AN EXPERIMENT IN MULTIVARIATE ERROR ANALYSIS AND LEAST-SQUARES PRINCIPLES USING NUMERICALLY GENERATED DATA

by

SHUN DER KO

B.S.E., National Cheng Kung University, 1966

M.S., National Central University, 1969

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September, 1977

Signature of Author_____
Department of Meteorology, August 25, 1977

Certified by_____
Thesis Supervisor

Accepted by_____
Chairman, Departmental Committee on Graduate Students

AN EXPERIMENT IN MULTIVARIATE ERROR ANALYSIS AND LEAST-SQUARES PRINCIPLES USING NUMERICALLY GENERATED DATA

by

Shun Der Ko

Submitted to the Department of Meteorology on 25 August 1977
in partial fulfillment of the requirements for the
degree of Master of Science

ABSTRACT

Least-squares principles applied to prediction and objective analyses are discussed. In prediction analysis we investigate the linear, linear weighted and nonlinear methods. Error weights have been taken into account in the linear weighted method. Numerical data are generated by a set of nonlinear equations. Averaged reductions of variance are used to test the advantages of all the schemes. The goodness of the prediction formula depends on the data interval involved but is nearly independent of data length ranging from 1500 to 15000 time steps.

Multivariate and univariate methods for objective analysis are studied by using data generated from an 8 4-cycle scheme. We find the former is superior to the latter. A general theory of multivariate error analysis is also introduced and tested with numerical data.

Thesis Supervisor : Edward N. Lorenz

Title : Professor of Meteorology

ACKNOWLEDGEMENTS

I wish to express my deep gratitude to Professor Edward N. Lorenz for his invaluable encourgement and guidance during the course of this study.

I am indebted to Mrs. Diana Spiegel for the computer programming help. Special thanks are due to Mr. Ron Errico, Mr. Larry Marx and Mr. Ross Hoffman for several helpful, stimulating discussions and sincere help.

I would like to appreciate my host family, Professor F. Sanders and his wife, for their kind help and cheerful receptions during my stay at M. I. T.

My financial support at M. I. T. came from the fellowship of National Science Council of the Republic of China. Appreciations are extended to the President of Taipei Institute of Technology and Director of Institute of Physics, Academia Sinica for making available for the generous grant of special leave for my graduate study at M. I. T.

Finally I am grateful to my wife Lihseih and daughter Chei for their patience and moral support.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

1. INTRODUCTION

Weather prediction has, in recent years, been a fundamental problem of considerable attention. Although many improvements have been made, weather prediction still fails to attain a satisfactory accuracy. This failure can be ascribed to some inherent causes (Lorenz, 1965, 1969). In weather forecasting two basic methods are usually favored; the statistical and dynamical methods. In the statistical forecasting problem linear regression methods are often involved. Lorenz(1962) investigated the linear regression methods by a set of deterministic nonlinear differential equations. He suggested that linear regression methods give excellent forecasts one day ahead, but only mediocre forecasts more than three days in advance. He(1977) reexamined the linear scheme with real weather data by introducing suitable nonlinear functions of the original predictors as additional predictors and found the improvements due to the nonlinearity.

The dynamical methods are based on solving a system of governing equations. Since the governing equations are nonlinear it is hardly possible thus far to get exact solutions mathematically. Even if these equations could be solved strictly, perfect prediction would not be attainable in pratice because the governing equations are not perfectly known. One of the dynamical methods called numerical weather prediction requires solving the governing finite difference equations on a grid lattice. It is necessary to perform a process of interpolating observed values from unevenly distributed observation

points to regularly spaced grid-point values which are, to some degree, representative of the actual atmospheric state. This process has been referred to as objective analysis. Many analysis schemes require a rather subjective pre-specification of weight factors in the interpolation process.

Cressman(1959) suggested a bell-shaped weighting function which depends on the distance between the grid and the observation points. Gandin(1963) and Eddy(1964) indicated that weighting functions for objective analysis should be obtained through the autocorrelation procedures. Kruger(1969) found that the best relative results were obtained when the weight factors were based on the autocovariances of observed minus trial values. Thiebaux(1973, 1974) treated, extending the idea of Gandin's optimal interpolation method, the meteorological parameters as vector-valued variables and proposed a multivariate procedure by taking into account the covariances of observed variables. Schlatter(1975, 1976) applied the multivariate schemes to wind and geopotential height fields. He stated that a distinct advantage of a multivariate approach is that a height analysis based upon observed height and wind is significantly better than one based upon observed height alone. He(1976) also found univariate schemes fit the data as well as the multivariate schemes when the observations are plentiful, but forecasts based on the latter are consistently better. Therefore it seems that the goodness of the schemes depends on the amounts of observations and distribution of data points.

In this study we attempt to reexamine whether the multivariate analysis is only by chance better than the univariate one. This investigation is based on numerical data generated by a set of six nonlinear equations which are related to larger system of equations developed by Lorenz(1963c). Since the calculations of weight factors in both the linear regression and multivariate methods are based on the same principles - minimization of the error variance - we shall discuss both schemes under the same title of least-squares principles.

It is certainly true that errors exit in any observations or measurements owing to the limiting accuracy of measuring instruments or some unavoidable conditions. In general, error is defined as the difference between the observed or calculated value and the true value. However, in many cases we really do not know what the true value is, especially in meteorology due to the nonperiodic properties(Lorenz, 1963b) and complicated mechanisms in atmospheric systems. But we often do know approximately what it should be either from the theoretical approaches or from the earlier observations. Such approximation can be used as an indication of whether the result is of the right order of magnitude. In many meteorological research problems, we usually forecast or calculate a number of parameters simultaneously from an equal or different number of observation data through some physical laws or equations which are not exactly applicable to the complex behaviors of the atmosphere. In these cases to determine how much

confidence we can have in our predictions or calculations we need some systematic ways to estimate the errors of observations and how they will be propagated to the final parameters. These ways come under the heading of multivariate error analysis.

2. BASIC CONCEPTS

The least-squares principles are useful for extracting information from a set of observations or data points. The principles are best in the sense that the parameters are normally distributed about the true parameters with least possible variance or standard deviation. If the observations and parameters are linearly related the least-squares method is the best analytical technique, but for the majority of problems, which are nonlinear, the method of least-squares refinement suffers from a number of disadvantages.

Essential to the error analysis are the concepts of variance V, covariance C, correlation coefficient r and standard deviation σ. Variance is a measure of the dispersion of the observations about the mean value. The mean value is usually referred as the true value in such a case when the latter is not available. Instead of variance a more convenient measure of the dispersion of observations is the standard deviation which is defined as the square root of the variance. In the error analysis the covariance is a measure of the way in which the errors in two quantities vary together. The covariance is zero when the two quantities are physically independent. However zero covariance does not ensure independence. If the covariance is positive and

high, it implies that the factor which is causing one quantity to be assigned, say, a value higher than the true value is also causing the other quantity to be assigned too high a value. In such case covariance is therefore an important factor in error analysis, especially in a many-parameter problem, which will be incorrect if it is neglected. The correlation coefficient r between two quantities x_i and x_j is defined in terms of covariance C and standard deviations of these two quantities, i.e.,

$$r(x_i,x_j) = C(x_i,x_j)/\sigma(x_i)\ \sigma(x_j) \qquad (1)$$

The correlation coefficient refers to the extent to which the two errors are correlated. It greatly affects the way in which errors propagate. It must lie in the range of -1 to +1. If the measurements are connected physically in some way, the correlation coefficient may be nonzero. And it will be zero for two independently measured quantities.

