1. Classification *INPE-COM.4 (RPE)*
C.D.U.: 551.509.314

2. Period

4. Distribution Criterion
internal [ ]
external [X]

3. Key Words (selected by the author)
LONG RANGE WEATHER PREDICTION; ATMOSPHERIC CIRCULATION; STATISTICAL WEATHER FORECASTING

5. Report No.
INPE-1948-RPE/265

6. Date
November, 1980

7. Revised by
Vernon E Kousky
Vernon Edgar Kousky

8. Title and Sub-title
STEPWISE LINEAR REGRESSION EXPERIMENTS IN 14- AND 19- DAY HEMISPHERIC WEATHER PREDICTION

9. Authorized by
Parada
Nelson de Jesus Parada
Diretor

10. Sector *DME* Code

11. No. of Copies *06*

12. Authorship *Srinivasam Srivatsangam*

14. No. of Pages *25*

15. Price

13. Signature of first author

16. Summary/Notes

The zonal variance of the heights of constant pressure surfaces in the atmosphere is taken to be Markovian, i.e.,unpredictable over long times by its own history. It is shown through cross--correlation and multiple regression that the zonal variance is predictable over long times at a given location, through its lagged cross-correlation at other locations.

17. Remarks *Submitted to the International Symposium on statistical and Probabilistic Weather Prediction, Nice, France, 8-12 Sept. 1980.*

1948

INDEX

ABSTRACT

The zonal variance of the heights of constant pressure surfaces in the atmosphere is taken to be Markovian, i.e., unpredictable over long times by its own history. It is shown through cross-correlation and multiple regression that the zonal variance is predictable over long times at a given location, through its lagged cross-correlation at other locations.

LIST OF TABLES

LIST OF FIGURES

"Scientists have been little attracted by discussions on the limitation of science for the good reason that, so often, seemingly impossible things have been achieved."

Sverre Petterssen (11) on predictability.

1. INTRODUCTION

There has been a sea change in meteorology since Petterssen wrote that about a quarter of a century back. Indeed, the very year in which Petterssen's remarks were printed, a thesis published by Enger (3) commenced the process which has led to the abondonment of the index cycle as a subject of serious study, and to the questioning of the existence of any significant periodicities between the diurnal and annual periods in atmopsheric data. As the concept of the index cycle is associated with the name of Rossby (13, 14), whose other ideas have been so profitable to meteorology, we are made to wonder if that concept is really baseless.

Following Enger's thesis, Shapiro and Ward (15, 17), Lorenz (6) and Julian (5) have examined the zonal index as well as other weather-related variables, and have found no evidence for the existence of any significant cycles except the diurnal and annual ones. Enger (3) showed that the behaviour of the zonal index may be closely approximated by a first order, linear Markov process, (This process is described in the references cited in this paragraph.) Enger thus showed that the zonal index variations are mostly due to persistence and random fluctuations. Lorenz (6) confirmed this, demonstrating also that the kind of quasi-periodic fluctuations with periods of several weeks suffered by the zonal index may also exist in first order, linear Markov processes. In the following, as there are no grounds for ambiguity, I will call the first order, linear Markov process, simply the Markov process and any variable, which has a behaviour similar to this process, a Markov variable.

Shapiro and Ward (15), analysing a seven year, four month series of 500-mb data, found that "except for the annual period, the energy of the north-south flow has no preferred periods of oscillation. This is true not only for the total flow but also for its (spatial) spectral components. Furthermore, there is no evidende for the existence of the "index cycle". Ward and Shapiro (17) extended this conclusion to other variables, including the zonal index at several locations. They found the only distinghising property of their various spectra (other than the prominence of the daily and yearly periods) to be their resemblance to "red noise", i.e., generally higher variance at the lower frequencies (or longer periods). Thus (almost) all atmospheric variables have come to be characterised as Markovian, all producing "red noise" spectra. Accordingly, (almost) all atmospheric variables are lacking in predictability based on their past history, i.e, they lack auto-predictability.

Now, auto-prediction is a specialisation of multiple regressive prediction. It is also knows that multiple regression frequently constitutes a jungles. I would not jump into it lightly. The rest of this paper also discussed a specialisation of multiple regression, although the specialisation is not to the evidently fruitless auto-regression. The *via media* taken here may be explained and given motivation as follows.

