The assimilation of past data in dynamical analysis. II

O. TALAGRAND[1] and K. MIYAKODA, *Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, Princeton, New Jersey 08540*

(Manuscript received February 15, 1971; revised version April 2, 1971)

ABSTRACT

Systems for continuous data assimilation are presented and discussed. The balanced barotropic equation is used as an example of a dynamical (evolution) equation. In the first method, geopotential height data are assimilated into the model's computation, as the data become available, without any modification. It turns out that this leads to intolerable growth of error, suggesting the necessity of filtering the input information or alleviating discrepancies between the newly injected data and the prediction fields. In the second method, an optimal assimilation is attempted by applying adequate weights to the input data, which minimize statistically the mean-square difference between the estimate and the true solution. The analysis result is acceptable. But the magnitude of the error reduction is comparable to that of the previous synthesis analysis discussed in Part I of this paper. In the assimilation process we employed, the quality of the analysis seems to be determined by the characteristics of the measurements, i.e., density, distribution and magnitude of error, and also by the rate of inherited error growth in the dynamical prediction.

1. Introduction

Part I of this paper dealt with a method of data analysis based on the averaging of different estimates of the "present" state of the atmosphere obtained by prediction from sets of "past" data. In reference to the GARP report (GARP Study Group on Numerical Experimentation—Four-dimensional Assimilation, 1970), this method corresponds to the "intermittent assimilation" scheme. A basic requirement in this method is that measurement data be first processed through a spatial "objective analysis", and also through a suitable "initialization" technique.

In the immediate future, meteorological data will be furnished increasingly in non-synoptic form at relatively short time intervals. Already satellites are providing cloud pictures as well as vertical temperature profiles. The EOLE project (in which one of the authors is involved) or other similar programs of balloon tracking systems consist essentially of continuous or quasi-continuous measurements of atmospheric quantities. It appears, therefore, that there is a need to define a technique which allows the treatment of data that are scattered continuously or irregularly not only in three-dimensional space but also in time.

A general way of handling such data is what may be called "continuous assimilation" (Fig. 1). The data are inserted into the numerical prediction system at the times and places at which they are measured, and thereby the data fields are continually updated.

The "continuous" method has an advantage. It avoids the use of any kind of spatial "objective" analysis, which is intended to fill the spatial gaps of input data, and which introduces inevitably supplementary errors. It is not quite obvious, however, whether this method really works well in improving the accuracy of the data arrays. In the method described in Part I, it was almost certain that the synthesis technique tends to reduce random errors of measurement and analysis by the averaging of independent forecasts. Particularly, it was demonstrated that the error can be reduced to almost zero

[1] On leave from le Centre National de la Recherche Scientifique, Paris, France.

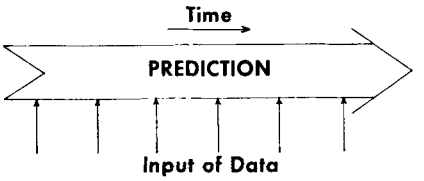

Fig. 1. Continuous data assimilation.

in the case of the linear evolution equation with constant coefficients.

In the continuous assimilation, on the other hand, the impact of the measurement errors on the final analysis is not clear *a priori*. There is no guarantee of a cancellation of random errors. It is the main objective of the present paper to discuss and ascertain the feasibility of increasing the accuracy of analysis with a continuous dynamic data assimilation.

2. Direct assimilation of data

The simplest way of performing a continuous assimilation is to replace the predicted values directly with measured values whenever they become available. We have examined this technique for the case in which the evolution equation is represented by a non-divergent, balanced, inviscid, barotropic model.

The equation is the same as that used in the previous experiment (section 3 of Part I). Namely

$$\frac{\partial \nabla^2 \psi}{\partial t} - J(\nabla^2 \psi + f, \psi) = 0 \qquad (1)$$

The numerical procedure is identical to that used in Part I. The domain and the data to be assimilated are taken in the same way. The domain is a two-dimensional channel consisting of 792 gridpoints with the rigid boundaries at the north and south ends and with the cyclic conditions at the east and west boundaries. Randomly, 352 of the points are designated as "observation stations".

The geopotential height is "measured" at each of the "stations", with a uniformly distributed random observation error. The error is such that the statistical mean is 0 and the standard deviation is 20 m. The initial state only was derived by Gandin's objective analysis from a set of synoptic data.