3. LEAST-SQUARES PRINCIPLES

In general the error variations in the measurements are not all equal because of different observation conditions or requirements. Thus it seems more reasonable to assume errors in the measurements of different variables have different weights. However in research problems many an author usually uses less sophisticated methods by neglecting the weighting factor. This neglect might causes the analysis to be incorrect. Therefore in this section some more general principles will be discussed by taking the error weight into cccount. The principle of least-squares states that it is the weighted sum of the

squares of the difference between the observed and the calculated(or predicted) values that must be minimized to obtain the best estimates of the parameters. The appropriate weighting method is to divide each element in the sum by a number proportional to the variance of the measurement from which it was calculated or predicted. The weight matrix W for related variables X is reasonably defined as the inverse of the variance-covariance matrix(hereafter denoted VC matrix), M_X, which can be obtained by using the definitions of variance and covariance, i.e.,

$$M_X = \frac{1}{m} \sum_{i=1}^{m} {}_iX' \; {}_iX'^T = \langle X' \; X'^T \rangle \tag{2}$$

or

$$M_X = \begin{bmatrix} V(x_1), \; C(x_1,x_2), \; \ldots, \; C(x_1,x_n) \\ C(x_2,x_1), \\ \ldots \\ \ldots \\ C(x_n,x_1), \; \ldots \qquad\qquad V(x_n) \end{bmatrix} \tag{3}$$

where left-hand subscript i indicates the ith set measurements, m total number of sets of measurements, superscript T the transpose of matrix, $\langle \; \rangle$ mean, and

$$X' = X - \langle X \rangle \tag{4}$$

$$X = (x_1, \; x_2, \; \ldots, \; x_n)^T \tag{5}$$

$$\langle X \rangle = (\langle x_1 \rangle , \; \langle x_2 \rangle , \; \ldots, \; \langle x_n \rangle)^T \tag{6}$$

$${}_iX' = {}_iX - \langle {}_iX \rangle \tag{7}$$

Since $C(x_i,x_j) = C(x_j,x_i)$, M_X is a symmetric matrix. If all the measurements are independent then M_X becomes a diagonal matrix. The VC matrix can also written in terms of standard deviation and correlation coefficient by using Eq.(1)

$$M_X = \begin{bmatrix} \sigma^2(x_1), & \sigma(x_1)\,\sigma(x_2)\,r(x_1,x_2), & \ldots, & \sigma(x_1)\,\sigma(x_n)\,r(x_1,x_n) \\ \sigma(x_2)\,\sigma(x_1)\,r(x_2,x_1), & \ldots & & \\ \ldots & & & \\ \ldots & & & \\ \sigma(x_n)\,\sigma(x_1)\,r(x_n,x_1), & \ldots & & \sigma^2(x_n) \end{bmatrix} \tag{8}$$

To simplify the notation we define a diagonal matrix of standard deviations S_x and a matrix of correlation coefficients R_x, whose diagonal elements are unity,

$$S_X = \begin{bmatrix} \sigma(x_1), & 0, & \ldots & \ldots, 0 \\ 0, & \sigma(x_2), & & \\ \ldots & & & \\ \ldots & & & \\ 0, & \ldots & & \sigma(x_n) \end{bmatrix} \tag{9}$$

$$R_X = \begin{bmatrix} 1, & r(x_1,x_2), & \ldots, & r(x_1,x_n) \\ r(x_2,x_1), & 1, & & \\ \ldots & & & \\ \ldots & & & \\ r(x_n,x_1), & \ldots & & 1 \end{bmatrix} \tag{10}$$

Then $M_X = S_X\, R_X\, S_X$. (11)

Suppose E is column matrix of the difference between the observed and calculated values. Then the sum G and weighted sum G_w to be minimized becomes respectively,

$$G = E^T E \tag{12}$$

$$G_w = E^T W E \tag{13}$$

where

$$E = (e(x_1), e(x_2), \ldots, e(x_n))^T \tag{14}$$

$$e(x_i) = x_i \text{ (observed)} - x_i \text{ (predicted)} \tag{15}$$

and

$$W = M_X^{-1}. \tag{16}$$

3-1. PREDICTION ANALYSIS

In this section we shall discuss the prediction analysis by linear and nonlinear methods.

3-1-1. LINEAR METHOD

Suppose there are p different-variable predictands and m predictors. If the predictands depend linearly on the predictors, and an additional constant predictor x_1 whose value is always unity is included(Lorenz, 1973, 1977), then the equations relating each predictand ${}_kY$ to m + 1 predictors can be given by

$${}_kY \text{ (predicted)} = {}_kB\, X \qquad k=1,2, \ldots, p \tag{17}$$

where

$${}_kY = ({}_ky_1, {}_ky_2, \ldots, {}_ky_n) \tag{18}$$

$$ {}_kB = ({}_kb_1,\ {}_kb_2,\ \cdots,\ {}_kb_{m+1}) \tag{19} $$

$$ X = \begin{bmatrix} {}_1x_1,\ {}_1x_2,\ \cdots,\ {}_1x_n \\ {}_2x_1,\ {}_2x_2,\ \cdots,\ {}_2x_n \\ \cdots \\ {}_{m+1}x_1,\ {}_{m+1}x_2,\ \cdots,\ {}_{m+1}x_n \end{bmatrix} \tag{20} $$

$$ {}_1x_i = 1, \quad i = 1, 2, \ldots, n \tag{21} $$

n = number of observations

That is, ${}_kY$, ${}_kB$ and x are 1 x n, 1 x (m+1) and (m+1) x n matrices respectively. To make best estimate of ${}_kB$, the mean-squares error

$$ {}_kG = {}_kE\ {}_kE^T \qquad k=1, 2, \ldots, p \tag{22} $$

should be minimized with respect to the elements of ${}_kB$. From Eqs. (17) and (22) we get

$$ {}_kG = ({}_kY - {}_kB\ X)\ ({}_kY - {}_kB\ X)^T \qquad k=1, 2, \ldots, p \tag{23} $$

To minimize ${}_kG$, $\partial_k G / \partial_k B_i$ or $\partial_k G / \partial_k b_i$ must be zero for all values of i

$$ \frac{\partial_k G}{\partial_k B_i} = -{}_iX\ ({}_kY - {}_kB\ X)^T + \ ({}_kY - {}_kB\ X)\ (-{}_iX)^T \tag{24} $$

$$ = 0 \qquad k = 1, 2, \ldots, p \qquad i = 1, 2, \ldots, m + 1 $$

where ${}_kB_i$ is the ith column of ${}_kB$, ${}_iX$ is an 1 x n matrix containing the elements of ith row of X. Since, on the right-hand side of Eq.(24), the transpose of the first term is equal to the second term, they are identical and are both equal to zero, i.e.,

$$({}_kY - {}_kB\ X)\ {}_iX^T = 0 \qquad k=1, 2,..., p \qquad i = 1, 2,..., m + 1 \tag{25}$$

Putting the (m + 1) terms on the left-hand side of Eq.(25) into a row matrix and setting it equal to zero, i.e.,

$$({}_kY - {}_kB\ X)\ ({}_1X^T, {}_2X^T, \ldots, {}_{m+1}X^T) = 0 \tag{26}$$

or
$$({}_kY - {}_kB\ X)\ X^T = 0 \qquad k = 1, 2, \ldots, p \tag{27}$$

that is,

$${}_kB\ X\ X^T = {}_kY\ X^T \qquad k =1, 2,...., p$$

or
$${}_kB = ({}_kY\ X^T)\ (X\ X^T)^{-1} \qquad k = 1, 2, \ldots, p \tag{28}$$

Similarly we can put the p's ${}_kB$ into a column matrix we obtain

$$B = Y\ X^T\ (X\ X^T)^{-1} \tag{29}$$

where
$$B = ({}_1B, {}_2B, \ldots, {}_pB)^T \tag{30}$$

$$Y = ({}_1Y, {}_2Y, \ldots, {}_pY)^T \tag{31}$$

Then Y, B and X are p x n, p x (m+1) and (m+1) x n matrices respectively. With the help of Eq.(29), the predictands can be predicted by

$$Y = B\,X\ . \tag{32}$$

On the other hand, we can estimate Y by taking into account the error weight. In this case we rewrite Eq.(32) as

$$X = A\,Y \tag{33}$$

where A is the inverse of B if p = n. In the estimation we minimize, with respect to Y, the weighted mean-squares error G_W given by

$$G_W = (X - A\,T)^T\,W\,(X - A\,Y) \tag{34}$$

To minimize G_W, we differentiate G_W with respect to the elements of Y and set it equal to zero. After some manipulation we get

$$Y = (A^T\,W\,A)^{1}\ \ A^T W\,X \tag{35}$$

which will be called weighted linear method in this study.

