

UNCLASSIFIED

AD 287 465

*Reproduced
by the*

ARMED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA



UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

CATALOGUED BY AD11A

AS AD 110.

287465

287 465



The Linear Predictability of Precipitation

Thomas M. Noel

Massachusetts Institute of Technology

Department of Meteorology

Cambridge, Massachusetts

Research Report Number 35

Prepared for the

Geophysics Research Directorate

Air Force Cambridge Research Laboratories

Office of Aerospace Research

United States Air Force

Bedford, Massachusetts

AF19(601) 5220

THE LINEAR PREDICTABILITY OF PRECIPITATION

by

Thomas M. Noel

Weather Radar Research

Department of Meteorology

Massachusetts Institute of Technology

Cambridge, Massachusetts

Contract No. AF19(604)-5220

Project No. 4641

Task No. 46414

RESEARCH REPORT NUMBER 35

September 1962

Prepared for

Geophysics Research Directorate

Air Force Cambridge Research Laboratories

Office of Aerospace Research

United States Air Force

Bedford, Massachusetts

Requests for additional copies by Agencies of the
Department of Defense, their contractors, and other
government agencies should be directed to the:

ARMED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA

All other persons and organizations should apply
to the:

U.S. DEPARTMENT OF COMMERCE
OFFICE OF TECHNICAL SERVICES
WASHINGTON 25, D. C.

THE LINEAR PREDICTABILITY OF PRECIPITATION

by

Thomas N. Noel

ABSTRACT

The problem of the short-term predictability of precipitation has remained rather unamenable to investigation by dynamic methods for reasons which are well known to meteorologists. The overall purpose here is to make a diagnostic study of this problem using a linear prediction scheme based upon statistics of precipitation intensity. The data necessary for the determination of such statistics were obtained by the aid of a horizontally scanning radar with the signal returned by the precipitation being a measure of its intensity. It is found that (1) the linear predictability of precipitation is relatively insensitive to reasonable variations in the spatial density of the data, however the spatial extent of the data may be a critical parameter; (2) the non-linear behavior of precipitation elements whose sizes range from one mile up to 100 miles in order of magnitude is important for the short term predictability; (3) the space resolution with which a forecast is made will influence to some extent the level of predictability; and (4) the quality of linear prediction decreases with time as expected but not in a systematic manner because of the presence of different sizes of precipitating elements in the storms from which the data were obtained.

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	THE DATA	2
III.	THE PREDICTION METHOD: MULTIPLE LINEAR REGRESSION	4
IV.	REDUNDANCY OF INFORMATION AND USE OF MORE THAN ONE TIME LAG	7
	A. Spatial density of data	7
	B. Spatial extent of data	9
	C. Two time lags	11
V.	THE METHOD OF MEASURING STATISTICAL SIGNIFICANCE	12
VI.	THE RESULTS OF THE REDUNDANCY AND TWO TIME LAG EXPERIMENTS	17
	A. Predictability in the mean as a function of prediction time	17
	B. Predictability as a function of space	21
	C. Further speculations	23
	D. Summary and criticism	24
VII.	METEOROLOGICAL INTERPRETATION OF THE RESULTS	25
	A. Non-meteorological factors	26
	B. Meteorological factors	27
	C. The nature of the precipitation	30
	D. Interpretation of predictability as a function of lag time	32
	E. The interpretation of the results of the redundancy and two time lag experiments	36
	F. An alternative method of measuring predictability	38
VIII.	CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH	40
	REFERENCES	41
	FIGURES AND TABLE	42

I. INTRODUCTION

This report presents a study of the short term predictability of precipitation based upon a linear scheme. The precipitation pattern and its changes are determined from photographs of the plan position indicator (PPI) of a weather radar system taken during storms occurring in both the winter and summer seasons but not including thunderstorm situations. The linear prediction scheme involved the use of a multiple regression method. The investigation reported here is a continuation of the research which was initiated by Noel and Fleisher (1960).

The overall purpose of this study is to make a diagnosis of the problem of short term predictability of precipitation rather than to make the best possible linear forecast. To achieve this purpose, the study is primarily directed toward finding (1) the quality of a forecast based upon a linear prediction scheme, (2) where the linear predictive information is located with respect to the station for which the precipitation intensity (radar signal) is being predicted, (3) how much predictive information is necessary, and (4) what meteorological factors control the answers to (1), (2), and (3).

The work reported earlier dealt with linear predictability as a function of prediction time at different stations in a grid of fixed size and fixed data density; (1) and (2) above were among its primary concerns. Some results indicated that predictability of the radar signal did not require the spatial density of data to be as high nor the spatial extent of data to be as large as were used, and that predictability might be improved if, instead of looking only at contemporary data to make a prediction for some time in the future, data from other

times in the past were also used. These two indications can be referred to more concisely as a redundancy of information and the use of more than one time lag, respectively.

One of the two purposes of this present investigation is to determine to what extent the data are redundant and whether linear predictability might be improved if two time lags were used. To isolate these parameters, the data, the stations, and the methods are the same that were used in the earlier work. The other purpose is to present a unified meteorological interpretation of the results of the linear prediction experiments of the study. The meteorological factors which these results indicate to be the important ones to the whole problem of forecasting precipitation on a short term basis will form the framework of the interpretation. This new work is therefore primarily concerned with (3) and (4) above.

II. THE DATA

The data were received by the AN/CPS-9 radar located at M.I.T. They are the plane distributions of signal intensity recorded every ten minutes within a circle sixty miles in diameter.

The intensities are read by eye from a 35 mm film of the PPI, coded on a scale of 8 and assigned to 5-mile squares. To obtain this scale, the squares are divided into quarters, the quarters are weighted by the degree of film blackening (nothing - 0, moderate - 1, saturated - 2) and summed.

The data were drawn from the photographic records of ten storms. At each station (i.e. each 5-mile square) records for the ten individual

storms were placed back to back to form a sample of reasonable length (406 entries). It is assumed thereby that the statistics of each storm are the same.

These particular ten storms were chosen because they all produced good coverage of the PPI scope throughout their duration (excluding the beginning and ending of the individual storm). It is important to have good coverage because otherwise the prediction scheme would be operating mostly on zeros and about the only thing the results would indicate would be a confirmation that precipitation is a rare event. Therefore the storms had to be the type that produce many large areas of precipitation. Squall lines and cold fronts are not of this type and were not included.

Figure 1 reproduces the working grid. Its longest dimension is 85 miles. While the data were being read from the film, the grids were always oriented such that the wind at the 700 millibar (mb) pressure level blew across the grid from top to bottom. The direction of the 700 mb wind varied among the storms so this procedure partially standardized the drift of the echo patterns across the grid: only partially because the varying wind speeds were not taken into account. It is a common observation on the PPI of a weather radar that the small precipitation areas tend to move with a direction and speed that resembles the 700 mb wind.

At the time the photographs of the ten storms were taken, the receiver of the radar had a limited dynamic range. Thus all the signals above a certain critical strength saturated the PPI scope. Considering this limitation along with the obvious subjectivity involved in reading the film by eye, it must be concluded that the quality of the data is not

very good. It was the best available in the necessary amounts. Since the purpose of this study is primarily an attempt at the diagnosis of the problem rather than to make the best possible prediction, the data quality may be sufficient for a first try.

III. THE PREDICTION METHOD: MULTIPLE LINEAR REGRESSION

The signal intensity at a station i at time $t + n\Delta t$ is expressed as a linear combination of the signal intensities X_j at stations j of the grid at time t . The prediction equation is

$$Y_i(t + n\Delta t) = A_0 + \sum_{j=1}^N A_j X_j(t) + \epsilon_i(t) \quad (1)$$

Y_i is the predictand and there are N predictors $X_j(t)$; $n\Delta t$ is the prediction time with $\Delta t = 10$ minutes. The coefficients A_0, A_j comprise the prediction operator and are determined by the method of least squares. $\epsilon_i(t)$ is the prediction error. It must be noticed that this equation applies only when the difference in time between the predictand and all the predictors is equal to the prediction time - that is, there is only a single time lag. For two time lags, equation (1) will also contain another summation of the form

$$\sum_{K=1}^N A_{N+K} X_K(t - n\Delta t).$$

Applied here the method of least squares requires that $\overline{\epsilon_i^2}$ be minimized with respect to A_0, A_j - that is, the expression

$$\partial \overline{\epsilon_i^2} / \partial A_j = 0$$

must be satisfied; the bar over ϵ_i^2 refers to an average over the data

sample. From this expression are obtained $N + 1$ equations in $N + 1$ unknowns the solutions of which are the prediction operator A_0, A_j .

To measure the usefulness of a particular prediction equation on the dependent data (the ten storms), the ratio of the variance of the prediction to the variance of the predictand is computed, this ratio being called the reduction in variance (RV). The variance of the prediction is the variance of the predictand minus the variance of the error ϵ_i .

$$RV = \frac{\sigma_{Y_1}^2 - \overline{\epsilon_i^2}}{\sigma_{Y_1}^2} = 1 - \frac{\overline{\epsilon_i^2}}{\sigma_{Y_1}^2} \quad 0 \leq RV \leq 1 \quad (2)$$

The higher the RV the better the prediction equation.

For two reasons it is desirable that the number of predictors used in the prediction equation be many fewer than the number of stations in the grid (120): first, simply to reduce the computations; secondly, to avoid difficulties which arise because the data sample is finite. If all 120 stations were used as predictors on the dependent data sample, the prediction operator no doubt would be almost worthless on independent data because of chance relationships (between the predictand and the predictors) which were incorporated into it.

