A UNIFIED ANALYSIS-INITIALIZATION TECHNIQUE

David L. Williamson and Roger Daley
National Centre for Atmospheric Research*
Boulder, Colorado 80307

ABSTRACT

A unified analysis-initialization technique is introduced and tested in the framework of the shallow water equations. It consists of iterating multivariate optimal interpolation and nonlinear normal mode initialization. For extropical regions it is shown that such a technique produces an analysis consistent with observational errors and in nonlinear balance. The linear errors of optimal interpolation associated with geostrophically related covariances are eliminated.

* The National Centre for Atmospheric Research is sponsored by the National Science Foundation.

1. Introduction

There is mounting evidence that the specification of the initial atmospheric state is as much to blame for errors in numerical weather forecasts as the mathematical formulation of the model or the parameterization of physical processes. The complete forecast system is comprised of four components--observations, analysis, initialization, and numerical model--with three corresponding interfaces. Analysis schemes produce an estimate of atmospheric state variables distributed uniformly in space and time from the irregular and incomplete set of observations. Initialization schemes modify the analyses to make them acceptable as initial states for models so that the model forecasts do not contain spurious waves which obscure or even interfere with the development of meteorologically important features.

To a large extent, these four components have been developed independently by specialists in each area with limited but not complete understanding of the principles involved in all the others. However, there are notable instances of collaboration among specialists in problems related to the interface between observations and analysis, and the interface between initialization and models. For example, statistical objective analysis or optimal interpolation takes into account the error characteristics and distribution of various observing systems and first guess to produce an estimate of the analysis error. Dynamic initialization and nonlinear normal mode initialization both use the forecast model itself as a basis of determining the necessary changes to the analysis to make it acceptable to that model. The interface between analysis and initialization has received less attention. Although four-dimensional data assimilation is an attempt to bridge this gap, a unified approach is still lacking.

Because of a mismatch between the analysis and numerical forecast models, forecasts produced by the models from the original analyses often exhibit high frequency noise. Nonlinear normal mode initialization successfully eliminates these spurious high frequency oscillations from the forecasts (Daley, 1979; Williamson and Temperton, 1981) but, in doing so, it often produces modifications to the analyses which exceed the expected analysis error even over data-rich regions such as the continental United States. As we will show below with a more concrete example, these changes are in part due to shortcomings of both aspects (analysis and initialization) of the system. The analysis has errors due to, for example, adopting a linear geostrophic assumption while the model (and atmosphere) actually satisfies a more complicated nonlinear gradient relationship. The unconstrained nonlinear normal model initialization does not recognize the specific nature of this analysis error and thus does not eliminate it.

Williamson et al. (1981) showed that multivariate optimal interpolation introduces an analysis error by the use of the geostrophic relationship to obtain height-wind covariances from height-height covariances. Such covariances produce analysis increments that satisfy the linear geostrophic relationship, but the atmosphere is actually closer to nonlinear gradient wind balance. Nonlinear normal mode initialization modifies the analysis to satisfy gradient balance, and therefore in regions of large curvature where the difference between geostrophic and gradient is greatest it can make substantial changes. This effect was observed in their tests with good quality, complete height and wind observations over a reasonably dense network. In this case the initialization still made 30 m changes in the height field in trough regions. The problem is more serious given just height data or just wind data, and, as we will show later, the better the observations are relative to the first guess, the more serious the problem.

For example, with height observations only, the analysis procedure draws for the heights and produces corresponding geostrophic wind increments. In regions of large curvature, this geostrophic increment will be in error since the atmosphere is more in gradient balance. However, except for the largest scales, the nonlinear normal mode initialization modifies the originally correct height analysis to be in nonlinear balance with the erroneous geostrophic wind (Daley, 1978, 1980), producing a relatively large and, in this case, incorrect change in the height analysis.

In the complementary case of wind observations only, the analysis draws to the winds and produces a corresponding geostrophic height increment. For the largest scales, the initialization modifies the correct wind field to match the erroneous height field. For smaller scales, the initialization modifies the height field to be in nonlinear balance with the correct wind field. In this case, the change made to the smaller scales by the initialization is appropriate and desirable.