Consider a set of Markov variables as the set evolves. At each interval of time, the set forms a new pattern.

If the lag between two patterns is small, persistence would imply a great similarity between the two patterns. By this I mean all os the following three properties:

That the two patters have locally nearly the same values or that local changes in the variables are small.

That the linear correlation coefficient obtained from local values in the two patters is large.

That, the local changes being so small, the similarity between the two patterns can be grasped by the eye.

Therefore, in the terminology of Lund (9), the two patterns would be of the same "type", and in the terminology of Lorenz (7, 8), the patterns would constitute a pair of "analogues". I will say that the two patterns at small lag belong to an "explicitly related type.

Consider next the same series, but pairs of patterns separated by a long, fixed span of time. The pairs will not be influenced by persistence, and, therefore, there will be little explicit similarity. However, it is possible to extend the correlation-coefficient criterion given above. Let us relate the value of one Markov variable, at the later times, not just to its own values at the earlier times, but to the values of the other variables at the earlier times as well. Thus we form linear combinations of predictors. If these linear combinations were to relate the earlier and later time patterns well, we evidently have a resemblance, but not of the explicit type. That is, the resemblance may not be perceived by the eye, as the changes in local values may be very large in any pair of patterns. I will call this "implicit resemblance". Two patterns at large lag will rarely, if ever, belong to a stringently defined "explicitly related type" (7), but they may be more closely related through "implicit resemblance".

My aim here is to show that there is much implicit resemblance, at large lags, in one meteorological variable closely related to the zonal index.

2. THE MODEL AND THE DATA

Let two parallel vertical lines denote zonal averaging and an asterisk the departure from the zonal mean so t at a variable z may be represented by

$$z = | z | + z^* \qquad (1)$$

It follows from Eq. (1) that $|z^{*2}|$ would represent the zonal variance of z. Hence $|z^{*2}|$ would be a measure of the deviation from purely zonal structure or flow. And if z were the height of an isobaric surface, then $|z^{*2}|$ would be negatively correlated with the zonal index. The prediction of the zonal variance of isobaric surface height is the subject of this study.

The study uses a 120-month set of daily 1200 GMT data of the heights of five isobaric surfaces. It covers the period from 1963 to 1972. It is a collection of the U.S. National Meteorological Center (NMC), made available for this study by the National Center for Atmospheric Research (NCAR) at Boulder, Colorado. The data are for the 700-, 500-, 300-, 200- and 100- mb levels, in each instance from about 20°N to the North Pole. These data were reduced to $|z^{*2}|$ values at 20°N, 25°N, 30°N, ..., 85°N. In ther number, in $|z^{*2}|$ (j,k), j will represent one of these latitudes and k one of the isobars. For each day, then, there are seventy values of $|z^{*2}|$.

The 14-day predictions were made for the second half of each calendar month and the 19-day forecasts for four mid-season priods. I will describe the procedure adopted for the 19-day forecasts. The one for the 14-day predictions may be inferred from the description given here or seen in (16).

From each of the ten years, I separated the data for January 1 to February 9. This set of 10 x 40 patterns of $|z^{*2}|$ was used in making the 19-day winter forecasts. Similar subsets of data were extracted for the other seasons, commencing on April 1, July 1 and October 1. In each case, the set of 400 patterns was divided into two halves, viz., 10 x (day 1 to day 20) and 10 x (day 21 to day 40). The first half was used as the source of independent variables and part of the second as the source of dependent variables in computing regression coefficients. In order to keep the two halves of data statistically independent for later use, the ensemble neam value of $|z^{*2}|$ were computed from the first half only as

$$\widetilde{|z^{*2}|}\ (j,k) = \sum_{d=1}^{20} \sum_{y=1}^{10} |z^{*2}|\ (j,k)\ /\ 200 \qquad (2)$$

In Eq. (2), the tilde denotes ensemble averages, d the day from the first of the month as above and y the year. The departure from this ensemble mean is given by

$$Z(j,k) = |z^{*2}|\ (j,k) - \widetilde{|z^{*2}|}\ (j,k) \qquad (3)$$

for either half of data. This new variable Z will be used here rather than $|z^{*2}|$.