However, if the number of predictors were arbitrarily limited to a certain number, some significant predictive information might be missed. Therefore we borrowed a procedure applied by Robert Miller of Travelers Research Center called screening or stepwise regression, which permits the successive scrutiny of single predictors.

Roughly, this is how it works. The square of the correlation coefficient between the predictand $Y_j(t + n\Delta t)$ and each of the possible

predictors $X_j(t)$ is computed. The highest value designates the first predictor X_a . The least squares estimate of all the remaining predictors $X_j(j \neq a)$ based on X_a is obtained and then a new value of each $X_j(j \neq a)$ is computed by removing the least squares specification. Let these be called the residues $X_{j(a)}$. They now replace the predictors. The residues are all independent of X_a , that is $\overline{X_{j(a)}X_a} = 0$. This means that all the predictive information contained in X_a has been removed from each $X_j(j \neq a)$ and the remaining predictors (now the residues) are independent of X_a . To get the second predictor, the correlations between the predictand Y_1 and the predictor residues $X_{j(a)}$ are computed. The highest correlation squared designates the second predictor $X_{b(a)}$. The least squares specification is repeated, only this time based on $X_{b(a)}$. This whole procedure is repeated to pick the third, fourth, and etc., predictors. As each predictor is chosen, the reduction in variance accumulated by all the predictors chosen so far is computed. This process is continued until one of two possible things occur: (1) an F test which tests each predictor chosen for statistical significance is no longer satisfied or (2) an arbitrarily set upper limit on the number of predictors which may be chosen is reached. Whichever occurs first stops the process. The details of this procedure are contained in papers by Miller (1958) and Shorr (1957). Mr. Miller was kind enough to provide us with a computer program embodying his screening procedure. A note of caution: it is important to realize that this procedure does not pick the optimum set of predictors, for each being individually the best at some stage in the process is no assurance that they will remain the best in linear concert.

The computations of the earlier work included all the 120 stations

as possible predictors and 60 from among them were chosen as predictands (Fig. 1). For prediction times of 10, 30, 60, 120, 180, and 240 minutes and a single time lag, these quantities were obtained:

- (a) The reduction in variance accumulated with each additional predictor picked;
- (b) The distribution of reduction in variance over the grid;
- (c) The predictors;
- (d) The time and space correlation fields;
- (e) The prediction operator A_0, A_j .

IV. REDUNDANCY OF INFORMATION AND USE OF MORE THAN ONE TIME LAG

A. Spatial density of data

As a result of the earlier computations, the correlations between all the possible pairs of stations (a predictand and a predictor) were tabulated as a function of the separation of the stations in space and time. The correlations for all pairs, whose relative position (without regard for geographical location) and relative time difference were the same, were summed and an average value determined. From all the average values determined, fields of mean correlations were constructed, one for each prediction time. As an example, Figure 3 shows the field for the 30 minute prediction time. The value in the center box of each field is the mean correlation for all pairs of stations with a zero space separation - this is the mean autocorrelation. The value in each box out away from the center is the mean correlation for all pairs of stations whose relative positions are as the center box is to the particular box being looked at, with the stations corresponding to the center box being the predictand in each pair. Each square here is 15

miles on a side - this was the result of averaging over nine adjacent 5-mile squares in order to obtain a smoother field.

These correlation fields have a good deal of circular symmetry about the center. The field for 30 minutes shows the symmetry (Fig. 3). The circular symmetry provides an argument for there being a redundancy of information within the grid because what it means is that no matter in which direction you look from any station (predictand) there is a predictor which is equally as good as any other at the same distance from the predictand. Therefore, the implication is that, if the density of stations in the grid were to be appreciably reduced, the predictability would not be affected because there would still be enough stations at a given radial distance from the predictand to provide just as "good" predictors as before. If the fields were not circularly symmetric about the center, this would imply that one or, at most, several directions were preferred by the predictors. If the station density were reduced by very much in this case, some of the "good" predictors would very likely be eliminated in the process.

To test for the existence of this indicated redundancy of information within the grid and to obtain an idea of how much information is redundant becomes the first problem here. New computations have been made with the station density reduced to one-half the original density by using only 60 of the original 120 stations. The stations used are the 60 chosen as predictands for the earlier computations with the original station density and original grid size (Fig. 1). There are no a priori grounds for deciding how much of a reduction to try, so since it had to be a stab in the dark the decision was to make a fairly bold one. Forty of the 60 stations were chosen as predictands (Fig. 1); the

computations were made for prediction times of 10, 30, and 60 minutes. Hereafter, those computations will be referred to as those with the reduced station density (abbreviated RSD).

B. Spatial extent of data

The spatial distribution of the reduction in variance for each prediction time was presented as a result of the earlier computations. Figure 11B shows this distribution for the 30-minute prediction, for example. The pattern in the distributions for all the other prediction times is essentially the same as the one for 30 minutes even though the mean reduction in variance over the grid deteriorates with increasing prediction time. This consistence of pattern is an argument for there being a redundancy of information outside the grid. A redundancy of information outside the grid means that all the events outside the grid give predictive information which at best just repeats the information that can be obtained inside the grid. When we say that there is a redundancy of information outside the grid, we imply that the grid is large enough to contain the necessary predictive information for each station (predictand) in this grid regardless of its location. If, at a certain prediction time, the grid were no longer large enough to contain all the necessary predictive information for some stations in a certain section of the grid, these stations would suffer an additional loss in reduction in variance that would be superimposed upon the deterioration all the stations would suffer due to increasing prediction time. This localized loss would cause the pattern at this particular prediction time to be different from the pattern at the next shorter time.

But a single reduction in variance pattern is consistent for all

the prediction times so the grid must be entirely large enough at least for prediction times up to and including 240 minutes. If it is entirely large enough is it possible that the predictive information provided by the stations within the grid but near the outside boundaries is redundant with respect to the stations nearer the center? If this is true, the size of the grid could be reduced and the remaining stations would retain the same predictability.

The consistency of the reduction in variance patterns argues in an additional way for there being a redundancy of information outside the grid. Let us assume that advection is the prime contributor to the predictive information. If the grid were not large enough to include the necessary advective information for each station at a certain prediction time, we would expect the reduction in variance pattern to show a definite prejudice toward high values at the stations along the leeward boundary of the grid and low values along the windward boundary. This sort of pattern was not observed at any of the prediction times. In fact, along the right side of the grid where there is the longest unbroken dimension of the grid along the 700 mb wind and thus the best place to discern this pattern if it were to occur, the values were uniformly low at all the prediction times. The pattern for 30 minutes is representative of this uniformity (Fig. 11B) The implication is that the grid is large enough to include the necessary advective information even for stations located on the windward boundary. Therefore, events outside the grid are redundant with respect to advective information.

Admittedly, both these arguments are loose ones and the second may even succumb to a faulty initial assumption (that advection is the primary contributor to the predictive information) so new computations are

in order to see if they are valid. Computations were made with the size of the original grid reduced so as to include only 60 of the 120 stations. Figure 2 shows this reduced-size grid. The computations were made for a 30 minute prediction time; 30 of the 60 stations were chosen as predictands. Of these 30 predictands 10 were not among the 60 predictands chosen for the comparable computations with the original station density and grid size. Hereafter, these computations will be referred to as those with the reduced size grid (abbreviation RSG). In order to compare reduction in variance patterns for consistency, computations for at least 10 and 60 minutes in addition to 30 minutes would be necessary. These would have been made but for a long shutdown of the computer.

C. Two time lags

The mean autocorrelations, resulting from the earlier computations and mentioned in connection with the argument for redundancy within the grid, provide an argument for the use of two time lags. The mean autocorrelations for each prediction time form an autocorrelation function. Figure 4 shows this function. Now if this function were of the form $\exp(-k|t|)$, t being the separation time, it could be concluded that the signal intensity follows a simple Markoff process. Simple Markoff processes operate at any given time with no "memory" of events occurring prior to the last one. What this means for predictability of signal intensity is that a second time lag would add no information. However, in this case the autocorrelation function is roughly of the form $\exp(-k|t|^{\frac{1}{2}})$. Thus a simple Markoff process is not indicated here and a second time lag may add some predictive information. To see if this is so some new computations are needed.

Computations were made for a 10 minute prediction with the first

lag at 10 minutes in the past and a second lag at 20 minutes in the past and for a 30 minute prediction with the first lag at 30 minutes in the past and a second lag at 60 minutes in the past. The stations used for these computations were the 60 used for the computations concerned with redundancy within the grid.

These new computations were made with the same data and the same methods as the earlier computations. Any significant differences between the results here and the comparable earlier results should be due only to the changes made in the experiments. From all of the new computations, the same quantities listed for the earlier computations were obtained except the time and space correlation fields which are the same so there was no need to repeat them.

V. THE METHOD OF MEASURING STATISTICAL SIGNIFICANCE

The accumulated reduction in variance with each predictor picked has been plotted against the number of predictors for each station (predictand) in each of the new experiments. As a typical example Figure 6 shows these curves for one station. First it may be noted that the reduction in variance always increases as the number of predictors increases. This occurs because the data sample - that is, the record length - is finite. Even though another predictor may not add any more information, it does admit another adjustable constant and thus decreases the degrees of freedom by one. This makes for a better fit and therefore a better reduction in variance. Because of this characteristic, we need to know the point beyond which the predictors picked are no longer significant, which is to say they add no more information.