Before proceeding to the result of this assimilation experiment, let us describe the prediction solution obtained from this initial state. The solid curve in Fig. 2, labelled "predictability decay", is the time development of the mean-square error between this prediction and the true solution. The error invariably grows with time due to dynamic instability and the presence of initial error (Lorenz, 1963).

Now the other two curves, labelled "12 hours" and "24 hours", show the error growths obtained by the direct assimilation technique. In one case measured data are injected every 12 hours and, in the other, every 24 hours at all observation stations. The injection is performed at the time level corresponding to the measurement time of the new data, and does not include any process of carrying the data forward or backward in time. Whenever data are inserted, the mean-square error from the "true" state is reduced instantaneously, and this causes the saw-tooth curves. Yet in both cases, the error grows with time from the very beginning of the process, at a rate somewhat smaller than the predictability decay rate. The result for the injection of data at 6 hour intervals is not shown in Fig. 2, but the error growth is similar to the above. In the case of the 12 hour interval, the growth continues until the computation blows up.

Very interesting results with this type of technique have already been published by Charney, Halem & Jastrow (1969) and Jastrow & Halem (1970), though the evolution equations are based on the 2 level primitive equation model. By imitating observational conditions defined for the GARP project, they showed that with only temperature observations and with this simple replacement process, it is possible to reconstruct the wind and the surface pressure field within the degree of error acceptable for prediction purposes.

They found that the root-mean-square error of wind approaches a low asymptotic value. (Presumably the final analysis of the temperature field contains errors comparable in size to measurement errors.)

Our results are quite different from those reported by Jastrow & Halem. Why is the difference large? It is likely that the main variance comes from the difference in the dynamics of the evolution equations. The balanced barotropic model does not have

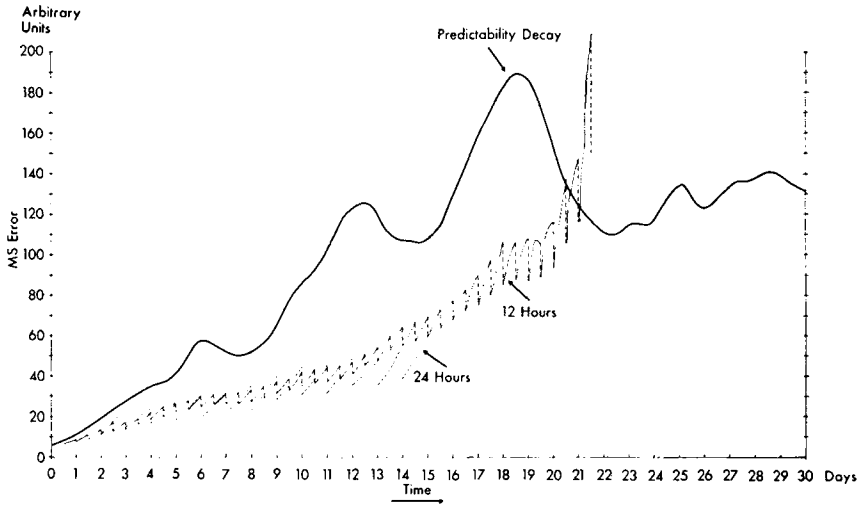

Fig. 2. Time variation of the mean-square error of geopotential height for the "direct" assimilation method (12 hour and 24 hour intervals) and in the decay of predictability.

enough degrees of freedom to assimilate data as dense as those used in this experiment effectively. This may be aggravated by the large random measurement errors, which render the data initially inconsistent and, consequently, more difficult to assimilate by the model. As demonstrated by Smagorinsky, Miyakoda & Strickler (1970), in the baroclinic primitive equation model, there is a certain degree of dynamical redundacy. Even if some initial information is not right, adequate values are produced during the course of the prediction. This is due to the "adjustment" process which acts only in a dynamical system with multi-freedom of the variables.

There is one other important reason for the large difference. In our case, the injection of data was made at the same time for all observation stations. On the other hand, in the case of Jastrow & Halem, imitating the input of infra-red vertical soundings by a satellite that scans the global atmosphere, the data were added piecemeal at one time but frequently. This is an important aspect that we have to consider. It will be discussed in the next section. (A reason which may play an important role and perhaps a determining one, in the difference between both series of experiments has been pointed out by the referee: in the present case, the time integration is performed with the "leap-frog" scheme, a three-level time integration scheme, while the injection of a given set of data is performed at one time level only. Moreover, the "leap-frog" scheme has no damping properties as the Euler-backward scheme used by Jastrow & Halem, which damps high-frequency noise, as the one which may be created by the injection of new data.)