3-1-2. NONLINEAR METHOD

If the predictands Y depend nonlinearly on the predictors, we may assume the deviations of predictands and predictors from their respective means are so small that linear dependence is also valid, that is,

$$Y' = B'X' \tag{36}$$

where

$$Y' = Y - \langle Y \rangle \tag{37}$$

$$X' = X - \langle X \rangle \tag{38}$$

$$X = ({}_1X,\ {}_2X,\ \ldots,\ {}_{m+1}X)^T \tag{39}$$

$${}_kX = ({}_kx_1,\ {}_kx_2,\ \ldots,\ {}_kx_n) \qquad k = 1,\ 2,\ \ldots,\ m+1 \tag{40}$$

$$Y = ({}_1Y,\ {}_2Y,\ \ldots,\ {}_pY)^T \tag{41}$$

$${}_kY = ({}_ky_1,\ {}_ky_2,\ \ldots,\ {}_ky_n) \qquad k = 1,\ 2,\ \ldots,\ m+1 \tag{42}$$

Following the same principles as that described in the above section we obtain B'

$$B' = Y'\ X'^T(X'\ X'^T)^{-1}\ . \tag{43}$$

3-2. OBJECTIVE ANALYSIS

Univariate and multivariate methods will be investigated for the linear case in this study. The former is used to estimate the grid-point individual variable by the same kind of variables obtained at the nearby observation points. The latter is used to produce grid-point estimates through the autocovariance and covariance of the observed variables.

3-2-1. UNIVARIATE METHOD

In this case the equations used are the same as those in prediction analysis. Suppose there are m observation points around a grid point and p different variables at each point. If the estimated variables

at grid point Y depend linearly on the observed variables, then

$$Y(\text{estimated}) = B\ X \tag{44}$$

where

$$Y = ({}_1Y,\ {}_2Y,\ \ldots,\ {}_pY)^T \qquad (p \times n) \tag{45}$$

$${}_kY = ({}_ky_1,\ {}_ky_2,\ \ldots,\ {}_ky_n) \qquad k = 1,\ 2,\ \ldots,\ p \tag{46}$$

$$B = ({}_1B, {}_2B,\ \ldots,\ {}_pB)^T \qquad (p \times (m+1)) \tag{47}$$

$${}_kB = ({}_kb_1,\ {}_kb_2,\ \ldots,\ {}_kb_{m+1}) \qquad k = 1,\ 2,\ \ldots,\ p \tag{48}$$

$$X = \begin{bmatrix} {}_1x_1, & {}_1x_2, & \cdots, & {}_1x_n \\ {}_2x_1, & {}_2x_2, & \cdots, & {}_2x_n \\ \cdots, & & & \\ {}_{m+1}x, & {}_{m+1}x_2, & \cdots, & {}_{m+1}x_n \end{bmatrix} \qquad ((m+1) \times n) \tag{49}$$

$${}_1x_i = 1, \qquad i = 1,\ 2,\ \ldots,\ n \tag{50}$$

n = number of observation

Although the p different variables are combined into a compact matrix equation, they are treated independently. Following the arguments given in section 3-3-1, we have

$$B = Y\ X^T\ (X\ X^T)^{-1}. \tag{51}$$

3-2-2. MULTIVARIATE METHOD

First we assume the parameters at grid points are linearly dependent on the nearby observed values. If there are m sets of observations and p different variables at each observation point, then the equation relating the grid-point estimates Y to the nearby observations X is

$$Y = B\,X \tag{52}$$

$$Y = ({}_1Y,\ {}_2Y,\ \ldots,\ {}_pY)^T \qquad (p \times n) \tag{53}$$

$${}_kY = ({}_ky_1\ ,\ {}_ky_2,\ \ldots,\ {}_ky_n) \qquad k = 1, 2, \ldots, p \quad (1 \times n) \tag{54}$$

n = number of observations

$$B = ({}_1B,\ {}_2B,\ \ldots,\ {}_mB) \qquad (p \times mp) \tag{55}$$

$${}_kB = {}_kb_{ij} \qquad i, j = 1, 2, \ldots, p;\ k = 1, 2, \ldots, m \qquad (p \times p) \tag{56}$$

$$X = ({}_1X^T,\ {}_2X^T,\ \ldots,\ {}_mX^T)^T \qquad (mp \times n) \tag{57}$$

$${}_kX^T = ({}_kX_1,\ {}_kX_2,\ \ldots,\ {}_kX_p)^T, \qquad k = 1, 2, \ldots, m \quad (p \times 1) \tag{58}$$

$${}_kX_i = ({}_kx_{i1},\ {}_kx_{i2},\ \ldots,\ {}_kx_{in}) \qquad i = 1, 2, \ldots, p \quad (1 \times n)$$

Since the form of Eq.(55) is the same as that of Eq.(17) we can follow the similar procedures and get

$$B = Y\,X^T\,(X\,X^T)^{-1}\ . \tag{59}$$

4. MULTIVARIATE ERROR ANALYSIS

4-1. ERROR CALCULATION AND PROPAGATION

It is inherent that small deviations from the true value will be obtained in any observation or measurement. Therefore error analysis is important in scientific problems. As mention above, the standard deviation describes the dispersion of observations and the correlation coefficient affects the way in which errors propagate. Therefore it is an essential work in error analysis to investigate the standard deviation and correlation coefficient of measurements and of parameters.

Assume that there are measurements of n different variables and p parameters. The errors of measurement X and of parameters Y are defined as the deviation of any measurement x_i and of any parameter y_j from their corresponding mean $\langle x_i \rangle$ and $\langle y_j \rangle$, that is

$$x_i' = x_i - \langle x_i \rangle \quad , \qquad i = 1, 2, \ldots, n \tag{60}$$

$$y_j' = y_j - \langle y_j \rangle \quad , \qquad j = 1, 2, \ldots, p \tag{61}$$

From Taylor expansion, neglecting the highorder terms (linear assumption), the above expressions can be written as follows:

$$x_i' = \sum_{j=1}^{p} \left(\partial x_i / \partial y_j \right) y_j' \quad , \qquad i = 1, 2, \ldots, n \tag{62}$$

$$y'_j = \sum_{i=1}^{n} (\partial y_j / \partial x_i)\, x'_i , \qquad j = 1, 2, \ldots, p \tag{63}$$

The linear assumption is based on the validity of Taylor expansion of each parameter around the mean values of the measurements and moreover on the fact that the linear terms in the expansions are the most significant in the error region. To simplify the notation and manipulation we use the matrix forms,

$$X' = A' Y' \tag{64}$$

$$Y' = B' X' \tag{65}$$

where

$$X' = X - \langle X \rangle \tag{66}$$

$$Y' = Y - \langle Y \rangle \tag{67}$$

$$X' = (x'_1, x'_2, \ldots, x'_n)^T \tag{68}$$

$$Y' = (y'_1, y'_2, \ldots, y'_p)^T \tag{69}$$

$$X = (x_1, x_2, \ldots, x_n)^T \tag{70}$$

$$Y = (y_1, y_2, \ldots, y_p)^T \tag{71}$$

$$\langle X \rangle = (\langle x_1 \rangle, \langle x_2 \rangle, \ldots, \langle x_n \rangle)^T \tag{72}$$

$$\langle Y \rangle = (\langle y_1 \rangle, \langle y_2 \rangle, \ldots, \langle y_p \rangle)^T \tag{73}$$

$$A' = \begin{bmatrix} \frac{\partial x_1}{\partial y_1}, \frac{\partial x_1}{\partial y_2}, \ldots, \frac{\partial x_1}{\partial y_p} \\ \frac{\partial x_2}{\partial y_1}, \ldots \\ \ldots \\ \ldots \\ \frac{\partial x_n}{\partial y_1}, \ldots \quad \frac{\partial x_n}{\partial y_p} \end{bmatrix} \tag{74}$$

$$B' = \begin{bmatrix} \frac{\partial y_1}{\partial x_1}, \frac{\partial y_1}{\partial x_2}, \ldots, \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1}, \ldots \\ \ldots \\ \frac{\partial y_p}{\partial x_1}, \ldots \quad \frac{\partial y_p}{\partial x_n} \end{bmatrix} \tag{75}$$