It would seem that the F test mentioned earlier would do this job for us by stopping the screening process when this point is reached. However it does not do this because the F test is not a valid test when applied to radar signal intensities. The F test is based upon the assumption that the data are distributed normally. Our signal intensities are far from being distributed normally. Figure 5 shows their distribution. Even with a "perfect" radar, their distribution would essentially be the same. Therefore, some other test of significance is needed.

An empirical test for significance was devised in the following manner. The entire data record at two of the stations was randomized. This did not disturb their statistics but it removed all predictability. In most of the experiments these two random records were included among the predictands and treated as such. Figure 8 shows the accumulated reduction in variance versus the number of predictors for the random records included in the new experiments plus those included in the 10 and 60 minute prediction experiments with the original station density and grid size. The increase in reduction in variance here is due entirely to the decrease in the degrees of freedom as each predictor is picked. The point where the slope of the accumulated reduction in variance curve for a predictable record matches the slope of the corresponding curve for a random record should mark the last significant predictor. Beyond this point the increase in reduction in variance of the predictable record should be due only to the decrease in the degrees of freedom. These points are marked by the solid black circles in Figure 6.

Since the evaluation of such a critical part of the results depends upon these random records, their behavior and use are worth considering

in more detail. There are two statistical parameters which should determine their behavior (with respect to accumulated RV as a function of the number of chosen predictors) in the mean over a large sample of such records included as predictands in the screening process. These parameters are the record length (sample size) and the number of possible predictors. The effective length of all records included in the screening process varies according to the prediction time in an inverse fashion, this reduction in length being necessary in order to compute the time lagged cross correlations. Since the record length is finite, when it is decreased from the length for one prediction time to the length for the next greater prediction time, the averaged spectrum of the cross correlation values between a large sample of random records and all of the possible predictors (predictable records) should widen. Thus on the average higher and higher values of correlations between a large sample of random records and all of the predictable records should occur as the prediction time increases. The reduction in variance by a predictor chosen by the screening process is proportional to the value of the correlation between it and the predictand (a random record in this case). Therefore, a set of parametric curves, each of which is the accumulated RV averaged over a large sample of random records as a function of the number of chosen predictors and with record length (prediction time) the parameter which varies among them, should have RV axis intercepts and slopes which increase with decreasing record length (increasing prediction time). In the preceding discussion the other parameter, the number of possible predictors, is being held constant. If, on the other hand, the record length were held constant and the number of possible predictors was the varying parameter, a similar set

of curves should have RV axis intercepts and slopes which decrease with decreasing numbers of possible predictors. The width of the averaged spectrum of values of cross correlations between a large sample of random records and all of the possible predictors should be a function of the number of possible predictors. Therefore, the upper boundary of the averaged spectrum will fall to lower correlation values and, ergo, lower reductions in variance as the number of possible predictors is decreased.

In summary then, we see that the width of the averaged spectrum of cross correlation values between a large sample of random records and all of the possible predictors should be an inverse function of the record length and a direct function of the number of possible predictors. It must be pointed out that the behavior of individual random records may deviate from the mean behavior as determined by these two parameters. Figure 7 shows the behavior of an individual random record. This record was included in the screening process in the earlier work with original station density and grid size. Here is an example of an individual random record whose behavior follows the expected mean behavior of a large sample of random records. The varying parameter is the record length. In Figure 8 we see an example of an individual record whose behavior deviates from the expected mean behavior. Consider the curves for random record #1 included in the 30 and 60 minute lag computations with the reduced station density (RSD). The varying parameter is again the record length. Since the record length is shorter for the 60 minute curve than it is for the 30 minute one, the 60 minute curve should be higher than the 30 minute curve if both were averaged over a large sample of random records. But here just the

opposite is observed.

We have only two random records instead of a large sample with which to determine significant predictability. The results may be affected by variability in their behavior, therefore it would be desirable to minimize the effect of this variability as much as possible. If the randomized signal intensities were distributed normally, then for each combination of the parameters, it would be possible to establish confidence limits on the variability of the accumulated RV as a function of the number of chosen predictors for an individual random record. But the randomized intensities are no more Gaussian than the intensities of a predictable record since the distribution of values is the same for both. A way to accomplish approximately the same thing would be to repeat a given experiment many times using different (independent) data each time. A single random record would be included in the screening process each time. The highest and lowest RV curves for the random record from among the repetitions would form the confidence limits. In practice this procedure will have to wait until many independent sets of data are available.

In the present study, a substitute procedure was carried out where possible. If both random records were included in an experiment, their RV curves were thought of as approximations to the true limits of the variability. The average of the two was taken and the resulting curve was smoothed. This curve became the gauge of significance for that particular experiment. If only one of the random records was included in an experiment, its RV curve was used alone (smoothed).

No doubt there is some variability remaining among mean curves which are the average of the RV curves for at best only two random records.

In an effort to find out how much this variability may affect the determination of the significant RV of the predictable records, a test was made as follows. For the 30 minute experiment with two time lags, three different RV curves for random records (each having the same combination of the two statistical parameters) were successively used as gauges of significance in the manner described earlier. The curves were (1) the average of the RV curves of the two random records included in the 60 minute experiment with the original station density and grid size, (2) the RV curve of random record #2 included in the 30 minute experiment with two time lags, and (3) the average of the RV curves of both of the random records included in the experiment mentioned in (2) (Fig. 8). The values of the mean RV over all the stations (predictable records), as separately determined using the three RV curves, were at the most only 3% apart (50% was the mean RV).

There were no random records included in the 10 and 30 minute experiments with the reduced station density. For the 10 minute experiment on RV curve was constructed using as a guide the curve for the random record included in the 10 minute experiment with the original station density and grid size. For the 30 minute experiment, the RV curve for the random record included in the 30 minute experiment with the reduced grid size was used.

VI. THE RESULTS OF THE REDUNDANCY AND TWO TIME LAG EXPERIMENTS

A. Predictability in the mean as a function of prediction time

A comparison of predictability using the original station density and grid size with predictability using the reduced station density, the

reduced grid size, and two time lags at the several corresponding prediction times should confirm or dash the contentions of redundancy of information and improvement with a second time lag. Let us consider predictability in the mean over all the predictands in each experiment. Figure 9 shows this comparison graphically as a function of prediction time and a tabular comparison is made in Table 1 on page 46. Each value is the reduction in variance averaged over all the predictands included in that particular experiment. The circled points in Figure 9 are the mean reductions in variance with the original station density and grid size; the X's mark the mean reductions in variance of the new experiments.

First notice in Figure 9 that the experiments at 10, 30, and 60 minutes with the reduced station density cause predictability to suffer but very little. This comparison confirms that most of the information removed by reducing the station density is redundant. Notice, too, that the difference between the mean reduction in variance with the original station density and grid size and with the reduced station density decreases slightly but uniformly as the prediction time increases from 10 to 60 minutes. Whether or not this seemingly systematic improvement with increasing time has any meaning is quite open to question. Since all the differences are so small, it could easily have happened strictly by chance because of the same type of statistical variability discussed in connection with the random records. But assuming that it is real (not due to chance), the fact that the quantitative improvement is small would not detract from its importance. It would indicate that the amount of redundant information is at least somewhat a function of the prediction time. What is the probable source of this relationship? It would seem to be the effect of station density on the location of

the possible predictors relative to any predictand. Some findings of the experiments with the original station density and grid size give the clue. At ten minutes prediction time and to a lesser degree at 30 minutes the significant predictors clustered together in a relatively small region close to the predictand and in the upwind direction from it. At greater prediction times (up to and including 240 minutes), neither did they cluster nor were they so close to the predictand. They were "scattered" about over the grid in a somewhat random fashion with the randomness becoming more and more pronounced as the prediction time increased. Keeping in mind these findings we may ask how predictability will be affected by station density. Reducing the density has the effect of isolating the remaining stations from each other. In other words, the mean distance from a certain station to the next nearest station is now greater. Therefore, with a given reduced density the predictability at any particular station should suffer the most when the necessary predictive information is in a small region located in one specific direction from the predictand and the least when this information is not restricted by distance or direction from the predictand. Thus with a given reduced density the loss in predictability will be the greatest at 10 minutes and it will decrease as the prediction time increases.

Fig. 9 indicates that predictability at 30 minutes with the reduced grid size is not only as good but actually better than with the original grid size. The indication that the predictability is at least as high as it was with the original grid size confirms that the information removed by reducing the grid size is redundant. The stations in the reduced grid can at best only duplicate the predictive information available

in the original size grid so the indication that the predictability is actually better with the reduced grid size can only be due to chance and/or the fact that ten of the predictands included in the experiment with the reduced grid size were not among the ones included in the experiment with the original grid size. The latter could have an effect because we are considering predictability in the mean over all the predictands included in a particular experiment.

Finally the comparison of results (Fig. 9) also indicates that a second time lag seems to give a small improvement in predictability at 10 and 30 minutes. The predictability could be higher due to chance rather than the second lag. However, at 10 minutes most or all of the increase can be accounted for by something other than either of these. For this experiment the computations for only 23 out of the 40 predictands were usable because of trouble either with the computer or the screening program. Most of the 23 are predictands that, in the experiment with the original data density and size grid at 10 minutes, had relatively high reductions in variance and nearly all of the other 17 had relatively low reductions in variance. The mean over the 23 predictands in the two lag experiment is no doubt weighted toward the high side.

