias of TOVS temperatures with respect to objectively-analyzed rawinsonde data. None of the previous studies of TOVS biases used precisely this standard of comparison. Phillips et al. (1979) and

Gruber and Watkins (1982) used nearly-colocated rawinsondes as a comparison; while Schlatter (1981) used NMC Final Analyses. We decided, therefore, to recalculate the biases.

3. Data and analysis

The case study on which the variational analysis model was first run is 10-11 April 1979 (Achtemeier et al., 1986b). To do the calibration study, we acquired Tiros-N soundings and rawinsonde data for the period 26 March through 11 April 1979. This is the same period (plus three days) analyzed by Schlatter.

Layer mean virtual temperatures, derived from rawinsonde thicknesses, each 12 hr for the period 0000 GMT 26 March through 1200 GMT 11 April 1979 were objectively analyzed on a 21 x 21 grid (260 km grid spacing at 45°N) covering most of North America (Fig. 1). Biases were estimated by calculating the difference between satellite-estimated mean virtual temperatures and rawinsonde values interpolated in both time and space from the analyses to the satellite data.

Figure 2 shows the 12 hr average biases as a function of time for each layer. The dashed lines represent the mean biases for the periods 26 March through 8 April and 10-11 April. The three sounding types (clear, partly cloudy, cloudy) have been kept separate. The error bars represent 95% confidence intervals assuming that the biases are normally distributed about the 12 hr mean, which proved to be a good assumption upon examination. Two aspects are disturbing: (1) For a large number of points the error bars do not include the mean represented by the dashed line. Only one in 20 points should not include the mean if the long term average is representative. (2) There seems to be a 24-hr

oscillation, which indicates that daytime and nighttime biases may be different.

Day and night soundings were separated, and the mean biases for the period 26 March through 8 April were calculated. These results are plotted in Figure 3 and tabulated in Table 1. Again, the error bars represent 95% confidence intervals. It is clear that biases for day and night soundings are statistically different (95% confidence) at most levels for clear and partly cloudy soundings, and at several levels for cloudy soundings. Day-night differences are particularly evident for clear soundings. In the mid-troposphere, nighttime soundings have little bias, while daytime soundings have a large cold bias. Schlatter's results are plotted for comparison. Because Schlatter did not separate day and night soundings, his biases tend to split the difference between the day and night biases. In the upper troposphere, Schlatter's biases tend to be colder than the biases calculated with respect to rawinsonde analyses.

4. Conclusions

It is concluded that at least for the time period 26 March through 11 April 1979 there was a significant day-night variation in TOVS mean layer virtual temperature biases with respect to analyses of rawinsonde data over the United States. Day-night variations may exist for other time periods and for other locations.

Acknowledgments. This work was supported by the National Aeronautics and Space Administration under grant NAS8-34902.

REFERENCES

Achtemeier, G.L., H.T. Ochs III, and J. Chen, 1986a: A Variational Assimilation Method for the Diagnosis of Cyclone Systems. Part I: Development of the Basic Model. To be submitted to Mon. Wea. Rev.

Achtemeier, G.L., S. Q. Kidder and R. W. Scott, 1986b: A variational assimilation method for the diagnosis of cyclone systems. Part II: Case study results with and without satellite data. To be submitted to Mon. Wea. Rev.

Gruber, A., and C.D. Watkins, 1982: Statistical assessment of the quality of TIROS-N and NOAA-6 satellite soundings. Mon. Wea. Rev., 110, 867-876.

Phillips, N., L. McMillin, A. Gruber, and D. Wark, 1979: An evaluation of early operational temperature soundings from TIROS-N. Bull. Amer. Meteor. Soc., 60, 1188-1197.

Schlatter, T.W., 1981: An assessment of operational TIROS-N temperature retrievals over the United States. Mon. Wea. Rev., 109, 110-119.

Table 1

Biases and standard deviations of operationally-retrieved Tiros-N layer mean virtual temperatures (K).

Layer (mb)	Day			Night		
	Clear	Partly Cloudy	Cloudy	Clear	Partly Cloudy	Cloudy
			Biases			
200-100	-0.02	-0.05	0.54	0.19	0.18	0.51
300-200	0.98	0.65	1.20	0.61	1.49	1.62
400-300	-0.04	0.20	-0.41	0.41	0.26	-0.02
500-400	-0.87	0.33	-1.14	0.14	-0.25	-1.11
700-500	-0.90	0.29	-1.35	-0.10	-0.23	-1.37
850-700	-0.16	0.67	-0.65	0.09	1.08	-0.31
1000-850	0.32	1.03	1.87	-0.42	1.35	2.06
			Standard Deviations			
200-100	1.29	1.63	1.48	1.38	1.12	1.47
300-200	1.57	1.40	1.74	1.18	1.34	1.45
400-300	1.50	1.37	1.83	1.38	1.40	1.99
500-400	1.43	1.43	1.70	1.59	1.30	1.89
700-500	1.62	1.72	1.77	1.25	1.44	1.77
850-700	1.84	2.42	2.31	1.84	2.28	2.60
1000-850	2.11	2.52	2.60	2.64	3.12	3.39

Figure Captions

Fig. 1. Asterisks indicate the location of the 101 rawinsonde stations used to construct the objective analyses for comparison with satellite soundings. The dashed line encloses the satellite soundings. (This is the same area chosen by Schlatter, 1981.) Note that the satellite soundings are well within the boundaries of the rawinsonde objective analysis area; thus edge effects should be minimal.

Fig. 2. 12 h average biases of satellite soundings: (a) clear soundings, (b) partly cloudy soundings, (c) cloudy soundings. The error bars represent the 95% confidence interval. Dashed lines represent the average of the biases. The temperature scale on the right is in kelvin.

Fig. 3. Average biases for the period 26 March through 8 April 1979: (a) clear soundings, (b) partly cloudy soundings, (c) cloudy soundings. Day and night biases have been kept separate. The biases published by Schlatter (1981) are plotted for reference. The error bars represent the 95% confidence interval.

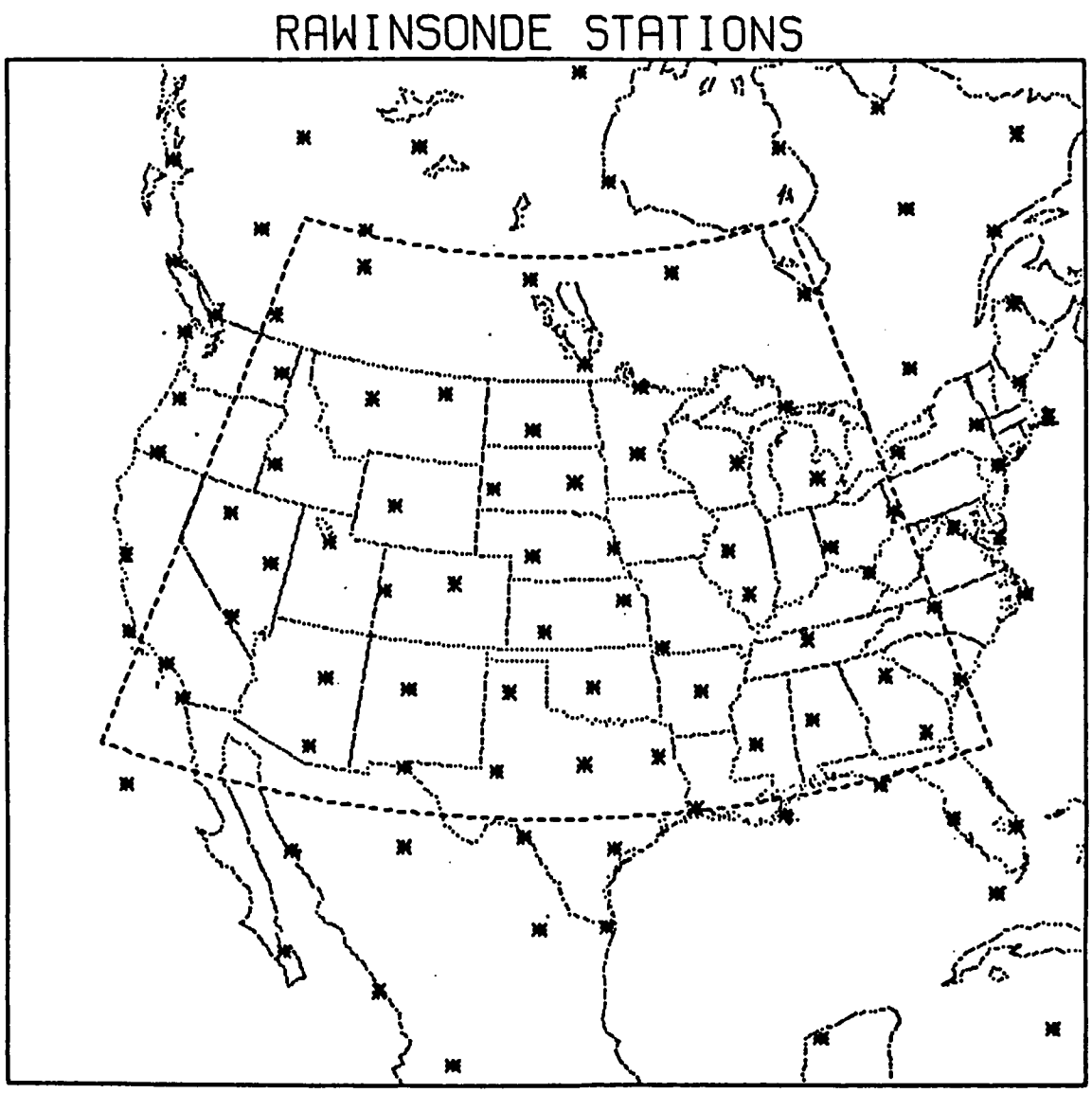


Figure 1

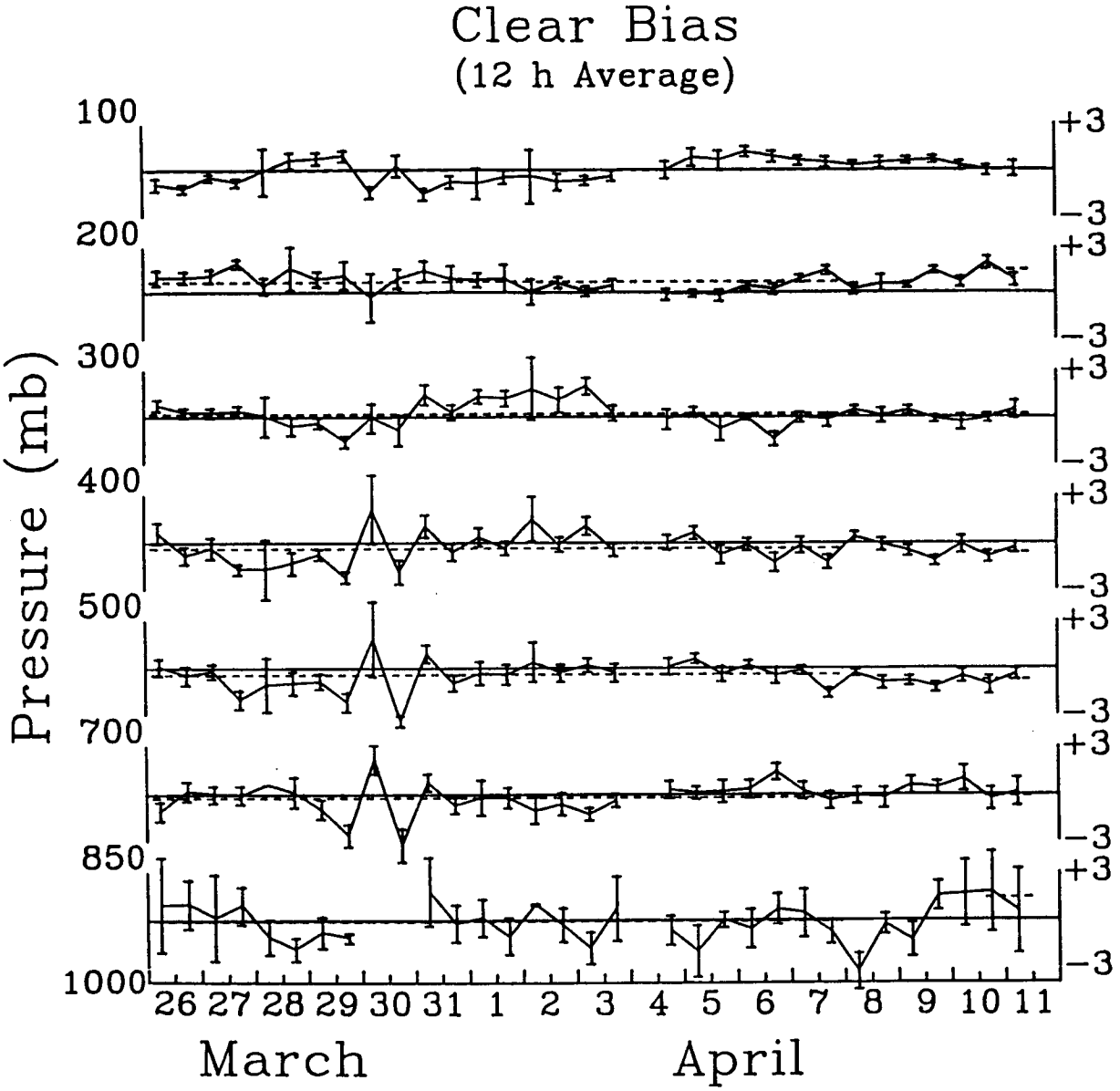

Clear Bias
(12 h Average)
Pressure (mb)
100
200
300
400
500
700
850
1000
+3
−3
26 27 28 29 30 31 1 2 3 4 5 6 7 8 9 10 11
March
April

