Stochastic dynamic prediction[1]

By EDWARD S. EPSTEIN, *University of Michigan, Dept. of Meteorology and Oceanography, Ann Arbor, Mich.*[2,3]

(Manuscript received April 14, 1969)

ABSTRACT

Stochastic dynamic prediction assumes the laws governing atmospheric behavior are entirely deterministic, but seeks solutions corresponding to probabilistic statements of the initial conditions, thus recognizing the impossibility of exact or sufficiently dense observations. The equation that must be solved is the continuity equation for probability. For practical reasons only approximate solutions to this equation are possible in general. Deterministic forecasts represent a very low order of approximation. More exact methods are developed and some of the attributes and advantages of stochastic dynamic predictions are illustrated by applying them to a low order set of dynamic equations.

Stochastic dynamic predictions have significantly smaller mean square errors than deterministic procedures, and also give specific information on the nature and extent of the uncertainty of the forecast. Also the range of time over which useful forecasts can be obtained is extended. However, they also require considerably more extensive calculations.

The question of analysis to obtain the initial stochastic statement of the atmospheric state is considered and one finds here too a promise of significant advantages over present deterministic methods. It is shown how the stochastic method can be used to assess the value of new or improved data by considering their influence on the decrease in the uncertainty of the forecast. Comparisons among physical–numerical models are also made more effectively by applying stochastic methods. Finally the implications of stochastic dynamic prediction on the question of predictability are briefly considered, with the conclusion that some earlier estimates have been too pessimistic.

List of symbols

a_{ijk} coefficients of quadratic terms in the deterministic prognostic equations (11)

A_i coefficients of harmonic terms used to specify the vorticity (18)

$\hat{A}_i$ least squares estimates of A_i (33)

$\hat{\mathbf{A}}$ $=(\hat{A}_1, \ldots, \hat{A}_D)$

b_{ij} coefficients of linear terms in the deterministic prognostic equations (11)

c_i constant terms in the deterministic prognostic equations (11)—(in sec. 3 only)

c_i coefficients in the approximate stochastic equations (23–31)—(in sec. 4 only)

D the number of parameters used to define the state of the atmosphere; equivalently the dimensionality of the phase space

$E(\ )$ expected value operator (5)

k minimum wave no. in the x-direction (18)

l minimum wave no. in the y-direction (18)

n number of observing stations

N sample size for Monte Carlo experiments

s standard error of each observation

$\mathbf{S}$ $=(\mathbf{Z}'\mathbf{Z})^{-1}$, proportional to covariances of $\hat{A}_i$

t time

t_0 time of initialization of the forecast

x, y orthogonal coordinates defining the plane of motion

x_i, y_i coordinates locating the ith observing station (in sec. 6)

x_i the general time-dependent parameters defining the state of the atmosphere; coordinates in phase space (in sec. 2, 3)

---

[1] It has been brought to my attention that V. I. Tatarskiy (Izvestia, Atmospheric and Oceanic Physics, 5 (3), 1969, 293–297) has independently derived the approximate stochastic dynamic equations.

[2] Contribution No. 156 of the Department of Meteorology and Oceanography. The University of Michigan, Ann Arbor, Michigan.

[3] Contribution No. 204 from the Institute of Meteorology, University of Stockholm.

$\mathbf{X}$ $=(x_1, ..., x_D)$
$\mathbf{X}_f$ a forecast for $\mathbf{X}$
$\hat{\mathbf{X}}$ an estimate of $\mathbf{X}$ from observations
Y_i observation made at the ith station
$\mathbf{Y}$ $=(Y_1, ..., Y_n)$
z_{ij} terms in eq. (32) by which the A_i are multiplied
$\mathbf{Z}$ $=((z_{ij}))$
α $=k/l$
μ_{ij} expected values of x_i or of A_i (6)
ϱ_{ij} expected values of $x_i x_j$ or of $A_i A_j$ (7)
σ_{ij} covariances of x_i or of A_i (8)
σ_i $=\sigma_{ii}^{\frac{1}{2}}$
τ_{ijk} third moments of x_i about their means
φ probability density, a function of the state of the atmosphere and of time
ψ streamfunction for two-dimensional flow
$(\dot{})$ $d(\ )/dt$, a total derivative with respect to time

1. Introduction

The present study is an attempt to deal explicitly with the problems imposed on meteorological prediction by the patent impossibility of observing the atmosphere either in sufficient detail or with sufficient accuracy to consider the initial state as known with certainty.

It has long been recognized that observational errors contribute substantially to errors in prediction. Much effort has gone into "analysis" as a procedure for determining that single atmospheric state which, in terms of the available data, is a "best" approximation of the unknown true state of the atmosphere. There are a dozen or so weather services which daily produce their own analyses, each different from all the others, but each consistent with all the observational data. One cannot say that the product of any such system is right or wrong; each represents an individual member of an infinite ensemble of atmospheric states which are consistent with the data. The different analyses will yield different forecasts, even if each were submitted to the same forecast procedure. If there is no way of determining which, if any, analysis is right, and since none is known to be wrong, there is no way of knowing, in any instance, which forecast to believe.

The best that we can do is to assign to each analysis some number which is proportional to the amount of confidence we wish to place in that result. There exist formal statistical procedures for this, which amounts to assigning a probability number to every possible analysis, or better, to describing a probability density function in a multidimensional phase space. This notion will be pursued further, but for the present let us simply note that the notion of a probability density function is equivalent to that of an infinite ensemble of initial states with relative frequencies within the ensemble proportional to the probability densities. In either case, this is the only accurate way of stating the initial conditions for a forecast. This point has been clearly stated by Gleeson (1968). Since the probabilistic nature of the initialization is unavoidable, then so also is the probabilistic nature of the prediction.

It has been shown earlier (Epstein, 1969) that even when each member of an ensemble obeys identical deterministic laws, the center of the ensemble will in general follow a different set of laws.

Thus we propose an approach to forecasting that treats the atmosphere as deterministic, obeying the fundamental laws of hydrodynamics, but recognizes that the state of the atmosphere can be known only in a probabilistic form. We refer to this approach as "stochastic dynamic". It is important to emphasize that the physical aspects of this approach dominate, but that application of the physics to prediction must be modified in the light of elementary probability theory. The use of the term "statistical" is avoided here to make it clear that the method does not depend on the accumulated records of past weather, other than to the extent that these can verify the physical models used to describe atmospheric behaviour.

The present study was stimulated largely by the work of Gleeson (1966, 1967, 1968) and Lorenz (1965, 1968*a*). In an important contribution not sufficiently known in meteorological circles, Freiberger & Grenander (1965) anticipate much of the framework we shall discuss. Other than the work of these authors, however, and the earlier contributions by Thompson (1957) and Novikov (1959), very little has been done to deal explicitly with the general problem of uncertainties in the initialization of the forecast.

The approach being proposed here actually offers more than simply improving the forecast. As will be shown it is essential to the method

that there be given clear expression to the eventual uncertainties (or degrees of certainty) in the forecast. By examining some of the aspects of stochastic dynamic prediction many of these advantages will, it is hoped, become apparent. Necessarily, because the procedures to be described provide more, they also cost more in terms of the computing effort they require. Only by examining both aspects of the problem can the merits of this approach be judged, but just as experience and study is likely to reduce the production cost, so will time and effort point out increasingly the benefits of stochastic predictions. In what follows only brief suggestions of the possibilities for carrying out and using stochastic predictions can perforce be given.

In the next section we will discuss the primary mathemathical basis for stochastic dynamic prediction, which is the continuity equation for probability (Gleeson, 1966, 1968). Since this equation cannot be readily dealt with numerically, some approximate methods will be introduced in sections 3 and 4. Various applications of these methods, and examples designed to throw some light on their attributes and the potential benefits from the stochastic approach are described in sections 4–8. These examples are based primarily on Lorenz's (1960) maximum simplification of the dynamic equations. In the light of these examples some of the implications of the stochastic method for determining the predictability of the atmosphere are presented in section 9. The last section presents a few concluding remarks.

2. The continuity equation for probability

Let us represent the state of the atmosphere at any time t by a finite number of parameters, x_i, $i = 1, \ldots, D$. One can think of these as defining a vector $\mathbf{X}$ in a D-dimensional phase space. The dimensionality of the phase space may be quite large and will depend on the particular model chosen to represent the atmosphere. The models that will be dealt with here will be relatively simple and contain only a few (< 10) parameters. In any case the models should have fewer parameters than there are observations to define them, so that in a practical situation dimensions of the order of several hundred, but no larger, may be encountered.

The changing state of the atmosphere is represented by the motion through phase space of the end-point of the vector $\mathbf{X}$. The principles of physics may in general be stated in such a way that the instantaneous velocity of the vector in phase space depends only on its current value. In other words the deterministic physical laws may be stated as

$$\dot{\mathbf{X}} = G(\mathbf{X}(t)) \tag{1}$$

where the dot ($\cdot$) implies a total derivative with respect to time.

The observations that are available to define the value of $\mathbf{X}$ at a given time may also be described by a vector $\mathbf{Y}$ (of somewhat larger dimensionality than $\mathbf{X}$). In effect, deterministic procedures are to define a single "best" initial condition $\mathbf{X}$ in terms of the observations $\mathbf{Y}$, and then to integrate (1), starting with $\mathbf{X} = \hat{\mathbf{X}}$ at $t = t_0$, to obtain the predicted state of the atmosphere $\mathbf{X}_f$ at some $t > t_0$. Symbolically one can write

$$\hat{\mathbf{X}} = A(\mathbf{Y}) \tag{2}$$

$$\mathbf{X}_f - \mathbf{X} = \int_{t_0}^{t} G(\mathbf{X}(t))\, dt \tag{3}$$

where A is the analysis operator by which $\hat{\mathbf{X}}$ is chosen for $\mathbf{Y}$.

We have argued that a more reasonable interpretation of the observations is that they imply a probability density $\varphi(\mathbf{X}; t_0)$ defined over the entire phase space, where φ must have the properties that $\varphi \geqslant 0$ for all $\mathbf{X}$ and t and $\int\int \ldots \int \varphi(\mathbf{X}; t)\, dx_1 \cdot dx_2 \ldots dx_D = 1$ (see footnote). In place of (2), then, one must consider an analysis $B(\mathbf{Y})$ which accomplishes this. One example of such a procedure is described in section 6.