This leaves us with little or no difference at 10 minutes between the one and the two lag experiments but with a small increase at 30 minutes. If we assume that this pattern is real, it would indicate that the improvement in predictability provided by a second lag could be a function of the prediction time with no improvement being discernible at the shortest possible prediction time (10 minutes). The relationship in time of the first and second lags in the design of the experiments may explain why this could be true. The second lag at 10 minutes looks

only 10 minutes further back in time than the first lag, whereas the second lag at 30 minutes looks 30 minutes further back. The second lag at 10 minutes is therefore not really much different in time than the first lag so we should expect very little new predictive information to be added. But the second lag at 30 minutes is enough different in time from the first lag to expect that a significant amount of information might be added.

Certain of the deductions drawn from this comparison of predictability in the mean depended entirely upon the revealed details being real and not due to chance. Of course, an additional difficulty could be the reverse of this obvious one - that is, there could be some real detail which is obscured by the chance variability. In any case the validity of these deductions remains in question until the variability of the results due to chance can be determined. The problem is the same as determining the statistical variability of the random records. If all the experiments were repeated many times, each time using a different data sample, and if in the mean the detail previously noticed is still evident then it would be safe to say that it is real. The practical difficulty in actually doing this was mentioned in connection with the random records.

B. Predictability as a function of space

The reduction in variance patterns over the grid may be compared in two ways. The pattern for each class of experiments at the same prediction time may be compared or the patterns at the several prediction times for one certain class of experiment, say the experiments with the reduced station density, may be compared. Neither of these comparisons can be made here as adequately as desired because of the limited number

of new experiments which were undertaken.

It was mentioned earlier that the reduction in variance pattern at each prediction time in the experiments with the original station density and grid size was essentially the same thus indicating that the grid was entirely large enough to contain the necessary predictive information for all the stations at all the prediction times. A similar comparison of RV patterns with the reduced grid size would be very interesting but computations only at 30 minutes were made. The only comparison that can be made is with the pattern for the experiment with the original station density and grid size at 30 minutes. Figures 10b and 11b afford such a comparison. The patterns are essentially the same. The mean predictabilities for these two experiments are also essentially the same. These two facts leave the impression that the reduced size grid is large enough to contain the necessary predictive information at least at 30 minutes. To speculate a bit, it is a possibility that, if computations with the reduced grid size were made at prediction times greater than 30 minutes, we might find that the pattern changes at one of these times thus indicating that the reduced grid is not large enough for all the prediction times.

Figures 10a and 11b compare the reduction in variance patterns for the experiments with the reduced station density and the original station density and grid size at 30 minutes. There is little difference between the two. The patterns at 10 and at 60 minutes are also the same. Since reducing the station density to half its original value causes only a small loss in predictability which apparently decreases with increasing prediction time, it seems likely that the patterns for experiments with the reduced station density at prediction times greater

than 60 minutes would also be the same as the pattern at 30 minutes.

C. Further speculations

Consider again the results of the experiments with the reduced station density and reduced grid size as presented by the comparison of their predictability in the mean with that of the experiments with the original station density and grid size (Figure 9). Let us assume that the indicated decrease in the difference between their corresponding mean reductions in variance with increasing prediction time is real and that the explanation for this relationship is correct. Based upon the reality of this indication and the correctness of its explanation it seems quite likely that if we want to eliminate redundant data from the original working grid with the least possible expense to predictability, reducing the grid size will be the best method at 10 and possibly 30 minutes and reducing the station density will be the best method at prediction times greater than 30 minutes.

By separately reducing the station density by half and the grid size by about a third we have found no significant loss of predictability thus confirming that much of the data is redundant for prediction times up to 60 minutes with respect to spatial density and up to 30 minutes with respect to spatial extent. But are the data redundant at prediction times greater than these? If so, how much more of the data are redundant? How much further can the station density or the grid size be reduced without a significant loss in predictability? The only way to answer these questions is to try some further experiments.

Experiments with the present reduced station density at prediction times greater than 60 minutes would be worthwhile as a check. The next step might be to reduce the station density again by a half by using

only 30 of the original 120 stations. An interesting attempt would be to see what would happen if we were to reduce the density such that the number of stations would be equal to the mean number of significant predictors as indicated by the experiments so far. Figure 9 shows that the mean number of significant predictors varies from only 3 at 10 minutes to 9 at 60 minutes.

Experiments with the present reduced size grid at prediction times greater than 30 minutes are necessary before an even smaller grid size is tried. As was mentioned earlier, the present reduced grid size might well not be large enough to include the necessary predictive information at one or more of these prediction times and this must be ascertained.

Experiments using a second lag at prediction times greater than 30 minutes are needed for further indications of whether or not the information added by a second lag is a function of prediction time. Of course in these experiments the second lag must bear the same time relationship to the first lag as in the experiments at 10 and 30 minutes.

D. Summary and criticism

We have asked certain pointed questions about the data and have been returned the answer that much of it is redundant. Before this conclusion is extended to other cases we might pause to wonder whether the questions might have been too pointed for the kind and quality of the data.

Reducing the density of the stations within the grid seemed to incur no significant loss in predictability. As was mentioned earlier, the approximate circular symmetry of the mean correlation field (Figure 3) hinted at such a result. But the circular symmetry also implies that by lumping ten different storms together to make up the data we have

obliterated almost all nuances of structure. Individual storms are rarely amorphous. The statements of redundancy therefore need the qualification that they apply only to generalized circumstances. The same qualification should be attached to the results for two time lags because combining the storms together may have blurred differences in the time structure as well as in the space structure.

Reducing the size of the grid seems also to incur no significant loss in predictability. We may yet detect large losses when this test is extended to prediction times greater than 30 minutes. However, it would be well to realize that the same results would have been obtained if the data discarded - that is, the data at the stations eliminated by reducing the grid size - were of a poorer quality than the data at the stations which remained. This possibility is not entirely unlikely. Figure 11b shows that generally the stations nearer the edge of the grid do not do as well. The reduction in variance patterns for the other prediction times with the original station density and grid size show the same thing.

VII. METEOROLOGICAL INTERPRETATION OF THE RESULTS

The objective is to present a meteorological interpretation of the results reported here and earlier by Noel and Fleisher (1960). The results of the experiments with the original station density and grid size which are summarized in Fig. 12 will be discussed first because they include the greatest prediction time spectrum with all other experimental parameters constant. This analysis will be followed by an interpretation of the results of the experiments with the reduced station density, reduced

grid size, and two time lags. These latter experiments should be able to be interpreted just as variations on the experiments with the original station density and grid size. Finally a different method of measuring linear predictability which has meteorological implications will be discussed.

A. Non-meteorological factors

The results are influenced by both meteorological and non-meteorological factors. Therefore, first of all, we must consider and separate the effects on the results from non-meteorological factors.

- (1) Regardless of the type of prediction scheme which might be used, linear or otherwise, the poor quality of the data no doubt causes scatter about the regression curve. With predictability being measured by the reduction in variance (RV), which is inversely proportioned to the mean-squared-error over the sample, scatter produces a $RV < 1$ regardless of how well the particular regression curve fits the data in the mean. Therefore the poor quality of the data accounts in part for the low average level of predictability over the range of the lag times as shown in Fig. 12.
- (2) The relatively small data sample (record length) may cause higher cross-correlations to occur as a result of chance alone than due to physical relationships. These chance correlations give a reduction in variance which is too high. The possibility of high chance correlation occurring becomes even more likely as the prediction time increases because the effective record length becomes shorter. For the 240 minute prediction, the effective length is less than half what it was for 10 minutes.

The results as shown in Fig. 12 give evidence that such chance correlations did occur. The evidence is the nearly constant level of predictability beginning at 60 minutes and continuing on out to 240 minutes where there is an unexpected rise. It would seem most probable that the rise in RV can be attributed to the chance correlations.

- (3) There is another possible non-physical explanation of the apparent constant predictability beyond 60 minutes. It is that there is really no predictability at all beyond 60 minutes with the approximately constant level of mean RV in the vicinity of 40% representing only an upwardly displaced zero reference. It must be expected that the mean RV will be greater than zero because of chance correlations which will always occur with a finite data sample size. However, a 40% mean RV is higher than the RV from a random record whose correlations with the predictors are all due to chance (see Figures 7 and 8). Consequently it is considered that the 40% level represents a significant mean RV and the above theory must be discarded.

B. Meteorological factors

There are two factors which are the keys to this meteorological interpretation. In order of importance they are

- (1) Non-linear behavior of the precipitating elements over periods of time which correspond to the prediction times used in the experiments, and
- (2) the effect of the ratio of the horizontal size of the precipitating elements to the space resolution of the stations.

It can be assumed that the ten storms were made up of separate precipitating elements, both small and large, rather than a horizontal, layer-type structure. The interpretation is based upon this assumption. Evidence that the storms did in fact have the former type structure rather than the latter will be presented later. We will first consider the effect of these two factors on linear predictability in a general way.

A linear prediction method will filter out only linear behavior and thus the average level of predictability over the prediction times, indicated in Fig. 12, is in part a result of the inherent inability of the prediction method to account for the non-linear behavior of the precipitation elements. The degree of non-linearity in physical behavior is usually a function of the time or distance over which the behavior is considered. This characteristic applies to the behavior of precipitation elements because if we mentally subtract that portion of the RV in Fig. 12 which is believed to arise from high chance correlations, it is observed that linear predictability decreases monotonically as a function of prediction time.