Figure 2a

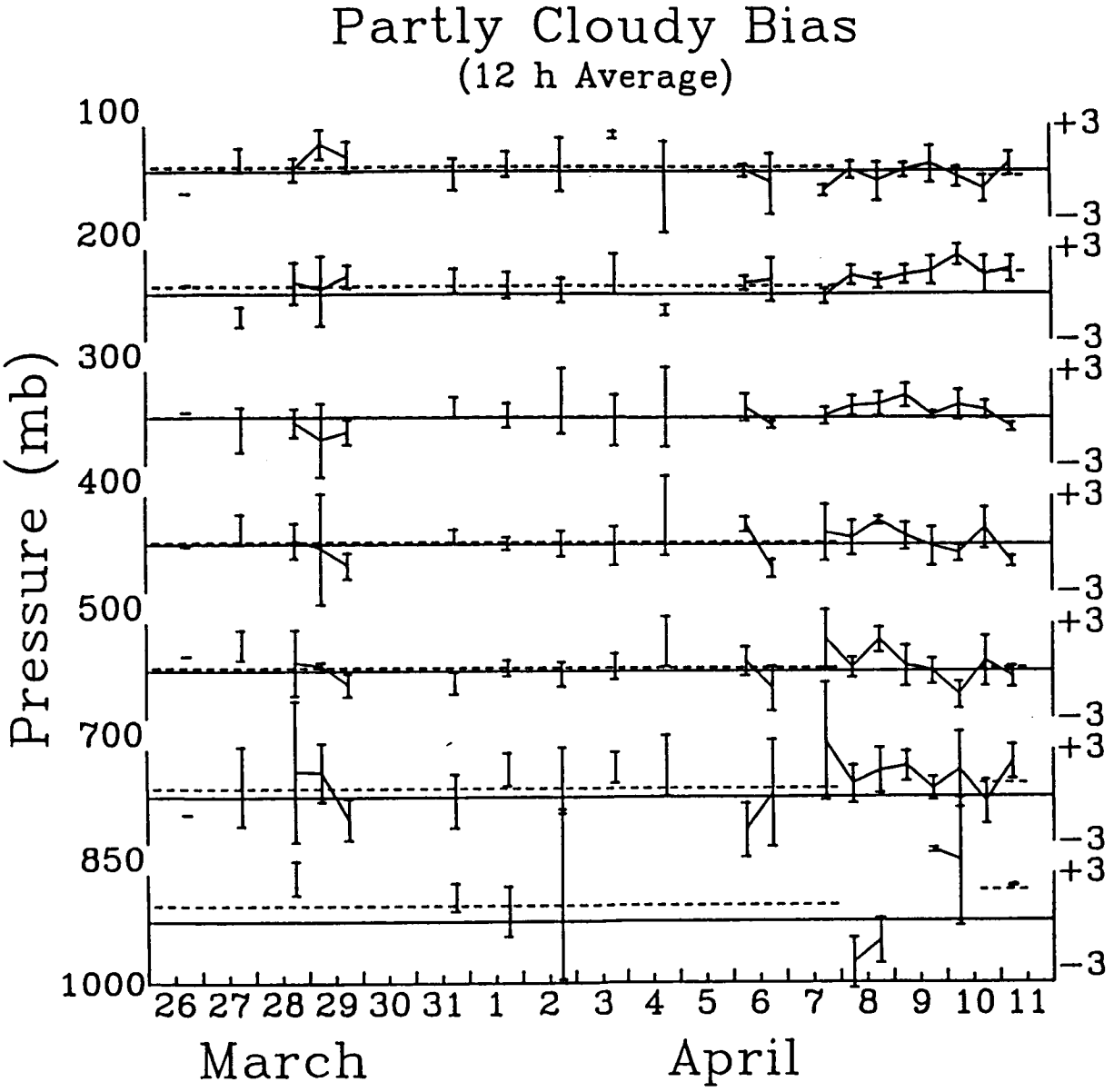

Partly Cloudy Bias
(12 h Average)
Pressure (mb)
100
200
300
400
500
700
850
1000
+3
−3
26 27 28 29 30 31 1 2 3 4 5 6 7 8 9 10 11
March
April

Figure 2b

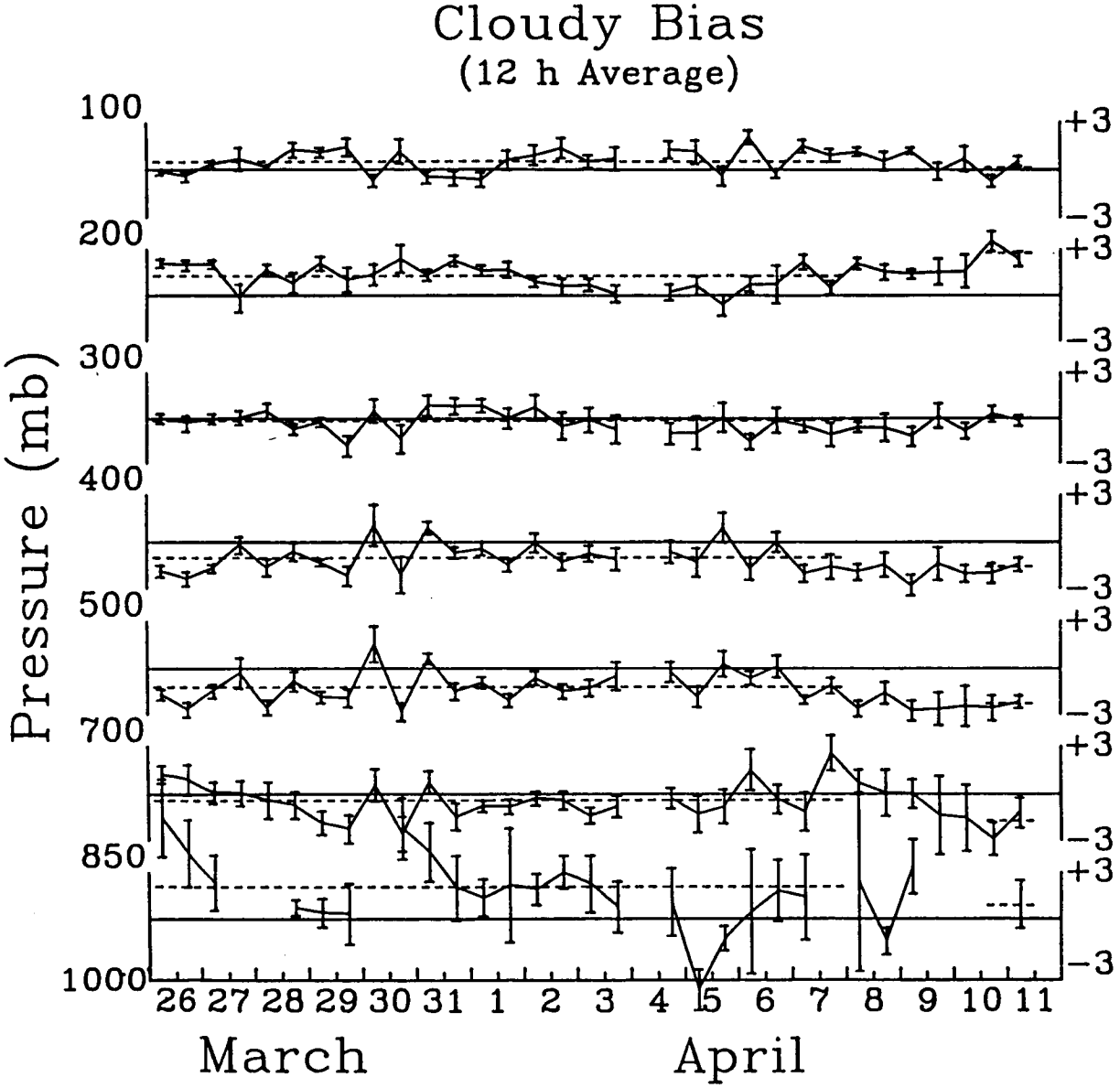

Cloudy Bias
(12 h Average)
Pressure (mb)
100
200
300
400
500
700
850
1000
+3
−3
26 27 28 29 30 31 1 2 3 4 5 6 7 8 9 10 11
March
April

Figure 2c

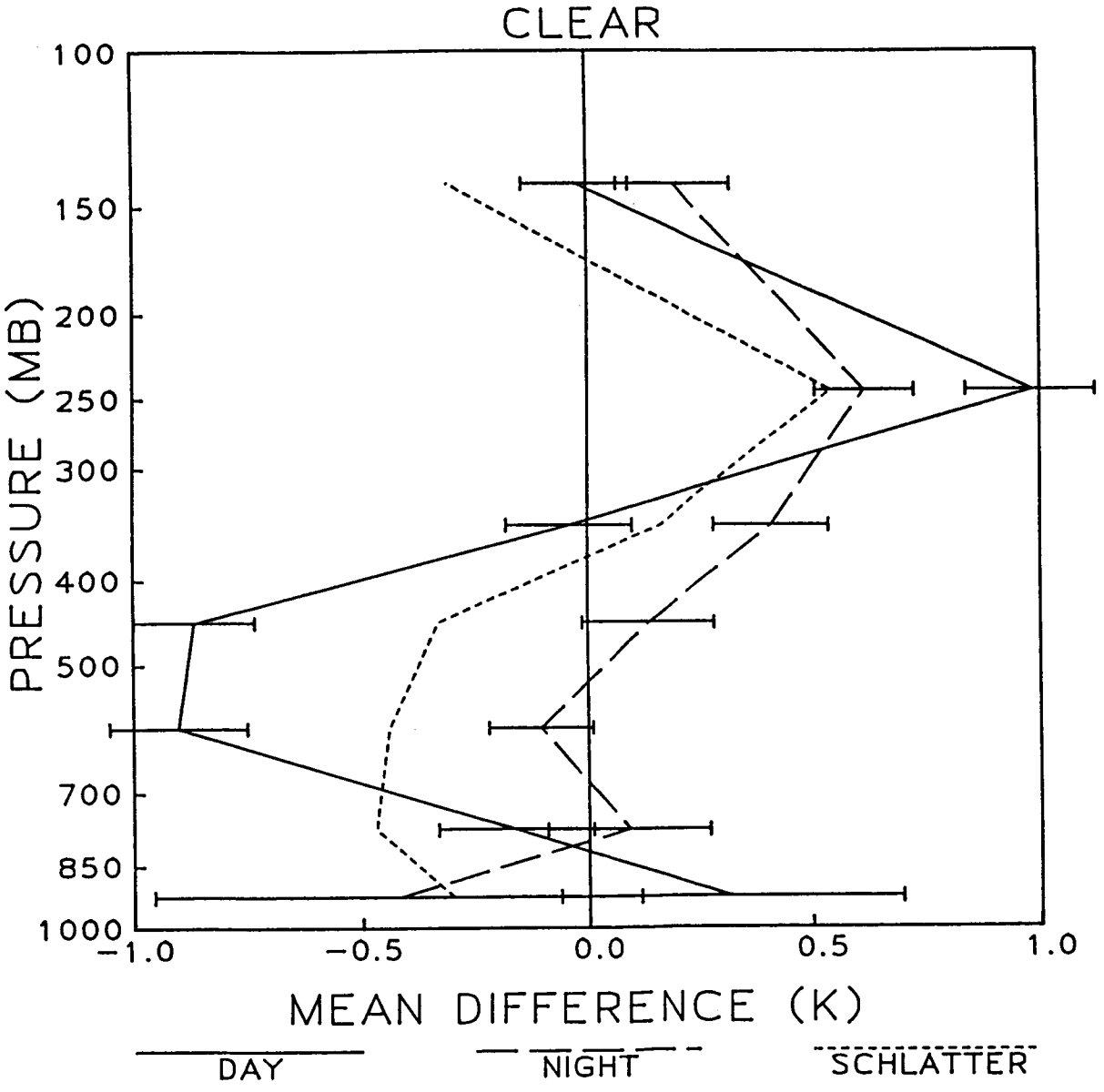

CLEAR
100
150
200
250
300
400
500
700
850
1000
PRESSURE (MB)
−1.0
−0.5
0.0
0.5
1.0
MEAN DIFFERENCE (K)
DAY
NIGHT
SCHLATTER

Figure 3a

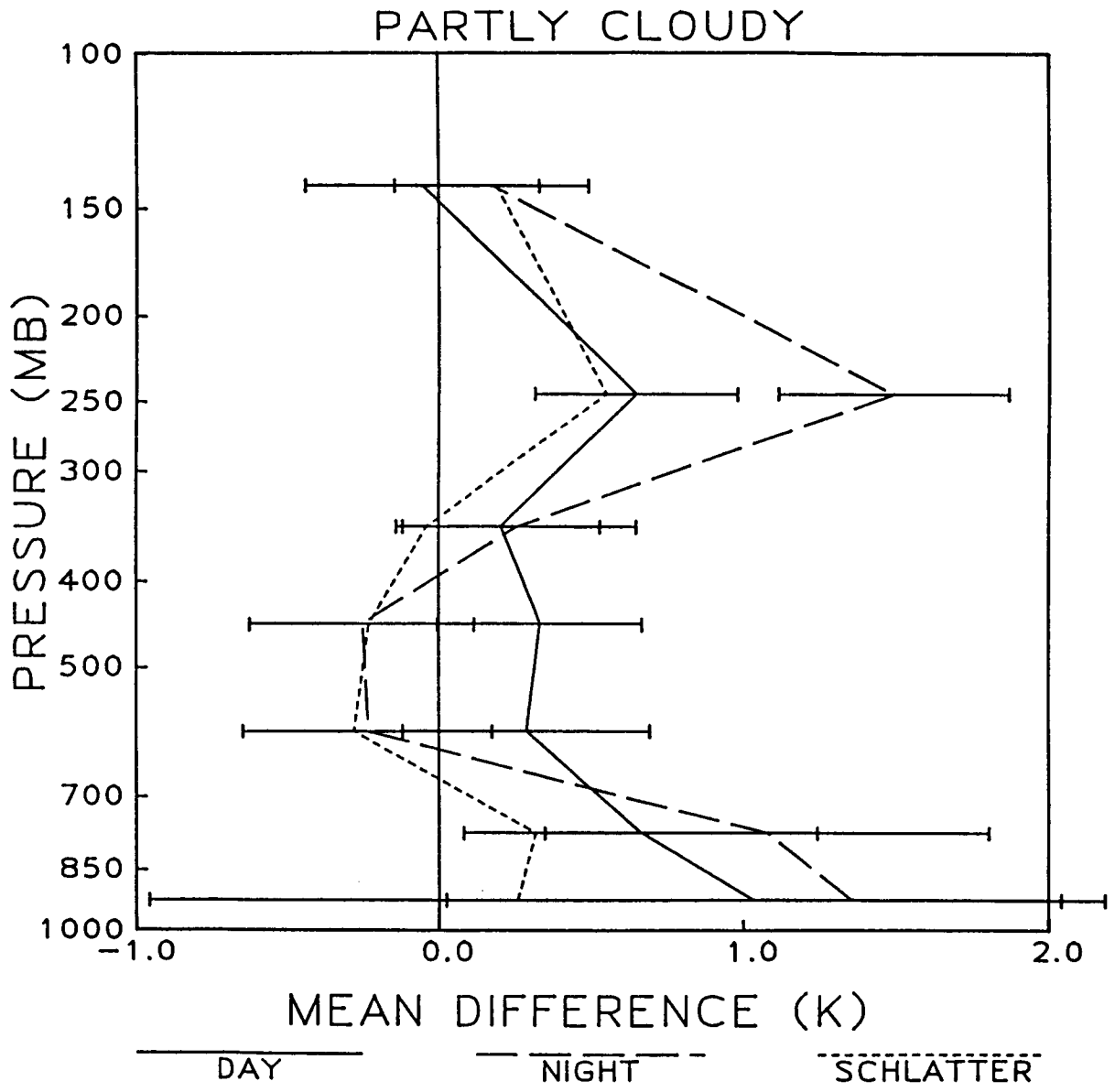

PARTLY CLOUDY
100
150
200
250
300
400
500
700
850
1000
PRESSURE (MB)
−1.0
0.0
1.0
2.0
MEAN DIFFERENCE (K)
DAY
NIGHT
SCHLATTER

