Statistical Forecasting Operators Based on Dynamical Equations

By DUANE S. COOLEY, Air Force Cambridge Research Center

(Manuscript received August 19, 1957, revised February 14, 1958)

Abstract

Statistical forecasting operators are derived on the basis of a simple two-parameter model of the atmosphere. In the course of this investigation, the predictability of pressure-height changes due to the various physical processes specified by this dynamical model is investigated. It is found that the contribution of the non-linear processes to the 12 and 24 hr predictability is small.

1. Introduction

In recent years, dynamical forecasting by the numerical integration of partial differential equations describing the fluid motion in certain simplified models of the atmosphere has become a common practice. Among the many mathematical models which have been investigated is that expressed by the equations of thermotropic flow as derived by THOMPSON and GATES (1956). A statistical study of the physical processes expressed by these equations has been undertaken with a view to deriving suitable statistical forecasting operators based directly on the dynamical equations and also with a view to examining the importance of these physical processes for prediction purposes. By this procedure, a forecasting technique is synthesized from the statistical and the dynamic approaches. The statistical approach is similar to those discussed by WIENER (1949), WADSWORTH (1951), WHITE and PALSON (1955), MALONE (1955) and MILLER (1956). The form of the statistical solution to this prediction problem will be shown to be loosely analogous to that of the solution of the dynamic equations by Green's method.

2. The Dynamic Equations

The thermotropic equations derived by THOMPSON and GATES (1956) characterize a two-parameter model of atmospheric flow. In this model along with the usual assumptions of hydrostatic, adiabatic and geostrophic flow, the vertical shear of the horizontal wind, or the so-called thermal wind, is assumed to remain constant with height. For flat terrain the equations take the form

$$\nabla^2 \frac{\partial \overline{\psi}}{\partial t} + J(\overline{\psi}, \nabla^2 \overline{\psi}) + \beta \frac{\partial \overline{\psi}}{\partial x} + KJ(\overline{\varphi}, \nabla^2 \overline{\varphi}) = 0 \tag{1}$$

$$\nabla^2 \frac{\partial \overline{\varphi}}{\partial t} + J(\overline{\psi}, \nabla^2 \overline{\varphi}) + J(\overline{\varphi}, \nabla^2 \overline{\psi}) + LJ(\overline{\varphi}, \nabla^2 \overline{\varphi}) + \beta \frac{\partial \overline{\varphi}}{\partial x} - \mu^2 \left[\frac{\partial \overline{\varphi}}{\partial t} + J(\overline{\psi}, \overline{\varphi})\right] = 0 \tag{2}$$

where the bar represents an average with respect to pressure in the vertical, $\psi = gzf^{-1}$, $\varphi = RTf^{-1}$, J is the Jacobian operator, ∇^2 is the Laplacian operator, K and L are functions of pressure only, μ^2 is a positive constant determined by the static stability, R is the gas constant for air, and g, z, T, f and β have their usual meanings. In solving these equations numerically, $\bar{\psi}$ is taken as proportional to the 500 mb height and $\bar{\varphi}$ as proportional to the 500—1,000 mb thickness. These equations were solved numerically for a square grid of points about 280 km apart over the United States and Canada. For details of the work, see THOMPSON and GATES (1956).

The solution of partial differential equations (1) or (2) could, presumably, be carried out by Green's method (HILDEBRAND, 1952). Let equation (1) be represented as

$$L\frac{\partial z}{\partial t} + F(x, y) = 0 \tag{3}$$

Then the solution to the equation at any point x, y can be represented formally as

$$\frac{\partial z}{\partial t}(x, y) = \int_{y_1}^{y_2}\int_{x_1}^{x_2} F(\xi, \eta)\, G(x, y; \xi, \eta)\, d\xi\, d\eta \tag{4}$$

where G is the Green's function for the problem, ξ and η are dummy variables of distance in the x and y directions respectively. If the integral solutions are obtained in finite difference form over q grid points on the earth, one may write $\{\underline{X_D}\} = \underline{A}\{\underline{F}\}$ (5) where $\{\underline{X_D}\}$ is the column vector of the tendencies $\left(\frac{\partial z}{\partial t}\right)_i$ in this dynamic model at the q grid points, $\{\underline{F}\}$ is the column vector of the values of F at each point and A is the matrix of coefficients representing finite-difference forms of the appropriate Green's function for the problem. In the equations (5) the boundary values of $\frac{\partial z}{\partial t}$ have been incorporated in F for convenience.

3. The Statistical Equations

Consider the prediction of $\frac{\partial z}{\partial t}$ by statistical means. It may be treated as a linear least squares multiple regression problem in the multiple time series of $\frac{\partial z}{\partial t}$ and F at the q grid points. For simplicity we will consider that $\frac{\partial z}{\partial t}$ is a deviation from its mean value. For a given point i, $\left(\frac{\partial z}{\partial t}\right)_i$ can be approximated by

$$\left(\frac{\partial z}{\partial t}\right)_i = \sum_{j=1}^{q} b_{ij} F_j + e_i \tag{6}$$