If (2) is modified, so also must a substitute be found for (3); it is necessary to describe an algorithm by which $\varphi(X; t)$, $t > t_0$, can be determined once $\varphi(\mathbf{X}, t_0)$ is known. For this purpose the one-dimensional continuity equation for probability introduced by Gleeson (1966) may be extended to any number of dimensions and written

[1] It is possible to visualize the probability density in phase space, as analogous to mass density (usually ϱ) in three-dimensional physical space. Note that $\varrho \geqslant 0$ for all space and time, and $\int\int\int (\varrho/M)\, dxdydz = 1$ if M is the total mass of the system. The "total probability" of any system is, by definition, one.

$$\left.\begin{aligned} &\frac{\partial \varphi}{\partial t} + \sum_{i=1}^{D} \frac{\partial}{\partial x_i}(\dot{x}_i \varphi) = 0 \\ \text{or} \quad &\\ &\frac{\partial \varphi}{\partial t} + \nabla_D \cdot (\dot{\mathbf{X}} \varphi) = 0 \end{aligned}\right\} \tag{4}$$

where ∇_D is the del-operator in D-dimensional phase space. The derivation of (4) is entirely analogous to the derivation of the standard equation of continuity of mass. One derivation is based on the principle that mass is neither created nor destroyed, the other on the principle that the total amount of probability must always be one (no ensemble members may be created or destroyed).

Given appropriate boundary conditions (e.g. $\varphi \to 0$ as $x_i \to \pm\infty$) and the initial value $\varphi(X; t_0)$, the direct integration of (4) ... to obtain $\varphi(\mathbf{X}; t)$, $t > t_0$, offers no mathematical difficulties. Practically, however, any calculation of φ by direct numerical integration of (4) is a most ambitious undertaking. If one wishes to evaluate φ at L incremental values of each of the components of the phase vector, this corresponds to an $Lx\, Lx \ldots xL$ grid in phase space or D^L points, for each of which it is necessary to remember and update the numerical value at each time step. $D = L = 10$ would be a modest meteorological problem, but 10^{10} is too many quantities for any computer.

If the complete numerical evaluation of φ is an excessive problem, it could be added that this both requires and provides more information than is normally of interest. In general one is interested in only a relatively small number of functions of the form

$$E[f(\mathbf{X})] = \int f(\mathbf{X})\, \varphi(\mathbf{X}; t)\, d\mathbf{X} \tag{5}$$

where $dX = dx_1 \cdot dx_2 \ldots dx_D$ and the integration is over all of phase space. This is just the "expected value" of $f(\mathbf{X})$ at time t. If $f(\mathbf{X})$ is of the form $x_1^a x_2^b \ldots$ one is just defining the moments of the distribution. Usually it is the simplest moments that are of concern. These are the means

$$\mu_i(t) = \int x_i\, \varphi(\mathbf{X}; t)\, d\mathbf{X} \tag{6}$$

and the second moments[1]

$$\varrho_{ij}(t) = \int x_i x_j\, \varphi(\mathbf{X}; t)\, d\mathbf{X} \tag{7}$$

In place of (7) one is generally more concerned with

$$\sigma_{ij} = \varrho_{ij} - \mu_i \mu_j = E[(x_i - \mu_i)\,(x_j - \mu_j)] \tag{8}$$

which are the variances ($i = j$) and the covariances ($i \neq j$). The set of quantities comprising the first and second moments has only $D(D+3)/2$ members and contains as much of the information about φ as is usually of general interest.[2] It would therefore be useful to evaluate only these quantities at each time step and hopefully this is a more feasible undertaking than working with the entire function φ.

Unfortunately we know of no way of obtaining exact representations of just these integrals in a closed finite set of equations that can be evaluated conveniently. Consequently we have turned to approximate procedures. The following sections describe and discuss some of these.

3. The approximate stochastic method

We first describe a procedure by which it is possible to obtain directly the moments (6) and (8), through the application of certain assumptions concerning higher order moments. Note first that differentiation of (5) gives

$$\frac{d}{dt} E[f(\mathbf{X})] = E[df(\mathbf{X})/dt]$$

$$= \int \dot{f} \varphi\, d\mathbf{X} \tag{9}$$

Therefore $\mu_i = E(\dot{x}_i)$

$$\dot{\varrho}_{ij} = E(x_i \dot{x}_j + \dot{x}_i x_j) \tag{10}$$

[1] To continue the analogy between mass and probability densities, means and second moments correspond to centers of mass and moments of inertia. The variances are the moments of inertia about the center of mass.

[2] The common notation of an expected value plus or minus one standard deviation ($\mu \pm \sigma$) is an illustration of this. Also there are theorems in the theory of mathematical probability which permit one to make limiting statements about φ given the means and covariances.

According to Lorenz (1963), forced dissipative systems, of which the atmosphere may be presumed to be an example, are typified by prognostic equations of the form

$$\dot{x}_i = \sum_{j,k} a_{ijk} x_j x_k - \sum_j b_{ij} x_j + c_i \qquad (11)$$

where

$$\sum_{i,j,k} a_{ijk} x_i x_j x_k = 0 \quad \text{and} \quad \sum_{i,j} b_{ij} x_i x_j \geqslant 0$$

(cf. the prognostic equations of Platzman (1960), Baer (1964), Ellsaesser (1966), Robert (1966), Saltzman (1962), Bryan (1959) and others). Then from (10)

$$\dot{\mu}_i = \sum_{j,k} a_{ijk} \varrho_{jk} - \sum_j b_{ij} \mu_j + c_i$$

$$= \sum_{j,k} a_{ijk} (\mu_j \mu_k + \sigma_{jk}) - \sum_j b_{ij} \mu_j + c_i \qquad (12)$$

and

$$\dot{\varrho}_{ij} = E[\sum_{k,l} (a_{jkl} x_i x_k x_l + a_{ikl} x_j x_k x_l)]$$

$$- \sum_k (b_{ik} \varrho_{jk} + b_{jk} \varrho_{ik}) + c_i \mu_j + c_j \mu_i \qquad (13)$$

Introducing $\tau_{ijk} = E[(x_i - \mu_i)(x_j - \mu_j)(x_k - \mu_k)]$, the third moments about the instantaneous mean, (13) can be rewritten

$$\dot{\sigma}_{ij} = \sum_{k,l} [a_{jkl}(\mu_k \sigma_{il} + \mu_l \sigma_{ik}) + a_{ikl}(\mu_k \sigma_{jl} + \mu_l \sigma_{jk})]$$

$$- \sum_k (b_{ik} \sigma_{jk} + b_{jk} \sigma_{ik})$$

$$+ \sum_{k,l} (a_{jkl} \tau_{ikl} + a_{ikl} \tau_{jkl}) \qquad (14)$$

It can be assumed that initial values of the μ_i and the σ_{ij} are known. However, unless the third moments are also known, (14) cannot be solved and, then, neither can (12). One could assume initial values for the τ_{ijk} but to integrate (14) these values would have to be known for $t > t_0$. A prognostic equation for the τ_{ijk} could be derived by again applying (9), but the new equations would then contain the fourth moments.

As discussed very clearly by Freiberger & Grenander (1965) as long as the deterministic prognostic equations are nonlinear it is impossible to write a closed finite set of prognostic equations for the moments. In other words, to predict exactly the future behavior of even the mean of the distribution, all the moments of the distribution (or, equivalently, the entire distribution itself) must be known. There is good justification, however, for expecting that a useful approximate procedure may be found by making suitable assumptions concerning the higher order moments.

If one expresses $\mu_i(t)$ as a Taylor series in time, using repeatedly (6) and (9), one finds that the higher order moments enter only in the terms containing higher powers of $(t - t_0)$. In other words, these terms would be very important for very long-range predictions, but they should have very little significance if t is not too large. In any case the range of time over which the higher moments remain unimportant will depend on the higher order moments themselves. Since, in the physical situations being considered here, all moments must remain finite, it should always be possible to find some limiting order of moment that one must retain in order to assure that the cumulative error in the mean remain sufficiently small for a specified length of forecast. In general, unless one included very high order moments, one could not expect a forecast to remain valid beyond the time when the probability density begins to resemble climatological relative frequencies.

On the other hand there is no guarantee that truncation of the series at a low order, say after the second moments, will produce acceptable results. Yet the requirement of holding down the total number of dependent variables seems to dictate this limit. Also, for many operational purposes it is these moments and only these that one is required to know, and, moreover, a knowledge of the first two moments may be a reasonable limit as to what one can expect to derive, with confidence, from the observational data.

Finally, there are earlier results (Epstein, 1969) which suggest that if the first two moments are the same, and assumptions concerning the higher order moments are not unreasonable, then the forecasts are insensitive to the assumptions. Thus, we have made the pragmatic decision to attempt to predict the first and second moments by making appropriate assumptions about the third moments.

Specifically we will assume that

$$\sum_{k,l} (a_{jkl} \tau_{ikl} + a_{ikl} \tau_{jkl}) = 0 \qquad (15)$$

This allows us to write, instead of (14),

$$\dot{\sigma}_{ij} = \sum_{k,l} [a_{jkl}(\mu_k \sigma_{il} + \mu_l \sigma_{ik}) + a_{ikl}(\mu_k \sigma_{jl} + \mu_l \sigma_{jk})] - \sum_k (b_{ik} \sigma_{jk} + b_{jk} \sigma_{ik}) \quad (16)$$

Equations (12) and (16) now form a closed set which can be integrated subject to initial values of the μ_i and σ_{ij}. We refer to these as the approximate stochastic equations.

It is the very success of deterministic predictions which suggests that the approximate stochastic equations are a useful set. While (12) and (16) include terms for the covariances and for the rates of change of the covariances, deterministic prediction implicitly assumes that all variances are zero. Thus the approximate stochastic equations are higher order approximations to the general stochastic dynamic equations than have previously been used. They will be superior, but whether the extent of the superiority can justify the extra effort can only be judged on the basis of experimentation. Sections 5 through 8 will describe a few experiments of this nature.

The general condition on the deterministic prognostic equations, that $\Sigma_{i,j,k}\, a_{ijk} x_i x_j x_k = 0$, implies that $\Sigma_i x_i^2$ is conserved along each trajectory, in the absence of forcing or dissipative effects (represented by the b_{ij} and c_i). The physical nature of the conserved quadratic property (or properties if partial sums of the $a_{ijk} x_i x_j x_k$ are zero depends of course on the model considered. Typically these are various forms of energy or, in certain common problems in fluid dynamics, the mean square vorticity. The corresponding quantity which is conserved according to the stochastic equations (12) and (14), and also according to the approximate stochastic equations (12) and (16), is $\Sigma_i(\mu_i^2 + \sigma_{ii})$. The stochastic equations thus provide for a partitioning of the energy of the system (or other quadratic terms) between that which is associated with the ensemble mean (a "specifiable" component, $\Sigma_i \mu_i^2$) and that which is associated with the variances (an "uncertain" component, $\Sigma_i \sigma_{ii}$).