In both cases, however, the need for such changes is not indicated by the expected analysis error. This inconsistency arises because the expected analysis error is determined directly from the specified error covariances assuming they are correct. Thus, these expected errors assume the geostrophic relation is correct and do not recognize the error of this relation.

The variational approach coupled to nonlinear normal mode initialization was introduced by Daley (1978) and extended by Tribbia (1982) as a means of obtaining a balanced initial state faithful to the analysis, faithful meaning that the changes made to the analysis by the procedure are not inconsistent with the analysis errors. The definition of weights for the fidelity measure in this procedure which reflect the degree of confidence in the analysis at each point is not as straightforward as simply using the expected analysis error since, as pointed out above, this expected error does not reflect the error in the analysis incurred by not recognizing the nonlinear nature of the relationship between heights and winds.

This variational approach is also computationally a very large problem especially if the weighting functions vary with all dimensions. Daley's (1978) weight functions varied arbitrarily in the north-south direction, but because of the computational constraints, the longitudinal variation was treated in a perturbative manner. Tribbia (1982) extended the approach to allow complete variation of the weight functions in the horizontal directions and demonstrated the effectiveness of this approach with a spectral shallow-water model of about rhomboidal 20 resolution. With current computers, it would be difficult to deal with much finer resolution or to include nonseparable vertical variations. Therefore, even if the problem of the specification of the weight functions can be solved, this approach lends itself more to the large-scale global motions than to the smaller, more local aspects discussed above.

As an alternative to variational methods, Phillips (1982) suggested an approach in which only the linear slow mode fields are analyzed via optimal interpolation. The covariance structure functions are based on relationships derived from the structure of the linear slow modes only. Thus, the nonlinear gradient error should not affect the anlysis. The nonlinear component of the field is obtained from the slow modes via Baer-Tribbia (1977) nonlinear modal initialization. His approach has a few restrictive assumptions that make it impossible to apply in practice as formally proposed. For example, all observations must be used in the analysis of each gridpoint variable. Because of the large number of atmospheric observations, this is impractical with optimal interpolation. Simplifying assumptions must be introduced into Phillips' approach to make it operationally feasible. His approach has not been tested in any practical situation.

Both the variational technique and the proposal of Phillips (1982) are essentially global procedures. Because of their global nature they require the inversion of gigantic, full matrices, an almost impossible task on present-day computers. Our following proposed iterative approach essentially handles the global problem in two steps. The part requiring matrix inversion is done locally with much smaller matrices. The other part is done globally, but it only deals with matrix multiplications, rather than inverses, which can be done on present-day computers. Therefore, our proposed iterative approach is feasible today. This will be explained more fully after the iterative approach is described.

More details on this subject can be seen in Williamson and Daley (1983).

2. Proposed iterative approach

The variational approach of Daley (1978) and Tribbia (1982) minimizes the difference between the balanced initial state and an analysis. A more desirable approach might be to minimize the difference between the balanced initial state and the observations themselves rather than an analysis of these observations. In principal, one can do this using the normal mode approach of Flattery's (1970) analysis scheme, but incorporating the ideas of nonlinear normal mode balance. Difficulties arise in such an approach, however, because the very convenient and desirable orthogonality properties of Hough functions are lost when dealing with an arbitrary observational network. In addition, the formalisms of optimal interpolation which determine the relative weights for observations of differing quality is also lost.