The regression coefficients b_{ij} in (6) are to be determined from n ($\geq q+1$) simultaneous observations of $\left(\frac{\partial z}{\partial t}\right)_i$ and F_j's in such a manner as to minimize the error sum of squares at point i given by

$$\sum_n e_i^2 = \sum_n \left[\left(\frac{\partial z}{\partial t}\right)_i - \sum_{j=1}^{q} \partial_{ij} F_j\right]^2 \tag{7}$$

where $\sum_n$ denotes the summation over n observations. When $\sum_n e^2$ is minimized with respect to b_{ij}, the following general equation for the coefficients b_{ij} is obtained:

$$\sum_{j=1}^{q} b_{ij} \sum_n F_{ij} F_k = \sum_n F_k \left(\frac{\partial z}{\partial t}\right)_i \quad k = 1.2 \ldots q \tag{8}$$

This can be written as

$$\sum_{j=1}^{q} b_{ij} \overline{F_j F_k} = \overline{F_k \left(\frac{\partial z}{\partial t}\right)_i} \quad k = 1.2 \ldots q \tag{9}$$

where the bar indicates an average over the n observations and the barred terms now represent covariances. This system of linear equations can be solved for the b_{ij} by inverting the matrix of covariances $[\overline{F_j F_k}] \equiv A$ and multiplying by the covariance vector

$$\left[\overline{F_k \left(\frac{\partial z}{\partial t}\right)_i}\right] \equiv P.$$

The predicted values of $\frac{\partial z}{\partial t}$ at the q grid points are simply given by the set of equations

$$\{\underline{X_s}\} = \underline{B}\{\underline{F}\} \tag{10}$$

where $\{\underline{X_s}\}$ is the vector of the predicted tendencies $\frac{\partial \hat{Z}}{\partial t}$ in the statistical model at the q grid points, $\{\underline{F}\}$ is the vector of predictor functions and $\underline{B}$ is the matrix of regression coefficients.

Equation (10) is a statistical analogue of equation (5). The matrix A of the coefficients of equation (5) representing a set of theoretical Green's functions is determined by the form of the equation governing the model and the characteristics of the grid system used. All mathematical models used for weather prediction are highly idealized representations of the true atmosphere. They usually do not include the effects of heating, accelerating wind fields, topography, etc. Those physical processes which are represented by the model cannot be evaluated accurately due to a lack of data over large areas of the globe or due to insufficient accuracy of the available observations. To some extent, statistical prediction suffers from the same inaccuracies and incompleteness of information.

The matrix B of the coefficients b_{ij} in the matrix equation (10) represents a set of empirical Green's functions. Because of its empirical nature, it results from the effects of all physical processes acting on or in the atmosphere. The effects of physical processes which are linearly related to the function F of the initial conditions will be included. This is a result of the fact that in the least-squares sense the procedures by which the b_{ij} were derived result in the best linear combinations of the function F obtainable from the data used.

In actual practice it is necessary to stretch the analogy between the systems represented by the equations (5) and (10). This is because instantaneous values of the contour height tendencies are not available to derive the truly analogous empirical Green's functions of equation (10) and $\frac{\partial z}{\partial t}$ must be approximated by finite height changes over 12 or 24 hours. The Green's functions derived here were not intended primarily for preparing routine forecasts but for studying the relative importance of certain linear and non-linear information about the lower troposphere in making predictions.

4. Procedures

The effects on predictability of pressure-height changes due to the physical processes represented by certain non-linear terms of equations (1) and (2) were studied from the evaluations of finite difference approximations to these terms. First, the predictability of heights obtainable from certain linear combinations of heights was determined. Then in various ways the linear combinations of heights were used jointly with the non-linear functions as predictors and differences in predictability were noted.

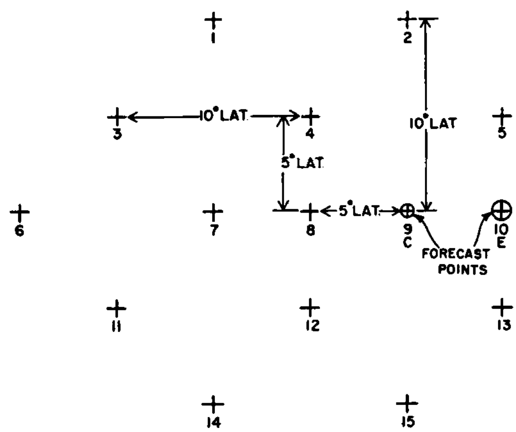

Fig. 1. Grid of 15 points at which values of the predictors were observed.

The predictor functions F were evaluated over a grid of 15 points of the form shown in fig. 1. The positions of the predictand points are circled. The data were gathered for various positions of the grid over the United States and southern Canada. However, the relative positions and orientations of the grid points were maintained.

The predictor functions F at the 15 grid points were of the following form:

F_1 through F_{15} : Z_5, 500 mb height
F_{16} through F_{30} : Z_{10}, 1,000 mb height
F_{31} through F_{45} : $J_1 \equiv J(Z_5, \nabla^2 Z_5)$
F_{46} through F_{60} : $J_2 \equiv J[(Z_5-Z_{10}), \nabla^2(Z_5-Z_{10})]$
F_{61} through F_{75} : $J_3 \equiv J(Z_{10}, \nabla^2 Z_{10})$
F_{76} through F_{90} : $J_4 \equiv J(Z_5, Z_5-Z_{10})$

These represent geostrophic approximations to the terms appearing in equations (1) and (2). J_1 expresses the advection of the vorticity of the 500 mb geostrophic wind by the 500 mb geostrophic wind. It is an approximation to the term representing the vorticity advection in a barotropic model. J_2 is a measure of the advection of the vorticity of the thermal wind from 1,000 to 500 mb by this thermal wind. J_3 expresses the advection of the vorticity of the geostrophic wind at 1,000 mb by the 1,000 mb geostrophic wind. Finally, J_4 is a measure of the vertically averaged temperature advection in the lower troposphere. It is evaluated from the 500 to 1,000 mb thickness gradient and the 500 mb geostrophic wind.