The stepwise regression method is a standard one and is described at lenght in many statistical texts, e.g., Chatterjee and Price (1) and Draper and Smith (2). Readable accounts of regression related problemas are Rao and Miller (12) and Haan (4). The levels of significance for the inclusion and exclusion of variables from the regression equations were set equal to each other at the one per cent level. The subroutine package used was the International Mathematical and Statistical Libraries (IMSL), available with many high-speed computing facilities.

In the stepwise regression scheme, any number of variables may be chosen from a set of independent variables. In the present problem, the seventy spatial values of Z constitute the sets of idependent as well as dependent variables. Therefore, to predict one Z (j,k), say Z (4,3) and any one or more of the remaining sixty-nine Z (j,k) as predictors. Usually, stepwise regression as used here leads to linear combinations of predictors. Since we need equations to predict the seventy Z (j,k), it follows that up to 70 x 70 predictors may be chosen to predict the Z (j,k) at a future time. Thus one normally selects from the of seventy independent variables, more than seventy predictors for the Z (j,k).

In the procedure I adopted, a new set of regression equations is developed for each interval of prediction for each season. As mentioned above, the independent variables are always derived from the first half of data, viz., days 1 to 20 of the months mentioned, of all the ten years. Let I denote the interval of prediction in days. Then, for the 1-day interval, the dependent variables are the data of days 1+i to 20+i or days 2 to 21. Proceeding in this manner, the dependent variables for the 19-day interval are the data of days 20 to 39. From these independent variables and dependent variables, the regression equations are developed for each of the 19 intervals.

3. VERIFICATION PROCEDURE AND RESULTS

The properties of the regression equations were studied with a nearly independent set of data. I assumed initial conditions to be the data of day 21. Then the successive sets of regression equations were used to make predictions for days 22 to 40. The observed data for these days were used to make forecast verification.

I said the verification data were nearly independent of the data used for computing the coefficients. This may be confimed by noting that the initial-conditions data, those for day 21, were used only as dependent variables in the computation of coefficients and that the forecast days 22 to 40 were never used in any manner in the computation.

Let $Z_r(j,k)$ be the forecast values of $Z(j,k)$. Then the root mean square (rms) error of the regression-model forecasts for an individual day is

$$E_r = \left(\sum_{j=1}^{14} \sum_{k=1}^{5} | Z(j,k) - Z_r(j,k)|^2 \ / \ 70\right)^{\frac{1}{2}} \tag{4}$$

The error of a model forecast is sometimes compared with the error of climatological forecasts. Since, however, "climate" may be defined variously, at least for quantitative purposes, there may be various climatological forecast errors as well. Here I wil compare E_r with a constant. This constant E_c is the average rms error of a model that always forecasts Z to be zero, and is defined by

$$E_c = \left(\sum_{y=1}^{10} \sum_{d=22}^{40} \left(\sum_{j=1}^{14} \sum_{k=1}^{5} z^2(j,k) \ / \ 70\right)^{\frac{1}{2}}\right)/190 \tag{5}$$

The forecast results are presented here in the form of the per cent ratio

$$R_1 = 100 \ E_r/E_c \tag{6}$$

Let $\overline{E}_r$ be the model forecast errors averaged for a given calendar day of all the ten years. $\overline{E}_r$ would be the mean error of the separate forecasts made at a given interval. The mean values of R_1 corresponding to $\overline{E}_r$ are presented in Table 1A for the 19-day forecasts and in Table 2A for the 14-day forecasts. I will consider here Table 1 only, leaving the reader to compare it with Table 2. The average of R_1 for the four seasons are also presented in Table 1A.

These are regressive forecasts and one know that if the forecastsreveal Ex $(E_r) > E_c$, where Ex is the statistical expectation, then the regression coefficients leading to E_r were inexact computed (although these might have been statistically significant for the sample used). Now the values of the errors in Table 1A reveal this, Ex $(E_r) >$ $> E_c$, in the interval from about day 25 to day 32, primarily. It follows that the regression coefficients computed for this interval were imprecise and that the cross-correlations between Z (j,k) at lags of 4 to 11 days were not stable enough to be indubitably determined.

In contrast to the errors over 4 to 11 lags is the remainder of Table 1A. At lags larger than 6 (day 27), we see a tendency for the errors to *diminish* with time, up to a lag of 17 (day 38).