Optimal interpolation can be looked at as a minimizational fit to the observations. In fact, Kimeldorf and Wahba (1970) show that for every covariance structure function in optimal interpolation, there is a variational problem that will give the same solution and vice versa. Therefore, in Tribbia's (1982) variational iterative approach, it seems reasonable to replace his variational minimization step with multivariate optimal interpolation. In other words, one could perform multivariate optimal interpolation followed by nonlinear normal mode initialization followed by multivariate optimal interpolation (using the first initialized analysis for the first guess or trial field), followed by initialization, etc., until, hopefully, the initialization and the analysis both make negligible changes indicating convergence. Of course, sequential optimal interpolation would offer no improvement over the first analysis if nothing was done in between, since no new information is added and the first optimal interpolation would minimize the expected analysis error within that formalism. However, the slow component of the analysis generally has a smaller error than the first trial field and therefore the subsequent geostrophic analysis increments will be smaller in the next optimal interpolation, resulting in smaller error from the geostrophic approximation. In such an iterative approach on successive iterations one could also introduce local variations in the structure functions which are derived from the previous iterate. We will not pursue that possibility further here, but rather concentrate on the geostrophic error.

As was mentioned above, in our proposed iterative techniques the essentially local optimal interpolation (as applied in practice) is separated from the global initialization aspect. The matrix inversion occurs in the optimal interpolation step. Because of the local nature of this operation in practice the matrices involved are reasonably small and are readily invertible. On the other hand, the global initialization requires the manipulation of very large matrices but these manipulations are basically matrix multiplications which can be handled readily on present-day computers.

Following Tribbia's (1982) examples (especially his Fig. 5) the iterative procedure introduced above can be interpreted geometrically using the slow manifold diagram introduced by Leith (1980) and further elaborated by Daley (1980). Figure 1 illustrates the procedure for an idealized case. The abscissa R is the Rossby manifold, the ordinate G the gravity manifold, the curve S is the slow manifold, and the lines D are