Statistical operators were derived for the prediction of the 12 and 24-hour contour height changes at 1,000 and 500 mb. Calculation of the linear regression equations and the associated multiple correlation coefficients was performed using data provided by the Joint Geophysics Research Directorate—Air Weather Service Numerical Prediction Project. These data were abstracted from sixty weather maps at the 1,000 and 500 mb levels, 0300 and 1500 GCT, 1 to 31 January 1953. Five or six sets of observations were taken from each of the maps, making a sample size of 336 observations. Dr. W. L. Gates made available his IBM 701 EDPM program for computing the initial values of the Jacobian terms of equations (1) and (2).

If the regression analysis were carried out using all of the predictors, a 90×90 predictor covariance matrix would have to be inverted. Because of the limitations of the machine storage capacity and the cost of such calculations, and also for reasons of statistical stability, no more than 60 predictors were used simultaneously. The following combinations of predictors were used:

1) : Z_5, Z_{10}	Linear case,	30 predictors
2) : Z_5, J_1	Barotropic case,	30 predictors
3) : Z_5, Z_{10}, J_1		45 predictors
4) : Z_5, Z_{10}, J_1, J_2		60 predictors
5) : Z_5, Z_{10}, J_1, J_3		60 predictors
6) : Z_5, Z_{10}, J_1, J_4		60 predictors

If the covariance matrix of any particular combination of predictors is designated as A_s then equation (9) can be expressed in simple matrix form as

$$A_s B = P \tag{11}$$

where B is the matrix of regression coefficients and P is the predictor and predictand covariance matrix. Then as indicated previously, the coefficients in B can be obtained by inverting A_s, since

$$B = A_s^{-1} P \tag{12}$$

The basic matrix inversion program used (employing the von Neumann—Goldstine method of matrix inversion, 1947) was capable of inverting 50th order matrices. Since three of the matrices (A_4 to A_6) are of 60th order, it was necessary to use a partitioning procedure and to obtain the inverse matrices by the escalator method (HOUSEHOLDER, 1953).

5. Results

The predictability of the dependent variables as a function of the various predictor combinations is readily measured by the percent reduction of the variance of the predictands as determined from the sample of data studied. (This percent reduction of the variance explained in the dependent data is equal to the square of the multiple correlation coefficient R between all the predictors and the predictand and will henceforth be designated as R^2.) From such comparisons we can deduce information about the relative importance of physical processes described by the terms of the thermotropic equations as they affect the statistical predictions.

The percent reductions of the variances of the forecast height changes at the central forecast point (point 9 in fig. 1) using various combinations of predictors are listed in Table 1. Comparison of the first two lines of this table indicates that the linear terms Z_5, Z_{10} yield better predictions of the 500 mb height changes than does the combination of information from Z_5 and the non-linear barotropic term J_1. The suggestion here is that linear baroclinic information is more important than non-linear barotropic information in predicting 500 mb height changes. When the predictors Z_5, Z_{10}, J_1 are combined as in A_3 some improvement in the 500 mb forecast is obtained indicating that the non-linear barotropic vorticity advection term does appear to add some useful information to that obtainable from the simple linear combination of heights at two levels.

In the fourth line of Table 1 are shown the effects of adding 15 values of J_2, the term representing the advection of the vorticity

Table 1. Reduction of variance (R^2, percent, where R is the multiple correlation coefficent) of predicted height changes at the central forecast point using various predictors.

Predictors	$\Delta_{12}Z_5$	$\Delta_{24}Z_5$	$\Delta_{12}Z_{10}$	$\Delta_{24}Z_{10}$
$A_1 : Z_5, Z_{10}$	72	69	68	70
$A_2 : Z_5, J_1$	66	58	50	51
$A_3 : Z_5, Z_{10}, J_1$	75	72	72	72
$A_4 : Z_5, Z_{10}, J_1, J_2$	77	73	73	73
$A_5 : Z_5, Z_{10}, J_1, J_3$	76	72	73	73
$A_6 : Z_5, Z_{10}, J_1, J_4$	78	76	79	77

(Predictands: columns 2–5)

of the thermal wind from 1,000 to 500 mb by the thermal wind in that layer, to the 45 predictors used in the third line. This term appears to add very little to the prediction of the 500 mb height changes. As might be expected, the term J_3, the advection of the 1,000 mb vorticity by the 1,000 mb geostrophic wind, which was used in arrangement A_5 contributed almost nothing to the prediction. The best set of 60 predictors appears to be that in the sixth row. The added predictors in A_6 are the values of the J_4 term which measure the advection of the mean virtual temperature in the lower troposphere by the 500 mb geostrophic wind.

The prediction of the 1,000 mb height changes as shown in the third and fourth columns of Table 1 follows a similar pattern for the most part. However, the predictors Z_5 and J_1 do quite poorly at the 1,000 mb level as might be expected, since they include information only from the 500 mb level. There does not appear to be the marked difference in predictability between 500 and 1,000 mb height changes usually found in numerical forecasting results (THOMPSON and GATES, 1956). Thus, the two non-linear terms which appear to contribute some information over and above that given by linear combinations of the heights of the two pressure surfaces are the barotropic vorticity advection term J_1 and the temperature advection term J_4.