One may be conclude from the above that there may be a reduction in the number of predictors chosen at lags of 4 to 11, in comparision with other lags. In Fig. 1, the four-season average of predictors chosen at each lag, called mean N, is plotted against the average R_1 of Table 1A. Unit-lag data have been omitted in Fig. 1. A clear correlation between mean N and lag can be inferred from Fig. 1 and Table 1A. At intermediate lags, when average R_1 is greater than 90, mean N tends to be small.

A further issue is the obliteration of forecast magnitude. By this I mean the tendendy if any of the model to forecast Z such that the rms value of Z diminishes with time. To formulate this problem quantitatively, I will write

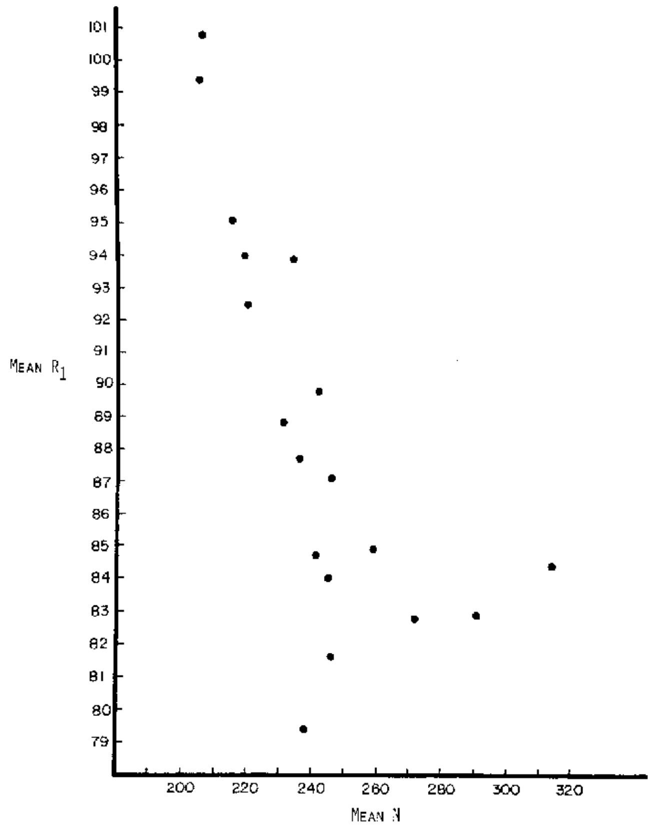


Figure 1 - The average of R_1 for all four seasons from Table 1A plotted against the corresponding average of predictors N. R_1 is a normalised error of forecasts. All forecast intervals except the 1-day interval are included.

TABLE 1

RESULTS FROM THE 19-DAY FORECASTS OF THE ZONAL VARIANCE OF ISOBARIC SURFACE HEIGHT.

FOR EACH DAY OF EACH MONTH, TEN FORECASTS WERE MADE.

THE TABULAR VALUES ARE THE AVERAGES FOR THE TEN FORECASTS.

Month	Day 22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
A. Errors of Model Predictions																			
Jan.	47.7	70.1	72.5	71.4	69.1	79.6	96.3	95.7	97.7	97.2	97.3	94.8	95.2	94.0	94.3	91.7	89.5	89.7	92.1
Apr.	48.8	72.0	82.1	86.3	91.2	91.6	86.2	82.4	84.4	82.8	75.2	73.4	69.8	68.4	67.6	68.0	64.4	64.0	63.6
Jul.	56.6	83.4	81.4	99.0	100.7	117.9	110.7	104.2	107.4	101.9	96.1	92.4	89.0	86.5	85.6	83.0	84.5	81.7	80.7
Oct.	61.0	92.0	90.4	98.3	109.0	113.9	104.2	98.1	86.3	93.8	90.4	90.3	94.4	90.5	91.2	93.1	92.6	96.3	101.3
Average	53.5	79.4	81.6	88.8	92.5	100.8	99.4	95.1	94.0	93.9	89.8	87.7	87.1	84.9	84.7	84.0	82.8	82.9	84.4
B. Rms Values of Model Predictions																			
Jan.	89	81	73	58	57	56	56	48	45	48	50	49	48	49	51	51	48	56	64
Apr.	82	74	71	64	62	61	67	65	64	69	69	69	68	73	74	77	78	83	84
Jul.	79	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	41	44	46	42	47	49	50	51	52	56	57	59
Oct.	96	90	86	86	82	74	66	57	50	51	54	60	65	60	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Average	86.5	78.3	71.8	66.3	62.0	57.5	56.0	52.8	50.8	53.3	53.8	56.3	57.6	58	58.3	60.0	62.0	67.4	[illegible]

TABLE 2

RESULTS FROM THE 14-DAY FORECASTS OF THE ZONAL VARIANCE OF ISOBARIC SURFACE HEIGHT FOR EACH DAY OF EACH MONTH, TEN FORECASTS WERE MADE THE TABULAR VALUES ARE THE AVERAGES FOR THE TEN FORECASTS.