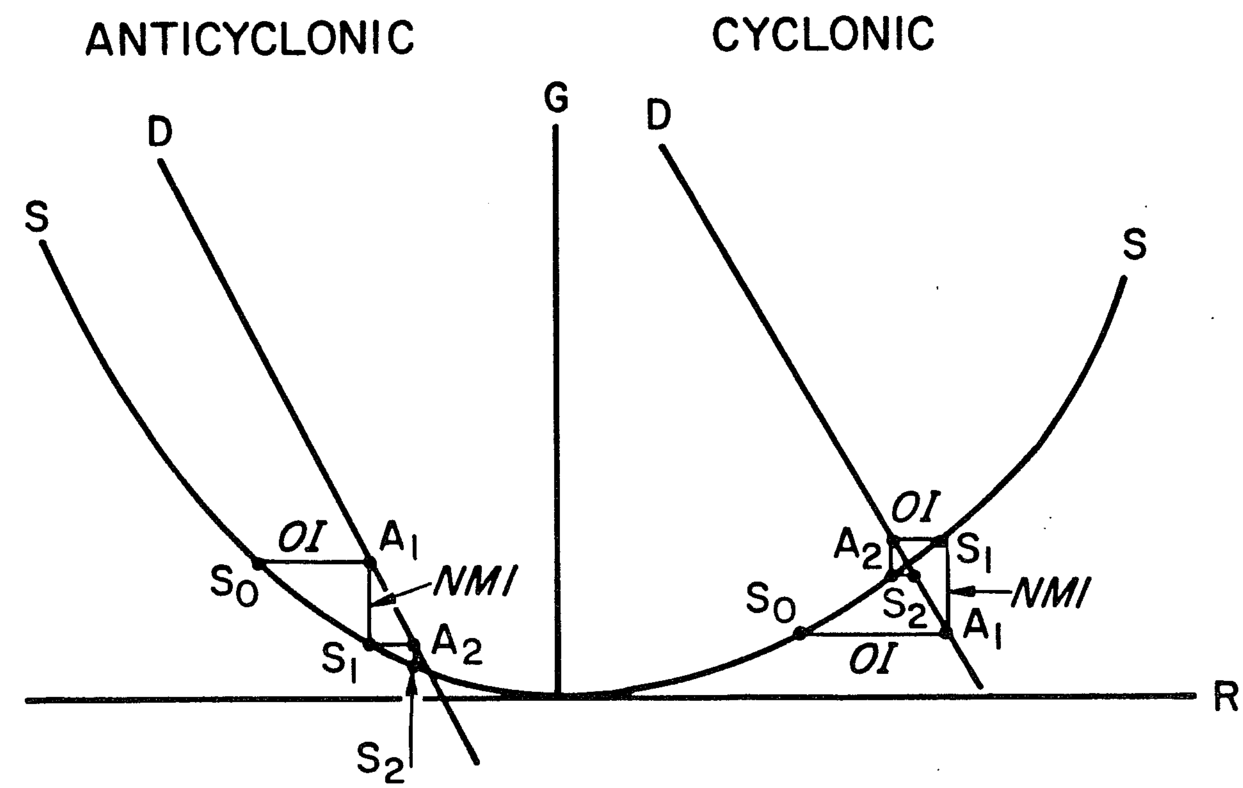

Figure 1 Schematic manifold diagram illustrating convergence of procedure.

data manifolds for a particular type of data. For the purpose of illustration, we take D to consist of height observations with no errors and further assume that the optimal interpolation scheme draws to the data exactly so the height analysis has no error. This, of course, is improbable in reality and we will illustrate how errors affect the procedure later.

Optimal interpolation requires a first guess or total field which is modified according to the observations. In our iterative approach there is a sequence of trial fields. The first trial field for the iterative procedure would normally be provided by a model forecast as is generally done in operational practice today and thus lies on (or can be made to lie on) the slow manifold and is denoted S_o in Fig. 1. The optimal interpolation with geostrophically related covariances modifies the Rossby modes without significantly altering the gravity modes of the first total field to produce the first analysis A_1 in Fig. 1, which under the above assumptions lies on the height data manifold D. The analysis procedure itself is then represented by a horizontal line denoted in the figure. Nonlinear normal mode initialization modifies the gravity waves of the analysis A_1 to produce its slow component S_1 lying on the slow manifold. The inititialization procedure is represented by a vertical line in Fig. 1 and denoted by NMI. A second analysis is then performed based on S_1 for the (second) total field to produce the second analysis A_2 after which its slow component S_2 is calculated. The procedure is continued with the slow componenet of the nth analysis serving as the total field for the (n+1)th analysis. Convergence is indicated by insignificant changes being made by both the analysis and the initialization. Figure 1 indicates that for cyclonic flows the error of the slow component height field changes sign with each analysis since it lies on the opposite side of the data manifold than the previous one, but for anticyclonic flows the error of the slow component height field always has the same sign as the error of the first trial field while its magnitude decreases.

The above example illustrates the procedure for an idealized case which will never exist in reality. Observations are not error-free and they are not collocated with analysis points so the analysis will have additional errors other than the geostrophic one which might affect the convergence properties of our proposed iterative approach. We have tested this approach in a more practical situation with a few simplifications to isolate the properties of this iterative approach from other problems associated with nonlinear normal mode initialization itself. We describe our experimental design in the next section

3. Experimental approach

a. Reference atmosphere

Our experimental approach follows that of Williamson *et al*. (1981) and we only review it here. We define a "reference atmosphere" or control to be an observed atmospheric state which has been balanced with respect to a particular prediction model. The model used is the spectral nonlinear shallow-water model of Bourke (1972), hemispheric with a rhomboidal 31 truncation and 2 km equivalent depth. The reference state is obtained from the 500 mb wind and geopotential NMC Northern Hemisphere Flattery-Hough analysis of 0000 GMT 29 January 1977 initialized for the model using Machenhauer's (1977) unconstrained nonlinear normal mode initialization technique with five nonlinear iterations. This reference state defines the atmosphere we want to analyze. It is defined as spherical harmonic coefficients so that "error-free observations" may be generated for any desired observational network configuration and errors with specified characteristics may be added. These observations serve as input to the iterative analysis-initialization scheme.

b. Slow and fast components

The slow component of the analysis consists of the Rossby modes plus the balanced portion of the gravity modes (i.e. the slowly varying component). The balanced portion of the gravity modes is obtained with one iteration of Machenhauer's unconstrained nonlinear normal mode initialization technique using the shallow-water spectral model. Thus, we have a perfect forecast model and the initialization technique is capable of reproducing the balance of our reference atmosphere. The convergence properties of the iterative analysis initialization procedure are examined without the additional errors that would be present in actual practice from an imperfect forecast model.