Actually, higher values of R^2 were obtained for predictions of height change at the easternmost forecast point of the network, indicating that this grid point arrangement relative to the central forecast point did not include enough information from the west. For the eastern predictand the level of accuracy reaches 84 percent using the predictors in A_6.

An attempt was made to test the significance of the improvement gained by additional predictors. Since the sample of 336 observations of each predictor was obtained from 60 consecutive weather maps at 12 hour intervals, the observations are not independent in either time or space. Because of this, the tests of significance applied must be considered as necessary but not sufficient conditions for establishing the reality of the difference in predictability.

Table 2 shows the improvement in R^2 obtained by increasing the number of predictors and the values of the F ratio calculated from the following formula given by ANDERSON and BANCROFT (1952):

$$F = \frac{R_r^2 - R_k^2}{\left(1 - R_r^2\right) \dfrac{r-k}{n-r-1}} \tag{13}$$

Using this, one can determine whether the last $r-k$ of the total of r variates made a significant contribution to R_r^2, the percent reduction of the variance of the predictand, above the percent reduction R_k^2 due to the first k variates. The F ratio then is used to test the null hypothesis that each of the last $r-k$ regression coefficients is equal to zero. Here n is the sample size and $r-k$ would be the number of degrees of freedom of the numerator, the increase in variance. The quantity $n-r-1$ is the number of degrees of freedom of the unexplained variance, 1-R_r^2, in the sample.

On the basis of this test, it would appear that the only combination of 30 linear and 30 non-linear predictors which makes a significant contribution toward improvement of the R^2 of the predictands at the central forecast point over all smaller combinations of predictors is that designated by A_6 (columns 7 to 9 of Table 2). For 11 of these 12 tests involving A_6, the calculated F ratio indicates that the additional $r-k$ predictors make a significant contribution to the reduction at better than the 1 percent level.

The regression coefficients for the simple linear case for a 24 hour forecast of the 500 mb and 1,000 mb height changes are shown in fig. 2. Note in the upper left, the empirical influence function relating the 500 mb height to the height change. As might be expected, the greatest influence on the forecast point (circled) is from the height at that point itself.

Table 2. F ratios as measures of significance of improvement ($\triangle R^2$) of per cent reduction of variance of predicted height changes at the central forecast point using various arrangements of predictors. 5 and 1 per cent significance levels of F are given at foot of table.

		Predictors										
		1	2	3	4	5	6	7	8	9	10	11
		A_4 vs. A_1	A_4 vs. A_2	A_4 vs. A_3	A_5 vs. A_1	A_5 vs. A_2	A_5 vs. A_3	A_6 vs. A_1	A_6 vs. A_2	A_6 vs. A_3	A_3 vs. A_1	A_3 vs. A_2
$\triangle_{12}Z_5$	$\triangle R^2$	5.3	10.9	1.7	4.5	10.1	0.9	6.8	12.4	3.1	3.6	9.3
	F	2.11	4.34	1.33	1.73	3.89	0.68	2.87	5.25	2.66	2.83	7.22
	$\triangle R^2$	3.9	15.1	1.5	3.2	14.4	0.8	7.2	18.4	4.7	2.4	13.6
$\triangle_{24}Z_5$	F	1.32	5.14	0.99	1.06	4.77	0.50	2.78	7.12	3.68	1.66	9.28
	$\triangle R^2$	4.3	23.0	1.3	4.3	23.0	1.2	10.3	29.0	7.2	3.0	21.7
$\triangle_{12}Z_{10}$	F	1.46	7.76	0.87	1.43	7.72	0.82	4.44	12.52	6.26	2.06	14.76
	$\triangle R^2$	3.4	22.6	1.3	3.4	22.7	1.4	7.5	26.7	5.5	2.0	21.3
$\triangle_{24}Z_{10}$	F	1.15	7.76	0.92	1.18	7.80	0.97	3.03	10.85	4.43	1.39	14.66
	F_5:	1.66	1.66	2.10	1.66	1.66	2.10	1.66	1.66	2.10	2.10	2.10
	F_1:	2.06	2.06	2.91	2.06	2.06	2.91	2.06	2.06	2.91	2.91	2.91