Month \ Day	17	18	19	20	21	22	23	24	25	26	27	28	29	30
B. Rms Values of Model Predictions														
Jan.	99	86	76	67	67	60	59	54	57	57	52	49	51	53
Feb.	92	79	69	63	59	63	57	56	64	70	71	73	73	74
Mar.	112	94	80	74	73	74	72	81	77	70	78	74	72	68
Apr.	97	84	79	66	63	61	51	52	57	61	65	64	72	77
May	76	61	57	57	54	52	54	52	53	57	58	60	62	66
Jun.	79	60	57	60	61	55	54	54	56	58	60	40	63	67
Jul.	73	55	43	42	39	37	36	41	56	56	52	53	55	50
Aug.	78	63	58	56	55	56	61	64	59	60	59	61	55	54
Sep.	83	76	72	64	59	56	52	50	57	67	72	71	69	71
Oct.	87	67	56	50	46	43	47	54	58	62	63	70	69	71
Nov.	77	65	60	63	63	65	64	64	63	59	60	60	64	66
Dec.	58	47	44	40	40	45	49	48	50	54	44	49	48	49
Average	84.3	69.8	62.6	58.5	56.6	55.6	54.7	55.8	58.9	60.9	61.2	62.0	62.8	63.8
A. Errors of Model Predictions														
Jan.	52.0	76.1	94.9	100.7	106.7	98.0	106.3	108.8	102.5	98.1	103.2	98.0	99.4	104.7
Feb.	42.2	64.4	84.7	93.4	103.3	116.5	115.8	109.7	104.7	107.3	118.0	115.9	99.3	88.5
Mar.	68.5	79.4	86.3	90.5	91.0	92.8	91.3	95.8	88.3	89.3	92.3	89.2	86.4	80.4
Apr.	51.5	69.7	80.9	84.2	88.5	85.0	82.9	85.9	90.1	89.8	91.1	78.9	75.3	71.1
May	54.3	73.8	76.6	81.5	87.9	89.9	82.4	92.5	95.6	109.5	101.8	104.7	109.9	114.0
Jun.	53.0	73.6	91.4	104.6	99.6	83.3	82.2	80.3	78.0	85.3	87.8	94.8	97.4	98.3
Jul.	65.9	86.8	97.9	85.1	92.5	99.7	96.9	88.4	103.3	86.7	97.0	105.5	99.6	98.4
Aug.	54.1	63.3	77.1	81.0	94.5	104.2	104.0	94.5	85.7	93.9	108.6	95.2	96.1	92.6
Sep.	54.1	73.2	86.9	91.8	97.6	89.0	78.5	79.6	89.5	98.8	95.7	112.3	110.1	105.5
Oct.	57.5	77.8	81.4	100.5	109.0	113.0	119.5	103.7	97.5	95.8	96.1	98.9	94.3	83.8
Nov.	40.9	62.1	71.6	82.0	85.7	70.3	79.2	89.2	90.6	91.7	98.8	104.0	97.0	94.0
Dec.	49.3	85.7	96.9	106.0	97.8	85.2	88.9	92.8	96.9	106.5	100.5	101.2	104.5	104.2
Average Lorenz	53.6	73.8	85.6	91.8	96.2	93.9	94.0	93.4	93.6	96.1	99.2	99.9	97.4	94.6

$$S_r = \left(\sum_{j=1}^{14} \sum_{k=1}^{5} Z_r^2 (j,k) \;/\; 70 \right)^{\frac{1}{2}} \tag{7}$$

and

$$R_2 = 100\, S_r \;/\; E_c \tag{8}$$

To determine the extent of obliteration, I then ask the question: Does R_2 decrease with time on the average?