From Eqs.(64) and (65) it can easily be shown that B' is equal to the inverse of A' if p = n. If either the measurements or the parameters are not all independent, then, A' or B', of the sets of derivatives will not be defined. When the parameters are overdetermined ($n > p$) by a greater number of measurements, matrix A' is used. When there are the same numbers of parameter and observation, either A' or B' may be used. For propagation errors in observations A' must be used, but for error propagation to non-independent parameter, B' must be used.

The problems of errors propagation lie in the calculation of VC matrix for the parameters. The VC matrix for Y, M_Y, can be deduced from M_X. From the definitions of variance and covariance, M_Y is given by

$$M_Y = \langle Y' \, Y'^T \rangle = \frac{1}{n} \sum_i Y' \, {}_i Y'^T \tag{76}$$

Assuming that Y' and X' are linearly related over the range of errors, Y' can be expressed by an equation of the form of Eq.(35)

$$Y' = (A'^T \, W \, A')^{-1} \, A'^T \, W \, X' \tag{77}$$

Substituting Eq.(77) into Eq.(76)

$$M_Y = \langle (A'^T W A')^{-1} A'^T W X' X'^T W^T A' ((A'^T W A')^{-1})^T \rangle \tag{78}$$

Since W and $(A'^T W A')^{-1}$ are symmetric,

$$W^T = W \tag{79}$$

and

$$((A'^T W A')^{-1})^T = (A'^T W A')^{-1}. \tag{80}$$

Thus

$$\begin{aligned} M_Y &= (A'^T W A')^{-1} A'^T W \langle X' X'^T \rangle W A' (A'^T W A')^{-1} \\ &= (A'^T W A')^{-1} A'^T W M_X W A' (A'^T W A')^{-1} \\ &= (A'^T W A')^{-1} A'^T W A' (A'^T W A')^{-1} \\ &= (A'^T W A')^{-1} \\ &= (A'^T M_X^{-1} A')^{-1} \end{aligned} \tag{81}$$

where we have used the relation $W M_X = I$. Then the propagation of erro rs from n measurements to p parameters may be expressed as

$$\sigma(y_i) = ((M_Y)_{ii})^{\frac{1}{2}} \tag{82}$$

$$r(y_i, y_j) = \frac{(M_Y)_{ij}}{\sigma(y_i)\,\sigma(y_j)} \tag{83}$$

4-2. ERROR ELLIPSOID

We recall that the correlation coefficient was a measure of the extent to which errors in two quantities are correlated. Therefore it is directly related to the probability distribution of errors in particular pairs. A typical distribution with fairly general significance

is the normal distribution. Thus throughout this section we will assume the distribution of errors is normal. In this case the probability of a measurement has a value between x_i and $x_i + dx_i$ is given by

$$P(x_i)\, dx_i = (2\pi)^{-\frac{1}{2}}\ \sigma^{-1}(x_i)\ \exp\left(-\frac{(x_i')^2}{2\,\sigma^2(x_i)}\right) dx_i\ ,\quad i = 1, 2, \ldots, n$$

If X are all measured independently, then the probability of having values between X and X + X' is given by the product of individual probabilities,

$$\begin{aligned} P(X) &= (2\pi)^{-\frac{n}{2}}\ (\sigma(x_1)\ \ldots\ \sigma(x_n))^{-1} \exp\left(-\frac{(x_1')^2}{2\,\sigma^2(x_1)} - \ldots - \frac{(x_n')^2}{2\,\sigma^2(x_n)}\right) \\ &= \prod_{i=1}^{n} (2\pi)^{-\frac{1}{2}}\ \sigma(x_i)\ \exp\left(-\frac{(x_i')^2}{2\,\sigma^2(x_i)}\right) \end{aligned} \tag{84}$$

Eq.(84) can also be written as

$$P(X) = (2\pi)^{-\frac{n}{2}}\ \frac{1}{|S_X|}\ \exp\left(-\tfrac{1}{2}\, X'^{T} (S_X^2)^{-1}\, X'\right) \tag{85}$$

where $|S_X|$ is the determinant of S_X. Eq.(84) or (85) is usually called multivariate normal density function (Morrison, 1976). To make a geometric interpretation of probability we define <u>error ellipsoid</u> as

$$\sum_{i=1}^{n} \frac{(x_i')^2}{\sigma^2(x_i)} = 1 \tag{86}$$

Such a ellipsoid represents the probability of $e^{-\frac{1}{2}}$ times the maximum probability $1/((2\pi)^{n/2}|S_X|)$. All points outside this ellipsoid have a smaller probability than this and all points inside a larger probability. The covariance and coefficient are both zero on the axes of the ellipsoid.

5. NUMERICAL EXPERIMENTS

5-1. EQUATIONS OF SPECIFIC MODEL

In order to test the schemes that are developed in previous sections we use numerical data which are generated by a set of nonlinear equations

$$\frac{dU}{dt} = - a (V^2 + W^2) - c (Y^2 + Z^2) - U + H \tag{87}$$

$$\frac{dV}{dt} = a U V - b U W - e X Z - V \tag{88}$$

$$\frac{dW}{dt} = b U V + a U W + e X Y - W \tag{89}$$

$$\frac{dX}{dt} = (e - f) (V Z - W Y) - X \tag{90}$$

$$\frac{dY}{dt} = c U Y - d U Z - f X W - Y \tag{91}$$

$$\frac{dZ}{dt} = d U Y + c U Z + f X V - Z \tag{92}$$

where U, V, W, X, Y and Z are six variables, a, b, c, d, e, and f are constant coefficient, H is external forcing. This set of equations is deduced from a larger system of equations suggested by Lorenz(1963c). Although the equations used in this study do not closely approximate the so-called governing equations of the atmosphere, they are similar

in mathematical form to the latter. The linear terms in these equations correspond to friction, quadratic terms nonlinear contributions of various scales. The behaviors of the variables in the equations depend on their initial conditions, the constant coefficients and the forcing constant. The numerical values of these factors used in this study are shown in Tables 1 and 2. These values are not totally arbitrary, most of them had been tested by Hoffman(1977). Some initial conditions lead to symmetric vacillation and some to unsymmetric vacillation after a quite large number of time steps. To avoid computational instability we reexamined these initial conditions. They are stable at least up to 30000 time steps.

For the objective analysis we use additionally the following three equations

$$\psi = \sqrt{2}\,(U \cos y + X \cos 2y) + (V \sin y + Y \sin 2y) \cos 2x + (W \sin y + Z \sin 2y) \sin 2x \tag{93}$$

$$u = \sqrt{2}\, U \sin y - (V \cos 2x + W \sin 2x) \cos y + 2\sqrt{2}\, X \sin 2y - 2\,(Y \cos 2x + Z \sin 2x) \cos 2y \tag{94}$$

$$v = 2\,((W \cos 2x - V \sin 2x) \sin y + (Z \cos 2x - Y \sin 2x) \sin 2y) \tag{95}$$

where ψ represents stream function, u and v flow field and where we have used the relations

$$u = -\frac{\partial \psi}{\partial y} \tag{96}$$

$$v = \frac{\partial \psi}{\partial x} \tag{97}$$

Table 1. Constant coefficients.