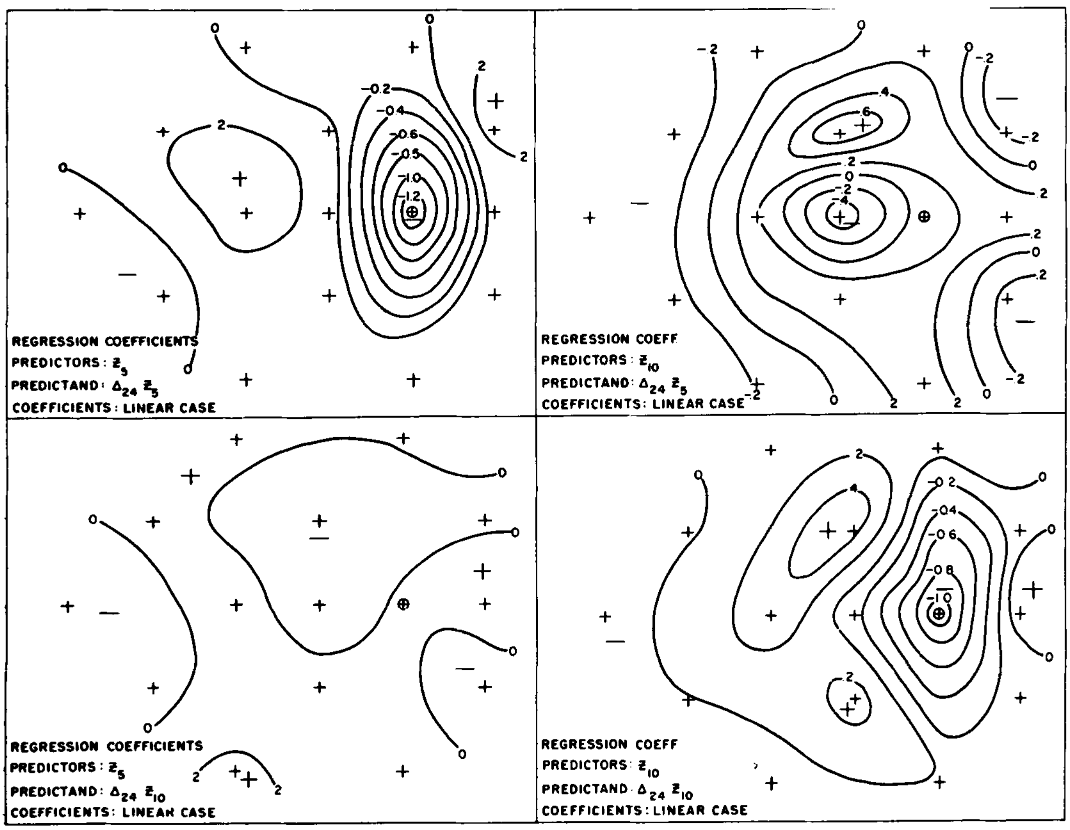

Fig. 2. Regression coefficients for the simple linear case for a 24 hour forecast of the 500 and 1,000 mb height changes using 15 values each of 500 mb and 1,000 mb height as predictors. The forecast is made for the circled point.

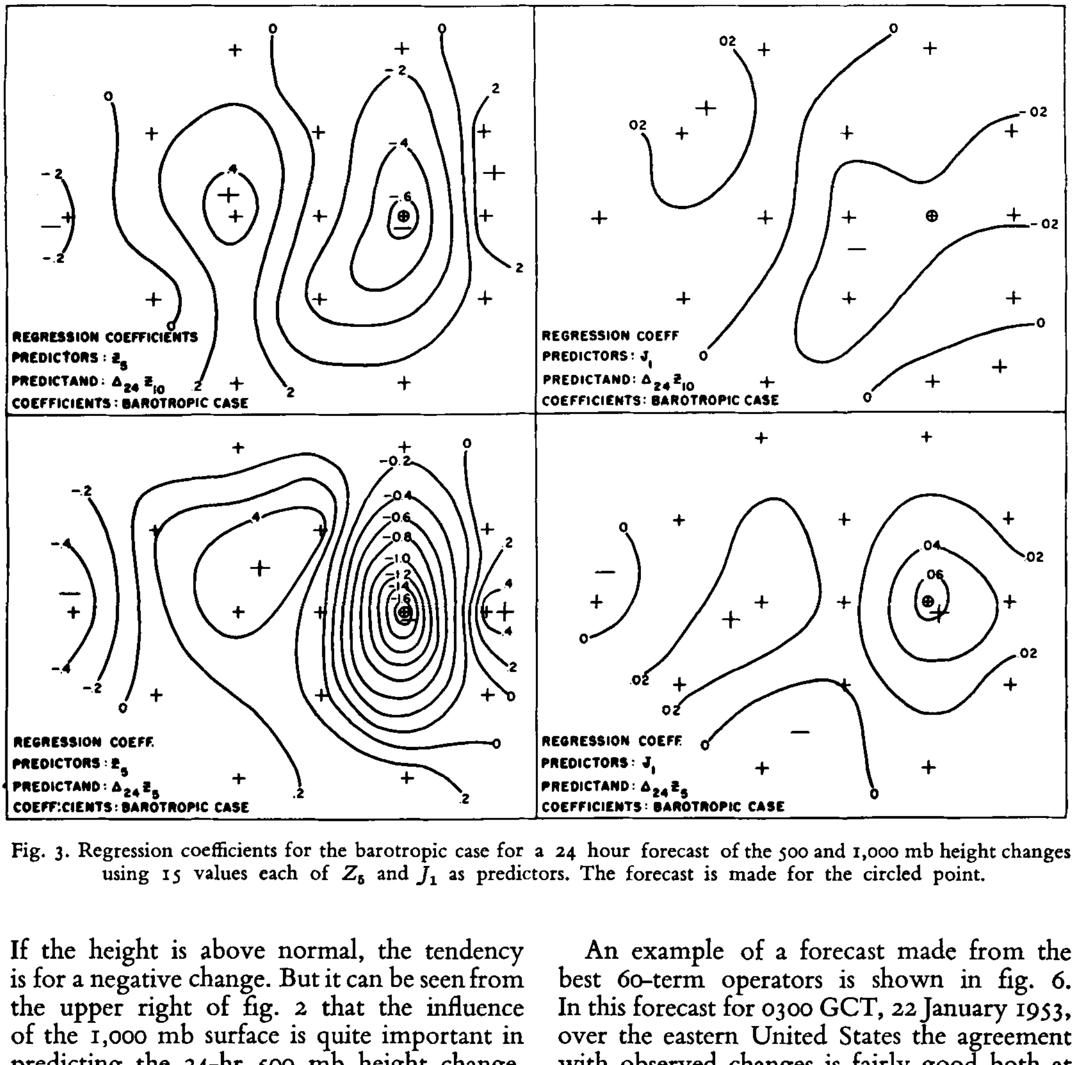

Fig. 3. Regression coefficients for the barotropic case for a 24 hour forecast of the 500 and 1,000 mb height changes using 15 values each of Z_5 and J_1 as predictors. The forecast is made for the circled point.