In this paper, we shall proceed with our study confined to the balanced dynamical equation. We intend to focus attention on the problem of pure data assimilation, leaving the problem of dynamical redundancy and adjustment for the time being.

3. General principle of a weighted assimilation

In the dynamical assimilation process, mere increase in the spatial or temporal densities of the data may not be useful by itself. This is partly because some information is redundant or because new information is sometimes inferior to the already analyzed data.

There is another reason. If the measured values are significantly different from the predicted values, the introduction of new data into the numerical model computation will create discontinuities with the environmental field which may generate spurious perturbations and may eventually lead to a break-down of the analysis.

Thus it is necessary, first, to consider the effective use of data so that the full information value is extracted from the data, secondly, to control contamination by unfavorable error from the dynamical solution as well as from the input data, and thirdly, to consider the alleviation of excessive "shock" due to the input of alien data.

For these purposes, some knowledge about the statistics of the observation errors and the prediction errors is desirable; this would allow more precise assessment of the reliability of these observations and predictions and to take these reliabilities into account in the assimilation scheme.

A method through which this is possible may be derived from a study on filtering and prediction problems by Kalman (1960), later extended to the meteorological problem by Petersen (1968). They considered a dynamical system $\Psi(t)$, varying with time, defined by N parameters

$$\Psi = \{\psi_1, \psi_2, \ldots, \psi_N\} \tag{2}$$

each of these N parameters varying itself with time. In the meteorological case, these N parameters will be the values of the meteorological variables at the gridpoints.

The time evolution of Ψ is assumed to be given by a known equation

$$\frac{d\Psi}{dt} = L(\Psi, t) \tag{3}$$

Kalman & Petersen treated the case, where L is linear with respect to Ψ; however, a part of their results is still valid without this hypothesis, which will not be made here. The differences it entails will be discussed further below.

The system Ψ is known only through observations; at successive times $t_1, t_2, \ldots t_n, \ldots$ (not necessarily regularly distributed), a certain number μ of the ψ_i's (with indexes $i_1, i_2 \ldots i_\mu$) are measured which provide the observation data

$$\hat{\psi}_j(t_n) = \psi_j(t_n) + \varepsilon_j(t_n) \quad (j = i, i_2 \ldots i_\mu) \tag{4}$$

$\psi_j(t_n)$ is the actual value of the parameter, $\varepsilon_j(t_n)$ is the measurement error, the statistics of which are assumed to be known for each j and n and which is assumed to be independent of the state of the system $\Psi(t_n)$ at measurement time, and of measurement errors at previous times. In formula (4), μ can vary with time, t_n.

Kalman considered the optimal estimate of Ψ for time t which is derived from the measurements up to time t_n. Let us denote it by $\Psi(t/t_n)$, its N components being $\psi_i(t|t_n)$, $(i = 1, 2 \ldots N)$. By "optimal estimate", Kalman meant the estimate which minimizes the statistical mean square difference from the true state of the system. Of course, this is one of many possible definitions of "optimal estimate". Let us, however, decide temporarily in this paper that, whenever the expression "optimal estimate" will be used, it will be with this precise meaning.

It can be easily shown that the optimal estimate is equal to the statistical mean of all the possible states of the system, knowing all the measurements up to time t_n. Kalman demonstrated that (under the only condition that the optimal estimates be restricted to linear functions of the observations), if the optimal estimate $\Psi(t_{n+1}|t_n)$ at observation time t_{n+1} is known together with the observations $\hat{\psi}_j(t_{n+1})$ at the same time, $\psi_i(t_{n+1}|t_{n+1})$ is then given for all i's $(i = 1, 2 \ldots N)$ by

$$\psi_i(t_{n+1}|t_{n+1}) = \psi_i(t_{n+1}|t_n) + \sum_j \alpha_{ij} \cdot [\hat{\psi}_j(t_{n+1}) - \psi_j(t_{n+1}|t_n)] \tag{5}$$

where the summation $\sum$ is to be executed over all the parameters observed at time t_{n+1}, and where the weights α_{ij} given to the differences between the observed and the estimated values, must be determined appropriately. For this purpose, Kalman used the fact that, according to the definition of an "optimal estimate", $\Psi(t_{n+1}|t_{n+1})$ must minimize the statistical mean square difference from the true state of the system. Let us consider this difference

$$F = \frac{1}{N}\overline{\left\{\sum_{i=1}^{N} [\psi_i(t_{n+1}|t_{n+1}) - \psi_i(t_{n+1})]^2\right\}} \tag{5}$$

where the bar denotes the statistical mean, taken over a large number of realizations of the process. Each of the terms in the summation in (6) being determined independently (the weights α in equation (5) are different for each i), minimizing F is equivalent to minimizing each of these terms.