Coefficient	a	b	c	d	e	f
Value	2.0	4.0	1.0	3.0	4.0	1.0

Table 2. Initial conditions of six variables and H .

Data set	H	U	V	W	X	Y	Z
1	2.020	0.55	0.01	0.01	0.01	0.01	0.10
2	2.024	0.58	0.01	0.01	0.01	0.01	0.10
3	2.020	0.01	0.01	0.01	0.01	1.01	0.01
4	2.020	0.01	0.01	1.01	0.01	0.01	0.01
5	2.020	1.01	0.01	0.01	1.01	0.01	0.01
6	2.020	1.01	0.01	1.01	1.01	0.01	0.01
7	2.000	0.01	0.01	0.01	0.01	0.01	1.01
8	2.000	0.01	0.01	0.01	0.01	0.01	0.10

Table 3. Dependence of averaged reductions of variance on the data length for data generated by the 4 4-cycle scheme. All the prediction steps are up to 1500.

	Data length							
Data set	1500	3000	4500	6000	7500	9000	10500	15000
1	0.8910	0.8896	0.8892	0.8887	0.8887	0.8883	0.8883	0.8880
7	0.8890	0.8889	0.8888	0.8888	0.8887	0.8887	0.8886	0.8886

Table 4. Same as Table 3 except all the prediction steps are the same as the data length

Data set	Data length 1500	3000	4500	6000	7500	9000	10500	15000
1	.8910	.8871	.8862	.8856	.8853	.8852	.8850	.8848
7	.8890	.8803	.8775	.8759	.8750	.8744	.8740	.8732

Table 5. Averaged reductions of variance in linear prediction by present predictors for data generated by 4 4-cycle and 8 4-cycle schemes, with indicated time steps of predictands.

Scheme	4 4-cycle				8 8-cycle			
	Time steps of predictands							
Data set	1	2	3	4	1	2	3	4
1	.891	.637	.351	.123	.973	.895	.779	.638
2	.895	.650	.379	.163	.973	.897	.783	.646
3	.896	.649	.372	.142	.968	.879	.745	.582
4	.888	.620	.315	.053	.968	.879	.745	.584
5	.893	.642	.362	.144	.968	.880	.749	.592
6	.888	.621	.316	.063	.965	.867	.721	.546
7	.889	.627	.334	.092	.969	.880	.748	.586
8	.890	.629	.338	.089	.965	.869	.725	.554

By Eqs.(93) - (95) combined with Eqs.(87) - (92), we can generate three different variables, which are function of time and space, at each point.

5-2. PREDICTION ANALYSIS

Three basic problems usually exit in the statistical prediction analysis; (1) whether the data length is long enough, (2) whether the data interval is proper and (3) whether the prediction formula fits the data well. From the theoretical point of view, it seems that the longer the data length involved the better the prediction. However, in practice it is important to make good forecasts by a method involving recent past or present data only. Therefore in our study we first investigate the dependence of the prediction errors on the length of data points. For simplicity only the present predictors are involved in this study and the averaged reductions of variance are calculated. Here we define the averaged reduction of variance(ARV) as

$$\text{ARV} = \frac{\text{variance - mean-squared prediction error}}{\text{variance}}$$

As shown in Tables 3 and 4, the data lengths range from 1500 to 15000. Two data sets are used in this test. In Table 3 the averaged reductions of variance are obtained by the prediction up to 1500 steps for all data length involved. In Table 4 the results are obtained by the prediction up to the same steps as the data length involved. It is noted that the averaged reductions of variance are nearly independent of the data length. Due to this fact, hereafter, we shall use the data

Table 6. Averaged reductions of variance in linear weighted prediction by present predictors for data generated by the 4 4-cycle and the 8 4-cycle schemes, with indicated time steps of predictands.

Scheme	4 4-cycle				8 8-cycle			
	Time steps of predictands							
Data set	1	2	3	4	1	2	3	4
1	.890	.628	.327	.077	.973	.895	.777	.635
2	.894	.642	.357	.122	.973	.896	.780	.641
3	.895	.642	.352	.104	.968	.878	.742	.577
4	.886	.609	.283	-.006	.968	.878	.742	.578
5	.892	.635	.341	.105	.968	.879	.746	.586
6	.887	.610	.284	.006	.965	.866	.718	.540
7	.888	.619	.310	.046	.968	.879	.745	.580
8	.889	.623	.318	.050	.965	.867	.721	.544

Table 7. Averaged reductions of variance in linear (L) and linear weighted (LW) prediction of every 2-time-step in advance for data generated by the 4 4-cycle scheme.

	Data set	1	2	3	4	5	6	7	8
	Time step								
L	1	.411	.441	.439	.374	.431	.384	.407	.415
L	2	-.341	-.257	-.263	-.450	-.254	-.424	-.335	-.302
LW	1	.405	.436	.435	.367	.427	.377	.402	.411

Table 8. Averaged reductions of variance in nonlinear prediction one time step ahead, for data generated by the 4 4-cycle scheme.

Data set	1	2	3	4	5	6	7	8
	.891	.895	.896	.888	.893	.888	.889	.890

points up to 1500 time steps only unless specified.

Next we shall examine whether the goodness of the prediction schemes depends on the data interval or not. In this case we use two different groups of data; one is generated by a 4 4-cycle numerical scheme and the other an 8 4-cycle scheme whose general principles are discussed by Lorenz(1971). The time-step interval in the 4 4-cycle scheme is twice as much as in the 8 4-cycle scheme. From Tables 5 and 6 we see that the averaged reductions of variance using the data produced by the 8 4-cycle scheme are much larger than those using the data generated by the 4 4-cycle scheme. This means that to predict one day ahead by using one-day interval data is much better than to predict two days ahead by using two-day interval data. However, it seems that to predict two days in advance by using one-day interval data is slightly worse than that by using two-day interval data. For the 8 4-cycle data, prediction one time step ahead is nearly perfect, prediction two time steps ahead is fair and predictions three and four time steps in advance are far from perfection. These results are almost agree well with those obtained by Lorenz(1962). However, for the 4 4-cycle data, the prediction one time step ahead is only mediocre. Therefore the dependence of prediction error on the data interval is quite evident. We also forecast every 2-time-step in advance by using 1-time-step data. The results are shown in Table 7. It is rather obvious that the averaged reductions of variance are very low and become negative

for the prediction 2 2-time-step in advance. These facts imply that if we want to predict one day ahead by linear regression method or near weighted method using present predictors it is better to use only the one-day-interval data during the calculation of B in Eq.(29) than to use 12-hr-interval data at the same time.

Finally we compare three prediction methods- linear, linear weighted and nonlinear methods - described in section 3. Their averaged reductions of variance are presented in Tables 5, 6 and 8 respectively. From these Tables, it is easy to find that the averaged reductions of variance obtained by all three methods are almost the same. This means the nonlinear method based on the data deviated from the mean does not improve the prediction error.

5-3. OBJECTIVE ANALYSIS

In order to test the univariate and multivariate methods for objective analysis, we use differently dispersed positions of grid and observation points which are shown in Figs. 1 and 2. In Fig.2, the distances between grid points and observation points are relatively longer than that in Fig. 1. In stead of changing the distribution of the observation points we change the grid point in six different ways, denoted as G_1, G_2, ..., G_6, for different distributions of observation points. The averaged reductions of variance for these methods are given in Table 9 for different grid-point positions in Figs. 1 and 2.