If the height is above normal, the tendency is for a negative change. But it can be seen from the upper right of fig. 2 that the influence of the 1,000 mb surface is quite important in predicting the 24-hr 500 mb height change. These empirical influence functions are of course more irregular than the Green's functions derived from the partial differential equations. In fig. 3, the maps show the regression coefficients for the case we have referred to as the barotropic case. Again the greatest weight is attached to the forecast point itself, but the barotropic vorticity advection term J_1 appears to affect the forecast only slightly.

Fig. 4 shows the coefficients for the forecast of $\Delta_{24}Z_5$ using 60 predictors which are values of the four variables Z_5, Z_{10}, J_1 and J_4. In fig. 5, the coefficients for forecasting $\Delta_{24}Z_{10}$ are shown using the same predictors as in fig. 4.

An example of a forecast made from the best 60-term operators is shown in fig. 6. In this forecast for 0300 GCT, 22 January 1953, over the eastern United States the agreement with observed changes is fairly good both at 500 and 1,000 mb. The correlation between forecast and observed 24-hour height changes at 500 mb is 0.91 and at 1,000 mb is 0.77.

6. Critical Comments

Some critical comments should be made with regard to the interpretation of the foregoing results.

1. In the process of solution of the partial differential equations, the field of the instantaneous height tendency $\dfrac{\partial z}{\partial t}$ is calculated and the values are assumed constant for a short time period, perhaps one hour. Using

4—803878

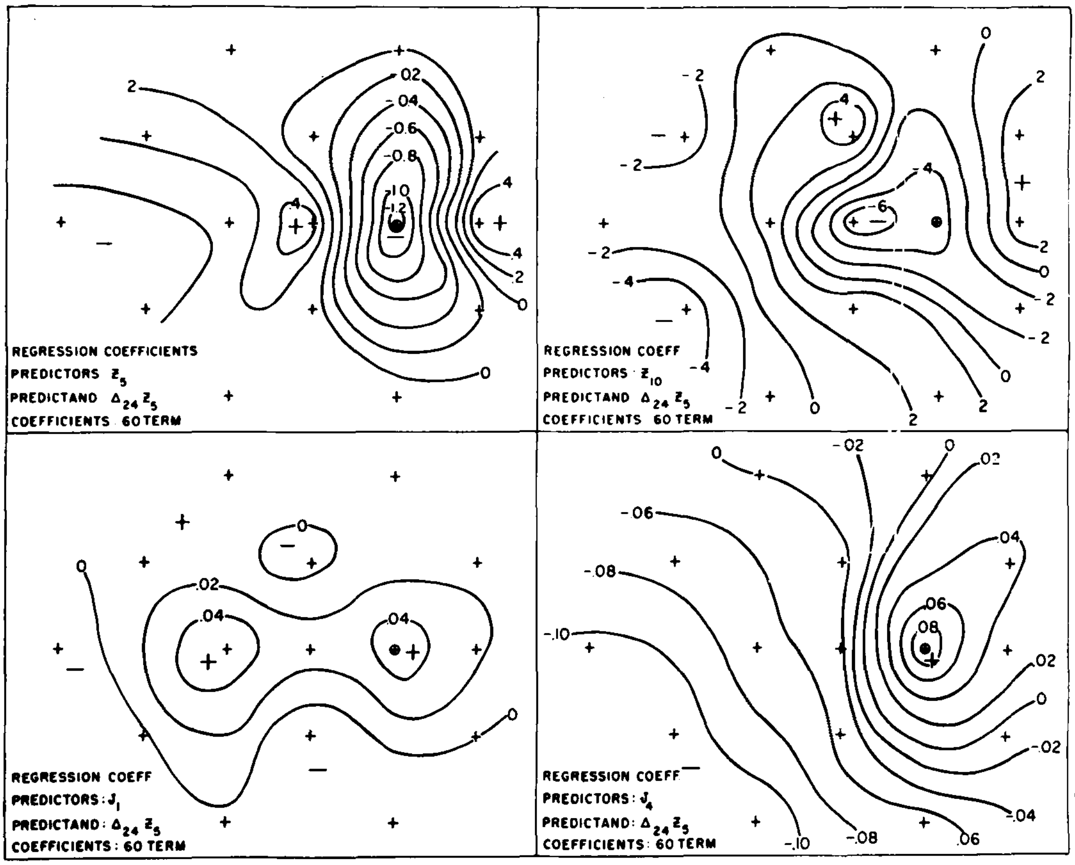

Fig. 4. Regression coefficients for a 24 hour forecast of the 500 mb height change using 15 values each of Z_5, Z_{10}, J_1 and J_4 as predictors. The forecast is made for the circled point.

these height changes, a new height field is obtained and the process is iterated to determine the height field at some later time. In the statistical process, the initial fields of the predictors F are related directly to the 12 or 24 hour height changes. Because of this difference, the conclusions drawn from the results of the statistical study do not permit one to make fully conclusive inferences about the effectiveness of non-linear terms in numerical weather prediction by purely dynamical methods.