Also note that the approximation (15), involving third moments, implies the derivatives of the individual variances are inexact. However the equation for the rate of changes of $\Sigma_i \sigma_{ii}$ does not contain any third moments. Therefore the approximate equations for the time rate of change of the uncertain component are of the same order of accuracy as those for the means.

4. Monte Carlo solutions

Monte Carlo methods for evaluating integrals have the disadvantage that they tend to require a considerable amount of computer time, but they also have the advantage that one is able to achieve any desired level of accuracy. Monte Carlo methods suggest themselves especially in the present instance because of the multidimensional nature of the integrals (6) and (7) that are to be evaluated. The computing time required for evaluation of integrals by Monte Carlo methods increases much more slowly as the dimension of the problem increases than do, say, quadrature methods.

There exist many sophisticated methods for Monte Carlo calculations (see e.g. Hammersley & Handsomb, 1965) but no effort has been made to exploit these. We will mention only one relatively crude procedure, and some applications of this method will be shown later. The particular technique employed belongs in the category of "importance sampling" and corresponds directly with the notion of an ensemble as the physical analog to our calculations. Discrete initial points in phase space are chosen by a random process such that the likelihood of selecting any given point is proportional to the given initial probability density. For each of these initial points (i.e. for each of the sample selected from the ensemble) deterministic trajectories in phase space are calculated by numerical integration of the particular version of (11) being used. Means and variances are determined, corresponding to specific times, by averaging the appropriate quantities over the sample.

In the applications of this method that we have made the initial probability densities have been taken to be multivariate normal. In some examples there are three parameters, and in others there are eight. In both cases the procedure for selecting the initial points is essentially the same. The first sample parameter (α_1) is selected from a normal distribution with mean μ_1 and variance σ_{11}. The second is selected as a normal variable with mean $\mu_2 + \alpha_1 \sigma_{12}/\sigma_{11}$

and variance $\sigma_{22} - \sigma_{12}^2/\sigma_{11}$. Each successive parameter is chosen as a normal variable with a mean conditional on the specific choices for the preceding parameters and a variance which is the total variance of that parameter (σ_{jj}) times $1 - R^2_{j\cdot 12\ldots(j-1)}$ where $R^2_{j\cdot 12\ldots(j-1)}$ is the multiple correlation coefficient between α_j and all those α's already selected (cf. Mood & Graybill, 1963). A total sample size N means N sets of D parameters, and requires ND references to the program for generating normal variables.

Our use of the Monte Carlo calculations, in what follows, is in the nature of a standard against which one can compare the results of the approximate stochastic calculations. It will be necessary to recognize that the "verification data" are themselves subject to some error. Standard statistical formulas may be used to judge the accuracy of the various Monte Carlo estimates. For example the standard deviations of the means will decrease in proportion to $N^{-\frac{1}{2}}$. However, many different quantities will be evaluated on the basis of each sample and the errors inherent in these quantities will not be independent of one another. This can create some difficulty in the interpretation of these results. Detailed statistical comparisons are avoided for this reason.

The selection of the multivariate normal as the initial probability density function is made because it is convenient and reasonable. In any particular real application the form of the initial probability densities will depend on both the analysis procedure and the error characteristics of the data. In section 6 the particular example used for illustration tends to emphasize the reasonableness of the multivariate normal.

Stochastic solutions of Lorenz's minimum hydrodynamic equations

As a first test of the stochastic methods they have been applied to the "minimum hydrodynamic equations" of Lorenz (1960). Lorenz considered two-dimensional horizontal flow governed by the vorticity equation

$$\frac{\partial}{\partial t}\nabla^2\psi = -\mathbf{k}\cdot\nabla\psi\times\nabla(\nabla^2\psi) \tag{17}$$

where ψ is the streamfunction, ∇ is the usual del-operator in two-dimensional physical space, and $\mathbf{k}$ is the unit vector normal to the plane of the flow. By considering the flow to be periodic over a rectangular region of dimensions $2\pi/k$ by $2\pi/l$, expanding the streamfunction in terms of a double Fourier series and then truncating the series at wave number 1, he was able to represent the vorticity as

$$\begin{aligned}\nabla^2\psi = {} & A_1\cos ly + A_2\cos kx + A_3\sin ly + A_4\sin kx \\ & + A_5\cos kx\cos ly + A_6\sin kx\sin ly \\ & + A_7\cos kx\sin ly + A_8\sin kx\cos ly\end{aligned} \tag{18}$$

where the coefficients $A_1 \ldots A_8$ are functions only of time. If, initially, $A_2 = A_4 = A_5 = A_7 = A_8 = 0$, then this condition, according to (17), must hold for all time. This reduces (18) to

$$\nabla^2\psi = A_1\cos ly + A_2\cos kx + A_6\sin kx\sin ly \tag{19}$$

Applying (17) to (19) one obtains the governing prognostic equations

$$\left.\begin{aligned}\dot{A}_1 &= -\tfrac{1}{2}\frac{1}{\alpha(\alpha^2+1)}A_2A_6 \\ \dot{A}_2 &= \tfrac{1}{2}\frac{\alpha^3}{\alpha^2+1}A_1A_6 \\ \dot{A}_6 &= -\frac{\alpha^2-1}{\alpha}A_1A_2\end{aligned}\right\} \tag{20}$$

where $\alpha = k/l$. These are Lorenz's "minimum equations" and they are a particularly simple three-component version of (11). Lorenz points out that the set (20) conserves both mean kinetic energy

$$E = \frac{1}{4k^2}\left(\alpha^2A_1^2 + A_2^2 + \tfrac{1}{2}\frac{\alpha^2}{1+\alpha^2}A_6^2\right) \tag{21}$$

and mean square vorticity

$$V = \tfrac{1}{2}(A_1^2 + A_2^2 + \tfrac{1}{2}A_6^2) \tag{22}$$

Equations (12) and (16), applied to (20) allow us to write down the approximate stochastic equations for this case:

$$\dot{\mu}_1 = c_1(\mu_2\mu_6 + \sigma_{26}) \tag{23}$$

$$\dot{\mu}_2 = c_2(\mu_1\mu_6 + \sigma_{16}) \tag{24}$$

$$\dot{\mu}_6 = c_3(\mu_1\mu_2 + \sigma_{12}) \tag{25}$$

$$\dot{\sigma}_{11} = 2c_1(\mu_2\sigma_{16} + \mu_6\sigma_{12}) \quad (26)$$

$$\dot{\sigma}_{22} = 2c_2(\mu_1\sigma_{26} + \mu_6\sigma_{12}) \quad (27)$$

$$\dot{\sigma}_{66} = 2c_3(\mu_1\sigma_{26} + \mu_2\sigma_{16}) \quad (28)$$

$$\dot{\sigma}_{12} = c_1(\mu_6\sigma_{22} + \mu_2\sigma_{26}) + c_2(\mu_6\sigma_{11} + \mu_1\sigma_{16}) \quad (29)$$

$$\dot{\sigma}_{16} = c_1(\mu_2\sigma_{66} + \mu_6\sigma_{26}) + c_3(\mu_2\sigma_{11} + \mu_1\sigma_{12}) \quad (30)$$

$$\dot{\sigma}_{26} = c_2(\mu_1\sigma_{66} + \mu_6\sigma_{16}) + c_3(\mu_1\sigma_{22} + \mu_2\sigma_{12}) \quad (31)$$

where $c_1 = -[2\alpha(1+\alpha^2)]^{-1}$, $c_2 = \frac{1}{2}\alpha^3(\alpha^2+1)^{-1}$, $c_3 = -(\alpha^2-1)/\alpha$, and the μ's and σ's are the means and covariances of the A's.

Note that (23)–(31) reduce to (20) if and only if the initial variances are zero. Even if the covariances which appear specifically in (23)–(25) are initially zero, they will not remain so, because of (29)–(31), unless the variances are also zero. If the variances are zero, then all the covariances must be zero ($\sigma_{ij}^2 \leqslant \sigma_{ii}\sigma_{jj}$) and (26)–(28) will indicate no change.

Although no explicit condition has been stated to the effect that $\sigma_{ij}^2 \leqslant \sigma_{ii}\sigma_{jj}$, or that $\sigma_{ii} \geqslant 0$, these are necessary conditions for $((\sigma_{ij}))$ to be considered a covariance matrix. In general the covariance matrix must be positive definite. The approximate stochastic equations (12) and (16), or in this example (23)–(31), will always maintain that condition if the initial conditions are properly chosen. There is always the possibility that the numerical solution, however, will violate that condition. It is necessary to take some care, therefore, to avoid both instabilities and excessive amplifications in the numerical solutions. A Runge–Kutta four-point integration scheme has been used. Regular checks on the solution (i.e. evaluation of the eigenvalues of the covariance matrix) have been made to assure that the matrix does indeed remain positive definite. Difficulties arose in an early stage of the study when the Heun method (as used by Lorenz, 1963) was employed. This experience corroborates the conclusions of Young (1968) regarding the choice of time differencing scheme.

The first situation studied by Lorenz was one in which there is a cyclical exchange of energy between the "zonal flow" (A_1) and the "eddies" (A_2 and A_6). He chose $\alpha = 2$ and initial conditions $A_1 = 0.12$, $A_2 = 0.24$, $A_6 = 0.00$. These coefficients have the dimension (time)$^{-1}$; Lorenz considered that the time unit was 3 hr and also scaled his problem by taking $2\pi/k = 5000$ km. For ease of discussion and comparison we will adopt the same scaling. For the stochastic solutions the initial conditions on the A's can only represent initial ensemble means. Thus, initially $\mu_1 = 0.12$, $\mu_2 = 0.24$ and $\mu_6 = 0.00$. Let us now assume also initial values $\sigma_{ij} = 0.0 (i \neq j)$ and $\sigma_{ii} = 10^{-4}(3\text{ hr})^{-2} (i = 1, 2, 6)$, i.e. that the initial parameter values are uncorrelated, but each has a standard deviation of 0.01 (3 hr)$^{-1}$. This corresponds to the maximum expected "zonal" wind of 65 km/hr having a standard deviation of 5.3 km/hr. As indicated earlier, for the Monte Carlo solution it is further assumed that the initial probability density is multivariate normal.

Fig. 1 shows the stochastic solutions for μ_2 and μ_6 and the deterministic solutions for A_2 and A_6. On the scale of the graphs it is not possible to distinguish between the approximate stochastic solution and the Monte Carlo solution which was based on a sample size $N = 500$. Partial tabulation of the solutions by the three methods is given in Table 1. Note that the difference between the two stochastic solutions is nowhere more than about 2.3 standard errors of estimate of the Monte Carlo mean $(\sigma_{ii}/N)^{\frac{1}{2}}$.

The differences between the deterministic and stochastic solutions become discernable only after about 50 hours, and remain small throughout the 150-hour period for which solutions were obtained. As one could have anticipated the initial uncertainties have the effect of reducing the amplitudes of the variations (Epstein, 1969).