What is the causal source of the non-linear behavior of the precipitating elements? Their behavior can be roughly divided into two modes: (1) persistence and advection by the wind, and (2) growth and dissipation. Growth and dissipation is most likely a non-linear process in time and is therefore the main reason for the overall non-linear behavior of the elements. There is no a priori reason why growth and dissipation has to be non-linear but it seems unlikely that the average level of predictability over prediction time as shown in Fig. 12 would be as low as it is on account of poor data quality alone.

If growth and dissipation of the precipitating elements are the

main source of their non-linear behavior over periods of time corresponding to the prediction times, it is automatically implied that the lifetimes of these elements are not too different from the range of time encompassed by the prediction times. If, on the other hand, their lifetimes were at least an order of magnitude greater than the average prediction time, their behavior would appear much more linear and a higher average level of predictability than is shown in Fig. 12 would occur.

Having classified the behavior of the precipitation elements into two modes, one linear and the other non-linear, we can deduce some additional physical insight about linear predictability where precipitating elements are involved. When the prediction time is short compared with the lifetime of the precipitating element, the non-linear mode (growth and dissipation) accounts for very little of its behavior while the linear mode (persistence and advection) accounts for most of its behavior. As the ratio of the prediction time to the lifetime of the elements increases, the relative importance of the two modes in accounting for their behavior reverses and therefore linear predictability must decrease with an increase in ratio of prediction time to lifetime.

The horizontal sizes of the precipitating elements are roughly proportional to their lifetimes - that is, as it is commonly observed on the PPI, the larger the precipitation area the longer it lasts. This connection between the horizontal size and the lifetime of precipitating elements serves as an introduction to the second of the meteorological factors - the ratio of the horizontal size to the space resolution of the stations in the grid. If the ratio is small, this means that a number of precipitating elements may be encompassed by a

single station at any given time. Although the signal intensity of the individual elements will vary rapidly with time on account of growth and dissipation, the variation of the station signal, being the average of the signals of many elements, will be smaller. The averaging makes the station signal more linear-like with time than it would have been if the ratio were unity or greater. Better linear predictability is the result.

C. The nature of the precipitation

In order to apply this somewhat generalized interpretation specifically to the results as shown in Fig. 12, we need to know with good certainty that the ten storms which composed the data were made up of separate precipitating elements and we need to know the sizes and corresponding lifetimes of these elements. As observed on the PPI, large precipitating areas whose greatest dimension was on the order of 100 miles were the dominant feature of each of the ten storms. The minimum observed lifetime of these large areas was about two hours. The area coverage by the radar did not permit a determination of the upper limit of the spectrum of their lifetimes. However, it is apparent that some persist for more than eight hours because they entered on one extremity of the PPI scope, crossed it, and left on the other extremity in this length of time. For the purposes of this interpretation, it will be assumed that the average lifetime of these large areas is on the order of five hours. They moved in a direction which was within 30° of that of the 700 mb wind; their speed varied considerably with time from almost no motion to speeds which were somewhat greater than the 700 mb wind speed (the mean 700 mb wind speed over the ten storms was approximately 28 mph). Except around their edges, they had signal

intensities strong enough to saturate the PPI scope much of the time they were observed.

There is good reason to believe that these large precipitating elements were made up of many small cells of the size observed by Newell (1959) and Battan (1953) among others. During January and July of both 1957 and 1958 Newell made observations of precipitation by means of the RHI scope of an AN/CPS-9 radar during nearly all of the storms of those months. He found during January that for approximately 60% of his observations, the precipitation was falling from convective cells whose horizontal size averaged about 1 mile; the remainder of the observations encountered a horizontal layer-type structure. During July all of his observations encountered convective cells whose horizontal size averaged about 2 miles. During January he found that sometimes precipitation which first appeared as a horizontal layer at a relatively high gain setting was actually made up of cells which appeared when the gain was turned down.

His statistics are based on a relatively small sample but nevertheless their indication is clear. Therefore there should be little doubt that the large precipitating elements of the ten storms used in this study were composed of these small cells a good majority of the time. Some RHI photos taken during several of the storms tend to confirm this - they showed the small scale cellular structure. These cells were not discernible on the PPI pictures of the ten storms most likely because of the high gain setting. It was usually set at maximum during the photo taking.

Unfortunately Newell's method of observation did not allow him to measure the lifetimes of these cells. However, Blackmer (1956), from

observations of four storms in 1948 (one each in February, March, May, and June) on the PPI scope of an SCR-615-B radar found an average lifetime of approximately one hour for small precipitating areas up to 20 miles in diameter. Many of these areas may have consisted of several smaller cells whose individual lifetimes were shorter. From observations of precipitation with the RHI scope of a 3 cm radar during the summer of 1947, Battan (1953) reported an average lifetime of 23 minutes for single convective cells. For the purposes of this interpretation, it will be assumed that the average lifetime of the small cells is 'less than one hour'.

To recapitulate, there were two basic sizes of precipitating elements that could be observed by the radar during each of the ten storms. One was on the order of 100 miles in horizontal extent with an average lifetime of approximately five hours and the other was between 1 and 2 miles in average horizontal extent with an average lifetime of less than one hour. The former moved with a speed which varied between zero and somewhat greater than the 700 mb wind speed and the latter have been observed by various researchers to move with approximately the speed of the 700 mb wind.

D. Interpretation of predictability as a function of lag time

This information on the nature of the precipitation of the ten storms will now be used to analyze the prediction curve shown in Fig. 12. The space resolution for the experiments represented by these results was 5 miles (the five mile squares). This is not much larger than the size of the small precipitating elements. The prediction times corresponding to the periods of time over which the behavior of the small elements can best be accounted for by a linear scheme must be less than

their average lifetime which is less than one hour. Therefore we can say that the mean reduction in variance (RV) at the 10, 30, and possibly the 60 minute prediction times is, in effect, the linear predictability of the small elements, with the rapid decrease in RV with increasing prediction time being due to the increasingly non-linear behavior of the elements as the prediction time-lifetime ratio increases to near unity. In the 10 and 30 minute experiments, the significant predictors for most of the predicted stations included either the predicted station itself or stations immediately upwind from the predicted station or both. At prediction times greater than 30 minutes, the significant predictors were spatially distributed in a rather random fashion, although a tendency to favor the upwind region continued out to 120 minutes; the predicted station itself was rarely chosen. This evidence corroborates the assumption that the linear mode of behavior of precipitating elements - persistence and advection - is important to linear predictability when the prediction time is small compared with the lifetime of the elements. With the size of the small elements being somewhat less than the space resolution, the linear scheme is really predicting on a scale that may include several of the small elements. The averaging effect of the resolution on the small elements explains why the first predictor picked at 10 minutes was almost always the predicted station - that is, why persistence was so evident. With a space resolution closer to the size of the small elements, the role of persistence should decrease somewhat and along with it the linear predictability.

Let us now consider the predictability from the vicinity of 60 minutes to 240 minutes. If we were, in effect, predicting the behavior of the small elements at these prediction times also, we should expect

the mean RV to continue to decrease in a smooth fashion beyond 60 minutes instead of leveling off. A logical answer is that at these prediction times we are no longer predicting the behavior of the small elements. Therefore it is possible that the linear behavior of some other precipitating elements is being detected by the prediction scheme. Most likely they are the large elements so evident in the PPI pictures of the ten storms. With these elements having an average lifetime of approximately five hours, over a prediction time of 60 minutes their behavior should be relatively linear. As the prediction time increases beyond 60 minutes their behavior will become less and less linear and hence the linear predictability should decrease. Again it is pointed out that part of the mean RV beyond 60 minutes in Fig. 12 is no doubt due to high chance correlations. Additional evidence that the prediction scheme is detecting something different at prediction times of 60 minutes and greater is shown in Fig. 4, the mean autocorrelation coefficient versus prediction time. Notice how the curve levels off in the vicinity of 60 minutes and decreases more gradually beyond.

What can be inferred concerning the linear predictability beyond four hours? Since we would be getting beyond the mean lifetime of the large elements, predictability would probably decrease sharply. If experiments were conducted at prediction times greater than four hours with the present data sample, there would be little chance of showing whether this is the case or not because the effect of high chance correlations would no doubt become even more pronounced and the mean RV would probably increase with prediction time. However, if a data sample large enough to reduce chance correlations to a minimum for prediction times as great as 6 or 8 hours were available, we should see the sharp decrease.

If, in addition, we repeated the experiments at all the prediction times from 10 minutes to 4 hours with the large sample also, the mean RV as a function of prediction time might be represented by the dashed curve marked 5 shown in Fig. 13.

The present data are made up of a variety of weather situations but contain no thunderstorm activity. Let us suppose that we had a set of data made up of thunderstorms only. It is generally observed that thunderstorms have lifetimes greater than one hour on the average. This is greater than the average lifetime which we thought was applicable to the small precipitating elements in the ten storms. Therefore for prediction times less than about two hours, the linear predictability should be better with thunderstorm data than it is with the present data. This improvement is due to the lower prediction time-lifetime ratio. A more linear-like behavior over the prediction times than with the present data is the result.

The present space resolution of the stations is such that only a few of the small precipitating elements may be encompassed by a station at a given time. Stations with a 30 mile resolution could encompass a large number of the small elements. A 30 mile resolution is also closer to the order of magnitude of the large precipitating elements than the 5 mile resolution. Thus a 30 mile resolution should result in a greater linearizing effect on the station signal than we now have with the 5 mile resolution. Therefore higher linear predictability should be expected at all prediction times less than approximately five hours and the mean RV as a function of prediction time might be represented by the dashed curve marked 30 in Fig. 13 (a large data sample is assumed). On the other hand, if the resolution were sharpened to

one mile, the effect on the linear predictability should be the opposite and the mean RV as a function of prediction time might be represented by the dashed curve marked 1 in Fig. 13.