The seasonal averages of R_2 are given in Table 1B. (See Table 2B for the 14-day forecast values.) It also contains the overall averages of R_2 for the four seasons. This information answer the quastion asked above in the negative. R_1 and R_2 are antiparallelly distributed in time. Generally speaking when one of them is large, the other is small. Therefore the worst forecasts, those made at an interval of about a week, are also the most obliterated ones. The later forecasts are not only less error-prone, but also well-predicted in amplitude. This is in complete contrast to the auto-regressive forecasts of Lorenz (8).

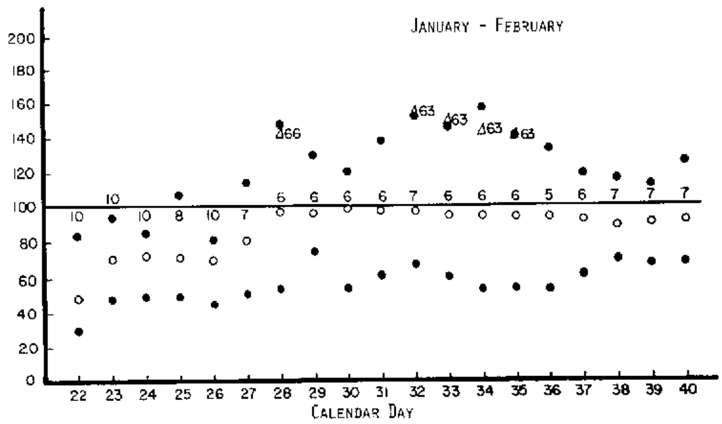
Figure 2 contains a few more details. It has the maximum and minimum values of R_1 for each day. I have also entered in Fig. 2 the number of instances, out of ten, for each day, of the forecasts for which $R_1 > 100$. Figure 2 also has some values of R_3 defined through

$$R_3 = 100\, E_{cd} \;/\; E_c \tag{9}$$

where

$$E_{cd} = \left(\sum_{j=1}^{14} \sum_{k=1}^{5} Z^2(j,k) \;/\; 70 \right)^{\frac{1}{2}} \tag{10}$$

Hence E_{cd} is just the daily rms value of Z, and R_3 is a normalised version of this rms value. In Fig. 2, if any value among the maxima of R_1 were greater than or equal to 140, the corresponding value of R_3 as observed that day is also given. This enables one to assess a



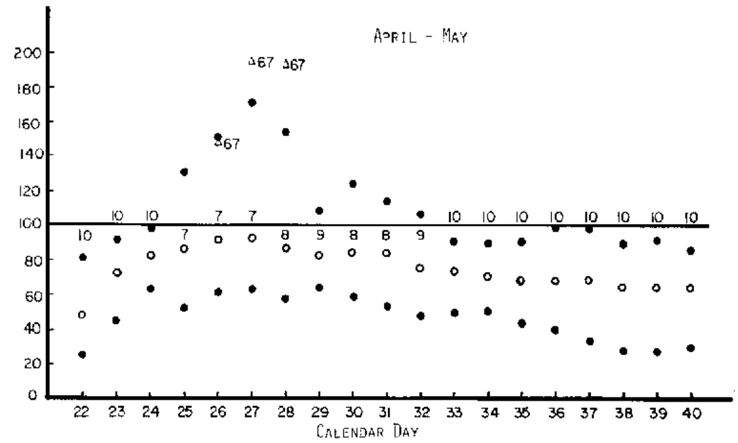
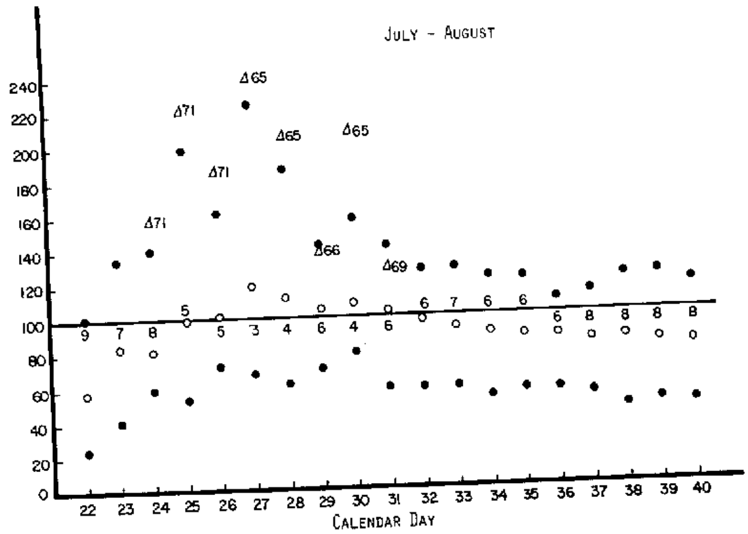