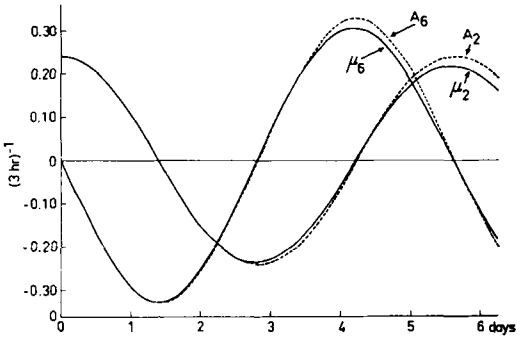

Fig. 1. Stochastic solutions for μ_2 and μ_6 and the deterministic solution for A_2 and A_6. Initial conditions for the deterministic solution are $A_1 = 0.12$, $A_2 = 0.24$, $A_6 = 0$. For the stochastic solutions all covariances are initially zero and the variances are each 10^{-4}. The Monte Carlo and approximate stochastic solutions essentially overlap.

Table 1. *Selected results of calculations of Lorenz's minimum equation*

For the Monte Carlo experiment the sample size was $N = 500$

t (hrs)	A_1 Deter-min-istic	$\mu_1 \pm \sigma_{11}^{\frac{1}{2}}$ Monte Carlo	$\mu_1 \pm \sigma_{11}^{\frac{1}{2}}$ Approx. stoch.	A_2 Deter-min-istic	$\mu_2 \pm \sigma_{22}^{\frac{1}{2}}$ Monte Carlo	$\mu_2 \pm \sigma_{22}^{\frac{1}{2}}$ Approx. stoch.	A_6 Deter-min-istic	$A_6 \pm \sigma_{66}^{\frac{1}{2}}$ Monte Carlo	$A_6 \pm \sigma_{66}^{\frac{1}{2}}$ Approx. stoch.
0	.120	.120 ± .010	.120 ± .010	.240	.240 ± .010	.240 ± .010	.000	.000 ± .011	.000 ± .010
12	.124	.124 ± .010	.124 ± .010	.210	.207 ± .011	.207 ± .011	−.167	−.167 ± .018	−.166 ± .017
24	.131	.132 ± .010	.131 ± .010	.111	.110 ± .021	.110 ± .021	−.291	−.291 ± .020	−.290 ± .019
36	.134	.134 ± .009	.134 ± .009	−.025	−.026 ± .033	−.025 ± .032	−.327	−.324 ± .015	−.324 ± .014
48	.129	.129 ± .007	.128 ± .007	−.151	−.151 ± .033	−.149 ± .032	−.255	−.250 ± .036	−.252 ± .036
60	.122	.123 ± .008	.123 ± .008	−.226	−.223 ± .020	−.222 ± .019	−.110	−.104 ± .062	−.107 ± .061
72	.121	.122 ± .011	.121 ± .011	−.236	−.228 ± .018	−.229 ± .014	.061	.065 ± .077	.060 ± .077
84	.126	.127 ± .013	.126 ± .013	−.180	−.168 ± .048	−.171 ± .047	.218	.214 ± .072	.210 ± .072
96	.133	.132 ± .012	.132 ± .012	−.065	−.056 ± .077	−.061 ± .080	.316	.299 ± .042	.297 ± .034
108	.133	.132 ± .008	.132 ± .006	.073	.072 ± .084	+.067 ± .087	.313	.288 ± .050	.290 ± .039
120	.126	.127 ± .006	.127 ± .004	.185	.172 ± .065	.169 ± .063	.210	.187 ± .099	.193 ± .101
132	.120	.123 ± .009	.123 ± .008	.237	.216 ± .037	.215 ± .019	.050	.037 ± .131	.045 ± .138
144	.122	.124 ± .013	.124 ± .014	.223	.195 ± .056	.196 ± .043	−.120	−.117 ± .133	−.110 ± .142

Plots of $\sigma_{22}^{\frac{1}{2}}$ and $\sigma_{66}^{\frac{1}{2}}$ are shown in Fig. 2. Here the two stochastic solutions give almost indistinguishable results over the first 50 hours. Thereafter some systematic differences begin to appear and grow. One would of course not expect the approximate stochastic solution to represent the variances quite as well as it does the means, since the approximation (15), while second order in the variances, is third order in the means.

In this first example, the initial uncertainties were relatively small, i.e. small enough that they did not have severe effects on the forecasts of the parameters. In cases such as this the utility of the stochastic procedure lies in the information it provides about the uncertainties. For example, after 100 hours the uncertainty of A_1 is not very different from its initial value, but the uncertainty of A_6 has increased more than threefold and that of A_2 is larger by a factor of 8. There is also information provided by the stochastic calculations on the correlations among the parameters. For example, after about 18 hours, and continuing until beyond 100 hours, A_1 and A_6 have a correlation coefficient less than -0.9.

This comparison of the deterministic and stochastic results is not necessarily typical. Let us now consider another situation discussed by Lorenz (1960), that of a large initial disturbance which gives rise to large oscillations in the zonal wind. If $A_6^2 > 30\, A_1^2$, such oscillations can occur

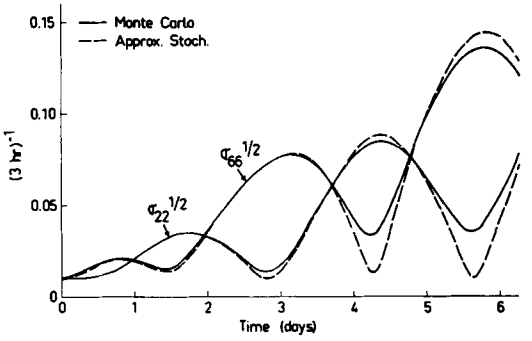

Fig. 2. Stochastic dynamic solutions for the standard deviations of A_2 and A_6. Initial conditions are as for Fig. 1.

and the sign of A_1 can alternate. For initial conditions, then, we now choose $\mu_1 = 0.12 \pm 0.02$, $\mu_2 = 0.00 \pm 0.02$ and $\mu_6 = 0.666 \pm 0.02$, again with all covariances initially zero. For the deterministic solution initial values are $A_1 = 0.12$, $A_2 = 0.00$, $A_6 = 0.666$ so that $A_6^2/A_1^2 = 30.8$ and indeed, as shown in Fig. 3, after about three days the sign of A_1 does become negative. The stochastic solutions for μ_1 shown in the same figure remain positive throughout the period. The differences between the two stochastic solutions are certainly real after day 3, i.e. attributable to the approximations (15). Yet the indicated behavior of the expected mean zonal wind is reasonably good through the fifth

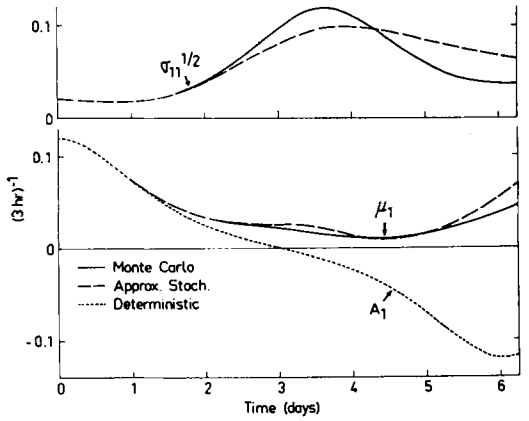

Fig. 3. Stochastic solutions for μ_1 and $\sigma_{11}^{\frac{1}{2}}$ and the deterministic solution for A_1. Initially $A_1(\mu_1) = 0.12$, $A_2(\mu_2) = 0.$, and $A_6(\mu_6) = 0.666$. For the stochastic solutions covariances were initially zero and variances were initially 4×10^{-4}.

day. The deterministic solution shows systematic departures from the Monte Carlo solution after 24 hours and is very misleading after about 3 days. Solutions for $A_2(\mu_2)$ and $A_6(\mu_6)$ shown in Fig. 4, tend to emphasize these differences.

Note that this does not mean the deterministic forecast is *wrong*, only that it is a *poor* forecast. It is possible that the deterministic solution would be verified in a given situation[1] but the

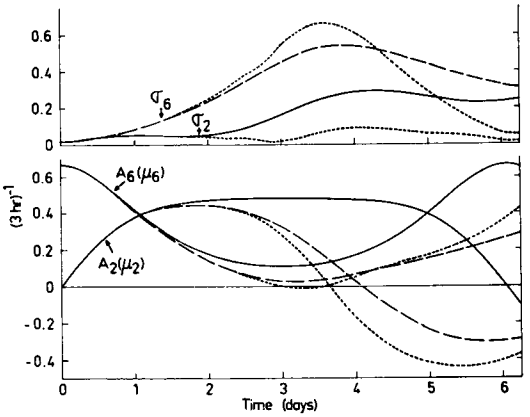

Fig. 4. Stochastic solutions (dashed lines) for μ_2, μ_6, $\sigma_{22}^{\frac{1}{2}}$, and $\sigma_{66}^{\frac{1}{2}}$, and the deterministic solution (solid lines) for A_2 and A_6. Initial conditions were as for Fig. 3. Shorter dashes refer to the approximate stochastic solution.

[1] In all examples based on Lorenz's minimum equations $d\varphi/dt = 0$. Therefore the deterministic solution will remain the modal solution if it started as such. In this case large differences develop between stochastic solutions would have better average verification scores.

Figs. 3 and 4 also show the stochastic solutions for the standard deviations ($\sigma_i = \sigma_{ii}^{\frac{1}{2}}$) of the three parameters. After three days, for example, the differences between the deterministic solution and the mean of the ensemble are about one standard deviation for A_6. It is apparent from Figs. 3 and 4 that the usefulness of the approximate stochastic results to indicate the individual variances diminishes very rapidly here after two days. This is the poorest comparison of the approximate stochastic and Monte Carlo solutions that has been calculated. The difficulties encountered here, due of course to the special circumstances chosen, however, also emphasize the advantages of the stochastic solutions over the deterministic solution for the parameter values themselves.

5. An extension including an example of analysis

The situations just discussed are peculiar in the sense that the model, (17) and (18), contains eight parameters, which were artificially divided into two groups: A_1, A_2 and A_6 were admitted to be subject to uncertainty, while the others were assumed known with certainty. A more realistic application of the stochastic method to this case should include the possible variability of all eight parameters. Indeed it is possible to assume the "true" streamfunction is given by Lorenz's simplifying initial conditions, but at the same time require that the "forecast" be initialized by making observations of the streamfunction.