Figure 3b

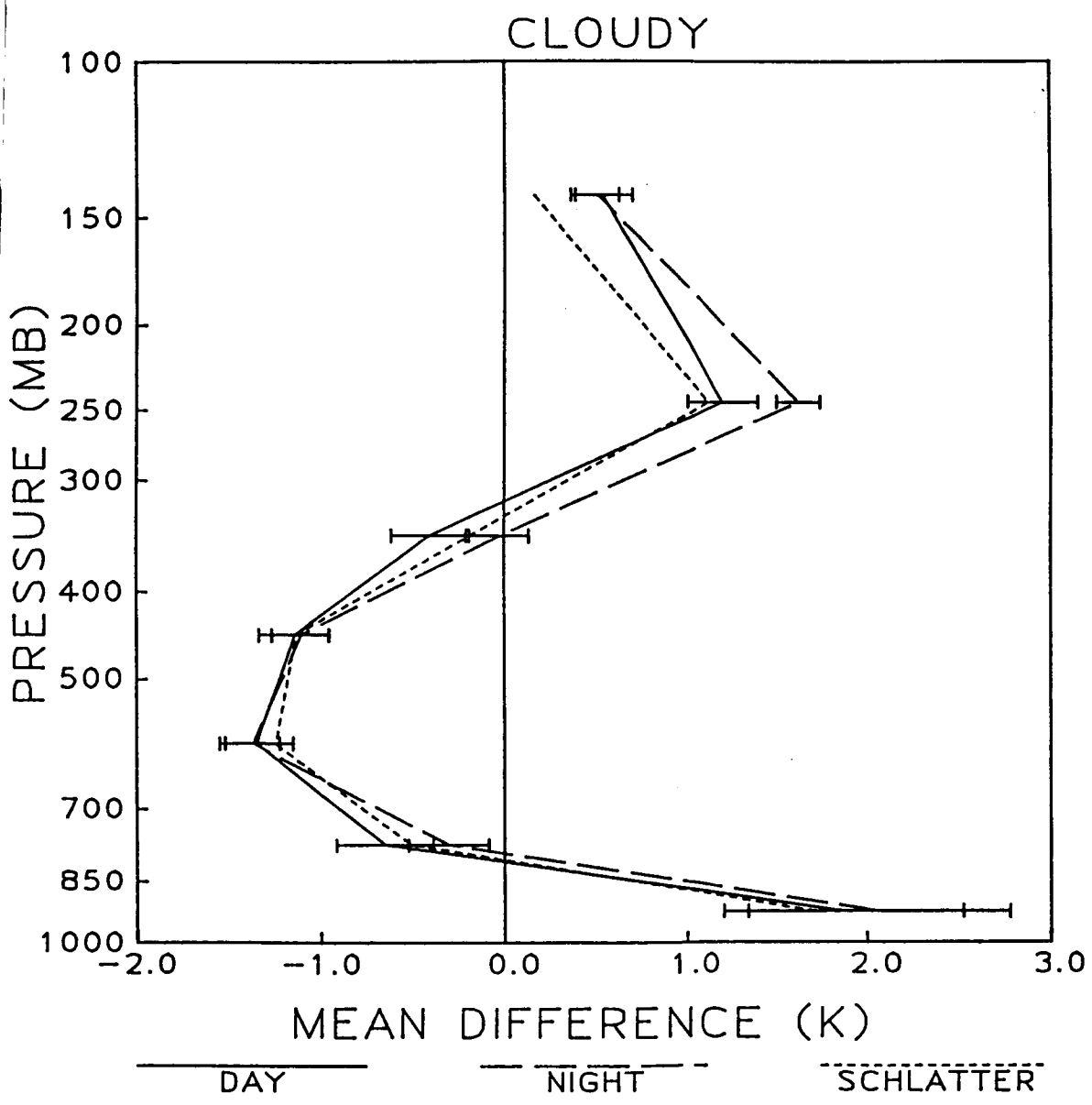

CLOUDY
100
150
200
250
300
400
500
700
850
1000
PRESSURE (MB)
−2.0
−1.0
0.0
1.0
2.0
3.0
MEAN DIFFERENCE (K)
DAY
NIGHT
SCHLATTER

Figure 3c

Chapter VI

The Impact of Data Boundaries upon a Successive Corrections Objective Analysis of Limited-Area Datasets

Gary L. Achtemeier

Climate and Meteorology Section

Illinois State Water Survey

Champaign, Il 61820

ABSTRACT

Successive corrections objective analysis techniques frequently are used to array data from limited areas without consideration of how the absence of data beyond the boundaries of the network impacts the analysis in the interior of the grid. This problem of data boundaries is studied theoretically by extending the response theory for the Barnes (1964, 1973) objective analysis method to include boundary effects. The results from the theoretical studies are verified with objective analyses of analytical data. Several important points regarding the objective analysis of limited-area data sets are revealed through this study.

1) Data boundaries impact the objective analysis by reducing the amplitudes of long waves and shifting the phases of short waves. Further, in comparison with the infinite plane response, it is found that truncation of the influence area by limited-area data sets and/or the phase shift of the original wave during the first pass amplified some of the resolvable short waves upon successive corrections to that first pass analysis.

2) The distance that boundary effects intrude into the interior of the grid is inversely related to the weight function shape parameter. Attempts to reduce boundary impacts by producing a smooth analysis actually draw boundary effects farther into the interior of the network.

3) When analytical tests were performed with realistic values for the

weight function shape parameters, such as the GEMPAK default criteria, it was found that boundary effects intruded into the interior of the analysis domain a distance equal to the average separation between observations. This does not pose a problem for the analysis of large data sets because several rows and columns of the grid can be discarded after the analysis. However, this option may not be possible for the analysis of limited-area data sets because there may not be enough observations.

The results show that, in the analysis of limited-area data sets, the analyst should be prepared to accept that most (probably all) analyses will suffer from the impacts of the boundaries of the data field.

1. Introduction

Field experiments often involve the collection of tropospheric data in networks of limited areal extent. The expense involved in obtaining upper air data usually restricts these networks to no more than 10-15 observing sites. Several limited-area tropospheric sampling networks have been operated during the last two decades to support meteorological research. The National Severe Storms Laboratory (NSSL) operated 8-10 rawinsonde sites during 1966-1970 (Barnes, et al., 1971) and the number of sites ranged from three to nine during the last decade (Alberty, et al., 1977; Doviak, 1981; Taylor, 1982). Other networks operated in the last 15 years included METROMEX: 10 pibal sites in 1971 (Changnon et al., 1971) and 11 sites in 1973 (ISWS, 1974), SESAME: 20 storm scale rawinsonde sites in 1979 (Hill et al., 1979), and CCOPE: five rawinsonde sites in 1981 (News and Notes, 1981).

In the analysis of upper air data from limited-area networks with eight or more measuring sites, the analyst may prefer an objective interpolation of the data from the irregularly spaced observation sites to points on a regular grid. An approach to the interpolation of the data would be the use of a multivariate statistical interpolation method (Gandin, 1963; Schlatter, 1975) found to be useful for the analysis of large data sets with several interrelated parameters. However such a technique requires both good first-guess fields and reliable models of the first-guess field error statistics. These are generally not available for limited-area networks. However, we can use the simpler successive corrections methods such as the techniques of Cressman (1959) or Barnes (1964, 1973).

Regarding the objective interpolation of meteorological data, Eddy (1964) suggested that the analyst take into consideration the data density, the significant wavelengths in the field, the best method for interpolating between observation points, and the noise level in the data. For limited-area data sets, the analyst should also consider the extent to which the absence of observations beyond the boundaries of the data field causes the method to degrade the analysis of the waves defined within the interior of the network. This latter problem is the subject of this paper. We are not concerned with extrapolation although extrapolation is often unavoidable in the transferral of information from irregularly shaped data fields onto a regular grid. Instead, we are concerned with how an objective analysis technique responds to the presence of boundaries in a limited-area data field. We seek answers to the following questions: What impacts are measured at various wavelengths? How far do the impacts extend into the grid interior and what can be done to confine adverse impacts to near the grid boundaries?

The spectral responses of several objective interpolation techniques that use distance-dependent weight functions have been derived with the assumption either that the data were distributed continuously (Barnes, 1964) or that they were distributed uniformly upon a plane within an "influence radius" from some point of interpolation (Stephens, 1967; Stephens and Stitt, 1970). This response theory will be extended to assess the impact of the boundaries of the data field upon an objective analysis.

We will use the successive corrections method developed by Barnes (1964) and extended by him in 1973. This method has found widespread use in the analysis of regional scale and mesoscale phenomena, studies that most often involve the analysis of limited-area data fields. It is also the objective

analysis technique in the GEMPAK program package (Koch et al., 1983) which is being distributed widely within the meteorological community. Throughout the theoretical discussion, it is assumed that a continuum of information exists within the data field. In the real world this is never achieved; the response is degraded further by the discrete data distribution and is beyond the scope of this paper. However, the continuum response provides a baseline for the best analysis achievable near data boundaries.

In Section 2 we formulate the problem and discuss the Barnes analysis technique in the context of first and second pass responses near data boundaries. Section 3 gives examples of the impact of data boundaries upon objective analyses of analytic data, and Section 4 presents a discussion of the results.

2. Impact of Data Boundaries Upon an Objective Analysis - Theoretical Studies

The Barnes (1973) report has become an unofficial instruction manual for those who use his objective analysis method. Therefore, we will adhere to the original nomenclature and developments where possible and will also use original examples to demonstrate the impact of data boundaries upon the analysis.

Suppose an atmospheric variable can be described by a horizontal function $f(x,y)$. Assume a continuum of observations regarding $f(x,y)$, and filter (weigh) these data according to their distance from an arbitrary point (x,y). We wish to determine the relationship between observed value, f, and weighted average value, g, at the same point (x,y),

$$g(x,y) = Q[f(x,y)], \tag{1}$$

where Q is a "response operator" and is wavelength dependent. If the relative locations between a grid point (x,y) and a data point ($x+r\cos\theta$, $y+r\sin\theta$) are as shown in Fig. 1, then the relationship between the true field and the filtered field may be expressed by

$$g(x,y) = \int_0^{2\pi}\int_0^{R_c} f(x+r\cos\theta,\ y+r\sin\theta)\ w(r,k)\ r\ dr\ d\theta \qquad (2)$$

where w(r,k) is a simple Gaussian low pass filter,

$$w(r,k) = [1/4\pi k]\ \exp(-r^2/4k). \qquad (3)$$

The 4k is an parameter which determines the shape of the weighting curve and thus the influence accorded to observations at distance r from (x,y).

Figure 2 shows how the limits of integration apply to (2) when part of the area of integration overlaps the boundary of the data field. We will assume that a continuum of data exists to the left of the data boundary. We also assume that the integration is carried out to some scan radius R_c, ($R_c < \infty$), beyond which the value of the weight function is some very small number so that truncation of the weight function at R_c will not noticeably affect the response characteristics. The interval of integration proceeds counterclockwise through the data-rich part of the scan area beginning at θ_1 and ending at θ_2. The remaining interval of integration covers the area where the scan area overlaps the data boundaries and includes the data-rich triangular area with two sides bounded by R_c and the third side bounded by the edge of the data field.

In the event that the grid point is far enough removed from the data boundary, the integral reduces to the equation for the response for a data

continuum over an infinite plane. Otherwise, a solution for (2) is difficult to obtain because the distance from the grid point to the data boundary is a function of the angle, θ. If f(x,y) is an idealized monochromatic data field of the form Asin(ax) where $a = 2\pi/\lambda$, then g (x,y) is determined by the weighted sum of the original function, f (x,y), with the original function shifted 90 degrees out of phase,

$$g(x,y) = D(a,k)\ f(x,y) + E(a,k)\ h(x,y) \qquad (4)$$

where h(x,y) = A cos(ax). The amplitude responses, D(a,k) and E(a,k) are integrals of higher-order Bessel functions. Barnes (1964) presented the analytical solution for D(a,k) for data distributed over an infinite plane. In his 1973 paper, he showed that E(a,k) vanishes under the same conditions. These conditions are not satisfied near data boundaries and both integrals are non-zero. We have solved them numerically.

a) Theoretical Response for a Single Data Boundary

Throughout this development, a continuum of data within the boundaries of a finite data field (Fig. 2) is assumed. If f(x,y) is specified, then the weight function shape parameter and the length of the wave are all that are required to find the responses D(a,k) and E(a,k). However, to better relate the theoretical results from (4) to distances measured from the edge of the data field, we introduce a length scale $S = \lambda^*/2$ where λ^* is a reference wavelength. The reference wavelength is chosen to equal the minimum resolvable wave, the final response of which must be prespecified in the GEMPAK method (Koch et al., 1983). Thus, if the methods described here are applied to the analysis of real data, S is equivalent to the average spacing between

discretely distributed observations.