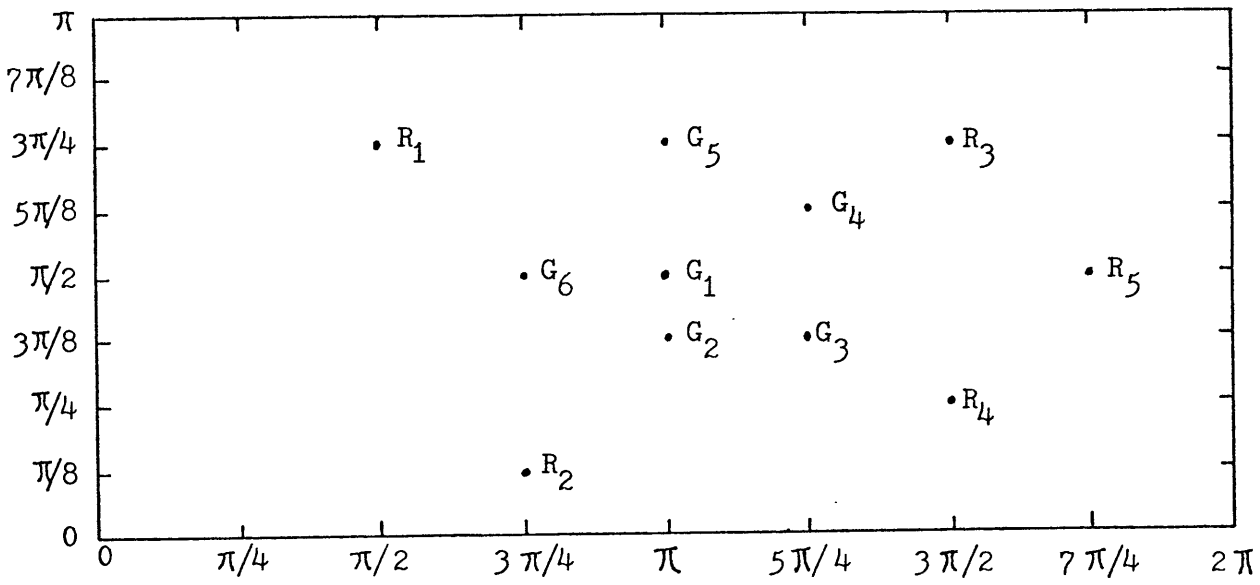

Fig. 1. The domain of grid points and observation points

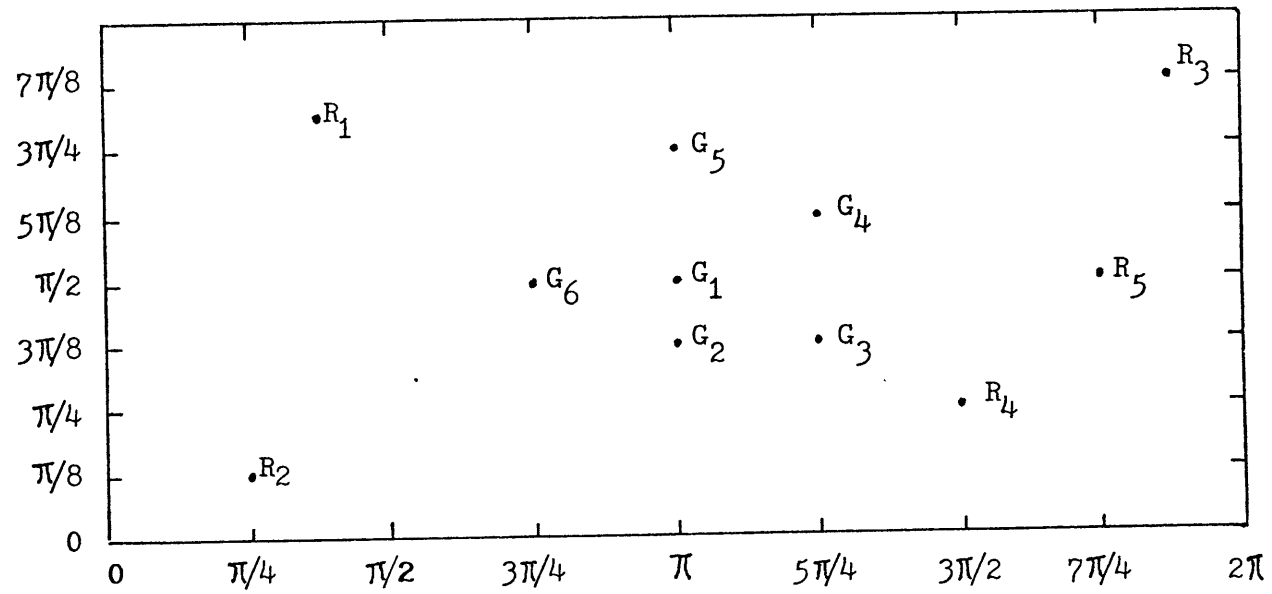

Fig. 2. The domain of grid points and observation points

Table 9. Averaged reductions of variance calculated by univariate (UM) and multivariate (MM) methods for data generated by the 8 4-cycle scheme, with indicated grid-point positions in Figs. 1 and 2.

Figure	scheme	Grid-point position					
		1	2	3	4	5	6
Fig.1	UM	.911	.899	.911	.886	.917	.991
	MM	1.000	1.000	1.000	1.000	1.000	1.000
Fig.2	UM	.958	.838	.931	.893	.976	.888
	MM	1.000	1.000	1.000	1.000	1.000	1.000

Table 10. Total prediction errors by univariate (UM) and multivariate (MM) methods for data generated by the 8 4-cycle scheme, with indicated grid-point positions in Figs. 1 and 2.

Figure	Scheme	Grid-point position					
		1	2	3	4	5	6
Fig.1	UM	414.87	445.16	404.82	567.88	371.58	420.06
	MM	.001	.001	.003	.001	.000	.001
Fig.2	UM	196.83	715.04	315.32	533.10	107.95	523.44
	MM	.035	.039	.001	.015	.035	.026

It is remarkable that the results obtained by the multivariate method are perfect and almost independent of the distributions of the nearby points. For the univariate method, however, good results are obtained when the grid points are at point 6 in Fig. 1 and ponits 1 and 5 in Fig. 2. These points are likely located around the centroid of the observation points. The squared prediction errors for these two methods are compared in Table 10. The prediction errors by multivariate analysis can hardly be detected, while those by the univariate method are relatively large. We do not claim that our cases are the general one. It is, however, reliable to say that the multivariate scheme developed in this study is much better than the univariate one. That is, for objective analysis to estimate the grid-point values from those at unevenly distributed observation points, it is much better to take into account the covariance between the variables than to involve only one variable.

5-4. MULTIVARIATE ERROR ANALYSIS

As mention above the central work of multivariate error analysis lies in the calculation of the standard deviation and correlation coefficient of both observations and parameters. In this study we take U, V, W, X, Y, Z in Eqs.(87) - (92) as observations and their respective derivatives $\dot{U}$, $\dot{V}$, $\dot{W}$, $\dot{X}$, $\dot{Y}$, $\dot{Z}$ as parameters. Here we shall investigate how the standard deviation and correlation coefficient of observations change and how they propagate to those of parameters

Table 11. Average (A) and standard deviation (σ) of six variables for data generated by the 4 4-cycle scheme.

	Variables					
	U	V	W	X	Y	Z
A	0.8020	-0.0311	-0.0140	-0.1498	0.0040	0.0031
	0.2817	0.4844	0.4717	0.3307	0.3722	0.3672

Table 12. Correlation coefficient of six variables.

	U	V	W	X	Y	Z
U	1	0.068	0.053	-0.024	-0.017	-0.016
V		1	0.017	-0.036	0.401	-0.142
W			1	-0.040	0.143	0.383
X				1	0.117	0.035
Y					1	-0.009
Z						1

Table 13. Standard deviations of six variables after increasing 1 percent of initial condition of a variable.

Variable changed	Standard deviation of variables					
	U	V	W	X	Y	Z
U	.2708	.4894	.4934	.3593	.3486	.3532
V	.2794	.5254	.5288	.3285	.2832	.2852
W	.2826	.4685	.4716	.3588	.3798	.3809
X	.2839	.5138	.5135	.3340	.3105	.3097
Y	.2669	.4868	.4889	.3580	.3560	.3580
Z	.2616	.4905	.4895	.3582	.3550	.3549

Table 14. The same as Table 13 except for 5 percent.