$$F_i = \overline{[\psi_i(t_{n+1}|t_{n+1}) - \psi_i(t_{n+1})]^2} \tag{7}$$

Using equations (4) and (5), and setting for all $l (l = 1, 2 \ldots N)$

$$E_l = \psi_l(t_{n+1}) - \psi_l(t_{n+1}|t_n) \tag{8}$$

one gets from (7)

$$F_i = \sum_{jk} \alpha_{ij} \cdot \alpha_{ik} \overline{(E_j E_k + E_j \varepsilon_k(t_{n+1}) + E_k \varepsilon_j(t_{n+1}) + \varepsilon_j(t_{n+1}) \varepsilon_k(t_{n+1}))} - 2 \sum_j \alpha_{ij} \overline{(E_i E_j + E_i \varepsilon_j(t_{n+1}))} + \overline{E_i^2} \quad (9)$$

where j and k run over the μ values $(i_1, i_2 \ldots i_\mu)$. The measurement errors $\varepsilon_j(t_{n+1})$ being independent of the state of the system and measurement errors at previous times, are also independent of the prediction errors E_j. So, the covariances between the measurement errors and the E_j's are equal to zero, i.e.,

$$F_i = \sum_{jk} \alpha_{ij} \alpha_{ik} \overline{(E_j E_k + \varepsilon_j(t_{n+1}) \varepsilon_k(t_{n+1}))} - 2 \sum_j \alpha_{ij} \overline{E_i E_j} + \overline{E_i^2} \quad (10)$$

The α_{ij}'s which minimize F_i must satisfy each of the μ following equations obtained by differentiating (10) with repsect to each of the α_{ij}'s

$$\sum_k \alpha_{ik} \overline{(E_j E_k + \varepsilon_j(t_{n+1}) \varepsilon_k(t_{n+1}))} = \overline{E_i E_j} (j = i_1, i_2 \ldots i_\mu) \quad (11)$$

So, for each i, the α_{ij}'s are given by a system of μ simultaneous linear equations, which can be written in matrix form as follows:

$$\{P(t_{n+1}|t_n) + X\} A_i = P_i^* \quad (12)$$

A_i is a p-column vector, whose elements are α_{ij}'s, i.e.,

$$a_{ij} = \alpha_{ij}, \quad j = 1, 2 \ldots \mu. \quad (13)$$

$P(t_{n+1}|t_n)$ is the *prediction error covariance matrix*, consisting of $p \times p$ elements. The element for the kth row and the lth column is

$$p_{kl} = \overline{E_k E_l} = \overline{[\psi_k(t_{n+1}) - \psi_k(t_{n+1}|t_n)] \times [\psi_l(t_{n+1}) - \psi_l(t_{n+1}|t_n)]} \quad (14)$$

P_{kl} is, therefore, the covariance of the prediction errors for the parameters k and l at time t_{n+1}.

$P(t_{n+1}|t_n)$ is a symmetrical matrix. X is the *matrix for the measurement error covariance*, consisting of $p \times p$ elements. The element is written as

$$\varkappa_{kl} = \overline{\varepsilon_k(t_{n+1}) \cdot \varepsilon_l(t_{n+1})} \quad (15)$$

X is the channel through which the measurement errors are taken into account. P_i^* is a p-column vector of the prediction error covariance. The element is, for i and j, i.e.,

$$p_{ij}^* = \overline{E_i E_j} = \overline{[\psi_i(t_{n+1}) - \psi_i(t_{n+1}|t_n)] \times [\psi_j(t_{n+1}) - \psi_j(t_{n+1}|t_n)]} \quad (16)$$

If the parameter ψ_i is observed at time t_{n+1}, P_i^* is identical to a column of $P(t_{n+1}|t_n)$. If not, there is no common term between P_i^* and $P(t_{n+1}|t_n)$. If $\Psi(t_{n+1}|t_n)$, p_{kl} for all parameters* and $\varkappa_{kl}$ are known, system (11) provides a way of computing $\Psi(t_{n+1}|t_{n+1})$ with use of the measurement data $\hat{\psi}(t_{n+1})$.