The analysis error for any iterate is obtained exactly by comparing it to the reference atmosphere. Similarly the slow component error is calculated as the slow component of the analysis minus the reference atmosphere. The fast component of the analysis is the difference between the analysis and its slow component.

c. Optimal interpolation

The optimal interpolation procedure used for the following tests was developed and described by Schlatter (1975) and Schlatter et al. (1976). We summarize its salient features here. Poleward of 20° latitude, the analysis is trivariate in geopotential height (h) and the horizontal wind components (u, v). The height-wind covariances are obtained from the height-height covariances via the geostrophic relationship. Each trial field error standard deviations are calculated by comparing the actual trial field to the reference atmosphere. The zonal averages of these calculated values are input to the analysis program. The observational error standard deviations are specified in the program corresponding to the actual observational errors. The observational network and errors will be described shortly. The scheme uses all data from the ten closest observation locations with a maximum search radius of 1500 km beyond which observations are not accepted.

Equatorward of 20° the analysis is bivariate in u and v and univariate in h. We do not examine the analysis in this region as we are concerned with eliminating the particular error arising from the geostrophically related covariances in mid latitudes. The analysis is performed on a uniform hemispheric grid with 96 points in longitude and 35 in latitude, giving grid intervals of 3.75° and 2.56°, respectively.

The observational points are chosen at random on the hemisphere with a mean specified spacing D as follows. The hemisphere is first covered with uniform boxes with sides as close to D in length as possible in the manner of the Kurihara (1965) grid. One observation point is then located randomly in each box from a rectangular distribution. Examples of such networks are shown in Williamson et al. (1981) for D = 250 km, 500 km, and 1000 km. Both height and wind are computed from the reference atmosphere. Random errors are added to these observations with a normal distribution and a specified standard deviation.

The first trial field needed to start the iterative analysis-initialization procedure is obtained from a 12 hr forecast from the initialized Flattery analysis 12 hr previous to that used to define the reference atmosphere. This forecast, which is balanced initially, remains in balance. The hemispheric rms differences between this trial field and the reference state are 24.44 m in the geopotential and 6.10 m sec^{-1} in the wind field.

As mentioned before, after each iteration the analysis error can be calculated by comparing it to the reference atmosphere and can be divided into its slow and fast components. These errors provide measures of the convergence of the technique. The goal is to obtain an analysis with a negligible fast component and with minimum error in its slow component. We have tested the iterative approach on several examples from purely hypothetical extremes of error-free observations of one type (heights or winds only) to more realistic configurations corresponding to dense radiosonde observations.

4. Conclusions

A unified analysis-initialization technique consisting of iterating multivariable optimal interpolation and nonlinear normal mode initialization is introduced. Arguments based on manifold diagrams are presented to illustrate in simple cases the nature of the convergence properties of the technique. The scheme is tested for more practical situations in a shallow water model framework. Several examples are considered from purely hypothetical extremes of error-free observations of one variable (heights or winds) to more realistic configurations corresponding to dense radiosonde observations.

With either error-free height observations or error-free wind observations the procedure converges to an analysis with very small error and minimal fast component. The linear geostrophic error is eliminated. In the case of height observations only, for cyclonic flow the error in the slow component of the height field changes sign with each iteration as it approaches zero while for anticyclonic flow the error retains the same sign as it decreases. In the case of wind observations only, the error in the slow component of the height field is primarily of large scale and decreases with successive iterates. These slow component errors are the relevant measures of the quality of the analysis as initial conditions for a forecast because they are the components which affect the time evolution of the forecast.