2. It is of interest to speculate on the small increase of predictability due to the non-linear terms. Presumably, if a set of values of one of the non-linear terms were used alone as a predictor, it would give a significantly good prediction of the height change. But this prediction might not be significantly better than one could obtain from a linear combination of the heights. This implies that much of the information contained in the quadratic terms is already included in the linear terms. Since the non-linear terms involving the Laplacian are computed from third order differences of heights or thicknesses, errors in the data could be amplified and cause large errors in these terms so that they fail to aid the prediction significantly. This point may be illustrated by the fact that in this study the simplest Jacobian, J_4, gave more significant results than the Jacobians representing higher order differences.

3. If the statistical operators are applied to

independent data it can be expected that the percent reduction will be reduced. A discussion of the expected reduction of R^2 is given by LORENZ (1956).

4. The significance tests applied as described above must be considered as necessary but not sufficient conditions for establishing the reality of the differences in predictability. Since the 336 observations of each variable came from only 60 weather maps at each level and the maps are 12 hours apart, correlations among the predictor data must be rather high in both space and time. Consequently, the number of degrees of freedom used in the significance tests is not realistic and the high significance indicated for some of the cases is seriously questionable.

Although this study indicates that the non-linear terms add only slightly to the linear statistical prediction, one should not infer directly that their value is small in dynamical forecasting. However, it is not believed that the statistical and dynamical forecasting techniques are so greatly different that there are not some serious implications about the relative importance of these non-linear terms for dynamical forecasting also.

7. Acknowledgments

The writer is indebted to Dr. Robert M. White for his considerable advice throughout the course of this work. Thanks also go to Mrs. Florence A. Mulkern who assisted in the

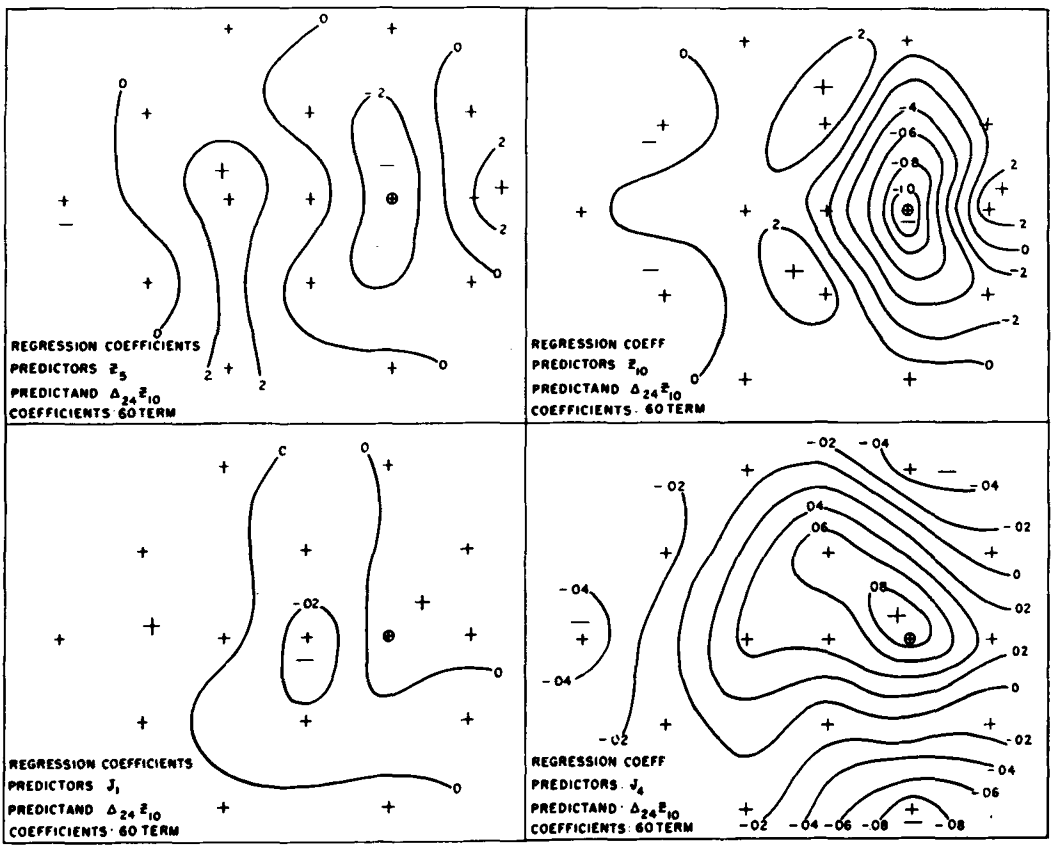

Fig. 5. Regression coefficients for a 24 hour forecast of the 1,000 mb height change using 15 values each of Z_5, Z_{10}, J_1 and J_4 as predictors. The forecast is made for the circled point.