The first pass responses for the first term of (4) at selected distances from a data boundary are shown in Fig. 3a. Distances are given in fractions of S (S = 10 km). The shape of the weighting curve is 4k=64, a value used by Barnes for the first pass analysis of data distributed on the 1970 NSSL surface mesonetwork. The response for a grid point removed a distance, 2S, from the data boundary is unchanged from the response for an infinite plane of continuous data for the range of wavelengths in Fig. 3a. At distance S from the boundary, the responses for the medium and long wavelengths are slightly less than the infinite plane responses - an indication that these waves receive additional damping due to boundary effects. Additional smoothing is clearly implied for all wavelengths when the grid point is located at distances less than 0.67S from the data boundary. The Gaussian filter degrades the spectrum of waves to the extent that less than 50 percent of the amplitudes of the long waves are restored at the data boundary. Further, upon extending this analysis to very long waves, it is found that D(a,k) approaches 0.5 in the limit as $\lambda \to \infty$.

Figure 3b shows how the second term of (4) shifts the phases of the waves near the data boundaries. Phase shifts are negligible at distances greater than S from the data boundary. Maximum phase shifts occur at the data boundary and for the short but resolvable waves in the range 20-60 km. Approximately 30 percent of the amplitude of the 30-km plane wave appears as a phase shifted wave at the data boundary.

The magnitudes of the impacts that the absence of data beyond the boundaries of a data field have upon filter fidelity at any location within an analysis grid are also dependent upon the shape factor 4k. Figure 4 shows

response curves for four values of 4k at a grid point located at a distance of 0.67S from the data boundary. The response for 4k=16 shows no significant impact because that part of the scan area where relatively large weights are accorded to the data does not overlap the data boundary. Thus, in a sense, the filter does not "see" the data boundary when 4k=16. A value of 4k=205 produces first pass response characteristics in both wave amplitude and phase shift at 0.67S similar to the response 4k=64 would produce at about 0.5S from the boundary (Fig. 3b). Since the larger 4k increase the effective scan areas, the deleterious impacts of the data boundary upon filter fidelity must increase in magnitude and must appear at greater distances into the grid interior because a greater percentage of the scan areas will overlap the data boundary.

The interpolation method may be modified to obtain the desired response at small wavelengths by applying correction pass(es) through the initial interpolation field. In application, we perform the nth pass by finding $g_{n-1}(x,y)$ through bilinear interpolation and then adding to the previous (n-1) pass field the smoothed residual difference between the observed data values and the (n-1)th pass estimated values at the data location. Thus,

$$g_n(i,j) = g_{n-1}(i,j) + Q_n\,[f(x,y)-g_{n-1}(x,y)]\ , \qquad (5)$$

where the general response operator, Q_n, may or may not take on the same value as for the previous pass.

For reasons of computer economy, Barnes (1973) modified the original analysis technique so that only one correction pass through the data is required to achieve the desired response at small wavelengths. By this method, the filter is made to return more of the amplitude of the short waves through a reduction of the shape factor by a fraction, γ. This procedure is analogous to decreasing the influence radius for the Cressman (1959) analysis technique except that the number of observations within the scan area remains the same. Instead of reducing the number of observations, the weights are adjusted so that the relative importances of the observations closest to the grid point are increased on the correction pass. By this method, the estimated values at the grid points are given by

$$g(i,j) = D' f(x,y) + E' h(x,y) \tag{6}$$

where the final responses, D' and E', for the modified analysis method are given by

$$D' = D_0 + D_1 - (D_0 D_1 - E_0 E_1) \tag{7}$$

$$E' = E_0 + E_1 - (D_0 E_1 + E_0 D_1) \tag{8}$$

If a grid point is located at greater distance from data boundaries, then $E_1 = E_0 = 0$, $E' = 0$ and D' reduces to the form given by Barnes (1973). Otherwise, the amplitudes of the phase shift term are non-zero where data boundaries influence the analysis. E_0 and E_1 are related through γ, both are of the same sign, their product is always positive, and therefore the phase shift excited at the first pass always increases the final response. The solid curves in Fig. 5a are examples of the final responses, D', after one correction pass at a point located at the data boundary. The value of 4k is 205 and

is equal to the value of 4k for the first pass of the GEMPAK objective analysis. The correction pass value for γ is 0.2. The curve labeled E = 0 is the final response calculated with E = 0 and it serves as a baseline for evaluating the impact of the phase shift upon the final analysis. The data boundaries cause the method to restore only about 70 percent of the original long waves. A comparison of the two curves illustrates the importance of the phase shift terms in increasing the response for the short wavelengths in the range 30 to 60 km, the increases for these waves ranging from six to ten percent. This is the range of waves for which the first pass response E_0 is greatest (see Fig. 4b for reference).

The tradeoff is that the data boundaries also excite large amplitude phase-shifted waves for the same range of waves near the data boundary (Fig. 5b). The maximum amplitude of the phase shifted waves occurs at the 30 km wave and is 43 percent.

It is expected that the impacts of the data boundaries will vary among objective analyses obtained by other methods or from the same method with different control parameters. The dashed curves in Fig. 5a are examples of the final responses, D′, after three correction passes with the Barnes (1964) method. We set the shape factor 4k = 205. A comparison between the dashed curve labeled E = 0 and the two solid curves, shows that this method substantially improves the fidelity of the Barnes filter near data boundaries for wavelengths greater than 40 km. Approximatley 85 percent of the amplitudes of these waves are restored. Inclusion of the phase shift terms improves the fidelity of the filter still more with the greatest increases in the 30 to 60 km range. This is the range of waves for which the first pass response E_0 is

greatest. These increases range from only six percent for the 20 km wave to 38 percent for the 40 km wave.

However, this analysis also produces large amplitude phase shifted waves for the same range of waves near the data boundary (Fig. 5b). Seventy-five percent of the amplitude of the original 30 km wave appears as a phase shifted wave at the data boundary. More than 50 percent of the amplitudes of the 20 and 40 km waves are returned out of phase. Phase shifting is a lesser problem with the longer waves. Further, the amplitudes of these phase shifted waves and those obtained with the single correction pass method become negligibly small for all waves where the distances from the boundary of the data set exceed S.

b) Theoretical Response for Limited Area Data Sets

The previous discussions have focused upon the impacts the single boundary of a data field have upon the filtering characteristics of the Gaussian weight function. We now turn to the limited area data set and consider that the response at <u>all</u> points within the small grid network may be impacted to some degree by one or more data boundaries. Keeping the numerical approximation to the general response equation, we modify the geometry of the data field by assuming that the data are distributed uniformly within a circle with a diameter equal to 2S.

The final response curves for limited area data sets (Fig. 6) are labeled in distances measured in fractions of S from a data boundary. They begin at the data boundary (0) and terminate at the center of the circle (S). The response curves are for the low pass filter designed to produce the infinite plane response identical to the GEMPAK default criteria (D_0=0.0064 and γ=0.2)

at the 2S wavelength. When applied to the limited area data set, this weight function modifies the character of all of the waves studied (Fig. 6a). All of the longer waves are filtered. The filtering is most extensive at the data boundary; however, the amplitudes of long waves at the center of the data area also are reduced. Short waves are amplified by this analysis. Responses for the waves in the range from 20 to 40 km wavelength are increased above the infinite plane response calculated with identical parameters (dashed line). Maximum increases at the 2S wave (20 km) approach 10 percent at the center of the limited-area data field. These increases cannot be explained by the addition of the phase shift term in (14) because the data are distributed symmetrically about the central point. This satisfies the condition for the phase shift to vanish and the phase shift does vanish (Fig. 6b line labeled S). Instead, the short waves amplify because the influence area is truncated at the data boundaries. Moreover, the magnitude of the amplification depends upon the extent of truncation and hence upon the size of the limited area data field. Figure 7 shows the differences between the truncated final response and the infinite plane final response for the 2S wave if R_c varies in the range from zero to 2S. The differences increase from $-e^{-1}$ (the truncated final response is equal to zero if there is only one data point) to +0.12 if R_c is equal to approximately 0.67S. The truncated final response approaches the final response for the infinite plane for $R_c \geq 1.67$ S.

3. Examples of Impact of Data Boundaries Upon Objective Analyses

In this section, we use objective analyses to show that the impacts of data boundaries extend for significant distances into the grid interiors. The

data locations are colocated with grid points on a 21 by 21 grid with a 3.175 km grid spacing so that there is no need for any additional interpolation to estimate values of the gridded fields at off-grid data locations. It also allows the direct comparison of the objectively filtered fields with the predictions of the response theory in Section 2. The final filtered value at each grid point after L correction passes through the data is the weighted average of M*N observations plus the sums of the correction passes according to

$$g_L(i,j) = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} w_{0,m,n}\, f(m,n)}{\sum_{m=1}^{M}\sum_{n=1}^{N} w_{0,m,n}} + \sum_{\ell=1}^{L} C_\ell \,, \qquad (8)$$

where

$$C_\ell = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} w_{\ell,m,n}\,[f(m,n) - g_{\ell-1}(m,n)]}{\sum_{m=1}^{M}\sum_{n=1}^{N} w_{\ell,m,n}} \qquad (9)$$

The weight function, $w_{\ell,m,n}$ is given by

$$w_{\ell,m,n} = \exp\,[-r^2(m,n)/4k\,]. \qquad (10)$$

The data are taken from analytic functions which include sloping plane surfaces and monochromatic waves that range from 20 km through 80 km. We use either three correction passes with $w_{\ell,m,n} = w_{0,m,n}$ or the GEMPAK default criteria, one correction pass with $w_{1,m,n} = 0.2\, w_{0,m,n}$.

Figure 8 demonstrates the impacts of data boundaries on a three-correction pass analysis and upon a one-correction pass analysis for a 20 km monochromatic wave (Fig. 8a). The initial value for 4k was 205. Filtering the wave field by (8) with three correction passes does not restore this 2S wave (Fig. 8b). However, a phase-shifted wave of amplitude comparable with the amplitude of the original wave appears near the boundaries in accordance with the response theory developed in the previous section (compare with Fig. 8a). We subtract from the analysis in Fig. 8b a filtered wave determined from response theory for data distributed over an infinite plane. This leaves the phase shifted wave as a remainder located near the boundary (Fig. 8c). The residual of 10 units corresponds to approximately 63 percent of the amplitude of the original wave. This compares favorably with theory which predicts a phase shifted wave with amplitude equal to 56 percent of the original wave (Fig. 5b). The one-correction pass analysis run with the GEMPAK default criteria restores e^{-1} of the amplitude of the original wave (Fig. 8d). Subtraction of the infinite plane component of this filtered wave also leaves a reversed phase wave nearly identical to the wave in Fig. 8c.

The analysis modelled after a limited area data set demonstrates that the absence of observations beyond the boundaries of the data field can have a significant impact over the whole analysis domain. We use a monochromatic 60 km wave for this part of the study. The wave is filtered with the one-correction pass method subject to GEMPAK default criteria which assumes that this wave is equivalent to the minimum resolvable wave. We then subtract a filtered wave determined from infinite plane response as was done in developing Fig. 8c. The line (curve 1) in Fig. 9 shows that this filter draws boundary effects into the interior of the analysis. (If this wave is equivalent

to the minimum resolvable wave then one station separation is equivalent to the length scale S.) The magnitudes of these boundary effects are in percentages of the amplitudes of the original wave and are in general agreement with predictions of theory. Decrease the initial 4k by a factor of four and (7) reduces the magnitude of the phase-shifted wave and concentrates it nearer the data boundary (curve 2). Increase the initial 4k by a factor of four to smooth out the undesirable boundary effects and (7) reduces the amplitude of the phase-shifted wave (at least for this 60 km wave) but draws the boundary effects into the grid interior (curve 3).