Tests with realistic observational error standard deviations show that overall the iterative procedure converges more rapidly in these cases than in the preceding error-free cases. The faster convergence occurs because with larger observational errors the analysis procedure gives less weight to the observations and more to the trial field resulting in smaller geostrophic analysis increments. The convergent analysis is an optimal combination of the trial field and observations based on the observational error and the first trial field error. The geostrophic error is eliminated by the iterative procedure.

The iterative optimal interpolation--nonlinear normal mode initialization--provides an analysis consistent with the observational errors and in nonlinear balance. The errors associated with geostrophically related covariances are eliminated. This occurs because the slow component of each successive analysis is a better approximation to the atmospheric state than the first guess used for that analysis. Thus, when the slow component is used as the first guess for the next analysis, the analyzed geostrophic increments are smaller and thus the geostrophic component of the analysis is smaller until eventually it is negligible.

The iterative approach need not be prohibitively expensive as successive analyses need only be performed in regions which are not in balance, rather than over the whole domain. The fast component of the analysis provides a local and global measure of convergence. In regions where it is small, the procedure has converged and the analysis need not be repeated. It is only in regions with a relatively large fast component that further analyses will be beneficial. Given height data only, these regions are limited to areas of large curvature which generally cover only a small fraction of the globe.

Our tests were all performed in the context of a reference atmosphere equivalent to the shallow water equations for which we have a perfect forecast model and a nonlinear normal mode initialization procedure which provide exactly the slow manifold component. We showed that the procedure produces an analysis satisfying horizontal nonlinear gradient flow. However, in reality, the flow in the baroclinic atmosphere is more complicated than simple horizontal gradient flow, our forecast models are not perfect and nonlinear normal mode initialization does not always converge when diabatic processes are present. The extent to which the iterative analysis-initialization procedure will produce an analysis with the general nonlinear balances of the atmosphere will depend on what extent we can get on the slow manifold of the atmosphere. In addition, the linear constraints in the analysis procedure should be at least partially descriptive of the atmosphere so that the multivariate aspect of the procedure will provide information over a univariate analysis. The multivariate aspect is a key in providing an analysis whose slow component is better than the trial field on successive iterations. As each component (analysis, model and initialization) is improved, the iterative analysis-initialization approach will become more successful in the operational environment.

REFERENCES

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

Bourke, W., 1972: An efficient one-level primitive equation spectral model. Mon. Wea. Rev., 100, 683-689.

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

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

-----, 1980: On the optimal specification of the initial state for deterministic forecasting. Mon. Wea. Rev., 108, 1719-1735.

Flattery, T. W., 1970: Spectral models for global analysis and forecasting. Proc. Sixth AWS Tech. Exchange Conf., U.S. Naval Academy, Air Weather Service Tech. Rep. 242, 42-53.

Kimeldorf, G., and Wahba, G., 1970: A correspondence between Bayesian estimation of stochastic processes and smoothing by splines. Ann. Math. Statist., 41, 495-502.

Kurihara, Y., 1965: Numerical integration of the primitive equations on a spherical grid. Mon. Wea. Rev., 93, 399-415.

Leith, C., 1980: Nonlinear normal mode initialization and quasi-geostrophic theory. J. Atmos. Sci., 37, 958-968.

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

Phillips, N. A., 1982: On the completeness of multivariate optimum interpolation for large-scale meteorological analysis. Mon. Wea. Rev., in press.

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

-----, G. W. Branstator, and L. G. Thiel, 1976: Testing a global multivariate statistical objective analysis scheme with observed data. Mon. Wea. Rev., 104, 765-783.

Tribbia, J. J., 1982: On variational normal mode initialization. Mon. Wea. Rev., 110, 455-470.

Williamson, D. L., and C. Temperton, 1981: Normal mode initialization for a multilevel grid-point model. Part II: Nonlinear aspects. Mon. Wea. Rev., 109, 744-757.

-----, R. Daley, and T. W. Schlatter, 1981: The balance between mass and wind fields resulting from multivariate optimal interpolation. Mon. Wea. Rev., 109, 2357-2376.

Williamson, D., R.Daley, 1983: A unified analysis-initialization technique. Submitted to Monthly Weather Review.