Let us therefore consider that n observing stations are spread over the region $0 \leqslant y \leqslant 2\pi/l$, $0 \leqslant x \leqslant 2\pi/k$, and that at each an observation Y_i is made. The observations, it will be assumed, are random variables with expected values $E(Y_i) = \psi(x_i, y_i)$ where

$$\psi(x, y) = -\frac{A_1}{l^2}\cos ly - \frac{A_2}{k^2}\cos kx - \frac{A_3}{l^2}\sin ly$$

$$-\frac{A_4}{k^2}\sin kx - \frac{A_5}{k^2 + l^2}\cos kx \cos ly$$

the mode and the mean. Therefore the distributions of A_2, A_1 and A_6 must be highly skewed.

$$-\frac{A_6}{k^2+l^2}\sin kx \sin ly - \frac{A_7}{k^2+l^2}\cos kx$$

$$\times \sin ly - \frac{A_8}{k^2+l^2}\sin kx \cos ly \qquad (32)$$

and (x_i, y_i) are the coordinates of the ith observing station. We will assume also that each observation is subject to the same variance, s^2. It then becomes convenient (and in a statistical sense optimum) to obtain estimates $\hat{A}_i$ of the parameters A_i by least squares; this procedure also provides estimates of the covariances among the $\hat{A}_i$.

Defining

$$z_{i1} = -\frac{1}{l^2}\cos ly_i$$

$$z_{i2} = -\frac{1}{k^2}\cos kx_i$$

$$z_{i3} = -\frac{1}{l^2}\sin ly_i, \quad \text{etc.}$$

and writing in matrix form $\mathbf{Z} = ((z_{ij}))$ we have

$$\hat{\mathbf{A}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y} \qquad (33)$$

and

$$\text{cov}\ (\hat{\mathbf{A}}) = (\mathbf{Z}'\mathbf{Z})^{-1}s^2 \qquad (34)$$

(Mood & Graybill, 1963).

Note in particular that in general the matrix $\mathbf{Z}$ is not orthogonal (unless all observations are made at specific grid points) and that the covariance matrix will therefore not be diagonal. The variances and also the covariances depend on the locations of the observing stations, and also on the model (32), but not on the specific initial conditions.

The $\hat{A}_i$ are random variables whose values depend on the observations (themselves random variables) and the unknown true values of the parameters A_i. The required starting point for a stochastic prediction is a probability statement about the A_i. The formal statistical procedure by which the sampling distribution of the $\hat{A}_i$ given A_i can be related to a probability density distribution for the A_i (given the data) is by applying Bayes' theorem (cf. Epstein, 1966).

Avoiding details, there are two inputs to the evaluation of $\varphi(\mathbf{A}\,|\,\hat{\mathbf{A}})$: the sampling distribution of the $\hat{A}_i$, and a "prior" probability density of the A_i. This latter, in the meteorological context, is the result of the previous day's forecast. Thus it is natural and consistent with stochastic methods to use the previous day's forecast as direct input into today's analysis. Here, however, no artificial weights or arbitrary coefficients are necessary. In a sense the relative "weights" of the observations and forecast are self-determining and are measurable in terms of the relative uncertainties, i.e. the variances of the two density functions.

In the present example, to keep things as simple as possible, we will assume that the uncertainty in the forecast is very large compared to the uncertainty in the estimates based on the observations. Then we can write $\mu_i = E(A_i) = \hat{A}_i$ and $\sigma_{ij} = \text{cov}\ (\mathbf{A}) = \text{cov}\ (\hat{\mathbf{A}})$.

For the present experiment a random set of $n = 20$ observing stations has been selected. Their locations are plotted in Fig. 5. The covariance matrix (34) corresponding to this array of stations, and also the implied correlation coefficients, are given in Table 2. For these calculations we have chosen $k(km^{-1}) = 2\pi/5000$, $\alpha = 2$, and $s^2k^4 = 0.004\ (3\ \text{hr})^{-2}$. This last figure corresponds, if one relates ψ to a geostrophic stream function gz/f, to a standard error of about 35 m in the determination of the height of a pressure surface in middle latitudes. Note that changing α, or k, or s changes the variances and covariances in Table 2, but not the correla-

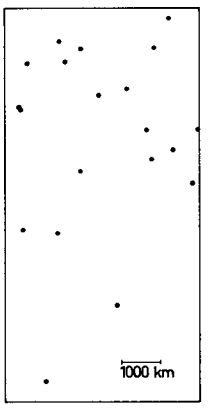

Fig. 5. Location of the 20 stations as determined by a random process such that all locations were equally likely for each station. The shape and scale of the region are determined by $2\pi/k = 5000$ km, $\alpha = 2$.

Table 2. *Variances (along the diagonal), covariances (above the diagonal), and correlation coefficients (below the diagonal) among the least square estimates of the parameters, based on the array of stations shown in Fig. 5, and with* $s^2k^4 = 0.004$ (3 hr)$^{-2}$

Units of the covariances and variances are (3 hr)$^{-2}$

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
A_1	$.3792 \cdot 10^{-4}$	$-.1550 \cdot 10^{-4}$	$.2775 \cdot 10^{-5}$	$.3068 \cdot 10^{-5}$	$-.1126 \cdot 10^{-3}$	$.1199 \cdot 10^{-3}$	$-.7669 \cdot 10^{-4}$	$-.6526 \cdot 10^{-4}$
A_2	$-.078$	$.1040 \cdot 10^{-2}$	$.3498 \cdot 10^{-4}$	$-.3085 \cdot 10^{-3}$	$.4583 \cdot 10^{-3}$	$-.7365 \cdot 10^{-3}$	$.1237 \cdot 10^{-2}$	$.2199 \cdot 10^{-3}$
A_3	.086	.208	$.2720 \cdot 10^{-4}$	$-.1752 \cdot 10^{-4}$	$-.1590 \cdot 10^{-4}$	$-.1080 \cdot 10^{-4}$	$.2961 \cdot 10^{-4}$	$.4389 \cdot 10^{-4}$
A_4	.178	$-.342$	$-.103$	$.7840 \cdot 10^{-3}$	$-.1513 \cdot 10^{-3}$	$.1139 \cdot 10^{-2}$	$-.6330 \cdot 10^{-3}$	$-.2051 \cdot 10^{-3}$
A_5	$-.359$	.279	$-.060$	$-.106$	$.2591 \cdot 10^{-2}$	$-.3776 \cdot 10^{-3}$	$.6912 \cdot 10^{-3}$	$.6525 \cdot 10^{-3}$
A_6	.338	$-.396$	$-.036$	.705	$-.129$	$.3327 \cdot 10^{-2}$	$-.1594 \cdot 10^{-2}$	$-.2029 \cdot 10^{-3}$
A_7	$-.237$	.730	.108	$-.430$	.258	$-.526$	$.2759 \cdot 10^{-2}$	$.3846 \cdot 10^{-3}$
A_8	$-.286$	.184	.227	$-.198$	.346	$-.095$	.198	$.1369 \cdot 10^{-2}$

tion coefficients. Based on the particular station locations that were selected (and Fig. 5 does resemble a map of aerological stations) some of the estimates of the initial parameter values will have considerable correlation. For example, if A_4 is overestimated, one would also expect a positive error in the estimate of A_6.

Returning to the first situation considered in the previous section, the initial values will be taken as $\mu_1 = 0.12$, $\mu_2 = 0.24$, $\mu_j = 0.00$ $(j > 2)$, as before, but with covariances as given in Table 2. With eight parameters the approximate stochastic equations are a set of 44 simultaneous equations similar to (23)–(31), and readily derived from (12), (16), (17) and (18). A Monte Carlo solution has also been obtained, with again the added assumption that the initial parameter estimates are multivariate normal. The last assumption would have been implied by the analysis procedure if the observations had been assumed to be Gaussian; the converse does not follow. The calculated initial conditions for the Monte Carlo experiment, based on a sample size of 1000, are given in Table 3. This should be compared with Table 2. The Monte Carlo run required about 30 times as much computer time as the approximate stochastic calculation, although neither program was designed to be particularly efficient.

We can examine here only a portion of the numerical results, trying to select those few quantities that provide the greatest insight into the various attributes of the results. Table 4 contains the calculated mean values and standard deviations of the eight parameters after 2, 4 and 6 days. These results are reasonably typical of those at other intermediate time steps. The differences between the two stochastic solutions remain within the statistical uncertainty of the Monte Carlo results. For example at $t - t_0 = 144$ hrs, the standard deviation of the estimate of μ_6, based on the Monte Carlo results, would be $0.104/\sqrt{1000} \cong 0.003$, while the difference between the two stochastic estimates of μ_6 is 0.004.

It is of some interest that the stochastic solutions for those parameters whose values are identically zero in the deterministic solution remain very close to zero. A situation where this is not the case will be discussed later. Also, here, the indicated changes in the variances of these terms are small, but these changes are similarly indicated by both methods.

Table 3. *Initial variances (along diagonal), covariances (above diagonal), and correlation coefficients (below diagonal) calculated from the sample of N = 1000 used in the Monte Carlo experiment*

Units of the variances and covariances are (3 hr)$^{-2}$

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
A_1	$.3850 \cdot 10^{-4}$	$-.1251 \cdot 10^{-4}$	$.2131 \cdot 10^{-5}$	$.2341 \cdot 10^{-4}$	$-.1071 \cdot 10^{-3}$	$.1042 \cdot 10^{-3}$	$-.7235 \cdot 10^{-4}$	$-.5802 \cdot 10^{-4}$
A_2	$-.062$	$.1059 \cdot 10^{-2}$	$.4207 \cdot 10^{-4}$	$-.3444 \cdot 10^{-3}$	$.5478 \cdot 10^{-3}$	$-.7947 \cdot 10^{-3}$	$.1244 \cdot 10^{-2}$	$.2469 \cdot 10^{-3}$
A_3	$.064$	$.241$	$.2878 \cdot 10^{-4}$	$-.1509 \cdot 10^{-4}$	$-.1517 \cdot 10^{-4}$	$.2409 \cdot 10^{-5}$	$.3538 \cdot 10^{-4}$	$.5497 \cdot 10^{-4}$
A_4	$.137$	$-.384$	$-.102$	$.7604 \cdot 10^{-3}$	$-.1445 \cdot 10^{-3}$	$.1098 \cdot 10^{-2}$	$-.6835 \cdot 10^{-3}$	$-.1867 \cdot 10^{-3}$
A_5	$-.334$	$.326$	$-.055$	$-.101$	$.2675 \cdot 10^{-2}$	$-.3138 \cdot 10^{-3}$	$.6920 \cdot 10^{-3}$	$.6640 \cdot 10^{-3}$
A_6	$.289$	$-.421$	$.008$	$.686$	$-.104$	$.3368 \cdot 10^{-2}$	$-.1699 \cdot 10^{-2}$	$-.1301 \cdot 10^{-3}$
A_7	$-.223$	$.731$	$.126$	$-.473$	$.255$	$-.559$	$.2745 \cdot 10^{-2}$	$.3460 \cdot 10^{-3}$
A_8	$-.258$	$.209$	$.282$	$-.187$	$.354$	$-.062$	$.182$	$.1316 \cdot 10^{-2}$