In order to solve system (11), it is required that $\Psi(t_{n+1}|t_n)$ and $P(t_{n+1}|t_n)$ both at t_{n+1} are determined from the previous solutions at t_n. Kalman and Petersen studied the problem when the dynamic (evolution) equation (3) is linear with respect to Ψ. In this case $\Psi(t_{n+1}|t_n)$ is determined simply by integrating (3) analytically. The solutions of (3), based on an optimal estimate, are also optimal estimates. Moreover, $P(t_{n+1}|t_n)$ is determined simply from $P(t_n|t_{n-1})$ together with the statistical information on the measurement errors at t_n.

4. Case of a non-linear model

The weighted assimilation technique is now applied to a non-linear balanced barotropic equation (1).

The data are taken in exactly the same way as in the non-weight experiment. Since the model is non-linear, the equations for $\Psi(t_{n+1}|t_n)$ and $P(t_{n+1}|t_n)$ cannot be solved in the simple manner mentioned above. In addition, even if the initial state of a forecast is an optimal estimate, the forecast produced is not necessarily optimal.

[1] Strictly speaking, it is necessary to know only the covariances of those couples of parameters, one of which at least is measured at time t_{n+1}.

In meteorological systems, such as in the example above, quadratic terms appear in the evolution equations (in equation (1), these quadratic terms make up the Jacobian). As said above, the optimal estimate knowing all observations up to time t_n is equal to the statistical mean of the state of the system knowing these observations. If we denote this statistical mean by a bar (different from the mean considered in the previous section), equation (1) will become, by taking mean

$$\frac{\partial \nabla^2 \bar{\psi}}{\partial t} - \overline{\frac{\partial(\nabla^2\psi + f)}{\partial x}} \cdot \overline{\frac{\partial \psi}{\partial y}} + \overline{\frac{\partial(\Delta^2\psi + f)}{\partial y}} \overline{\frac{\partial \psi}{\partial x}} = 0 \tag{17}$$

So, to integrate the optimal estimate $\bar{\psi}$, it is necessary to know the optimal estimates of the quadratic terms of the evolution equation. These terms are not necessarily the products of the optimal estimates of the factors, the statistical mean of a product not being necessarily the product of the statistical mean of its factors. So these terms should be integrated with time, which requires, as can be seen by taking their time derivatives, the knowledge of the third order moments, the integration of which requires fourth-order moments, and so on. Therefore, to perform the integration, a hypothesis on these moments is necessary to set up a kind of closure system. In the present case, as a first approximation, the simplest assumption was made: namely that the factors of the quadratic terms which appear in (17) are independent, and that the corresponding moments are simply the products of the optimal estimates of the factors. This is equivalent to computing $\Psi(t_{n+1}|t_n)$ from $\Psi(t_n|t_n)$ using equation (1), or more precisely the equation in its finite difference approximation.

The prediction error covariance matrix $P(t_{n+1}|t_n)$ is also very difficult, if not impossible, to determine from its value at the previous assimilation time because of the non-linearity. It was, therefore, derived statistically from the covariance errors computed for a large number of forecast samples beforehand.

For these reasons, the validity of the method in the case of non-linear equations is judged empirically, and tests are needed. In practice, the summation of gridpoints in formula (5) was not made over all the measurements, but only over those at stations within a 1 500 km radius from the point of assimilation. This should not affect the results in a significant way. In fact, the prediction error covariance between two gridpoints more than 1 500 km apart is negligible.

The observation errors are assumed to be independent of each other, and, as mentioned above, have a standard deviation σ of 20 meters. The matrix X in formula (12) is, therefore, written as

$$X = \sigma^2 \cdot I \tag{18}$$

where I is the unit matrix.

Formula (5) requires at each assimilation time an optimal estimate $\Psi(t_{n+1}|t_n)$ defined from the previous assimilation. At time t_1 when the first set of data is provided, there is no previous assimilation yet there is still an optimal estimate, which minimizes the root-mean-square error in a statistical sense. This optimal estimate is the climatological mean of the state of the atmosphere. In the numerical test, this climatological mean was taken as the mean state of the model over a long period. As a matter of fact, this assimilation performed according to equation (5), where $\Psi(t_{n+1}|t_n)$ is replaced with the climatological mean, is exactly identical to Gandin's technique of objective analysis (1963).