Variable changed	Standard deviation of variables					
	U	V	W	X	Y	Z
U	.2754	.4734	.4759	.3657	.3766	.3758
V	.2805	.5130	.5199	.3305	.3064	.3059
W	.2852	.4908	.4957	.3456	.3466	.3483
X	.2789	.4816	.4862	.3305	.3596	.3638
Y	.2800	.4986	.5072	.3477	.3304	.3322
Z	.2766	.5276	.5224	.3312	.2889	.2852

Table 15. The same as Table 13 except for 10 percent.

Variable changed	Standard deviation of variables					
	U	V	W	X	Y	Z
U	.2657	.4766	.4847	.3724	.3730	.3665
V	.2922	.4775	.4763	.3469	.3732	.3707
W	.2637	.4820	.4759	.3570	.3729	.3722
X	.2829	.5053	.5027	.3374	.3289	.3260
Y	.2764	.4857	.4861	.3335	.3606	.3607
Z	.2782	.4916	.4883	.3576	.3529	.3518

by changing the initial conditions of the observations.

Tables 11 and 12 give the standard deviations and correlation coefficients of six variables before the initial conditions are not changed. Tables 13-15 indicate the standard deviations of six variables after increasing respectively 1, 5 and 10 percents of the initial values of each variable. We see, from these Tables, when the initial condition of V or X increases one percent the standard deviations of V, W, Y and Z change obviously. However when the initial condition increases 5 and 10 percent the standard deviation of six variables almost keeps constant.

To calculate the error propagation we use Eqs.(82) and (83)

$$\sigma(y_i) = ((M_Y)_{ii})^{\frac{1}{2}}$$

$$r(y_i, y_j) = \frac{(M_Y)_{ij}}{\sigma(y_i)\ \sigma(y_j)}$$

where

$$M_Y = (A^T\ M_X^{-1}\ A')^{-1}$$

$$M_Y = B'\ M_X\ B^T \qquad (98)$$

A' and B' are given in Eqs.(74) and (75). According to Eq.(75), B' can be obtained by differentiating the parameters with respect to the observations in Eqs.(87)-(92), then

Table 16. Standard deviation of six parameters after increasing 1 percent of initial condition of a variable.

Variable changed	Standard deviation of parameters					
	$\dot{U}$	$\dot{V}$	$\dot{W}$	$\dot{X}$	$\dot{Y}$	$\dot{Z}$
U	1.4378	1.9201	1.8837	1.0949	0.9637	0.9516
V	1.6128	1.8737	1.8589	0.9573	0.7549	0.7501
W	1.3597	1.9277	1.9084	1.1177	1.0711	1.0676
X	1.5470	1.8846	1.8839	1.0077	0.8304	0.8325
Y	1.4326	1.8941	1.8781	1.0890	0.9728	0.9678
Z	1.4459	1.8733	1.8735	1.0848	0.9560	0.9563

Table 17. The same as Table 16 except for 5 percent.

Variable changed	Standard deviation of parameters					
	$\dot{U}$	$\dot{V}$	$\dot{W}$	$\dot{X}$	$\dot{Y}$	$\dot{Z}$
U	1.3650	1.9276	1.9246	1.1328	1.0509	1.0521
V	1.5679	1.8709	1.8623	0.9910	0.8122	0.8119
W	1.4533	1.9104	1.9059	1.0775	0.9553	0.9509
X	1.4672	1.8466	1.8248	1.0316	0.9805	0.9699
Y	1.4926	1.9097	1.8890	1.0575	0.9009	0.8953
Z	1.6016	1.8498	1.8758	0.9625	0.7554	0.7639

Table 18. The same as Table 16 except for 10 percent.

Variable changed	Standard deviation of parameters					
	$\dot{U}$	$\dot{V}$	$\dot{W}$	$\dot{X}$	$\dot{Y}$	$\dot{Z}$
U	1.3907	1.9117	1.9287	1.1290	1.0146	1.0290
V	1.3944	1.9042	1.9278	1.0988	1.0394	1.0456
W	1.5021	1.9022	1.9169	1.0406	0.8905	0.8974
X	1.4667	1.8380	1.8415	1.0398	0.9703	0.9706
Y	1.4310	1.9086	1.9189	1.0950	0.9678	0.9711
Z	1.3854	1.8849	1.9172	1.1220	1.0730	1.0792

Table 19. Correlation coefficient of six parameters after increasing 1, 5 and 10 percent of initial condition of U variable.

Percent	Parameters	$\dot{U}$	$\dot{V}$	$\dot{W}$	$\dot{X}$	$\dot{Y}$	$\dot{Z}$
	$\dot{U}$	1	-.023	-.021	-.050	-.001	-.009
	$\dot{V}$		1	.018	-.003	-.018	-.002
1%	$\dot{W}$			1	-.010	-.034	-.040
	$\dot{X}$				1	.011	-.009
	$\dot{Y}$					1	.009
	$\dot{Z}$						1
	$\dot{U}$	1	-.006	-.018	-.020	.005	-.002
	$\dot{V}$		1	-.006	.003	.011	.002
5%	$\dot{W}$			1	-.004	-.011	-.030
	$\dot{X}$				1	.019	-.019
	$\dot{Y}$					1	-.002
	$\dot{Z}$						1
	$\dot{U}$	1	-.001	-.005	-.016	-.005	.019
	$\dot{V}$		1	-.044	-.015	-.015	.010
10%	$\dot{W}$			1	.012	.004	-.012
	$\dot{X}$				1	.003	-.005
	$\dot{Y}$					1	-.022
	$\dot{Z}$						1

Table 20. The same as Table 19 except for V variable.

Percent	Parameters	$\dot{U}$	$\dot{V}$	$\dot{W}$	$\dot{X}$	$\dot{Y}$	$\dot{Z}$
	$\dot{U}$	1	-.019	-.014	-.000	-.003	.004
	$\dot{V}$		1	.002	-.008	.007	-.002
1%	$\dot{W}$			1	.006	-.008	.015
	$\dot{X}$				1	.016	-.007
	$\dot{Y}$					1	.004
	$\dot{Z}$						1
	$\dot{U}$	1	-.005	-.002	-.100	.010	-.001
	$\dot{V}$		1	-.018	.014	-.111	.028
5%	$\dot{W}$			1	.002	-.018	-.120
	$\dot{X}$				1	-.020	-.003
	$\dot{Y}$					1	-.014
	$\dot{Z}$						1
	$\dot{U}$	1	.009	.001	.103	.004	.003
	$\dot{V}$		1	-.016	.006	.147	.004
10%	$\dot{W}$			1	.009	.030	.133
	$\dot{X}$				1	-.020	.009
	$\dot{Y}$					1	-.002
	$\dot{Z}$						1

Table 21. The same as Table 19 except for W variable.

Percent	Parameters	$\dot{U}$	$\dot{V}$	$\dot{W}$	$\dot{X}$	$\dot{Y}$	$\dot{Z}$
	$\dot{U}$	1	.029	-.014	-.110	.001	-.007
	$\dot{V}$		1	.006	-.008	-.105	.024
1%	$\dot{W}$			1	-.021	-.019	-.128
	$\dot{X}$				1	.038	.001
	$\dot{Y}$					1	.007
	$\dot{Z}$						1
	$\dot{U}$	1	.010	.010	-.076	.005	-.002
	$\dot{V}$		1	-.014	.006	-.089	.010
5%	$\dot{W}$			1	.004	-.024	-.065
	$\dot{X}$				1	.021	-.019
	$\dot{Y}$					1	-.009
	$\dot{Z}$						1
	$\dot{U}$	1	.007	-.007	.165	.002	-.006
	$\dot{V}$		1	.007	.002	.177	-.040
10%	$\dot{W}$			1	-.000	.025	.175
	$\dot{X}$				1	.016	.003
	$\dot{Y}$					1	.008
	$\dot{Z}$						1

Table 22. The same as Table 19 except for X variable.