4. Discussion

Successive corrections objective analysis techniques have often been used to analyze (filter) data taken from limited area networks onto a regular mesh without regard for the impacts upon the analysis in the interior of the grid caused by the absence of data beyond the boundaries of the network. The response theory for the Barnes objective analysis methods was extended to include boundary effects and was compared with objective analyses of analytic data. The analytic data was distributed semi-continuously over grid points of a fine scale mesh. Several important points regarding the objective analysis of limited area data sets were revealed through this study.

a) Both the theoretical and the analytic studies showed that data boundaries can have a significant impact upon waves defined within the interior of an objective analysis. The most deleterious boundary effects were that the long waves were filtered and the short waves were phase shifted. Long waves suffered losses in amplitude of up to 50 percent. Up to 70 percent of the amplitudes of the short waves were restored out of phase. It was also found that, upon use of multiple-pass filtering, boundary effects amplified resolvable short wavelengths in the range from 2S to 6S relative to the responses

predicted by theory for points unaffected by data boundaries. The causes for these relative amplifications were feedbacks from waves shifted out of phase on the first pass and/or truncation of the influence area by limited-area data sets. The magnitudes of the feedbacks from phase shifted waves were sensitive to the wavelength. Relative amplifications ranged from about 6 percent for the 2S wave to 30 percent for both the 3S and 4S waves - an apparent increase in the filter fidelity of the Barnes methods near the data boundaries. The maximum relative amplification caused by the truncation of the influence area by small data sets occurred for the 2S wave and was approximately 12 percent.

b) The distance that boundary effects intruded into the interior of the grid was a function of the weight function shape parameter 4k. Attempts to decrease boundary effects through a smooth analysis obtained by using large initial 4k actually drew boundary effects farther into the grid interior. Reducing 4k decreased and concentrated the boundary effects to near the grid boundaries. However, the analyst should be aware that a reduction of 4k modifies the response characteristics to permit short wavelengths, a tradeoff that may cause phase changes and aliasing of waves within the interior of the grid if the observations are unevenly spaced.

c) After the analytic tests were performed with realistic values for 4k, such as the GEMPAK default criteria, it was found that boundary effects intruded into the interior of the analysis domain a distance equal to roughly one half the length of the wave. If the wave is the minimum resolvable wave, then this distance is equivalent to the average separation between observations. This poses no serious problem for the analysis of large data sets. The analysis area can be designed so that some data fall outside the grid or so that several rows and columns of the grid can be discarded after the analysis. This latter approach has been proposed by Koch et al. (1983). These options are not always possible for the analysis of limited area data sets; there may not be enough observations. The analyst must be prepared to accept that the data boundaries will modify the response characteristics within the interior of the analysis domain. For example, if a limited area data set consists of nine evenly spaced observations sited so that eight stations form the boundary and one station is at the center of the network, and if boundary effects penetrate to a distance equal to the average station spacing, then all of the domain will suffer to some extent from boundary impacts.

d) The analysis presented here is a "worst case scenario" as regards the phase of the original function f (x,y) in determining the final response for the Barnes filter near data boundaries. Our investigation of the phase-shift term of (4) revealed that, if f (x,y) = A cos (ax), then the phase-shifted wave is h (x,y) = A sin (ax). This wave vanishes at the data boundary where X = 0. In addition, it was found that the integral, E (a,k), changes sign when f (x,y) = A cos (ax). It follows, therefore, that E (a,k) must decrease to zero somewhere within the range of phases 0 to 1/2 for the original wave. Thus the

maximum magnitudes for both E (a,k) and h (x,y) are permitted where f (x,y) = A sin (ax).

e) Throughout this discussion, it has been assumed that a continuum of information exists within the limited area domain. The analysis of analytic data has been carried out with a densely distributed regularly spaced data field. A data continuum was approached for some waves. We have emphasized that the results herein are the best that can be expected for the Barnes analysis schemes. In the real world, the data are arrayed discretely and the data distribution further degrades the response to the filtering process. If the data are not evenly distributed, then phase changes and a higher "noise" level are inherent to the analyzed field. We have not emphasized the analysis of unevenly spaced data because these degrading factors are dependent upon the data - the phenomena represented by the data, the data distribution, and the boundaries of the data field. And, when the observation platform is suspended within the wind field, the data distribution and the boundaries of the data field are variables determined by the phenomena represented by the data.

In conclusion, consider the applications of limited area data sets from field programs designed for the investigation of mesoscale and/or regional scale phenomena. If the purpose of a limited area network is the acquisition of information on the spatial distribution of meteorological variables, including gradients of the wind field, the analyst should be prepared to accept that most (probably all) of the analyses will suffer to some extent from the impacts of the boundaries of the data field. It should be kept clear, however, that the above conclusion is based upon the results of an investigation with the Barnes objective analysis schemes. An improved objective analysis scheme that concentrates boundary effects to near the grid boundaries and produces what can be called a "good analysis" in the interior of a limited area domain is essential to provide an accurate description of the local structure of the atmosphere with a limited area data field.

Acknowledgements

This research was supported by NASA under contract NAS8-34902. Ms. Rebecca Runge is acknowledged for typing the manuscript.

References

Abramowitz, M., and I. A. Stegun, eds., 1965: Handbook of Mathematical Functions, (Dover Publications, Inc., New York).

Alberty, R., J. Weaver, D. Sirmans, J. Dooley, and B. Bumgarner, 1977: Spring Program '76. NOAA/Tech. Memo. ERL NSSL-83. 130 pp. (NTIS PB2807451AS).

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. J. Appl. Meteor., 3, 396-409.

Barnes, S. L., 1973: Mesoscale objective analysis using weighted time-series observations. NOAA Tech. Memo. ERL NSSL-62, National Severe Storms Laboratory, Norman, OK 73069, 60 pp. (NTIS COM-73-10781).

Barnes, S. L., J. H. Henderson, and R. J. Ketchum, 1971: Rawinsonde observation and processing techniques at the National Severe Storms Laboratory. NOAA Tech. Memo. ERL-NSSL-53, 246 pp. (NTIS COM-71-00707).

Changnon, S. A., Jr., F. A. Huff, and R. G. Semonin, 1971: METROMEX: An investigation of inadvertent weather modification. Bull. Amer. Meteor. Soc., 52, 958-967.

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

Doviak, R. J., 1981: 1980 Spring Program Summary. NOAA Tech. Memo. ERL NSSL-91, 124 pp. (NTIS PB81-234940).

Eddy, A., 1964: The objective analysis of horizontal wind divergence fields. Quart. J. Roy. Meteor. Soc., 90, 424-440.

Gandin, L. S., 1963: Objective Analysis of Meteorological Fields. (Translated (1965) by Israel Programme for Scientific Translations, Jerusalem, 242 pp.)

Hill, K. G. S. Wilson, and R. E. Turner, 1979: NASA's participation in the AVE-SESAME '79 program. Bull. Amer. Meteor. Soc., 60, 1323-1329.

ISWS, 1974: Project METROMEX: An overview of Illinois State Water Survey Projects. Bull. Amer. Meteor. Soc., 59, 89.

Koch, S. E., M. DesJardins, and P. J. Kocin, 1983: An interactive Barnes objective map analysis scheme for use with satellite and conventional data. J. Climate and Appl. Meteor., 22, 1487-1503.

News and Notes, 1981: CCOPE begins Spring 1981. Bull. Amer. Meteor. Soc., 62: 683-684.

Schlatter, T., 1975: Some experiments with a multivariate statistical objective analysis scheme. Mon. Wea. Rev., 107: 458-476.

Stephens, J. J., 1967: Filtering responses of selected distance-dependent weight functions. Mon. Wea. Rev., 95, 45-46.

Stephens, J. J., and J. M. Stitt, 1970: Optimum influence radii for interpolation with the method of successive corrections. Mon. Wea. Rev., 98, 680-687.

Taylor, W. L., 1982: 1981 Spring Program Summary. NOAA Tech. Memo. ERL NSSL-93, 97 pp. (NTIS PB82-244757).

Figure Captions

Figure 1. Coordinate system used in objective analysis expressed by (1). Point (x,y) is conveniently chosen as a grid point of a square mesh; point (x + rcos , y + rsin) represents one point where information is observed. Theoretically, these are continuously arrayed over the x-y plane, but in the practical application, they are discrete points, irregularly arrayed. (After Barnes, 1973).

Figure 2. Schematic showing the intersection of an influence area about a grid point with a data boundary.

Figure 3. a) Responses for 4k=64 for the first term of (4) at selected distances from the grid boundary. Distances from data boundary measured in fractions of S. b) Responses for the second term of (4) for the same distances from the grid boundary.

Figure 4. Response for four values of 4k at a grid point location two grid spaces from the data boundary.

Figure 5. a) Final response, D', for 4k=205 at the data boundary. b) Final response, E', for the same 4k. Solid lines are responses for the one-correction pass method and dashed lines are responses for the three-correction pass method.

Figure 6. Final response curves in fractions of S from the data boundary for a limited-area data set. a) Amplitude response, D' and b) phase shift response, E'. Response curves obtained with GEMPAK default criteria.

Figure 7. Differences between truncated and infinite plane responses for the 2S wave for different sized data areas. Response curve obtained with GEMPAK default criteria.

Figure 8. a) A 20 km 2S monochromatic wave used for analytical studies of boundary effects, b) A 3-correction pass Barnes analysis of the wave, c) Boundary impacts upon analysis and d) Analysis of same wave with GEMPAK default criteria.

Figure 9. Cross section along a 60 km wave used for three objective analyses of analytical limited-area data field. Boundary effects expressed as percentages of the amplitude of the original wave.