As mentioned above, the prediction error covariances necessary to compute the weights in each assimilation step were determined statistically. The computation of the forecasts used as statistical samples to determine the covariances at time t_{n+1} requires a knowledge of the assimilation weights at the previous times $t_1, t_2 \ldots t_{n-1}, t_n$. So, the covariances at different assimilation times are to be determined sequentially. The covariances for three successive assimilation times 12 hour apart (excluding the initial time) were computed, using respectively 337, 331 and 322 statistical samples of forecasts. The values found for these covariances are discussed further below (section 6).

From each of these three sets of covariances, the weights for all gridpoints were determined by solving equation (12). For a given gridpoint, the weight decreases rapidly with distance, being less than 0.10 at a distance of more than two grid lengths. The summation of the weights at one gridpoint varies between 0.4 and 0.8, depending mainly upon the number and the repartition of the neighboring observation

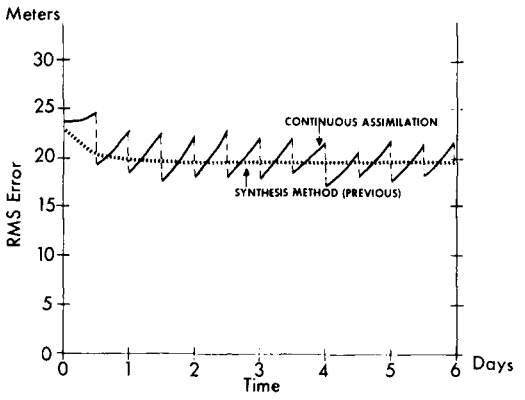

Fig. 3. Time variation of the root-mean-square error of geopotential height for the "continuous" assimilation method with adequate weights and excercising the "synthesis" method.

stations. The weight for the input data decreases with assimilation steps. For instance, the mean value of the weight given to new data at the location of this observation itself is 0.56 in the first assimilation, 0.45 in the second and 0.41 in the third. The decrease corresponds to the fact that, as the assimilation process goes on, the predicted state should become more and more exact, and that consequently less weight is given to the new data. (In the case of assimilation of actual data, there is a discrepancy between the evolution equation and nature, and the decrease of rms error would be slower.)

5. Results

In Fig. 3 the curve labelled "continuous assimilation" shows the variation of the root-mean-square error with time in a six day forecast, new "measured data" being assimilated every 12 hours.

Successive assimilations cause, at least in the beginning, a gradual decrease in the medial value of the root-mean-square error; the saw-tooth pattern of the curve is due to the fact that each new assimilation of data causes a sudden drop in the error. The three different sets of weights computed as mentioned above were used for the first three assimilations, and the third one for the following assimilations. Similar computations were performed with only one and two sets of weights; the results (not presented in the figure) are quite similar, the corresponding curves being situated slightly above that obtained with three sets of weights. The differences with the results presented in the figure are, after six days: 1.3 meter (for the assimilation using only one set of weights) and 0.3 meter (for the assimilation using two sets of weights). This shows that, as the assimilation process goes on, there is less and less utility in computing additional sets of weights; the gain obtained by computing a third set is already negligible. This is the reason why no more sets were computed beyond the third one.

Another test performed with the injection of "measurements" at 6 hour intervals (and one set of weights) yielded similar results: in particular the minimum value of mean-square error is the same.

It is interesting to note that in the results of the continuous assimilation the asymptotic level reached by the root-mean-square error coincides with the level obtained with the previous "synthesis" method studied in Part I (Fig. 3). The time necessary to reach this minimum level from the initial error is also the same. Moreover, as seen in Part I, the minimum level of the error is close to the systematic error occurring in Gandin's objective analysis under the assumption of no measurement error. So, it turns out that the three error levels are almost the same. It may also be noted that the residual error obtained in the "synthesis" and the continuous assimilation techniques are much closer to each other than they are to Gandin's analysis error. Thus with two fundamentally different technique s,we get very similar results. Does any dynamical analysis technique yield a lower residual error? As we have seen in Part I, it is certain that in the case of the linear evolution equation the error can be reduced to practically zero. In the case of the non-linear equation, the degree of possible improvement in the analysis largely depends on the predictability decay and also on the characteristics of the input data; that is, the spatial and temporal density of data (and relative randomness), and the size of the measurement errors. The relative importance of these characteristics on the residual error is not known. As for the spatial and temporal densities, there must be optimal values beyond which no gain is obtained. As for the degree of measurement error, whether it plays a dominant role in the determination of this residual error is not clear either. (Jastrow & Halem, 1970, discussed the evidence in this respect.)