Table 4. *Selected results from solutions with all eight parameters free to vary*

For the stochastic calculations the results are expressed as $\mu_i \pm \sigma_i$ and the uncertain component of the kinetic energy is expressed as a percentage of the total kinetic energy

	48 hrs			96 hrs			144 hrs		
	Deter-ministic	Monte Carlo	Approx. stoch.	Deter-ministic	Monte Carlo	Approx. stoch.	Deter-ministic	Monte Carlo	Approx. stoch.
A_1	.129	$.128 \pm .006$	$.128 \pm .006$	.133	$.133 \pm .008$	$.133 \pm .008$	.122	$.123 \pm .009$	$.123 \pm .009$
A_2	$-.151$	$-.155 \pm .052$	$-.153 \pm .051$	$-.065$	$-.057 \pm .059$	$-.059 \pm .056$	.223	$.213 \pm .036$	$.213 \pm .031$
A_3	.000	$.001 \pm .006$	$.001 \pm .006$	.000	$.002 \pm .008$	$.002 \pm .008$	.000	$-.000 \pm .006$	$-.000 \pm .005$
A_4	.000	$.002 \pm .029$	$.001 \pm .030$	.000	$-.001 \pm .026$	$-.000 \pm .026$	.000	$-.002 \pm .026$	$-.001 \pm .026$
A_5	.000	$-.001 \pm .052$	$.000 \pm .051$	.000	$-.001 \pm .051$	$.000 \pm .051$	.000	$-.002 \pm .051$	$-.001 \pm .051$
A_6	$-.255$	$-.250 \pm .037$	$-.251 \pm .037$	.316	$312 \pm .055$	$.311 \pm .056$	$-.120$	$-.128 \pm .104$	$-.124 \pm .104$
A_7	.000	$.002 \pm .056$	$-.001 \pm .056$	.000	$.009 \pm .063$	$.005 \pm .062$	.000	$.004 \pm .054$	$.001 \pm .053$
A_8	.000	$.000 \pm .033$	$-.000 \pm .034$	.000	$-.003 \pm .039$	$-.002 \pm .040$	.000	$.003 \pm .039$	$.002 \pm .040$
Uncertain KE		5.89 %	5.81 %		7.50 %	7.23 %		7.88 %	7.57 %

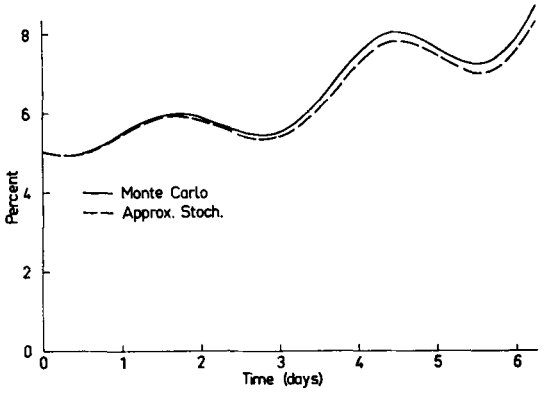

Fig. 6. Stochastic dynamic evaluations of the uncertain component of the kinetic energy assuming observations made at the stations of Fig. 4 and initial mean values $\mu_1 = 0.12$, $\mu_2 = 0.24$, $\mu = 0$ for $j > 2$. The ordinate is expressed as a fraction of the constant total kinetic energy.

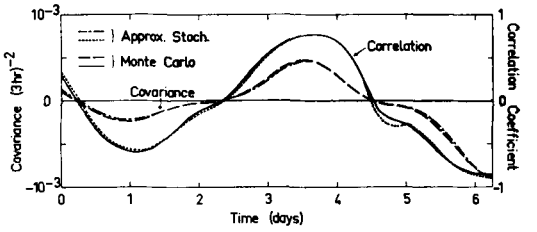

Fig. 7. Calculated correlation coefficient and covariance between A_2 and A_6 for the situation of Fig. 6.

It was mentioned earlier that the stochastic procedures would conserve the total mean kinetic energy, but separate that total into two components. Fig. 6 illustrates the changes in the uncertain component, as given by the two stochastic methods.

The general behaviour of these curves is reminiscent of, and undoubtedly typical of the variance of cyclical parameters of somewhat uncertain phase, amplitude and frequency, as described by Epstein (1969). Note in particular the general increasing trend. The amount by which these curves exceed their value at t_0 must be equal to the amount by which the deterministic solution overspecifies the kinetic energy of the predictable portion of the flow.

The covariance between A_1 and A_6 serves as a final example of the numerical results for this situation. This pair was chosen because the parameters themselves undergo significant fluctuations, and also because the correlation between them in this case becomes relatively large. Fig. 7 shows both the covariance and the calculated correlation coefficients. Although initially A_1 and A_6 are only weakly correlated (cf. Tables 3 and 2) their covariance undergoes successively more extreme fluctuations and the correlation coefficient reaches extremes outside the range ± 0.75.

Before leaving the subject of analysis entirely, it is worth remarking that the procedure sketched above for including the earlier forecast in the analysis pertains as well to the inclusion of asynoptic data or almost any form of supplemental information. It is only necessary to translate the new data into a likelihood function (i.e. the probability of the data given the atmospheric parameters) and to modify through Bayes' theorem the probability density otherwise determined for the appropriate time.

6. The value of more observational data

One of the most pressing problems facing meteorologists is the value of new data, or the cost of not having enough (Gandin *et al.*, 1967). On the one hand weather services juggle budgets to achieve compromises between observing stations and computers or other facilities. On the other planning proceeds for a World Weather Watch involving costly additional or novel observing capability. Stochastic methods can be used to assess the importance and value of new data. We will indicate here how the method can contribute to judgements concerning the optimum location of the new observing stations.

This discussion will be based on the premise that, within the context of a particular mathematical–physical model, the best forecast is the one with the least uncertainty. The uncertainty in a forecast will depend, in general, on the model and also on the data, the analysis and forecast procedures, and the particular meteorological situation. For the present, attention will be confined to the particular model, the particular analysis and forecasting procedure, and also the assumed initial conditions of the previous section. The questions to be studied are what improvement results from the addition of a single observation, and how does that depend on the specific location of the added station.

Adding a single observation at any specific location adds a row to the matrix **Z** and may

change all the elements of $\mathbf{S} = (\mathbf{Z}'\mathbf{Z})^{-1}$, which is proportional to the initial covariance matrix. In particular, though, the new $\mathbf{S}$ must always have diagonal elements which are less than or equal to those of the original $\mathbf{S}$. The variances (and consequently the uncertain component of the conserved quadratic quantities) will always be at least as small as they were originally.

Given the original 20 stations shown in Fig. 5, calculations have been made of the percentage reduction in the uncertain component of the initial kinetic energy for each of 49 possible additional observing sites placed on a regular grid over the region. Isopleths drawn from these data are shown in Fig. 8. The resulting pattern depends on both the analysis model (32) and the particular set of previously established stations, but not on the meteorological situation. Note that the greatest reduction in uncertainty is achieved by adding stations where there were "holes" in the previous network, and also that the isopleth pattern reflects the inclusion in the model of only wave numbers 0 and 1.

The 49 additional stations provided 49 different new covariance matrices with which to initialize the computations. Using the same set of initial mean values for the parameters as in section 6, approximate stochastic computations were carried out to produce 150-hour forecasts based on each of these initial covariance matrices. The forecasted expected values of the parameters were almost identical with one another,

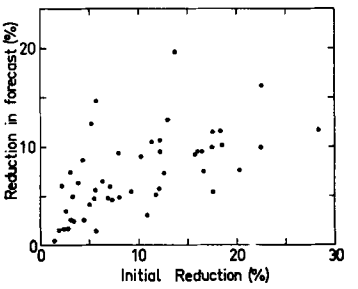

Fig. 9. Decrease in the uncertain component of kinetic energy after 150-hour forecast versus initial decrease. Each point represents the result for a different location of a single additional observation. Decreases are measured as the percent of the uncertain component of KE initially and after the 150-hour forecast, respectively, with only the original 20 observations.

and with the original run. However, changes in the variances were more noticable. Fig. 9 is a scatter diagram between the reduction in the uncertain component of mean kinetic energy due to each additional station initially, and after 150 hours. The pattern of the reduction in uncertainty, according to the location of the added station, is shown in Fig. 10.

There is clearly a correlation between the improvement in the initial conditions and the improvement in the forecast, but judging by the scatter in Fig. 9 that is far from the entire story. The shift in patterns between Figs. 8 and 10 is certainly dependent on the initial flow field

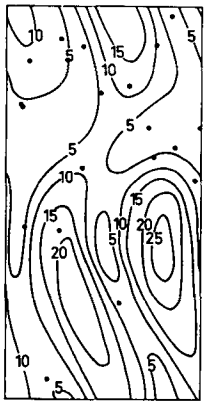

Fig. 8. Percent decrease in the initial uncertain component of the kinetic energy due to the addition of a single observation as it depends on the location of the observation. Locations of the original 20 stations are also shown.

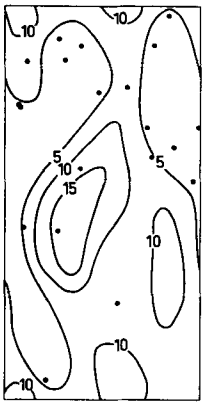

Fig. 10. Percent decrease in the uncertain component of the kinetic energy after 150-hour forecast, attributable to a single additional observation, as a function of the location of the observation. Locations of the original 20 stations are also shown.

and its relation to the locations of the pre-existing observations. No simple, fully consistent explanation of the particular changes is apparent.

If one were considering adding a station to a network, one would certainly not base his decision on the results from a single experiment involving one initial streamfunction. Instead one would try to maximize the improvement averaged over a wide range of initializations. I would speculate that if this averaging were carried out over an ensemble of streamfunctions that was isotropic and homogeneous, then the indicated best new location would be that which minimizes the initial uncertainty, in which case it would appear that the experiment might as well never have been carried out. If the averaging were over an ensemble which resembled a climatology of weather patterns, however, I then believe that there would be definite advantages to some locations other than those which look best in terms of initial uncertainty.

As is evident in Fig. 9, each additional station resulted in at least some improvement in the forecast. This need not be the case. The very peculiar situation can arise where an additional observation, although it must reduce the initial uncertainty, also changes the initial covariance matrix so as to produce, corresponding to some future time, a forecast which is more uncertain than it would have been if the observation had not been available. This, however, is an exceptional and transient phenomena. With that same additional station, but most other forecast periods, some improvement will be found.