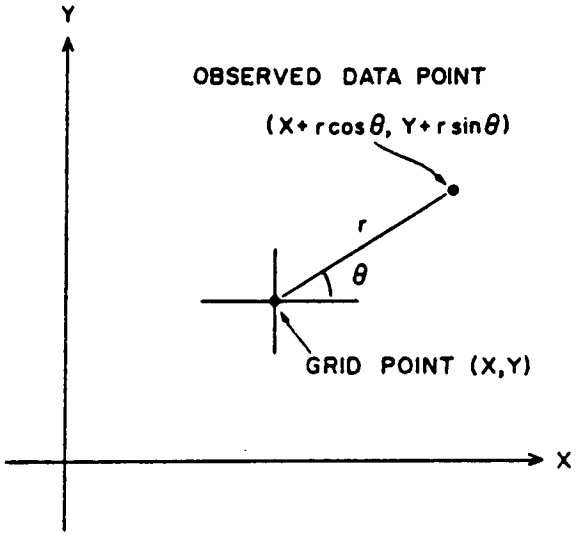


Figure 1

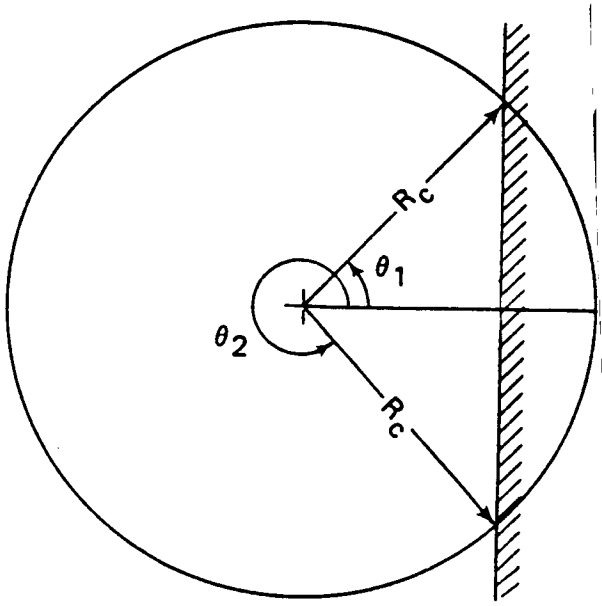


Figure 2

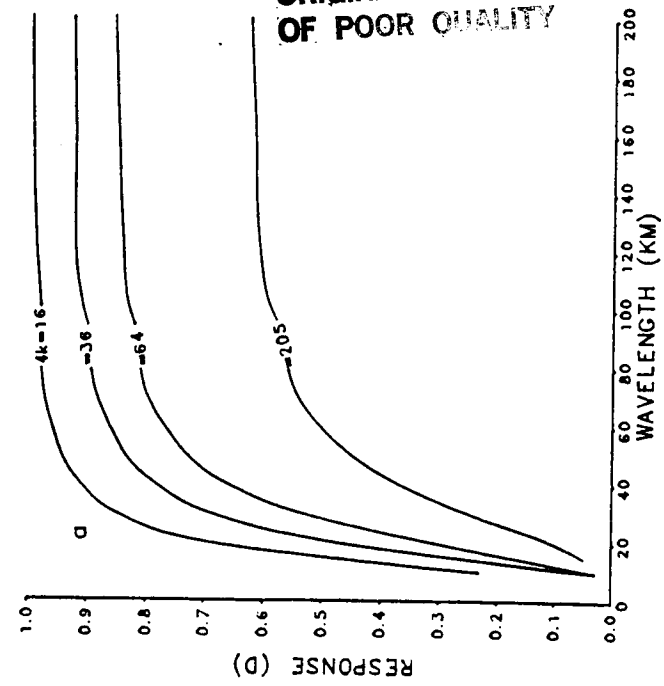


Figure 4

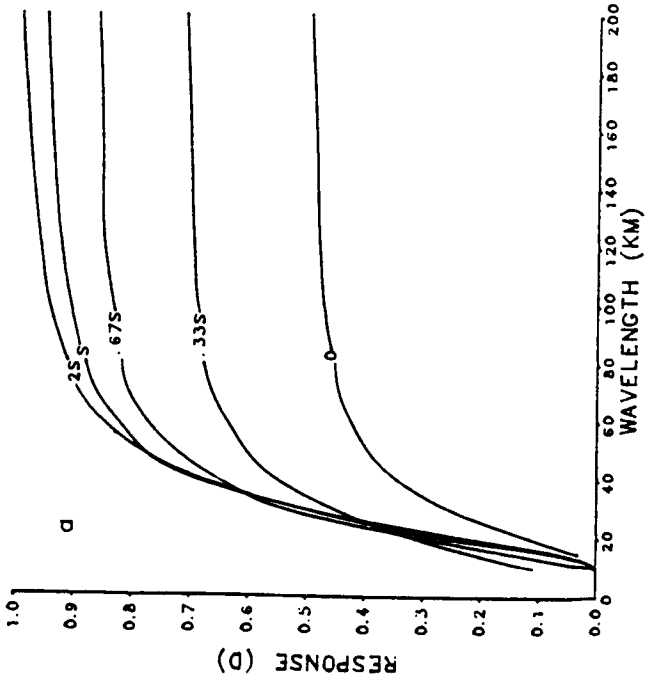


Figure 3

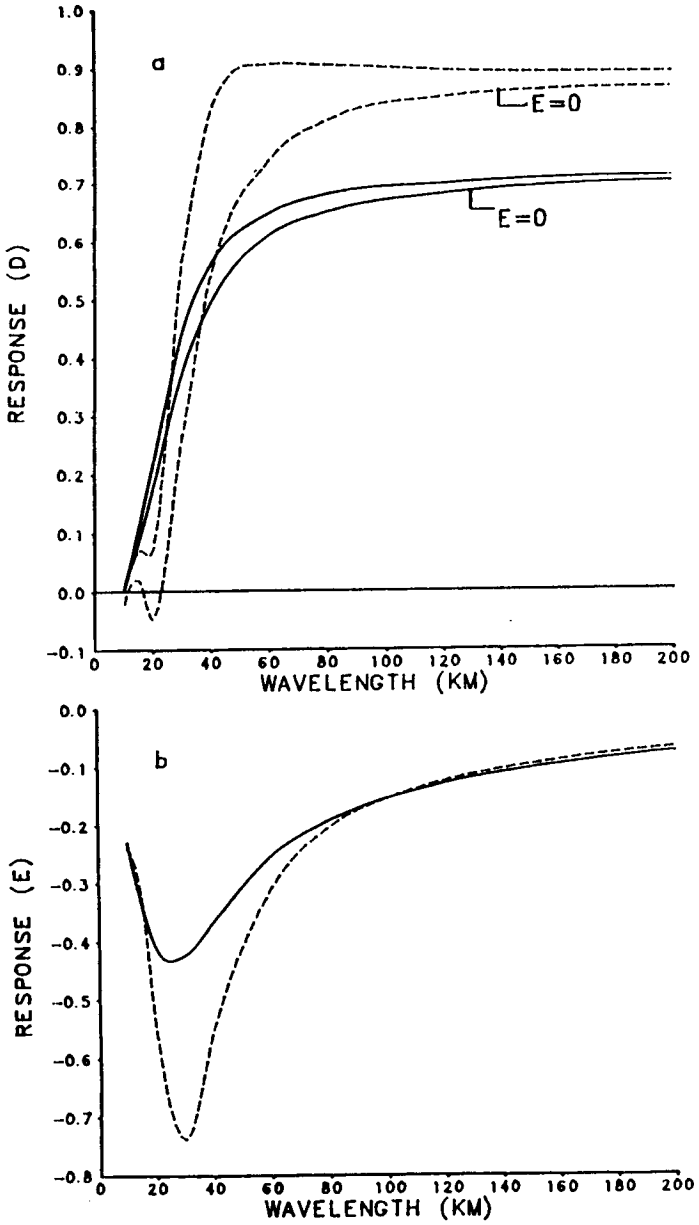


Figure 5

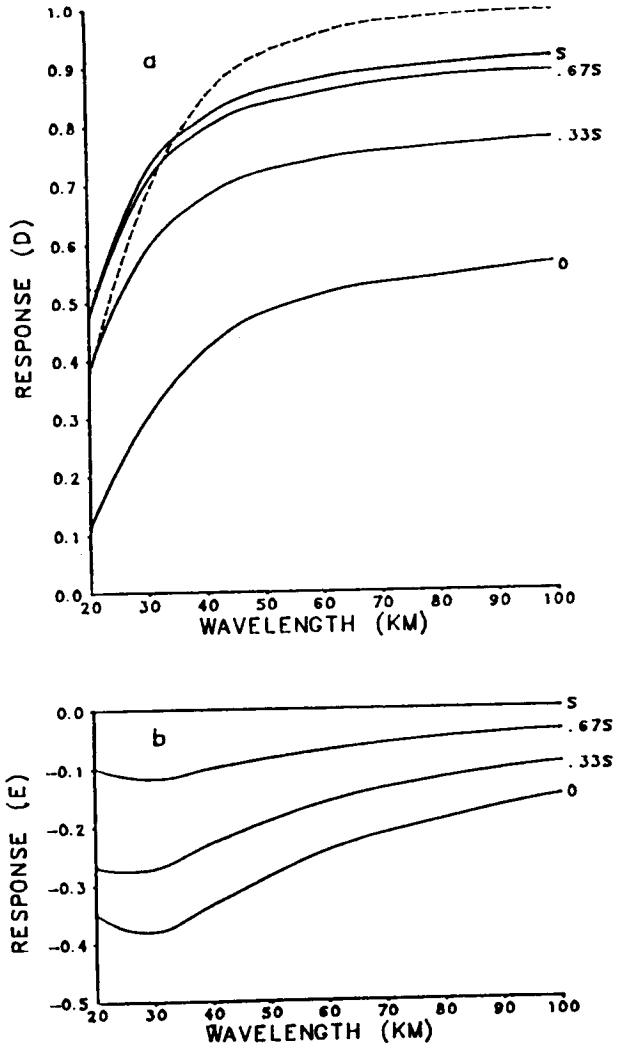
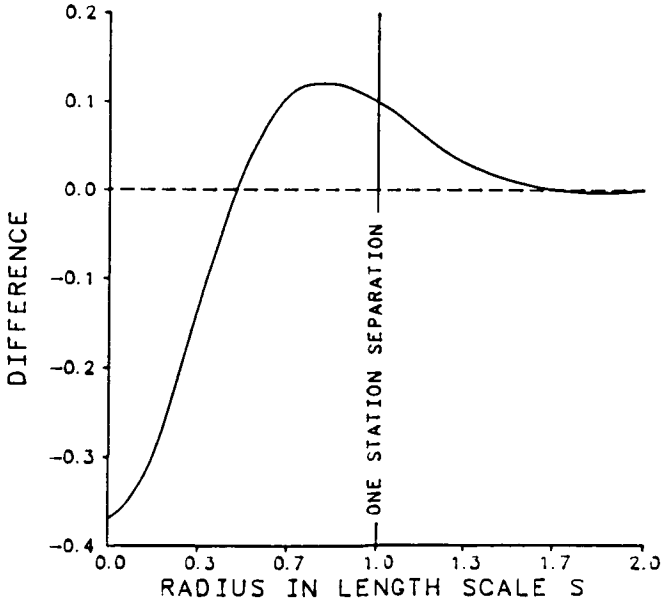


Figure 6

Figure 7


0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
DIFFERENCE
ONE STATION SEPARATION
0.0
0.3
0.7
1.0
1.3
1.7
2.0
RADIUS IN LENGTH SCALE S

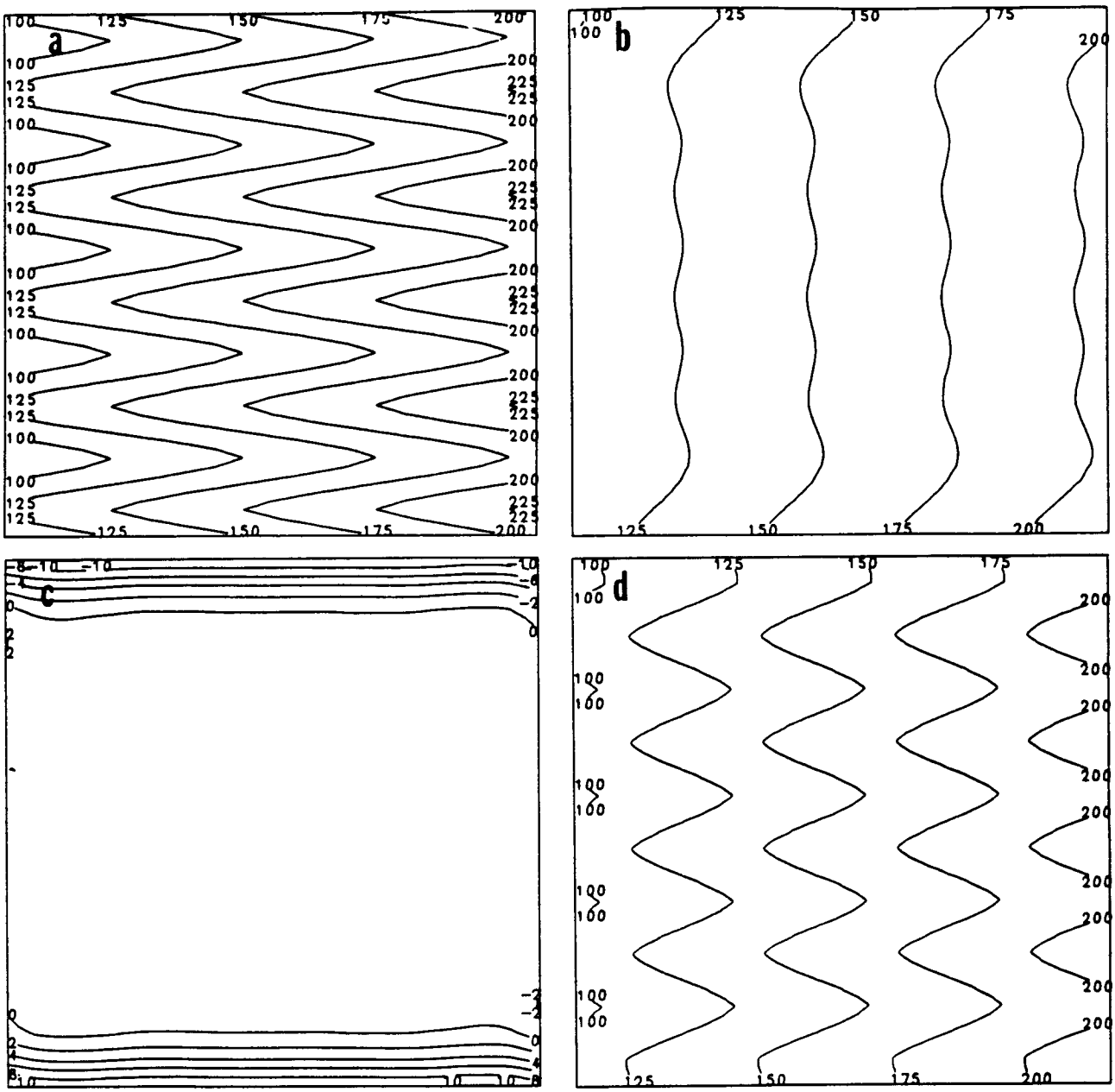


Figure 8

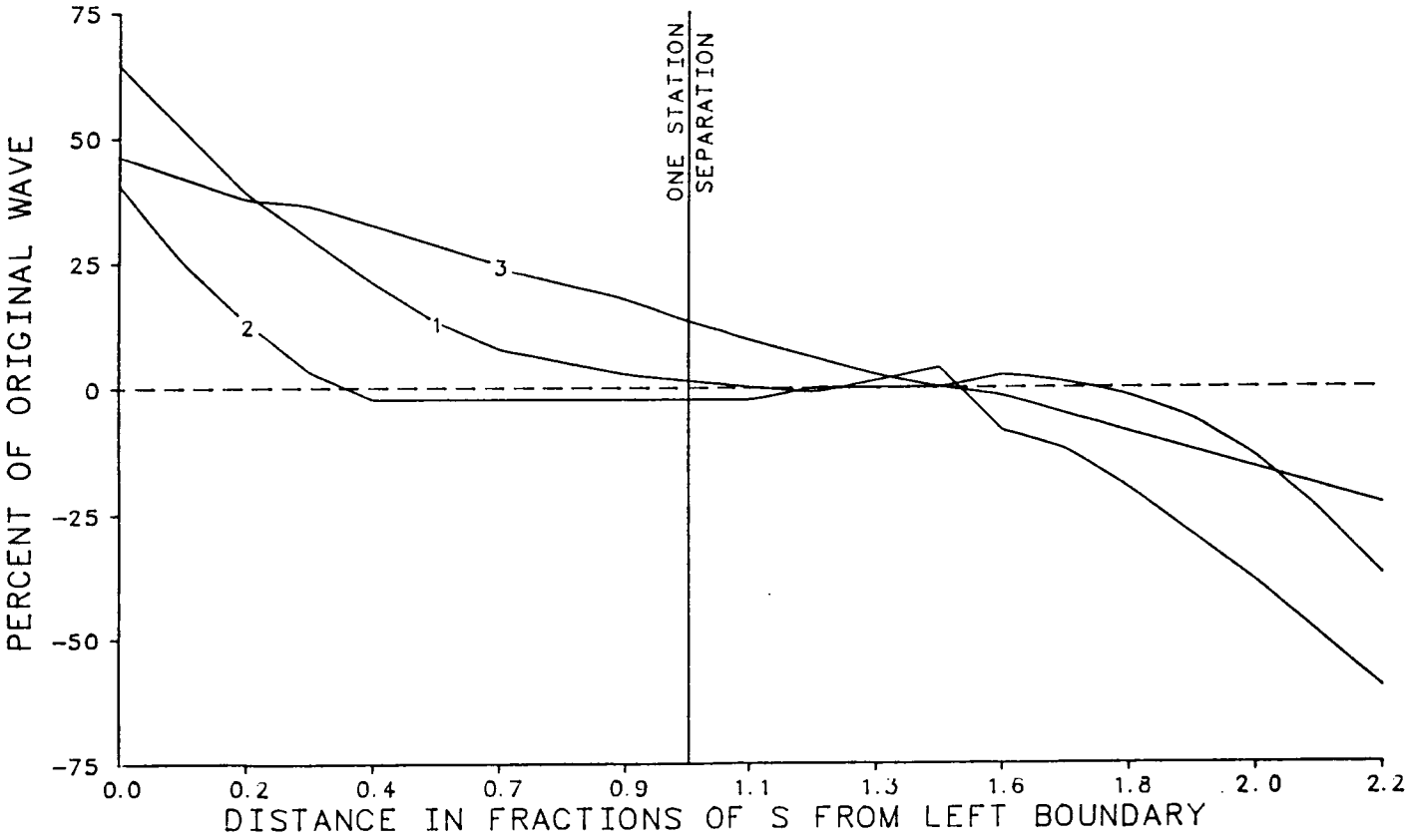

75
50
25
0
-25
-50
-75
PERCENT OF ORIGINAL WAVE
ONE STATION SEPARATION
1
2
3
0.0
0.2
0.4
0.7
0.9
1.1
1.3
1.6
1.8
2.0
2.2
DISTANCE IN FRACTIONS OF S FROM LEFT BOUNDARY

Figure 9

Chapter VII

On the Notion of Varying Influence Radii for a Successive Corrections Objective Analysis

Gary L. Achtemeier

Climate and Meteorology Section

Illinois State Water Survey

Champaign, Il 61820

ON THE NOTION OF VARYING INFLUENCE RADII FOR A SUCCESSIVE CORRECTIONS OBJECTIVE ANALYSIS

Gary L. Achtemeier

Climate and Meteorology Section
Illinois State Water Survey
Champaign, Illinois 61820

ABSTRACT

This study examines the NOTION that the best successive corrections objective analysis is obtained by first analyzing for the long wavelengths and then building in short wavelengths by successively reducing the influence radius for each correction pass. It is shown that the best objective analysis, as measured by filter fidelity (how well the objective analysis restores desired wavelengths and removes undesired wavelengths), is realized for the Barnes method if the effective influence area used for the correction pass is equal to the effective influence radius used for the first pass. The improvements are relatively small, ranging from a few percent for long wavelengths to about ten percent for short but resolvable waves. However, increased simplicity and potentially great reductions in computer time needed to analyze large masses of meteorological data advance these modest gains. Therefore, rather than attempt to build desired detail into an analysis, the analyst should determine the detail permitted by the data quality and distribution and analyze directly for these motion scales.