E. The interpretation of the results of the redundancy and two time lag experiments

The results of the experiments with the reduced station density, the reduced grid size, and two time lags as summarized in Fig. 9 will now be analyzed in an effort to find out to what degree they are consistent with the meteorological interpretation presented in the preceding sections. Consider first the results at prediction times of 10, 30, and 60 minutes with the reduced station density. A reasonable reduction in station density should cause a significant loss in linear predictability only when the prediction time-lifetime ratio is low. It is then that advection is an important contributor to predictability. As it was pointed out in an earlier discussion (Chapter VI), reducing the station density is more likely to cause a loss in predictive information when that information tends to be located in a specific region rather than randomly distributed. Advective information will always tend to be located in a specific region relative to the predicted station. As it has been pointed out, advective information was most important in the experiments with the original station density at the 10 minute prediction and to a lesser degree at 30 minutes. Therefore with the reduced density we should see the largest loss in predictability at the 10 minute prediction and a smaller loss at 30 minutes. This is just what happens (Fig. 9).

It was stated in an earlier discussion (Chapter VI) that if experiments using the reduced grid size were conducted at prediction times

greater than 30 minutes, we might very well see losses in predictability because the grid was not large enough to contain the necessary predictive information. Our meteorological interpretation of the results summarized in Fig. 12 may give us an idea of the critical prediction time at which the reduced grid size is no longer large enough. Predictive information on the large precipitating elements should be the first to be removed by a reduction in grid size. Since the large elements seem to be an important influence on the linear predictability as a function of prediction time beginning in the vicinity of 60 minutes, it would not be surprising to find a loss in predictability first occurring here.

It is shown in Fig. 9 that the addition of a second time lag produced a small increase in the mean RV for both the 10 and 30 minute predictions. In a discussion of these results in Chapter VI, the increase at 10 minutes was attributed to something other than the second lag. Also in that discussion an argument attempting to explain why a second lag should improve predictability was presented. The essence of this argument was that if the second lag was considerably longer than the first lag, the second lag would provide predictive information which would be supplementary to that given by the first lag. With the meteorological interpretation of the results of the experiments with just one lag as a guide, further consideration will now be given to a second lag. For any short prediction time, it is not evident why the inclusion of a second lag should provide significant information concerning the large precipitating elements (average lifetime approximately 5 hours). It has been assumed that the development aspect of the behavior of the small precipitating elements accounts in part for the quality of the

short time (10 and 30 minutes) linear predictability. It is not apparent that the introduction of a second lag should minimize the role of development. Also it is not apparent that additional information about the linear processes (advection) should be provided by a second lag. For the 30 minutes prediction, the second lag at 60 minutes may be providing some information (by detecting the large precipitating elements possibly) that is better than that provided by just one lag because for most of the stations their predictors included ones from the second lag. However, the resulting increase in predictability is small so the information concerning the large precipitating elements is hardly significant. With the increase being so small, it is possible that the second lag predictors were chosen as a result of chance relationships only. There seems to be no reason to expect that a second lag will provide any more improvement in predictability at longer prediction times than it has at 30 minutes. Nor can it be envisaged how a greater time difference between the first and second lag than was used with the 30 minute prediction would improve information from the second lag. On the contrary, it would be more likely to worsen it. The results which have been obtained here require further consideration in the light of meteorological processes because (as was noted in a discussion in Chapter IV) the character of the statistics of the data with time seems to indicate that a second lag would add significant information.

F. An alternative method of measuring predictability

It is appropriate in a meteorological interpretation of the results to consider an alternative method of measuring predictability.

The requirements of a good precipitation forecast usually are two:

- (1) the correct spatial pattern of precipitation intensity and (2) the

correct magnitude of the precipitation intensity at each point in the pattern. The reduction in variance measures the degree to which both of these requirements are attained because it is based upon the absolute magnitude of the intensity error. However, from the RV it is not possible to distinguish between pattern and intensity errors.

A much less stringent definition of a good forecast is that only the first requirement above be satisfied - that is, the correct pattern without regard for systematic errors in magnitude. In some meteorological situations, this accuracy may be adequate. For instance, in predicting the future of a fully organized squall line, the main interest is where it is going to be in a few hours. If the future pattern is predicted correctly, the forecast will be considered good.

It is interesting to note that there is good evidence that the linear scheme used in this study does better in predicting the pattern than in predicting the intensity. The prediction coefficients (A_0, A_j) generated by the step-wise regression incorporated in the screening program were used in the linear prediction equation (Eq. 1) to compute a small sample of forecasted maps for each prediction time of the experiments with the original station density and grid size. Each forecasted map was objectively compared with its corresponding actually observed map in two ways: (1) a spatial RV was computed over all the stations on the map (the definition of RV is the same as it was before but the intensity errors are averaged over the stations rather than time), and (2) a correlation coefficient was computed between the corresponding station intensities of the two maps. The latter determines the correctness of the pattern because it measures only the degree to which the errors in intensity are systematic and is not dependent on the magnitude of the

errors as will be the former. At each prediction time, the mean correlation coefficient over the sample of maps was higher than the mean spatial RV. Though the sample of maps was not large enough for the computed statistics to be very highly significant, the indication given by them is unmistakable nonetheless.

VIII. CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

Conclusions:

1. The short term linear predictability of precipitation in the type of storm included in this study is relatively insensitive to reasonable variations in the spatial density of information. However the spatial extent of information may be a critical parameter.
2. The non-linear behavior of precipitating elements whose sizes range from one up to 100 miles in order of magnitude is important for the short term predictability of precipitation.
3. The space resolution with which a forecast is made will influence to some extent the level of predictability.
4. The quality of linear prediction decreases with time as expected but not in a systematic manner because of the presence of different sizes of precipitating elements in the storms from which the data were obtained.

Recommendations for further research:

1. Investigate the linear predictability of precipitation patterns arising from weather situations different from those included

in the present data e.g., squall lines and cold fronts.

2. Investigate in a systematic manner, the effect of variations in the degree of space resolution on predictability.
3. Investigate further the spatial extent of information as a parameter influencing linear predictability.
4. Investigate further the significance of including a second time lag in a linear prediction scheme.

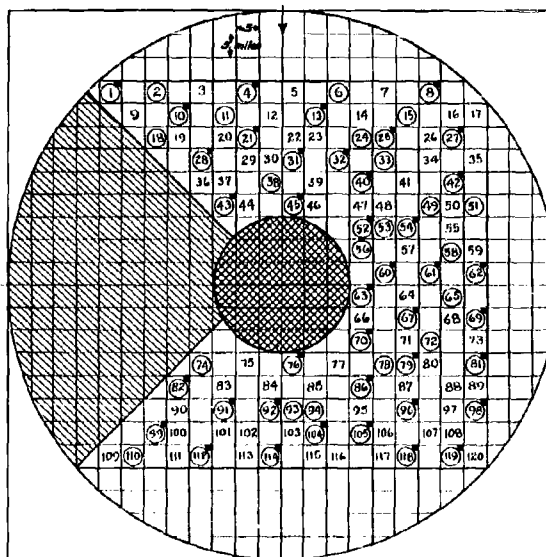


Fig. 1. The working grid of 120 stations. Circled stations were predictands in earlier computations and also comprised reduced density grid. Forty stations marked with small black squares are predictands for reduced density grid.

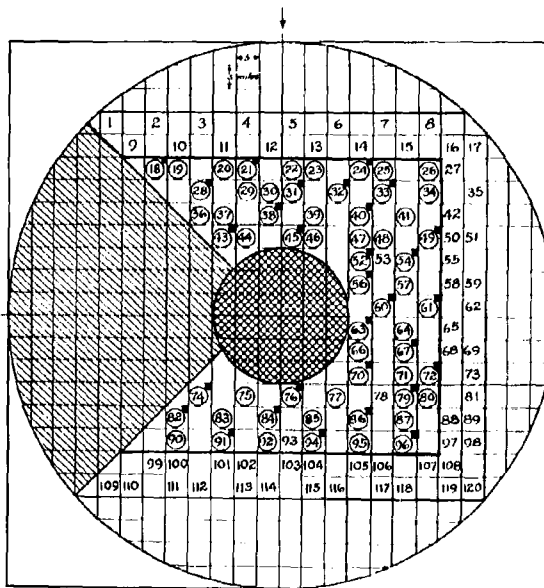


Fig. 2. The reduced size working grid. The 60 stations circled comprise the grid and the 30 stations marked by a black square are predictands.