This discussion has been intended to be exemplary of the approach that one might take to the question of the value of adding data. A similar approach could also provide a meaningful answer to the somewhat simpler, but nevertheless important question of the value of simply improving the present observations by reducing their individual errors. Also, one could develop comparable programs, in terms of more complex models, for comparing the benefits of various forms of novel measurements of winds, temperatures, densities, etc.

7. Prediction of unstable flow

The prediction of the development of instabilities rests on the recognition of a situation in which a small perturbation, if present, will grow at the expense of pre-existing "steady" flow. There are physical-numerical models now available which can simulate the non-linear interactions implicit in the growth and development of perturbations. Many different types of instability have been recognized. If conditions are ripe for instabilities to develop, however, it may not be possible to identify observationally the incipient disturbances. The prediction then becomes highly speculative and the specific forecast can depend very critically on random initial errors.

Stochastic methods cannot predict a development if there is insufficient initial information. They will, however, allow one to judge whether or not there is sufficient information, and they will indicate the amount of confidence, or lack thereof, one may place in a forecast made under conditions of instability.

The model for this study will again be (17) and (18), but this time we choose $\alpha = k/l = 0.95$. As discussed by Lorenz (1960) $\alpha < 1$ implies that small perturbations will grow at the expense of the zonal flow. He showed a solution in which a small perturbation grew to reach a maximum value after about 24 time units (3 days), at which time the zonal flow had decreased to zero and was beginning to reverse itself.

For this experiment we choose the same set of stations as previously, but a slightly different initial covariance matrix since $l = \alpha k$ is different. Initial mean conditions are taken to be $\mu_1 = 1.0$; $\mu_i = 0$, $i > 1$. Approximate stochastic and Monte Carlo ($N = 1000$) solutions have been calculated. A deterministic solution, corresponding to the specific initial mean conditions of the Monte Carlo sample has also been calculated. (The determininistic solution for the case $A_i = 0$, $i > 1$ is $A_i =$ constant for all i.) The forecasts of the A_i and μ_i are summarized in Table 5. Plots of the solutions for μ_1, μ_3 and μ_6 are shown in Fig. 11.

The instability is reflected in all three solutions. However, because the initial mean perturbations of the Monte Carlo solution were so small, the deterministic solution shows no appreciable decay in A_1 until the fourth day. By that time the stochastic solutions both clearly indicate an expected reversal of the mean zonal flow. It is evident from a comparison of these results that the dominating terms in the forecast of the means should be those representing the uncertainties, and not the small initial mean perturbations. Compare, for example the solutions for

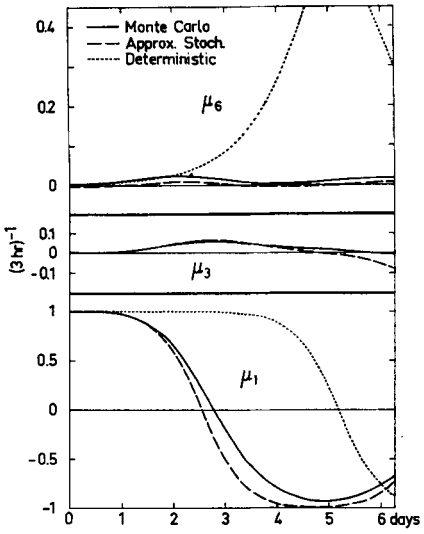

Fig. 11. Stochastic and deterministic solutions for selected parameters under conditions of unstable flow. The initial values for the deterministic solution were the same as the initial Monte Carlo sample mean and slightly different from the mean values for the approximate stochastic solution. The deterministic solution for A_3 is indistinguishable from zero. Note differences in scale.

$A_2(\mu_2)$ and $A_6(\mu_6)$. Since the largest initial mean perturbation is that of A_6, it grows most rapidly in the deterministic solution, while A_3 stays near zero. After 24 hours, however, the largest mean perturbation indicated by the stochastic solutions is μ_3. The Monte Carlo result for μ_6 is similar to the deterministic result up to 24 hours but this is deceptive. Some of the indicated increase must indeed be due to the initial mean perturbation, but much of it is the result of the uncertainties. Note how the continued growth of A_6 indicated by the deterministic results is not substantiated by the stochastic results.

The deterministic solution must at all times make a very specific allocation of the energy. On the other hand the stochastic solutions are able, in terms of the initial uncertainties and their subsequent developments, to indicate that only some of the energy can be specifically allocated to particular waves, while some is the uncertain component of the kinetic energy. This is shown in Fig. 12, in which the uncertain component of the kinetic energy, indicated by the two stochastic calculations, is plotted. Note the very large extent of the oscillation of this term.

Table 5. *Calculations of the coefficients (or their expected values) (units: $(3hr)^{-1}$) by the deterministic, Monte Carlo (with $N = 1000$), and approximate stochastic methods*

In this calculation $\alpha = 0.95$.

	Forecast period											
	$t = 0$			$t = 48$ hrs			$t = 96$ hrs			$t = 144$ hrs		
	Deterministic	Monte Carlo	Approx. stoch.	Deterministic	Monte Carlo	Approx. stoch.	Deterministic	Monte Carlo	Approx. stoch.	Deterministic	Monte Carlo	Approx. stoch.
$\mu_1(A_1)$	1.0006	1.0006	1.0000	.9997	.6495	.5807	.8937	−.7963	−.9535	−.7608	−.7567	−.7205
$\mu_2(A_2)$	.0009	.0009	.0000	.0362	.0328	.0128	.3933	.0366	−.0052	.5681	−.0249	−.0133
$\mu_3(A_3)$	.0002	.0002	.0000	.0001	.0414	.0473	−.0005	−.0031	.0309	−.0012	.0019	−.0791
$\mu_4(A_4)$	.0006	.0006	.0000	.0093	.0056	−.0055	.1009	.0126	.0010	.1457	.0092	.0318
$\mu_5(A_5)$	.0004	.0004	.0000	.0004	−.0001	−.0010	.0004	.0008	−.0020	.0008	.0025	−.0014
$\mu_6(A_6)$	.0037	.0037	.0000	.0247	.0223	.0083	.2654	.0036	−.0032	.3834	.0175	.0074
$\mu_7(A_7)$	.0011	.0011	.0000	.0011	.0012	.0005	.0011	−.0036	.0007	−.0005	−.0060	.0015
$\mu_8(A_8)$	.0007	.0007	.0000	.0063	.0036	−.0037	.0681	.0058	−.0024	.0984	−.0074	−.0213

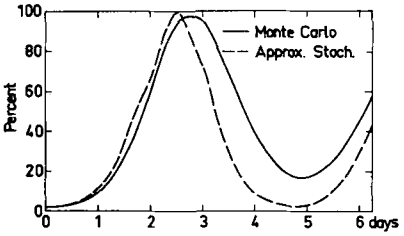

Fig 12. Stochastic calculations of the uncertain component of KE for the situation of unstable flow.

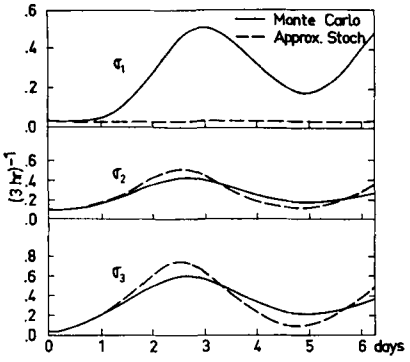

Fig. 13. Stochastic calculations of the standard deviations of some of the parameters for the situation of unstable flow. (Note differences in scale of the ordinates.)

The stochastic calculations indicate that after about three days, when the expected magnitude of the zonal component is least, it cannot be said how the energy that has been released will appear. A couple of days later, however, the uncertain component is much less because there is a relatively high probability that the zonal flow will be large and negative.

The estimates of the standard deviations of several of the terms are shown in Fig. 13. The discrepancy between the two results for σ_1 is relatively large (although further exaggerated by the scales of the graphs). Here again is a situation in which the approximations (15) have deficiencies which show up relatively early in the solution. Even so, in this case it has already been evident that approximate solution represents the expected flow very well, the allocation of the energy into specifiable and uncertain components with only slightly less fidelity, and even many of the individual variance terms are remarkably well represented.

8. Stochastic prediction and predictability

Considerable attention has been given, in recent years, to the question of atmospheric predictability (Charney *et al.*, 1966; Robinson, 1967; Lorenz, 1965, 1968*a*, 1969). To a large degree this interest is spurred by practical needs vis à vis forecasting: do the prospects for extended (and more accurate) forecasts justify expensive new observing programs. On the other hand predictability is an important and intriguing scientific problem, in its own right, and is deserving of even more attention. However, it has not been a very tractable problem, which explains why more has not been done in this area.

Because of its unique attribute of dealing specifically with the uncertainty of the forecast, stochastic dynamic prediction makes the investigation of predictability far more tractable. It is possible even now to outline a reasonable approach, based on the stochastic method, which should be a significant improvement on methods previously attempted. Also, we will point out, there is a hidden bias in the most recent estimates of predictability, a bias which only stochastic methods can remedy.

Consider an experimental program of extended stochastic forecasts based on a set of initial conditions carefully chosen to reflect climatological ranges and frequencies of observed states. This would lead directly, not only to estimates of the growth of uncertainty (dependent of course on the assumed observing program) on the average, but also it would reveal the extent to which predictability is itself variable and dependent on the initial state of the atmosphere. This last point has been raised previously by Lorenz (1965), but it has never really been possible to study it systematically. It may turn out to be a significant aspect of predictability studies.

This approach lies, in some respects, midway between the approaches of Charney *et al.* (1966) and Lorenz (1969). Lorenz (1968*b*) refers to the former as the "dynamical" approach and to the latter as "dynamical-empirical". In the dynamical approach one studies the rate of growth of small initial differences on the basis of extended integration with advanced models of the general circulation. One then extrapolates this "error" growth to the time when the resultant states will be sufficiently dissimilar from one

another. In the dynamical-empirical method Lorenz derives an equation for the growth of average "error kinetic energy". This average is taken over an ensemble of error fields (which one has to assume have certain restrictive statistical properties) and also over an ensemble of initial flow fields. The statistical representation of the flow fields is based on the atmosphere's observed spectral characteristics, but it is also necessary to assume these fields are homogeneous and isotropic. While these assumptions probably do not affect Lorenz's most important conclusion, that the growth of uncertainties from small scales to large may ultimately place a finite limit on the range of predictability of the atmosphere, they may seriously influence his estimates of that limit as applied to the atmosphere.