Figure 2 - Information from the 19-day forecasts. "Calendar day" is the integer count from the first day of a) January, b) April, c) July and d) October. The maximum and minimum values of the normalised error R_1 for each day are indicated by dots. The average value of R_1 by an open circle and the values of R_3, corresponding to those maxima of R_1 which equal or exceed 140, by triangles. The numbers beside the triangles indicate the year of occurrence. The numbers on either side of the line for ordinate value 100 are counts (out of ten) of forecasts for which R_1 was less than 100, i.e., the number of "better than average forecasts".



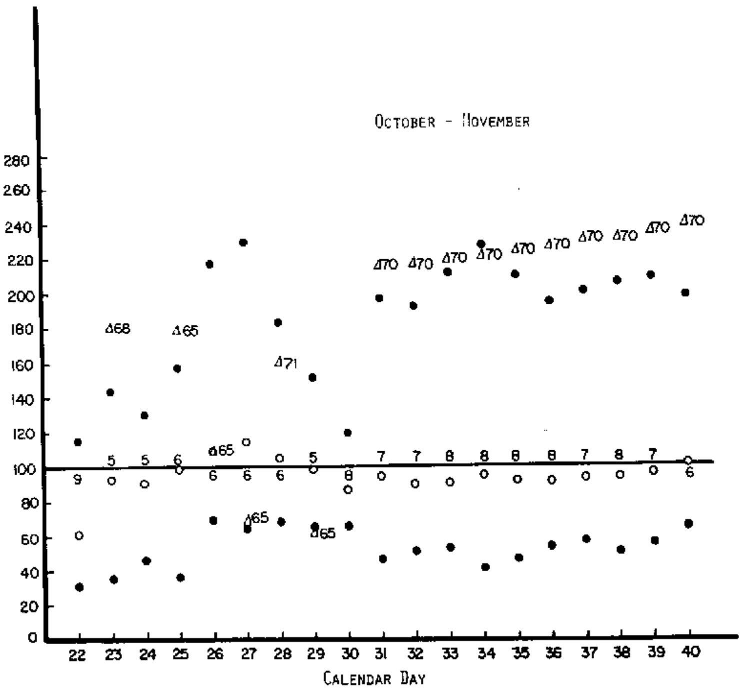


Figure 2 (Continuation)

partial cause of the worst forecasts. For instance, among the forecasts for January 28, the worst one has a value of R_1 in excess of 140. That forecast was for January 28, 1966. The associated value of R_3 that day, as indicated in Fig. 2, was also more than 140. So we see that one instance of a very deviant weather, as measured by R_3, was badly predicted. This is true of several other examples in Fig. 2, so that there is a tendency in the predictions *not* to anticipate the dratically deviant weather. Such a property is common to most regressive prediction models.

4. DISCUSSION

The strangest feature of these results, the one that demands reconciliation with our previous knowledge of the behaviour of the atmosphere, is, clearly, the diminution of the forecast error with time, after the passage of about ten days from initial conditions. Such a recovery has not been anticipated in any of the studies of atmospheric predictability, at least not as anything except mere hope.

Furthermore, the behaviour of the errors in these forecasts indicates a periodicity quite close to the periodic range attributed to the index cycle. Yet one must make note of what Namias (10) wrote of the usefulness of the index cycle: "While the index (cycle) concept has been of considerable help in classifying and studying general circulation patterns, it must be confessed that its usefulness in the practice of extended forecasting has been somewhat disappointing." If the results presented here are a reaffirmation of the existence of the index cycle, a new interpretation is required to connect them and the older, synoptic view.

It is nearly certain that the small values of the forecast error at small prediction intervals must be attributed to persistence, despite the formation of linear combination of predictors then also, as seen from the distribution of mean N in Fig. 1. Skills revealed by this models at a range less than about ten days are not substantially superior to the auto-regressive results of earlier studies, e.g., Lorenz (8).