1. INTRODUCTION

This article focuses on a class of objective analysis techniques known as "methods of successive corrections (SC)" first introduced to operational meteorology by Cressman (1959) and popularized for the analysis of mesoscale weather systems by Barnes (1964, 1973). Though recently replaced by a multivariate statistical technique as the operational interpolation method for some synoptic scale numerical forecasting models, the successive corrections techniques are widely used alone or in combination with other methods (Achtemeier, 1975; Ogura and Chen, 1977; Seaman et al., 1977) for gridding synoptic and regional scale data and, in particular, for gridding large quantities of high frequency data taken from special mesoscale networks (Achtemeier, 1983).

Cressman (1959) introduced the concept of building detail into an analysis by successively decreasing the number of observations that contribute to an estimated value at a gridpoint. Since the gridded data are weighted averages, a type of filtering takes place over the influence area defined as a circle centered on the gridpoint and having an "influence radius," R. The use of a series of scans with decreasing R allows the analysis of a spectrum of scales. From this approach has come the NOTION that the best analysis is obtained by making the method a successively higher pass filter, first analyzing for the large scales and then for the smaller scales. The NOTION has persisted with other SC methods designed to retrieve the maximum allowable detail from meteorological data (Endlich and Mancuso, 1968; Barnes, 1973; Ogura and Portis, 1982; Koch et al., 1983; Smith and Leslie, 1984).

We re-examine the NOTION for the following reasons: First. a theoretical study of the optimum influence radius for the Cressman method (Stephens and Stitt, 1970) found only small variations in the ratio of discrepancy variance to field variance for a rather large range of second pass influence radii. No comparable study has been put forth for the Barnes method, a technique that has been well-documented in the meteorological literature and has been made part of data assimilation and analysis packages (Koch et al., 1983; Smith and Leslie. 1984). Second, a reassessment of underlying concepts can lead to improvements over existing approaches to successive corrections interpolation. and third, in spite of larger and faster computing systems, computational economy is still an important factor in data analysis. Speed and size of computing systems have been offset by larger volumes of data and by more complex numerical models. Thus the prospect for increasing the computational speed of an objective analysis method is a reason for undertaking this study.

2. TEST OF THE NOTION

Barnes (1973) modified his (1964) method so that only one correction pass through the data was required to retrieve the desired detail at small but resolvable wavelengths. The filter was made to return more of the amplitude of the short waves through a reduction of the "effective influence radius", a procedure that is analogous to decreasing the influence radius for the Cressman (1959) analysis technique except that the number of observations within the scan area remains the same. Instead of reducing the number of observations, the weights are recalculated so that the relative importances of the observations closest to the grid point are increased on the correction pass.

The distribution of weights depends upon the theoretical response for the first pass D_0 and a multiplier γ for the correction pass to give a desired final filter response, D'. If the grid point is located so that the impacts of the data boundaries upon the final response are negligible, then for any wavelength λ and weighting parameter κ ,

$$D' = D_0 + (1-D_0)\, D_1 \ , \tag{1}$$

where

$$D_0 = \exp\, [-\kappa_0\, (\pi / \lambda)^2] \ , \tag{2}$$

and $D_1 = D_0$, if $\kappa_1 = \gamma\kappa_0$. Koch et al. (1983) reduced the subjectivity in the selection of the final response by requiring the final amplitude of the minimum resolvable wave λ^* (same as the 2S wave, where S is the average separation between data collection sites) to satisfy the constraint $D'(\lambda^*) = e^{-1}$. Then, upon specification of $D_0(\lambda^*)$ for the first pass, κ_0 and γ are determined uniquely for all λ and D' can be calculated for all wavelengths. We can repeat the above steps for different choices for D and find new D' as functions of the recalculated control parameters.

We can use the above methodology for finding the final response to formulate and test a "null hypothesis" which states that "The best objective analysis, as measured by filter fidelity (how well a filter restores desired wavelengths and removes undesired wavelengths), can be obtained if the effective influence radius used for the correction pass is equal to or larger than the effective influence radius used for the first pass." We can prove the NOTION false if it can be shown that there exists a $\gamma \geq 1.0$ for which the final responses for the resolvable spectrum of waves are equal to or greater

than the final responses obtained with all $\gamma < 1.0$.

a) Test of the null hypothesis for a data continuum

We first make the null hypothesis more specific by finding the value for γ in the range $\gamma \geq 1.0$ that maximizes the response of the Barnes filter for a set of resolvable wavelengths. We determine various combinations of D_0 and γ which satisfy the constraint upon D' and then calculate the final responses for waves in the range from 0.5 λ^* to 6 λ^*. Maximum final response occurs at $\gamma = 1.0$. Figure 1 shows the differences in D' ($D'(\gamma > 1.0) - D'(\gamma = 1.0)$) measured in percent of the original waves. Positive (negative) percentages mean that the use of a particular value for γ restores more (less) of a particular wave than does the method with $\gamma = 1.0$ used for the correction pass. All final responses for $\lambda = \lambda^*$ are identical by constraint. All longer waves receive greater filtering when $\gamma > 1.0$. Maximum reductions of amplitude occur for $\gamma = \infty$ and range from three percent for the long wavelengths to approximately eleven percent for the $2\lambda^*$ wave. A second scan with $\gamma = \infty$ is equivalent to adding the mean of the discrepancies between the first scan estimates and the observations to the first scan. The correction pass response is zero and the 2-pass Barnes method effectively becomes a single scan with $D_0(\lambda^*) = e^{-1}$.

The positive percentages for the waves shorter than λ^* are an indication that the Barnes filter amplifies these undesirable short waves as $\gamma \to \infty$. Figure 1 therefore shows that the use of a larger effective influence area for the correction pass leads to an increase of the amplitudes of the undesirable short waves and a decrease of the amplitudes of the desired longer waves in comparison with the amplitudes restored for the same waves if $\gamma = 1.0$ at the

second pass. These results can be used to state the null hypothesis more specifically: "The best objective analysis as measured by filter fidelity (how well the filter restores desired wavelengths and removes undesired wavelengths) can be obtained if the effective influence radius used for the correction pass is the same as the effective influence radius used for the first pass".

We can prove the null hypothesis is true if it can be shown that the final responses for the waves $\lambda > \lambda^*$ obtained with a smaller effective influence radius on the correction pass are degraded relative to the final response obtained with the effective influence radius unchanged on the correction pass. Figure 2 shows the differences in D' (D'($\gamma < 1.0$) - D'($\gamma = 1.0$)) measured in percent of the original waves for $0.2 < \gamma < 1.0$. The negative percentages for all waves $\lambda > \lambda^*$ mean that the final responses have been degraded by the use of $\gamma < 1.0$. Short but resolvable waves suffer the greatest losses in amplitude. The magnitudes of these losses vary inversely with γ.

Heuristically, the filter must be degraded for any choice of smaller influence area on the correction pass. If γ is to be made small on the correction pass, it is necessary to decrease D_0 on the first pass so that the constraint $D'(\lambda^*) = e^{-1}$ is satisfied. Equation (2) shows that the decreasing of D_0 is accomplished by increasing the shape factor κ_0. In the limit as $\kappa_0 \to \infty$, $D_0 \to 0$ and the filter only restores the mean of the data field on the first pass. Therefore it is necessary to choose κ_1 on the correction pass so that $D'(\lambda^*) = D_1(\lambda^*) = e^{-1}$. This is equivalent to a single pass objective analysis method since all waves are totally filtered on the first pass and any restoration of wave amplitudes is accomplished at the correction pass. The resultant final responses therefore must be identical to the maximum

degraded final responses found when an infinite influence area was used on the correction pass (Fig. 1).

The above reasoning is confirmed by Fig. 3 which shows the differences in D' with respect to $D'(\gamma = 1)$ for the 1.5 λ^* wave as γ ranges from zero to infinity. The differences are expressed in percent of the amplitude of the original wave. They are all negative and are distributed symmetrically about $\gamma = 1.0$. This means that the 1.5 λ^* wave subjected to the Barnes filter with the second pass effective influence area unequal to the first pass effective influence area suffers greater filtering in comparison with the same wave after application of the Barnes filter with the correction pass effective influence area unchanged. This result also applies to all other resolvable waves.

b) Test of the null hypothesis with analytical data

We now verify the theoretical results with analyses of a set of analytical waves. A 20 by 20 grid is sectioned into twenty-five 16-point arrays and four data points are randomly located into each array subject to the requirement that the data points are collocated with grid points. Locating the observations at the grid points eliminates any additional interpolation needed to estimate values of the gridded fields at the data locations and makes possible the direct comparison of the objectively filtered fields with the predictions of the response theory which was derived assuming a continuum of data.

The final filtered value at each grid point after one correction pass through the data is the weighted average of M*N observations plus the correction pass according to

$$g_1(i,j) = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} w_{0,m,n}\, f(m,n)}{\sum_{m=1}^{M}\sum_{n=1}^{N} w_{0,m,n}} + C_1 \,, \tag{3}$$

where

$$C_1 = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} w_{1,m,n}\,[f(m,n) - g_0(m,n)]}{\sum_{m=1}^{M}\sum_{n=1}^{N} w_{1,m,n}} \tag{4}$$

The weight function, $w_{\ell,m,n}$ ($\ell = 0, 1$), is given by

$$w_{\ell,m,n} = \exp\left(-r^2_{m,n}/\kappa_\ell\right) \tag{5}$$

The analytical fields, f(m,n), describe monochromatic sine and cosine waves that range from 2S to 12S in wavelength. The modeling approach for this study requires that the 100 data points are recalculated for each of 24 pairs of objective analyses with the Barnes method. The statistic of relative performance is the difference between the respective RMS errors (the RMS errors are between the analysis and the true field) for the analyses with $\gamma = 1.0$ and $\gamma = 0.2$ normalized by the amplitude of the analytical wave. It is a measure of the accuracy of the method in restoring the whole wave, not just the amplitude. However, this statistic is approximately comparable with the percent of the amplitude of the original wave used for the theoretical part of the comparative studies.

Figure 4 shows the results of this study for the grid interior. The solid line taken from Fig. 2 is the difference between the two methods as predicted by theory and expressed as percent of the amplitude of the original

wave. The results from the analyses compare well with theory for wavelengths between 2S and 4S. Negative values are indicative of smaller RMS errors (a measure of a better analysis) for the analyses with $\gamma = 1.0$. As expected, there is no significant difference between the methods for the 2S wave. The analyses depart from theory for wavelengths greater than 4S, the analytical results consistently show that the accuracy of the fixed influence area method is comparatively better than predicted by theory although the improvement is only several percent.

Achtemeier (1986) showed that the absence of data beyond the boundaries of a data network can have deleterious impacts upon waves defined within the interior of objective analyses in the form of smoothing of long waves and phase shifting of short waves. The distance these effects intrude into the interior of an analysis domain is a function of the influence area. This poses no serious problem for the analysis of large data sets. The analysis area can be set up so that some data fall outside the grid or so that several rows and columns of the grid can be discarded after the analysis. These options are not always possible for the analysis of smaller data sets, however, because there may not be enough observations. Then the choice for γ becomes a determining factor in the accuracy of the Barnes method near the boundaries of the data field.

Figure 5 compares the analysis methods at the grid boundaries. The solid line shows the relative accuracies for the theoretical amplitude responses for cosine waves in percent of the amplitude of the original wave. The theory predicts that the method with = 1.0 will restore about 5 percent more of the 2S wave and thus can be expected to give noisier gridded fields near data boundaries. The advantage of the = 1.0 method is that from five to eight

percent more of the other wavelengths will be restored by the $\gamma = 1.0$ method. The analyses with the irregularly spaced data (pluses) tend to verify the theoretical results for the amplitude response.

The dashed line in Fig. 5 compares the two methods for the phase shift responses for the sine waves. Because the phase shift responses are negative. the interpretation of the curve and the analytical results differs from the interpretation of the curves on other figures. The change of sign between 3S and 4S means that the theory predicts the method with $\gamma = 1.0$ to cause less phase-shifting of the original long waves and more phase-shifting of the original short waves.

The analysis results (boxes) for the phase shift response are in agreement with theory for the short wavelengths and the crossover point near 4S. The Barnes method with $\gamma = 1.0$ fares more poorly than expected in comparison with $\gamma = 0.2$ for the wavelengths between 2S and 4S and compares more favorably for the longer wavelengths. We found from an examination of the magnitudes of the RMS errors that analyses with both values for γ were quite acceptable for the longer wavelengths. However, $\gamma = 1.0$ produced better analyses with very small RMS errors at the data boundaries.

3. DISCUSSION

The method of successive corrections has found widespread use for the gridding of meteorological data collected at irregularly spaced locations. There has been for many years a NOTION among users of these methods that the best objective analysis is obtained by first analyzing for the long waves and

then building in the short waves through decreasing the radius of influence on correction passes through the data. Using a 2-pass successive corrections method developed by Barnes (1973), we have examined this notion with theoretical and analytical studies and have tested its validity in the grid interior and at the boundaries of the data field. The Barnes method was chosen for this study because it has found widespread use in the gridding of special network weather data and because it has been made part of comprehensive meteorological data acquisition and processing systems (GEMPAK by Koch et al., (1983); PROAM by Smith and Leslie (1984)).

The major finding from this study is that the NOTION is incorrect -- at least for the Barnes method. Using a classical hypothesis/null hypothesis approach, the analyses indicate acceptance of a null hypothesis which states that "the best objective analysis as measured by filter fidelity (how well the filter restores desired wavelengths and removes undesired wavelengths) is obtained when the influence radius for the second pass is the same as the influence radius for the first pass". The null hypothesis is valid at least for the range of waves ($2S \leq \lambda \leq 12S$) and over the entire analysis domain except for a few short wavelengths near the boundary of the data field which suffer from greater phase shifting.