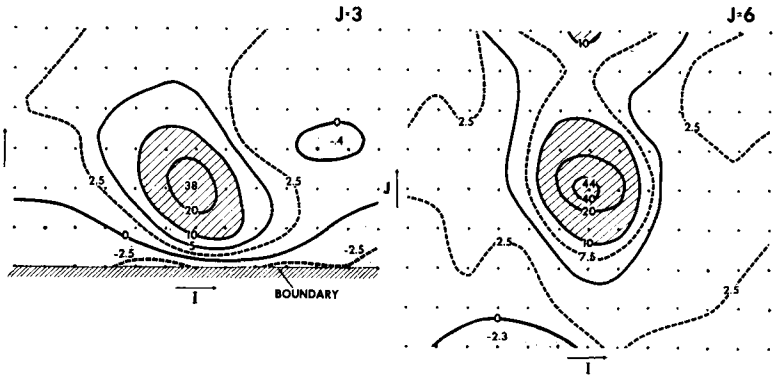

Fig. 4. Two-dimensional distribution of error covariance $E_k\ E_l$ at $J = 3$ and $J = 6$, where J is the grid-point index for y-direction. The units are arbitrary.

6. Prediction error covariance

The computation of the assimilation weight α_{ij} in equation (12) requires statistical knowledge of the prediction error covariance, $p_{kl} = \overline{E_k \times E_l}$ in (14). In the foregoing section, p_{kl} is computed at each gridpoint independently without any assumption about the influence of the surrounding area except, as was mentioned in section 4, that the prediction error covariance is neglected if the distance between two points is more than 1 500 km. Even under this restriction, however, the number of surrounding observation points used in the assimilation at a given gridpoint is, on the average, 12.5 with a minimum of 7 and maximum of 22. The total number of couples whose error covariance is to be determined is then 82 997 at one time. This number is reduced to 12 394 because of redundancies. This requires a considerable amount of computation. In the future, it may be possible to use a predetermined set of p_{kl} derived under appropriate assumptions. With this in mind, we examined the statistical properties based upon the materials we had in the experiments.

The assumption is that the distribution of p_{kl} is homogeneous in the x-direction but is a function of y. So $\overline{E_k \times E_l}$ computed at many gridpoints are averaged only with respect to x.

Fig. 4 shows two examples of $\overline{E_k \times E_l}$, one at $J = 3$, which is near the south boundary (J is the gridpoint index in the y-direction), and the other at $J = 6$, which is in the middle of the channel. The distribution is not isotropic.

Figure 5 is the profile of $\overline{E_k \times E_l}$ as a function of grid distance in the x-direction from the central point under consideration. The three profiles are for the 1st, 2nd and 3rd assimilation times. As the assimilation proceeds, the value at the central point becomes small but $\overline{E_k \times E_l}$ is spread over a wider range.

7. Remarks

Several points in the results presented in this paper call for further remarks.

First of all, the measure which we used to compare two different states of the model's solution was the mean-square difference. Among those which one can easily handle mathematically, the mean-square is probably one of the

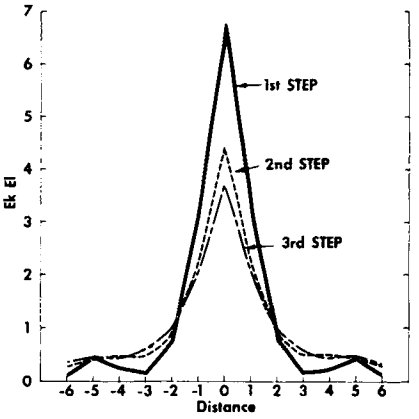

Fig. 5. Profiles of error covariance $E_k\ E_l$ as a function of gridpoint distance in the x-direction. The three profiles are for the 1st, 2nd and 3rd assimilation steps. (12-hour interval.)

most important measures. It fails, however, to reveal all aspects of the differences and in some cases it may even be deceiving, in that it does not provide much information about the dynamical similarity of the states that are compared.