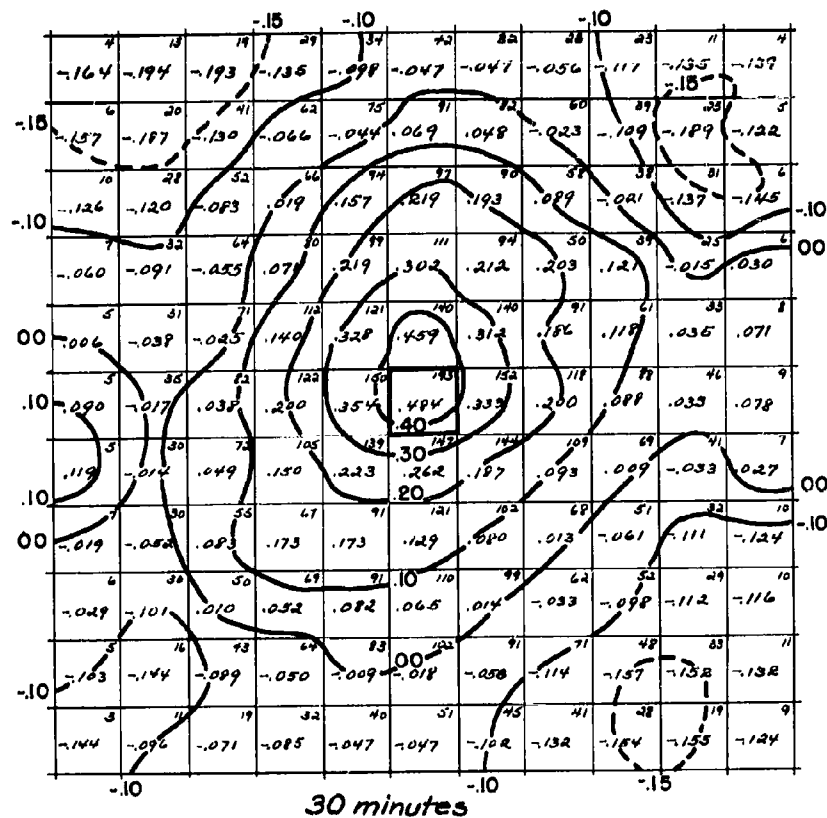


Fig. 3. The mean correlation field for 30 minutes prediction time with the original station density and grid size. Each box is 15 miles on a side. The number in the upper right hand corner of each box is the number of station pairs whose correlations were summed to compute the mean value appearing in the box.

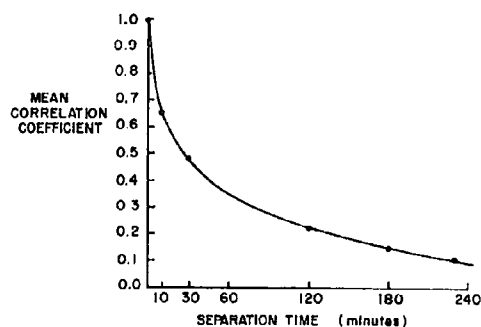


Fig. 4. The mean autocorrelation function of precipitation as determined from the computations with the original station density and grid size.

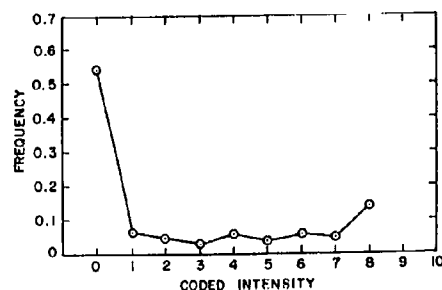


Fig. 5. The frequency distribution of coded signal intensity.

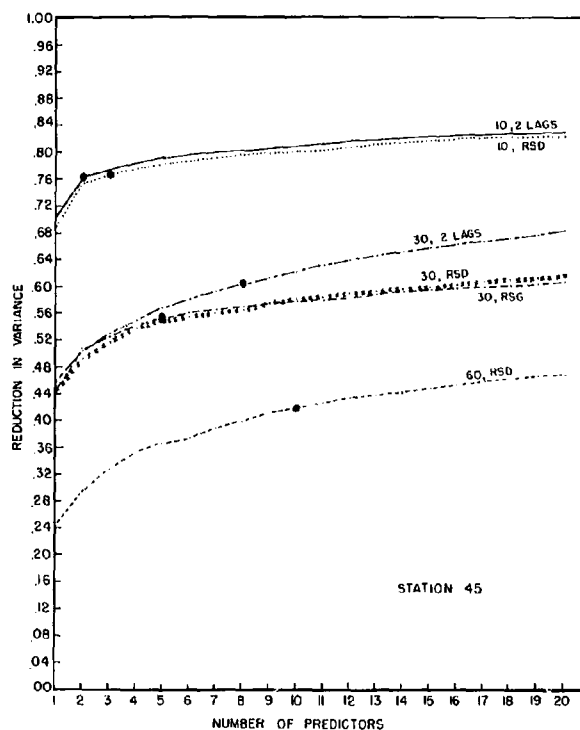


Fig. 6. The course of the reduction in variance (RV) for station 45 as predictors are accumulated. The solid black circles mark the last significant increase in the reduction in variance. Each curve is labeled with its prediction time and type of experiment (RSD, reduced station density; RSG, reduced size grid).

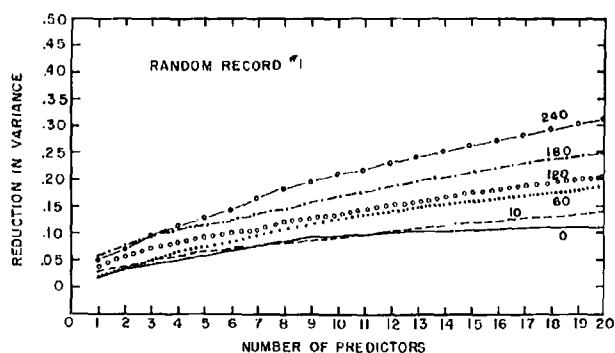


Fig. 7. The change in reduction in variance (RV) with accumulated predictors for random record #1 included in the experiments with the original station density and grid size. Each curve is labeled with the prediction time.

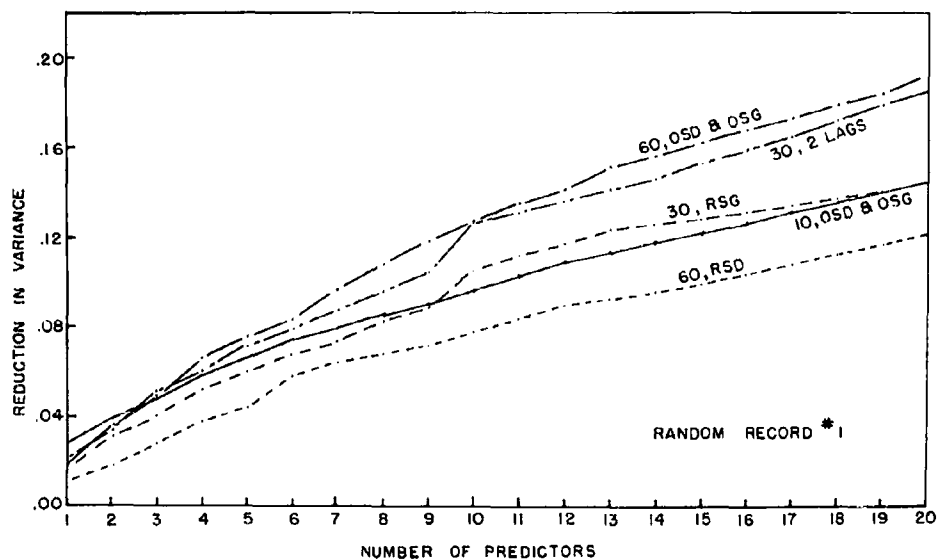
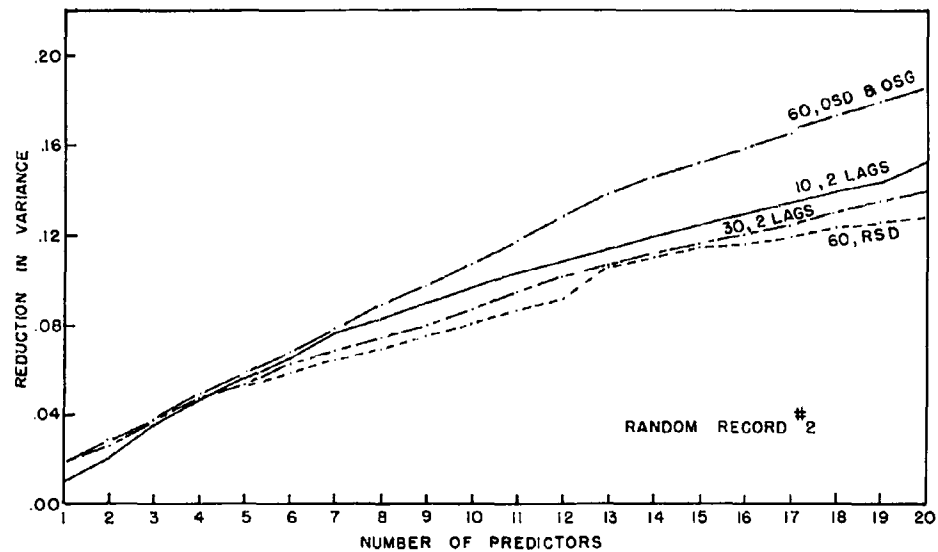


Fig. 8. The change in reduction in variance (RV) with accumulated predictors for the two random records. The curves are labeled with the prediction time and type of experiment (OSD and OSG, original station density and grid size; RSD, reduced station density; RSG, reduced size grid).

Table I. The mean percent reduction in variance for all the experiments for 10, 30 and 60 minute predictions. The mean is over all the predictands included in each experiment.