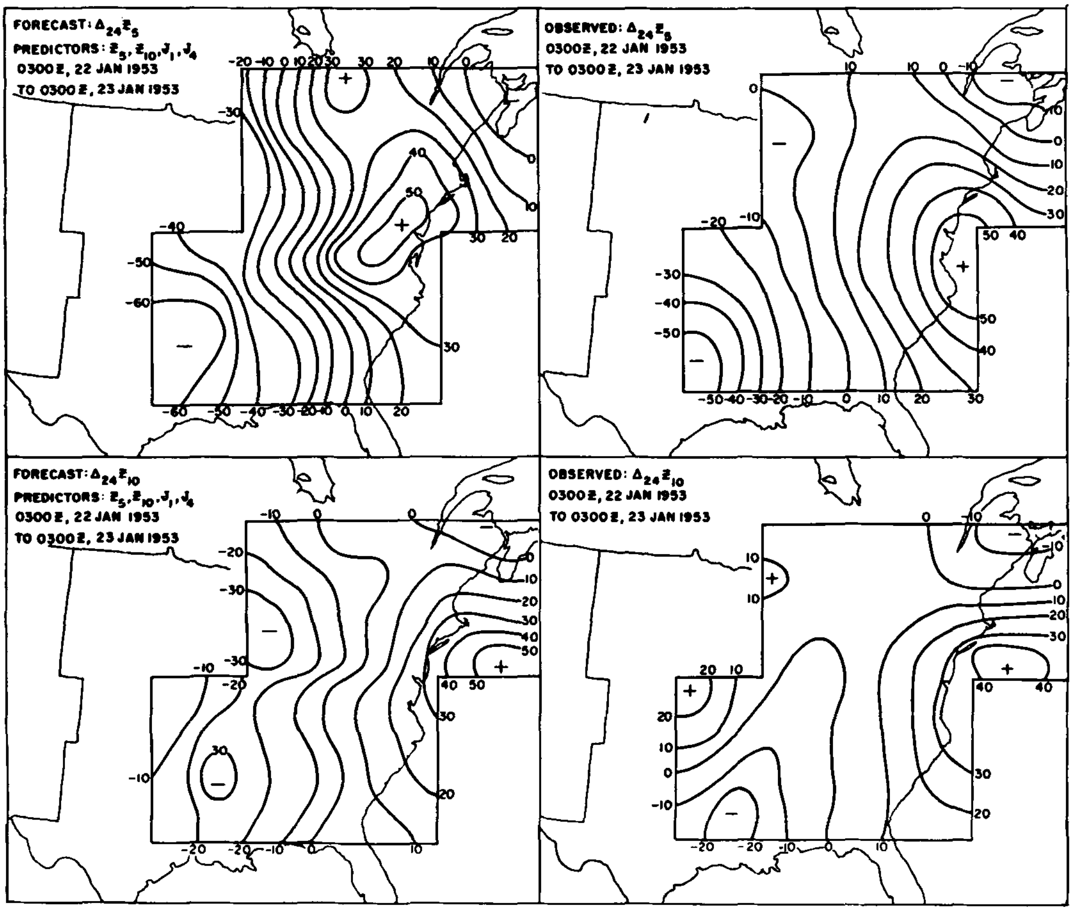

Fig. 6. Forecast and observed 24 hour height changes for the period ending 0300 GCT 23 January 1953. 500 mb above and 1,000 mb below. Forecasts were made using the best 60-term regression equations.

program preparation and in the production of computations on the IBM model 701 computer. Calculations were performed on an IBM 701 EDPM at International Business Machines Corporation in New York and on the IBM 701 at the Joint Numerical Weather Prediction Unit, U.S. Weather Bureau, Suitland, Maryland.

REFERENCES

ANDERSON, R. L., and BANCROFT, T. A., 1952: *Statistical theory in research*, New York, McGraw-Hill, pp. 168—172.

HILDEBRAND, F. B., 1952: *Methods of applied mathematics.* New York, Prentice-Hall, pp. 388—461.

HOUSEHOLDER, A. S., 1953: *Principles of numerical analysis.* New York, McGraw-Hill, 274 pp.

LORENZ, E. N., 1956: *Empirical orthogonal functions and statistical weather prediction.* Scientific Rep. No. 1, Contract AF 19(604)-1566, Cambridge, Massachusetts Institute of Technology, 49 pp.

MALONE, T. F., 1955: Application of statistical methods in weather prediction. *Proc. Nat. Acad. Sci.*, **41**, 806—815.

MILLER, R. G., 1956: Statistics and the predictability of weather. Abstract in *Bull. Amer. Meteor. Soc.*, **35**, 495—496.

THOMPSON, P. D., and GATES, W. L., 1956: A test of numerical prediction methods based on the barotropic and two-parameter baroclinic models. *J. Meteor.*, **13**, 127—141.

VON NEUMANN, J., and GOLDSTINE, H. H., 1947: Numerical inverting of matrices of high order, *Bull. Am. Math. Soc.*, **53**, pp. 1021—1099.

WADSWORTH, G. P., 1951: Application of statistical methods to weather forecasting. *Compendium Meteor.*, Boston, Amer. Meteor. Soc. 849—855.

WHITE, R. M., and PALSON, Jr., W. C., 1955: On the forecasting possibilities of empirical influence functions. *J. Meteor.*, **12**, 478—485.

WIENER, N., 1949: *Extrapolation, interpolation and smoothing of stationary time series.* Cambridge, Technology Press, Massachusetts Institute of Technology, 163 pp.