So the minimization of the mean-square difference does not necessarily produce the most optimal adjustment. The method of continuous assimilation presented in this paper is, therefore, not a unique procedure, and may not be the best. It would be desirable in further studies not to restrict oneself to a single criterion.

Another point is that we have not touched on the defects of the prediction model. Indeed this is the most critical factor. If the model is inferior, past data may not be usable at all. Perhaps the model's performance is more important in practice than natural predictability decay. But, be that as it may, the essential concept for a scheme of data assimilation may not be drastically altered from the one described here.

Another method that claims the authors' attention is that of Sasaki (1970), which seems attractive in the sense that various meteorological variables are used for an analysis of one variable through the dynamical constraints. This method is conceptually different from the continuous data assimilation, however.

We noticed that the 4-dimensional analysis, however, is being studied extensively at a number of institutes. Examples are Rutherford and Bradley in Montreal, Canada (personal communication) and Dixon in the United Kingdom and perhaps many in the USSR (Yudin, Mashkovitch, Gandin etc.) (see the report of Döös, 1969). We also know by personal communication that Morel et al. (1970) studied a method. It is based on feeding the measurements into successive forecasts and "hindcasts" performed with the assimilation model over the period of time in which the data are available. It will be interesting to compare our method with this.

8. Conclusions

A method of continuous data assimilation was tested with the balanced barotropic equation as the evolution equation. The result indicates that the assimilation method of the simple replacement of input data led to a blow-up of the analysis, where the measurement data were given simultaneously at a number of stations.

On the other hand, the method with adequate weights for the input data provided a reasonable solution for the analysis. The original error in the analysis based upon Gandin's analysis method was reduced with time by this continuous assimilation method. The reduction of error is, however, the same magnitude as that by the synthesis method described in Part I. So far as this method is concerned, there is an asymptotic level, beyond which the error could not be reduced as measured by the root-mean-square error.

Acknowledgements

The authors are thankful to Professor J. Smagorinsky, P. Morel, and S. Manabe, T. Matsuno for the useful suggestions and encouragement. The paper was carefully checked by Mr G. D. Hembree, and typed by Y. Towns.

REFERENCES

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

Döös, B. R. 1970. Numerical experimentation related to GARP. GARP Publication Series No. 6 WMO–ICSU Joint Organizing Committee.

Gandin, L. S. 1963. *Objective analysis of meteorological fields.* Gidrometeorologicheskoe Izdatel'stvo, Leningrad, 1963. Translated from Russian into English, Israel Program for Scientific Translations, Jerusalem, 1965, 242 pp.

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

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. *Journ. Basic Engineering 82*, 35–45.

Lorenz, E. N. 1963. The predictability of hydrodynamic flow. *Trans. New York Acad. Sci., Ser. 2, 25*, 409–432.

Miyakoda, K. & Talagrand, O. 1970. The assimilation of past data in a dynamical analysis, Part I. To be published.

Morel, P., Lefevre, G. & Rabreau, G. 1970. On initialization and non-synoptic data assimilation. To be published.

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

Sasaki, Y. 1970. Some basic formalisms in numerical variational analysis. *Mon. Wea. Rev. 98*, 875–883.

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

ИСПОЛЬЗОВАНИЕ ПРОШЛЫХ ДАННЫХ В МЕТЕОРОЛОГИЧЕСКОМ ДИНАМИЧЕСКОМ АНАЛИЗЕ ЧАСТЬ II

Рассматриваются системы для непрерывного использования данных. В качестве примера динамического уравнения берется баротропное уравнение. В первом методе данные геопотенциальной высоты используются в вычислениях по мере их поступления, без модификации. Это приводит к недопустимому росту ошибки, что говорит о необходимости фильтрации приходящей информации или смягчения расхождений между вновь поступающими данными и предсказанными полями. Во втором методе оптимальное использование достигается путем применения соответствующих весов к поступающим данным, что статистически минимизирует средний квадрат разности между оценкой и истинным решением. Результат анализа приемлем. Но величина уменьшения ошибки сравнима с аналогичной величиной в предыдущем анализе, разработанном в I части статьи. При используемой нами процедуре качество анализа, по-видимому, определяется характеристиками измерений, т. е. плотностью, распределением и величиной ошибки, а также скоростью роста ошибки в динамическом предсказании.