Percent	Parameters	$\dot{U}$	$\dot{V}$	$\dot{W}$	$\dot{X}$	$\dot{Y}$	$\dot{Z}$
	$\dot{U}$	1	-.013	.001	.055	-.012	-.000
	$\dot{V}$		1	.003	-.016	.065	-.014
1%	$\dot{W}$			1	-.002	.010	.066
	$\dot{X}$				1	-.020	.003
	$\dot{Y}$					1	.002
	$\dot{Z}$						1
	$\dot{U}$	1	.020	-.011	.237	-.004	.005
	$\dot{V}$		1	.001	-.002	.277	-.030
5%	$\dot{W}$			1	.024	.069	.263
	$\dot{X}$				1	.019	.008
	$\dot{Y}$					1	-.000
	$\dot{Z}$						1
	$\dot{U}$	1	-.019	.034	-.054	-.001	.001
	$\dot{V}$		1	-.004	.000	-.064	.002
10%	$\dot{W}$			1	.003	-.015	-.087
	$\dot{X}$				1	.013	.005
	$\dot{Y}$					1	.000
	$\dot{Z}$						1

Table 23. The same as Table 19 except for Y variable.

Percent	Parameters	$\dot{U}$	$\dot{V}$	$\dot{W}$	$\dot{X}$	$\dot{Y}$	$\dot{Z}$
	$\dot{U}$	1	-.031	-.007	-.097	-.004	.006
	$\dot{V}$		1	.008	-.000	.127	-.009
1%	$\dot{W}$			1	.006	.020	.124
	$\dot{X}$				1	-.002	.001
	$\dot{Y}$					1	.010
	$\dot{Z}$						1
	$\dot{U}$	1	.029	.026	.038	.003	-.003
	$\dot{V}$		1	-.012	-.000	.049	-.020
5%	$\dot{W}$			1	.005	-.002	.027
	$\dot{X}$				1	.013	.011
	$\dot{Y}$					1	-.005
	$\dot{Z}$						1
	$\dot{U}$	1	-.004	.001	-.221	-.009	-.004
	$\dot{V}$		1	-.004	-.015	-.254	.037
10%	$\dot{W}$			1	-.009	-.054	-.241
	$\dot{X}$				1	-.010	.003
	$\dot{Y}$					1	.005
	$\dot{Z}$						1

Table 24. The same as Table 19 except for Z variable.

Percent	Parameters	$\dot{U}$	$\dot{V}$	$\dot{W}$	$\dot{X}$	$\dot{Y}$	$\dot{Z}$
	$\dot{U}$	1	.031	.007	-.135	-.012	.011
	$\dot{V}$		1	.005	-.014	-.132	.028
1%	$\dot{W}$			1	.018	-.034	-.128
	$\dot{X}$				1	.021	-.006
	$\dot{Y}$					1	.005
	$\dot{Z}$						1
	$\dot{U}$	1	-.031	-.002	.018	-.009	.014
	$\dot{V}$		1	-.006	-.013	.031	.004
5%	$\dot{W}$			1	.019	.007	.032
	$\dot{X}$				1	-.001	-.012
	$\dot{Y}$					1	.000
	$\dot{Z}$						1
	$\dot{U}$	1	-.004	.027	.029	-.004	-.001
	$\dot{V}$		1	.004	-.008	.0027	.003
10%	$\dot{W}$			1	-.005	.015	.037
	$\dot{X}$				1	-.001	.008
	$\dot{Y}$					1	.010
	$\dot{Z}$						1

$$B' = \begin{bmatrix} -1 & -2aV & -2aW & 0 & -2cY & -2cZ \\ aV-bW & aU-1 & -bU & -eZ & 0 & -eX \\ bV+aW & bU & aU-1 & eY & eX & 0 \\ 0 & (e-f)Z & -(e-f)Y & -1 & -(e-f)W & (e-f)V \\ cY-dZ & 0 & -fX & -fW & cU-1 & -dU \\ dY+cZ & fX & 0 & fV & dU & cU-1 \end{bmatrix} \quad (99)$$

The propagation of errors is shown in Tables 18 through 24. We find from Table 18 when the V variable increases 1 percent of its initial condition the standard deviations of most parameter change apparently. No regular variation of standard deviations of parameters is observed in this study. Since the correlation coefficient between parameters is very small , the error popagations are nearly independent of each other. From the standard deviations of observations and of parameters we can estimate the error ellipsoids.

6. CONCLUDING REMARKS

In this study we applied least-squares principles to both prediction and objective analyses and got similar forms in the weighting function, B, of the prediction formula. We also found that the averaged reductions of variance depend on the data interval used but are nearly independent of the data lengths ranging from 1500 to 15000 time steps. The goodness of linear, linear weighted

and nonlinear methods decribed in this paper are almost the same. Both linear weighted and nonlinear techniques do not improve the prediction error.

In objective analysis we got an exciting result. To make grid-point estimates from the nearby irregularly distributed observation points, the multivariate scheme is superior to the univariate one.

As for the multivariate error analysis, we did not get obvious results. In our cases, the correlation coefficients between the observations and between the parameters are very small.

Further investigation using the real weather data seems necessary to make a more affirmative conclusion.

REFERENCES

Bengtsson, L., 1975: 4-dimensional assimilation of meteorological observations. GARP 15, ICSU/WMO, 76pp.

Bevington, P. R., 1969: Data reduction and error analysis for the physical sciences. McGraw-Hill, 336 pp.

Clifford, C. C., 1973: Multivariate error analysis. John willey & Sons, 112 pp.

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

Eddy, A., 1967: The ststistical objective analysis of scalar data fields. J. Appl. Meteor., 6, 597-609.

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

Gleeson, T. A., 1977: Comments on "Application of Stochastic Dynamic Prediction to Real Data". J. Atmos. Sci., 34, 1138-1139.

Hoffman, R. N., 1977: The 'climate' of a simple six component model. Personal communication.

Kruger, H. B., 1969: General and special approaches to the problem of objective analysis of meteorological variables. Quart. J. Roy. Meteor. Soc., 95, 21-39.

Lorenz, E. N., 1962: The statistical prediction of solutions of dynamic equations. Proc. Internat. Sympos. Numerical Weather Prediction. Tokyo, Japan, 629-635.

Lorenz, E. N., 1963a: The predictability of hydrodynamic flow. Trans. New York Acad. Sci., Ser. II, 25, 409-432.

———, 1963b: Deterministic nonperiodic flow. J. Atmos. Sci., 20, 130-141.

———, 1963c: The mechanics of vacillation. J. Atmos. Sci., 20, 448-464.

———, 1965: Astudy of the predictability of a 28-variable atmosphere model. Tellus, XVII, 3, 321-333.

———, 1971: An N-cycle time differencing scheme for stepwise numerical integration. Mon. Wea. Rev., 99, 644-648.

———, 1973: Predictability and periodicity: A review and extension. Preprints 3rd Conf. Probability and Statistics in Meteor. Boulder, Colo., Amer. Meteor. Soc., 1-4.

———, 1977: An experiment in nonlinear statistical weather forecasting. Mon. Wea. Rev., 105, 590-602.

Morrison, D. F., 1976: Multivariate statistical methods. McGraw-Hill, 415 pp.

Pitcher, E. J., 1977: Application of stochastic dynamic prediction to real data. J. Atmos. Sci., 34, 3-21.

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

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

———, G. W. Branstator and L. G. Thiel, 1976: Testing a global multivariate statistical objective analysis scheme with observed

data. Mon. Wea. Rev., 104, 765-783.

Thiebaux, H. J., 1973: Maximally stable estimation of meteorological parameters at grid points. J. Atmos. Sci., 30, 1710-1714.

———, 1974: Estimation of covariances of meteorological parameters using local time averages. J. Appl. Meteor., 13, 592-600.