To elucidate the initial increase of forecast error and its subsequente fall, I propose a new definition of the index cycle:

The index cycle is a quasi-periodic manifestation of several weeks duration arising from atmospheric wave-mean flow interaction modified by other forcing, such as the various forms of diabatic heating. It is characterised by strong asymmetry of eddy growth and decay phases, being essentially nonlinear. Weather patterns at the beginning and end of a cycle are only inplicitly related. This relationship, while meaning the visual dissimilarity of the same pahse of different cycles, nevertheless means that the past history of the atmosphere contains information - unrelated to its external forcing - on the future course of the atmosphere.

From this definition, it follows that the skill recovered by the forecasts after the first week of forecasting is due to the implicit similarities arising from the index cycle.

I attribute the worsening of forecasts over the first several days to the nonlinearity of the wave-mean flow interactions in the atmosphere.

5. SUGGESTIONS

I do hope that this study would lessen some of the gloom arising predictability studies and would prove to be of practical use. But,of course, to be effective it must be followed by other studies, at least partially confirming it.

I am fully aware that terms like "implicit resemblance", while conveying the implications of the results presented here, do not constitute a theory. It seems to me, in view of the development of theory in the past, that this complex set of results will not be explained easily or immediately. Without a credible theory, however, our understanding will continue to be nebulous as in the past.

Empirical studies bearing upon the issues raised here must also deal with the relationship between the local weather and the zonal statistical moments.

ACKNOWLEDGMENTS

This research was supported in parte by U.S. Naval Supply Center Grant 0228-76-C-3205 to Colorado State University. Free computer time was provided by the National Center for Atmospheric Research at Boulder, Colo., which is supported by the Atmospheric Science Division of the National Science Foundation.

This research was also supported by Financiadora de Estudos e Projetos (FINPE) through Convênio B28/79/002/00/00, Modelagem Atmosférica - FINEP/CNPq.

REFERENCES

CHATTERJEE, S.; PRICE, B., 1977: *Regression Analysis by Example.* Wiley.

DRAPER, N.R.; SMITH, H., 1966: *Applied Regression Analysis.* Wiley.

ENGER, I., 1957: Some attempts at predicting a meteorological time series from its past history. S.M. thesis, M.I.T., Cambridge, Mass.

HAAN, C.T., 1977: *Statistical Methods in Hydrology.* Iowa State Univ. Press, Ames.

JULIAN, P.R., 1966: The index cycle: A cross spectral analysis of zonal index data. *Monthly Weather Review,* 94, 283-293.

LORENZ, E.N., 1965: On the possible reasons for long-period fluctuations of the general circulation. In, WMO Tech. Note 66, WMO Secretariat, Geneva.

______, 1969: Atmospheric predictability as revealed by naturally occuring analogues. *Journal of Atmospheric Sciences,* 26, 636-646.

______, 1973: On the existence of extended range predictability. *Journal of Applied Meteorology,* 12, 543-546.

LUND, I., 1963: Map-pattern classification by statistical methods. *Journal of Applied Meteorology,* 2, 56-65.

NAMIAS, J., 1950: The index cycle and its role in the general circulation, *Journal of Meteorology,* 7, 130-139.

PETTERSSEN, S., 1957: Weather observations, analysis, and forecasting. *Meteorological Monographs*, 3, Nº. 15, American Meteorological Society, Boston, Mass.

RAO, P.; MILLER, R.L., 1971: *Applied Econometrics.* Wadsworth Publ., Belmont, Calif.

ROSSBY, C.G., 1939: Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. *Journal of Marine Research,* 2, 38-55.

______; WILLET, H.C., 1948: The circulation of the upper troposphere and lower stratosphere. *Science,* 108, 643-652.

SHAPIRO, R.; WARD, F., 1960: The time-space spectrum of the geostrophic meridional kinetic energy. *Journal of Meteorology*, 17, 621-626.

SRIVATSANGAM, S.; HSUAI, C.N.; REITER, E.R., 1977: Quasi-periodicities of atmospheric circulation and their application to long-range weather prediction. Environmental Research Paper 13, Colorado State University. Fort Collins.

WARD, F,; SHAPIRO, R., 1961: Meteorological periodicities. *Journal of Meteorology,* 18, 635-656.