The dynamical–empirical method is very general, and consequently requires sweeping assumptions. On the other hand the dynamical method is so specific to the two chosen initial states that the generalization of its results is subject to question. The stochastic dynamic method is in this sense a compromise. Averaging over error fields is intrinsic to the method, but the statistics of the error fields can be very general. Averaging over initial states can be done systematically, with full regard to the heterogeneity and anisotropy of the atmosphere.

One attribute of the stochastic dynamic results already presented bears significantly on this problem. In Fig. 6 and especially in Fig. 12 it was shown that the "uncertain component" of the kinetic energy underwent significant oscillations. The uncertain component of kinetic energy, as defined earlier, is entirely analogous to Lorenz's (1969) error kinetic energy, except that Lorenz considered an average over many initial states. If we were also to consider such an average it seems most likely that the oscillations in the individual curves would tend to cancel one another and the result would be a monotonically increasing function of time.

In the specific examples shown, however, it would be unreasonable to define predictability only in terms of current uncertainty. Certainly, during those periods when the uncertainty exceeds a value it will later attain, it would be much more reasonable to define predictability in terms of that eventual minimum value. Without being concerned with specific definitions, however, one can clearly see that in each individual case predictability will sometimes be greater, and never less than that indicated by the current uncertainty. Therefore the *average* uncertainty implies an underestimation of the predictability. Entirely on this basis one must judge that Lorenz's estimate of a limiting range of predictability for the atmosphere of about 2½ weeks is pessimistic.

One other point would seem pertinent here. When dealing with predictability one tacitly assumes the forecast procedure is the best possible. Deterministic procedures, coupled with realizable sets of observations are not optimum. Only by the use of stochastic procedures, which are designed to minimize the uncertainty, can one hope to approach the limits indicated by studies of predictability.

9. Summary and conclusions

The use of real observational data as the basis of forecasts requires that special care be taken in the interpretation and analysis of the data and the formulation of the prediction procedure. The state of the atmosphere can never be known with certainty, and by taking specific account of this uncertainty improvements in the forecasts can be achieved. One way of dealing with this uncertainty is to add simple correction terms to the deterministic prognostic equations, and thereby convert them to stochastic dynamic equations which predict the expected state of the atmosphere with the smallest possible mean square error. In order to evaluate the correction terms additional prognostic equations must be derived.

The total set of prognostic equations which are finally derived, the approximate stochastic equations, involve assumptions pertaining to the statistical nature of the uncertainties. Solutions of these have been compared, in a number of examples, with results of both the deterministic equations and Monte Carlo experiments to simulate the initial uncertainty of the initial conditions. Within the framework of the relatively simple dynamic systems studied, the results indicate that the approximate stochastic equations are excellent for forecasting the expected state of the atmosphere, and also give very good results in defining the amount and the nature of the uncertainty of the predictions.

Some of the applications of the method are

more obvious than others. Stochastic dynamic prediction can undoubtedly improve forecasts, and improvement in forecasts implies extended ranges for useful forecasts. Knowledge of the uncertainty of the forecast in each instance not only increases the value of the forecast, but also suggests tools for generating extended forecasts expressed in stochastic terms, and for studying the predictability of the atmosphere, not only in a highly idealized average sense, but also in terms of the specific initial situation. In this context, the notion of initial situation includes both the atmospheric state and the ways and means by which it is observed. This allows one to calculate how changes in the system for observing the atmosphere can influence predictability. Examples of this type of study are shown. Although idealized, they demonstrate that maximum reduction of the uncertainty in the analysis is definitely not the same criterion as minimum uncertainty in the forecast.

There are other advantages of the stochastic dynamic method which have been touched upon only lightly, for example the ease with which analysis can incorporate not only the previous forecast, but also additional asynoptic data. Neither has much attention been given to the added facility with which user requirements of weather forecasts can be met given specific information on uncertainties.

We can also point out that stochastic dynamic methods can be useful in studying and evaluating particular physical–numerical models of the atmosphere. A major difficulty in the past has been the isolation of the separate influences on verification data of errors in the initial data, errors in the prediction procedure itself, and errors in the verification data. However, stochastic dynamic methods permit a methodologically sound comparison of stochastically expressed forecasts with stochastically expressed verification analyses. Biases and distortions attributable to the model are estimatable by ordinary statistical procedures.

We have detected a single major drawback to the stochastic dynamic method. It requires much more computing than deterministic methods. This follows since one calculates not only the expected value of each parameter, but also the variances and covariances among all the parameters. It is possible that in more complex models than have been studied so far many of these terms will not change very much (or very rapidly). This would be useful in reducing somewhat the burden of the computations. Other possibilities for reducing the amount of computing also exist, but even at best the task will be formidable. For this reason it does not seem stochastic methods can now be given serious consideration as operational procedures. On the other hand we very firmly believe that they must be studied further and put into use as important tools of atmospheric research and in the development of improved weather analysis and prediction.

Acknowledgements

The author is very grateful to the University of Stockholm and especially the Meteorological Institute for providing support, facilities and the opportunity to pursue these studies. Professor Bert Bolin not only offered a good deal of encouragement, but also made the original suggestion to use Lorenz's minimum equations as an initial test of the method. Much benefit has also been derived from conversations with Dr Bo Döös and others on the MISU staff.

REFERENCES

Baer, F. 1964. Integration with the spectral vorticity equation. *J. Atm. Sci. 21*, 260–276.

Bryan, K. 1959. A numerical investigation of certain features of the general circulation. *Tellus 11*, 163–174.

Charney, J. G. *et al.* 1966. The feasibility of a global observation and analysis experiment. *Bull. Amer. Meteor. Soc. 47*, 200–220.

Ellsaesser, H. W. 1966. Evaluation of spectral versus grid methods of hemispheric weather prediction. *J. Appl. Meteor. 5*, 246–262.

Epstein, E. S. 1966. Quality control for probability forecasts. *Mon. Wea. Rev. 94*, 487–494.

Epstein, E. S. 1969. The role of initial uncertainties in prediction. *J. Appl. Meteor. 8*, 190–198.

Freiberger, W. F. & Grenander, U. 1965. On the formulation of statistical meteorology. *Review of the International Statistical Institute 33*, 59–86.

Gandin, L. S., Mashkovich, S. A., Alaka, M. A. & Lewis, F. 1967. Design of optimum networks for aerological observing stations. *World Wea. Watch Planning Report No. 21*, WMO Secretariat, Geneva.

Gleeson, T. A. 1966. A causal relation for probabilities in synoptic meteorology. *J. Appl. Meteor. 5*, 365–368.

Gleeson, T. A. 1967. On theoretical limits of predictability. *J. Appl. Meteor. 6*, 355–359.

Gleeson, T. A. 1968. A modern physical basis for meteorological prediction. *Proc. of the First Nat'l Conf. on Statistical Meteor.*, Amer. Meteor. Soc., 1–10.

Hamersley, J. M. & Handscomb, D. C. 1964. *Monte Carlo Methods*, 178 pp. Methuen and Co., London.

Lorenz, E. N. 1960. Maximum simplification of the dynamic equations. *Tellus 12*, 243–254.

Lorenz, E. N. 1963. Deterministic non-periodic flow. *J. Atmos. Sci. 20*, 130–141.

Lorenz, E. N. 1965. Study of the predictability of a 28-variable atmospheric model. *Tellus 17*, 321–333.

Lorenz, E. N. 1968*a*. On the range of atmospheric predictability. *Proc. of the First Nat. Conf. on Statistical Meteorology, Amer. Meteor. Soc.*, 11–19.

Lorenz, E. N. 1968*b*. Three approaches to atmospheric predictability. Abstract. *Bull. Amer. Meteor. Soc. 49*, 1041.

Lorenz, E. N. 1969. The predictability of a flow which possesses many scales of motion. *Tellus 21*, 289–308.

Mood, A. M. & Graybill, F. A. 1963. *Introduction to the theory of statistics*, 2nd ed., 443 pp. McGraw-Hill, New York.

Novikov, E. A. 1959. Contribution to the problem of the predictability of synoptic processes. *Bull. Acad. Sci. USSR, Geophysics Series* (English ed., AGU), 1209–1211.

Platzman, G. W. 1960. The spectral form of the vorticity equation. *J. Meteor. 17*, 635–644.

Robert, A. J. 1966. The integration of a low order spectral form of the primitive meteorological equations. *J. Meteor. Soc. of Japan 44*, 237–245.

Robert, A. J. 1968. Integration of a spectral barotropic model from global 500-mb charts. *Mon. Wea. Rev. 96*, 73–85.

Robinson, G. D. 1967. Some current projects for global meteorological observation and experimentation. *Q. Jour. Roy. Meteor. Soc. 93*, 409–418.

Saltzman, B. 1962. Finite amplitude free convection as an initial value problem, pt 1. *J. Atmos. Sci. 19*, 329–341.

Thompson, P. D. 1957. Uncertainty of initial state as a factor in the predictability of large scale atmospheric flow patterns. *Tellus 9*, 275–295.

Young, J. A. 1968. Comparative properties of some time differencing schemes for linear and nonlinear oscillations. *Mon. Wea. Rev. 96*, 357–364.

СТОХАСТИЧЕСКОЕ ДИНАМИЧЕСКОЕ ПРЕДСКАЗАНИЕ

Стохастическое динамическое предсказание предполагает, что законы, управляющие поведением атмосферы, полностью детерминированы, но ищет решения, соответствующие определенным статистическим характеристикам начальных условий, учитывая тем самым невозможность точных или достаточно плотных наблюдений. Уравнение, которое должно быть решено, является уравнением неразрывности для вероятности. По практическим причинам возможны только приближенные решения этого уравнения. Детерминированные прогнозы обнаруживают очень низкий порядок аппроксимации. Развиваются более точные методы. Некоторые свойства и преимущества стохастических динамических предсказаний иллюстрируются на пример системы динамических уравнений низкого порядка.

Стохастические динамические предсказания обладают значительно меньшими средними квадратическими ошибками, чем детерминированные методы, а также дают специальную информацию о природе и пределах неопределенности прогноза. Кроме того, расширяется интервал времени, для которого прогноз может быть полезным. Однако, эти методы требуют значительно большего объема вычислений.

Рассматривается вопрос об анализе начального состояния атмосферы, и здесь можно обнаружить значительные преимущества над настоящими детерминированными методами. Показано, как статистические методы могут быть использованы для оценки новых или улучшенных данных путем рассмотрения их влияния на уменьшение неопределенности прогноза. Путем применения стохастических методов более эффективно проводится сравнение различных физико-математических моделей. Наконец, кратко рассматривается на основе стохастического динамического предсказания вопрос о предсказуемости и делается заключение, что некоторые ранние оценки слишком пессимистичны.