Theoretical responses indicate that the use of a fixed effective influence radius will cause the Barnes filter to restore from about 6 to 7 percent more of the short waves in the range from 3S to 5S than will the same filter with the effective influence radius determined with the GEMPAK default criteria. These wavelengths are typical of the scales of meteorological phenomena sought through the use of limited-area special networks. There were no important differences in the final responses for the longer wavelengths. The

improvements from corresponding analytical studies with irregularly spaced (and also regularly spaced) data sets ranged from five to eight percent for the important short waves to about 2.5 percent for the longer waves. These improvements are relatively small but could be significant for studies designed to retain the maximum details permitted by the data without amplifying undesirable short wavelengths. Conversely, an analyst can choose the influence area parameters to reduce slightly more of the short 2S wavelength without loss of amplitude of important longer wavelengths.

These results suggest that, rather than attempt to build desired detail into analyses, the analyst should determine the detail permitted by the data quality and distribution and analyze directly for these motion scales.

Perhaps the major advantage of an objective data gridding technique that requires only one effective influence radius per analysis is that there can be an immense savings in computation time if large quantities of data are to be processed. Koch et al., (1983) noted that the most time consuming part of the objective analysis is the computation of exponentials. They recommended that the same calculated weights be applied to many different parameter fields or to the same parameter field over many different times as long as the data quality and distribution do not vary appreciably. The number of exponentials computed for some versions of the Barnes method is approximately $N_1=K(L)(M)(KX)(KY)$, where K is the number of data sets to be analyzed, L is the number of passes through the data for which the exponentials must be recalculated, M is the number of observations, and KX and KY are, respectively, the number of grid points in the X and Y directions.

For a more efficient objective analysis, it is best to compute the exponential array once and to use it for all subsequent analyses if the data

quality and distribution permit. Therefore, K=1. Further, L=2 for the 2-pass objective analysis. However, it is immediately apparent that the size of the exponential array can be cut in half if $\gamma = 1.0$ since the exponentials calculated for the first pass can be used for the second pass. Therefore, L=1 for the efficient objective analysis.

Achtemeier (1986) found for a point located at the center of an idealized circular limited-area data set that detrimental impacts of data boundaries are not significant unless the distance from the central point to the data boundary is less than about 1.6S times the average spacing between the observations. These results also apply to the filter response for a truncated circular influence area. The value of 1.6 was used by Barnes (1964) and was found by Stephens and Stitt (1970) to be the optimum influence radius for the Cressman (1959) successive corrections method. An influence area of radius 1.6S contains approximately 9 regularly spaced observations. Our studies with irregularly spaced data sets indicate that from 10 to 12 observations should be included within the influence area.

The total number of exponentials calculated for an efficient objective analysis is $N_2=10(KX)(KY)$. The ratio of the exponentials needed for the two methods is $r=N_1/N_2=0.1K(L)(M)$ if M>10. If, for example, there were to be objective griddings of 10 data sets consisting of 50 observations each, there would be a 100-fold reduction in the number of exponentials needed if an efficient form of the Barnes objective analysis method were used in place of some existing versions.

Acknowledgements

This research was supported by the National Aeronautics and Space Administration (NASA) under Contract NAS8-34902. Ms. Rebecca Runge is acknowledged for typing the manuscript.

REFERENCES

Achtemeier, G. L., 1975: On the initialization problem: A variation adjustment method. Mon. Wea. Rev., 103, 1089-1103.

Achtemeier, G. L., 1983: The relationship between the surface wind field and convective precipitation over the St. Louis area. J. Climate and Appl. Meteor., 22, 982-999.

Achtemeier, G. L., 1986: On the impact of data boundaries upon the objective analysis of limited-area data sets. Mon. Wea. Rev., 114, 40-49.

Baker, W. E., 1983: Objective analysis and assimilation of observational data from FGGE. Mon. Wea. Rev., 111, 328-342.

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. J. Appl. Meteor., 3, 396-409.

Barnes, S. L., 1973: Mesoscale objective analysis using weighted time-series observations. NOAA Tech. Memo. ERL NSSL-62, National Severe Storms Laboratory, Norman, OK 73069, 60 pp. (NTIS COM-73-10781).

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

Endlich, R. M., and R. L. Mancuso, 1968: Objective analysis of environmental conditions associated with severe thunderstorms and tornadoes. Mon. Wea. Rev., 96, 342-350.

Koch, S. E., M. DesJardins, and P. J. Kocin, 1983: An interactive Barnes objective map analysis scheme for use with satellite and conventional data. J. Climate and Appl. Meteor., 22, 1487-1503.

Ogura, Y., and Y. L. Chen, 1977: A life history of an intense mesoscale convective storm in Oklahoma. J. Atmos. Sci., 33, 1458-1476.

Ogura, Y., and D. Portis, 1982: Structure of the cold front observed in SESAME-AVE III and its comparison with the Hoskins-Bretherton frontogenesis model. J. Atmos. Sci., 39, 2773-2792.

Seaman, R. S., R. K. Falconer, and J. Brown, 1977: Application of a variational blending technique to numerical analysis in the Australian Region. Aust. Meteor. Mag., 25, 3-23.

Smith, D. R., and F. W. Leslie, 1984: Error determination of a successive correction type objective analysis scheme. J. Atmos. and Oceanic Tech. 1, 120-130.

Stephens, J. J., and J. M. Stitt, 1970: Optimum influence radii for interpolation with the method of successive corrections. Mon. Wea. Rev., 98, 680-687.

Figure Captions

Fig. 1. Departures between the fixed influence area final response and selected variable influence area final responses for which the influence areas are increased on the correction pass. Differences expressed in percent of the amplitude of the original wave.

Fig. 2. Departures between the fixed influence area final response and selected variable influence area final responses for which the influence areas are decreased on the correction pass. Differences expressed in percent of the amplitude of the original wave.

Fig. 3. Departures between the fixed influence area final response and the full range of variable influence area final responses for the 1.5 λ^* wave. Differences expressed in percent of the amplitude of the original wave (POW).

Fig. 4. Departures between analyses with the fixed influence area method and the variable influence area method compared with predictions by response theory. Analyses carried out with irregularly spaced data. Differences expressed in percent of the amplitude of the original wave (POW).

Fig. 5. Departures between analyses with fixed influence area method and the variable influence area method compared with predictions by response theory. Solid lines and pluses are for amplitude response for cosine waves. Dashed lines and boxes are for phase shift response for sine waves. Analyses carried out with irregularly spaced data and differences expressed in percent of the amplitude of the original wave (POW).

Figure 1

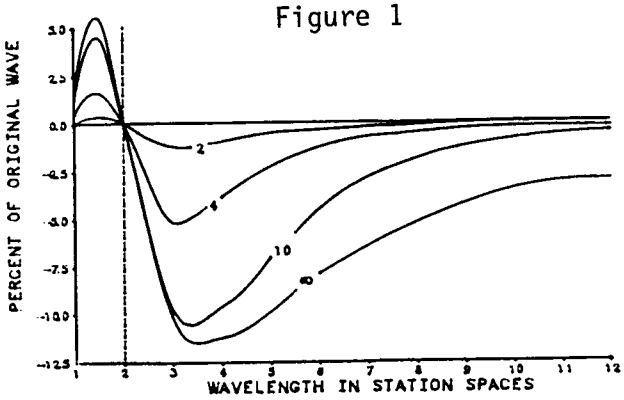


Figure 2

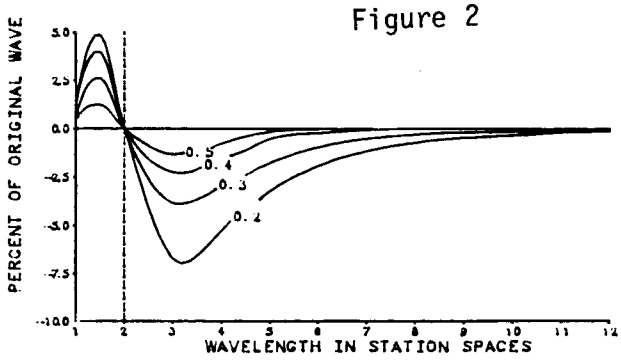


Figure 3

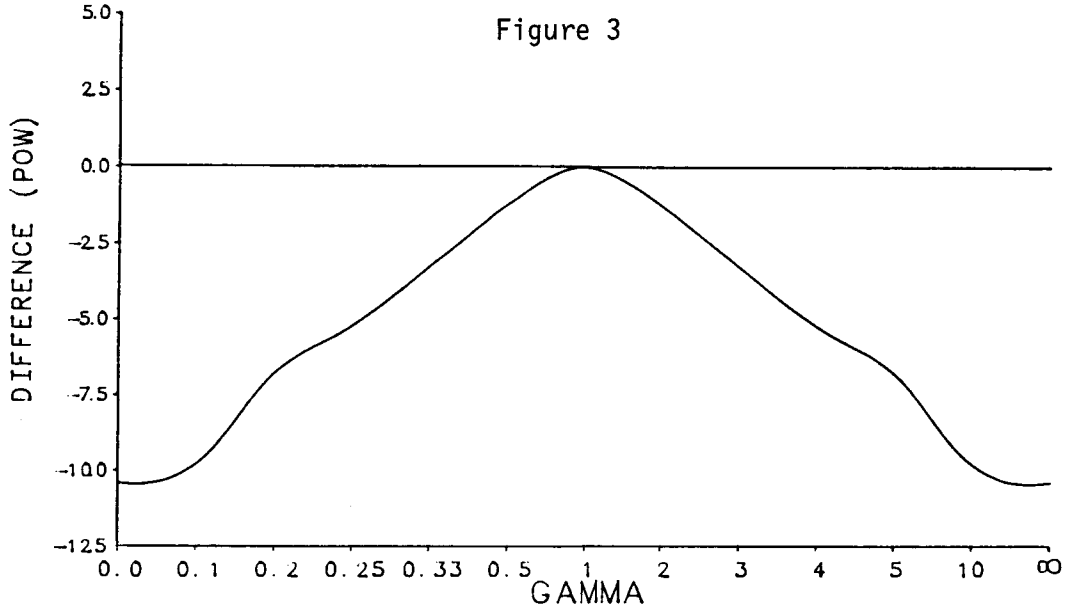


Figure 4

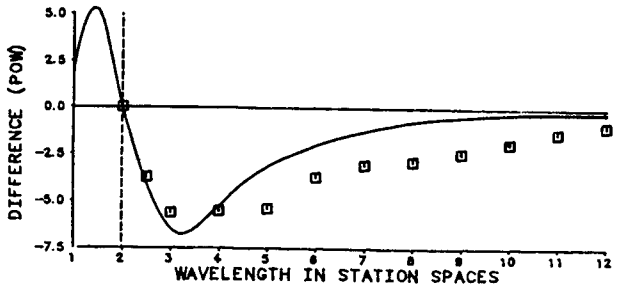
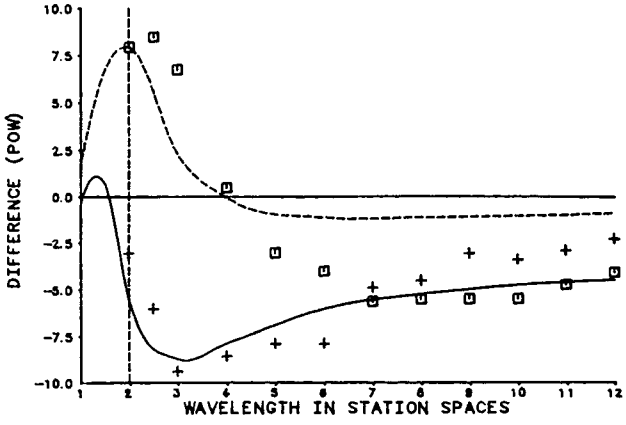


Figure 5



1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
NASA CR-3981		

4. TITLE AND SUBTITLE	5. REPORT DATE
A Variational Assimilation Method for Satellite and Conventional Data: Development of Basic Model for Diagnosis of Cyclone Systems	May 1986
	6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT
G. L. Achtemeier, H. T. Ochs III, S. Q. Kidder, R. W. Scott, J. Chen, D. Isard, and B. Chance	

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. WORK UNIT NO.
Climate and Meteorology Section University of Illinois Illinois State Water Survey 2204 Griffith Drive, Champaign, IL 61820	M-526
	11. CONTRACT OR GRANT NO. NAS8-34902

12. SPONSORING AGENCY NAME AND ADDRESS	13. TYPE OF REPORT & PERIOD COVERED
National Aeronautics and Space Administration Washington, D.C. 20546	Contractor Report
	14. SPONSORING AGENCY CODE ED42/MSFC

15. SUPPLEMENTARY NOTES

Contract Monitor: John W. Kaufman, Atmospheric Science Division, Systems Dynamics Laboratory, Marshall Space Flight Alabama 35812
Final Report

16. ABSTRACT

A three-dimensional diagnostic model for the assimilation of satellite and conventional meteorological data is developed with the variational method of undetermined multipliers. Gridded fields of data from different type, quality, location, and measurement source are weighted according to measurement accuracy and merged using a least squares criteria so that the two nonlinear horizontal momentum equations, the hydrostatic equation, and an integrated continutity equation are satisfied. The model is used to compare multivariate variational objective analyses with and without satellite data with initial analyses and the observations through criteria that were determined by the dynamical constraints, the observations, and pattern recognition. It is also shown that the diagnoses of local tendencies of the horizontal velocity components are in good comparison with the observed patterns and tendencies calculated with unadjusted data.

In addition, it is found that the day-night difference in TOVS biases are statistically different (95% confidence) at most levels. We also develop a hybrid nonlinear sigma vertical coordinate that eliminates hydrostatic truncation error in the middle and upper troposphere and reduces truncation error in the lower troposphere. Finally, it is found that the technique used to grid the initial data causes boundary effects to intrude into the interior of the analysis a distance equal to the average separation between observations.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
Global Circulation Planetary Scale Weather Tropospheric Dynamics Objective Analysis Satellite Weather Data	Unclassified - Unlimited Subject Categories 47, 64

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
Unclassified	Unclassified	225	A10

For sale by National Technical Information Service, Springfield, Virginia 22161
NASA-Langley, 1986