Prediction Time	10 min.	30	60
Original Station Density and Grid Size	67.3	49	41.5
Reduced Station Density	65.8	47.7	41.1
Reduced Size Grid	-	51	-
Two Time Lags	70.4	52.4	-

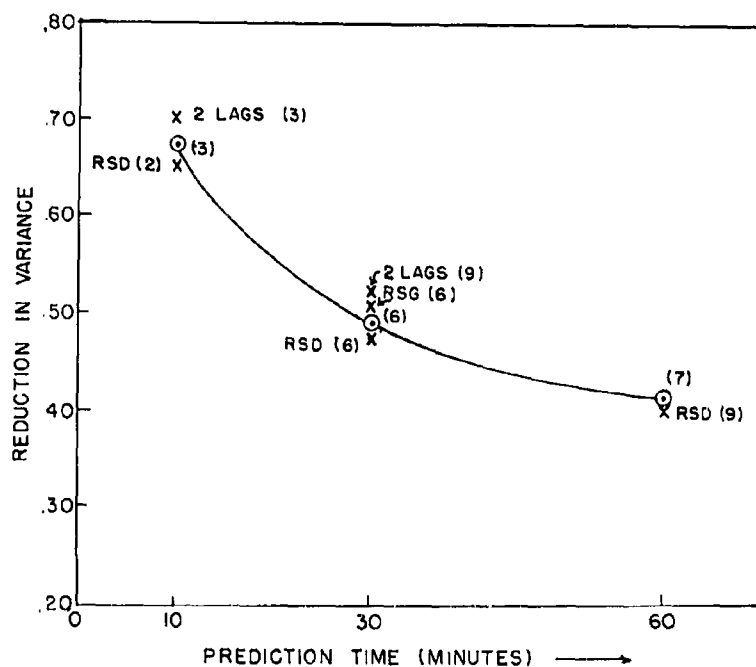


Fig. 9. The reduction in variance (RV) averaged over all the predictands in each experiment as a function of prediction time. The circled points are the values from the experiments with the original station density and grid size and the X's mark the values from the experiments with the reduced station density (RSD), the reduced size grid (RSG), and two time lags. In parentheses are the significant number of predictors averaged over all the predictands in each experiment.

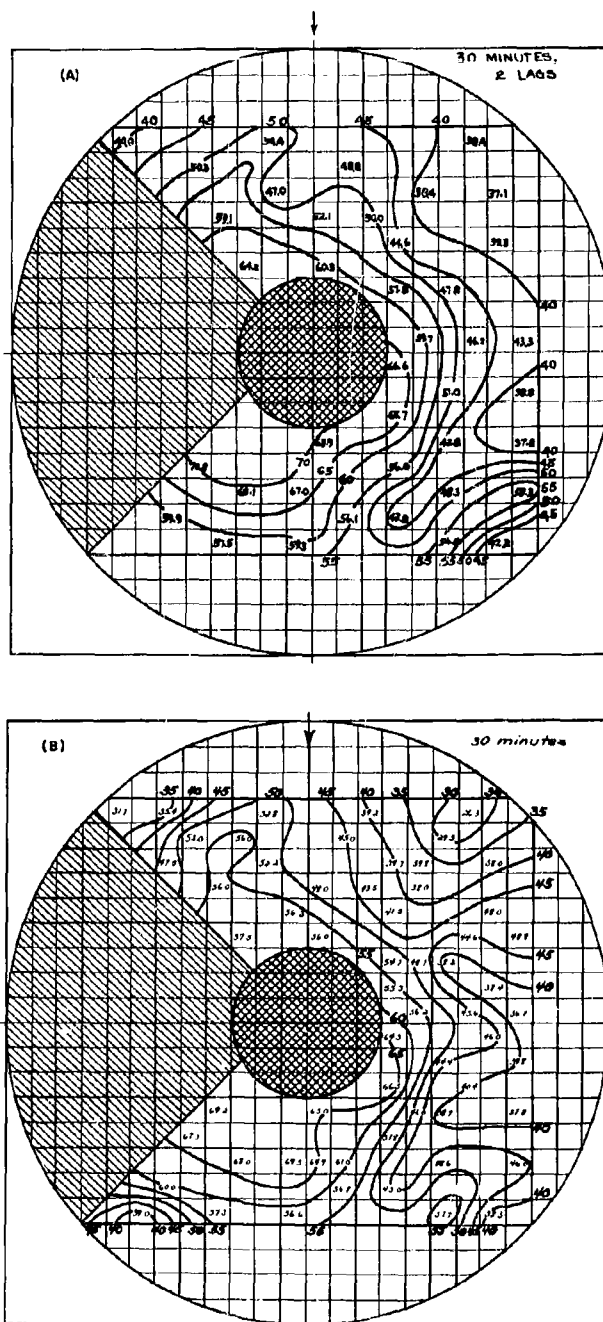


Fig. 11 (A, B). The space distribution of reduction in variance (RV) for the experiments with two time lags and the original station density and grid size for a 30 minute prediction.

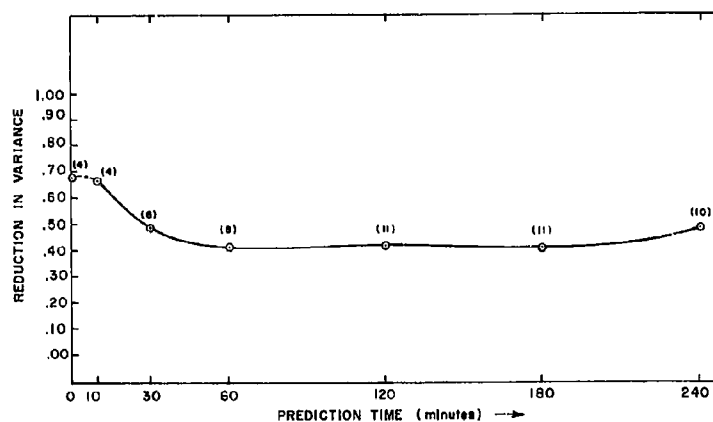


Fig. 12. The reduction in variance (RV) averaged over all 60 predictands for each prediction time in the experiments with the original station density and grid size. In parentheses are the significant number of predictors averaged over the 60 predictands.

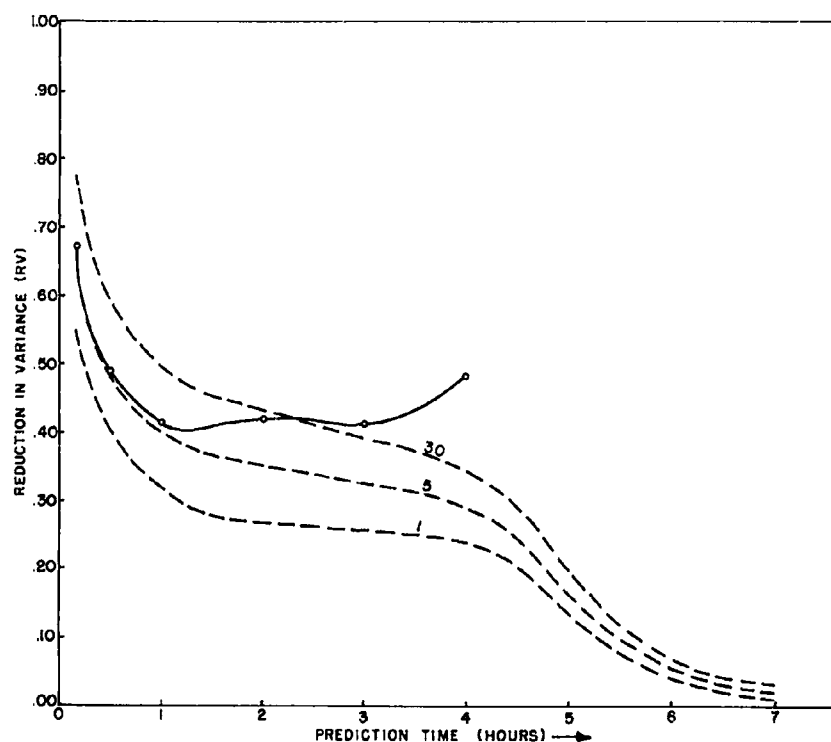


Fig. 13. Hypothesized mean reduction in variance for original station density and grid size but for very large data sample. Curves are for space resolution of 1, 5, and 30 miles respectively. Circled points and solid curve are reductions in variance for comparatively small sample in present study and space resolution of 5 miles.

REFERENCES

- Battan, L. J., 1953: "Duration of Convective Radar Cloud Units," Bull. Amer. Meteor. Soc., 34, p. 227.
- Blackmer, R. H., 1956: "The Lifetime of the Small Precipitation Echoes Observed on the PPI Scope of the SCR-615-B Radar," Research Report 25, Part II, Department of Meteorology, Massachusetts Institute of Technology.
- Miller, R., 1958: "Studies in Statistical Weather Prediction," Chapter II, Final Report, Contract Number AF 19(604)-1590, Travelers Weather Research Center.
- Newell, R. E., 1959: "Some Radar Observations of Tropospheric Cellular Convection," Research Report 33, Weather Radar Research, Department of Meteorology, Massachusetts Institute of Technology.
- Noel, T. M., and A. Fleisher, 1960: "The Linear Predictability of Weather Radar Signals," Research Report 34, Weather Radar Research, Department of Meteorology, Massachusetts Institute of Technology.
- Shorr, B., 1957: "Empirical Orthogonal Functions Applied to Prediction of the Sea-Level Pressure Field," Scientific Report 3, Statistical Forecasting Project, Department of Meteorology, Massachusetts Institute of Technology.