AN EXPERIMENT IN FORECASTING RAINFALL PROBABILITIES BY OBJECTIVE METHODS

HARRY R. GLAHN

U.S. Weather Bureau, Washington, D.C.

[Manuscript received August 30, 1961]

ABSTRACT

The regression screening and principal component techniques for developing forecast aids are investigated for their applicability to the objective forecasting of rainfall probabilities. The forecasting of summer rainfall in the Mississippi Delta is the particular problem studied. Subjective forecasts made for the area as well as objective forecasts are verified in terms of reductions of variance and saving over climatology. It is found that many of the forecast equations developed by regression screening and principal component techniques are not stable on test data. The results indicate that subjective screening of predictors is desirable before the regression screening is accomplished. It is found that useful aids can be developed with these linear techniques; at the same time the desirability of an approach that better integrates the physical processes of the atmosphere is indicated.

1. INTRODUCTION

For several years, summertime rainfall forecasts issued by the Weather Bureau in some areas of the country have included a quasi-quantitative indication of the areal coverage of rainfall. Certain terms used in the forecast, such as "risk of showers" and "widely scattered showers", have been given numerical meanings in terms of the percentage of the area which would receive rain. With only minor assumptions, such forecasts may be used in operational decision-making as if they were rain probability forecasts, and it is important to provide the forecaster with reliable objective techniques to aid in preparing the forecasts.

About two years ago, at the time the Weather Bureau began an expanded agricultural weather service in the Mississippi Delta region (see fig. 1), a research project was started in the Short Range Forecast Research Project of the Office of Meteorological Research to study this problem. The purpose of this paper is to present some results of an experiment performed as part of the study.

2. METHOD OF ATTACK

The problem of forecasting the areal coverage of summer rainfall has received little study as a physical problem. General rainfall throughout the Delta may often be attributed to an active low pressure system or to an instability line moving through the area. Intermediate and small values of areal coverage result from a wide variety of situations, in many of which the cause for the particular rain coverage is not clearly delineated by the available network of surface and upper-air observations. A statistical attack on the problem was indicated in view of the lack of physical and dynamical models which could be integrated to provide forecast estimates.

It is generally recognized that it is a long step between good or even perfect prognostic charts of large-scale circulation features and the forecasting or specification of actual weather parameters such as clouds and precipitation. (For example, see Sanders [1] and Sanders, Wagner, and

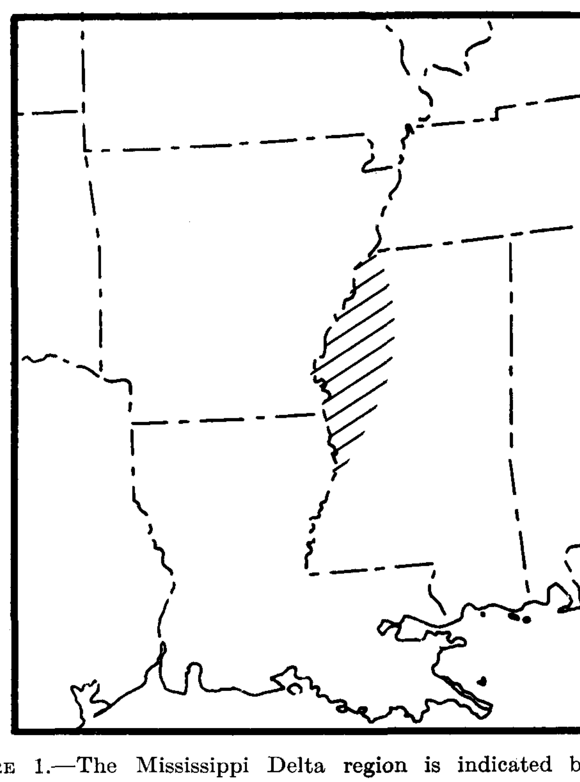

FIGURE 1.—The Mississippi Delta region is indicated by the hatched area.

Carlson [2].) A good method of analysis is needed to summarize and use information about future weather conditions from available observations or large-scale predicted parameters. The regression screening and regression with principal components techniques are designed to extract the linear predictive information from the input data made available to them. These techniques are feasible because of the availability of electronic computers to process large quantities of data.

3. TECHNIQUE DEFINITION

The screening regression technique is a method for picking predictors from a large group by stepwise least squares selection [3]. This procedure first picks the one best linear predictor and is equivalent to finding which of the M equations

$$Y_k = a_{0_k} + a_{1_k} X_k \qquad k=1, 2, \ldots, M$$

yields the highest reduction of variance, where X_k is the kth predictor, Y_k the predicted value of the predictand, a_{0_k} and a_{1_k} are constants, and M is the total number of possible predictors being considered.

Once this best X_k is found, the procedure is then to pick the best pair of predictors which includes the first predictor picked. This is the same as finding the best equation

$$Y_k = a_{0_k} + a_{1_k} X_1 + a_{2_k} X_k \qquad k=2, 3, \ldots, M$$

where the predictors are renumbered so that the first one picked has the subscript 1 and the remaining ones subscripts 2, 3, . . ., M.

This procedure can be carried out until P predictors have been chosen provided that $P \leq M$ and $P \leq N-1$ where N is the sample size.

The screening procedure does not necessarily yield the unique best set of predictors and it is possible that a type of screening in which pairs of predictors are considered at each selection step may be an improvement. These two variations will be referred to in this paper as "screening singly", or merely "screening", and "screening by pairs", respectively.

A technique that attempts to describe the linear information of a set of variables by a smaller set of variables is called principal component analysis. These components have been called by Lorenz [4] Empirical Orthogonal Functions. This analysis transforms the M time series of the M variables, or predictors, into M' new time series of M' new variables, $M' \leq M$, in such a way that these new variables are all mutually uncorrelated. Many times in meteorology 90 percent of the linear information in the M variables can be described by a set of M' variables where M' is only about 15 percent of M.

These new variables can now be used in a regression equation and again they can be screened to determine which ones to use. The use of the new orthogonal variables as predictors in screening regression will be called in this paper the principal component technique.

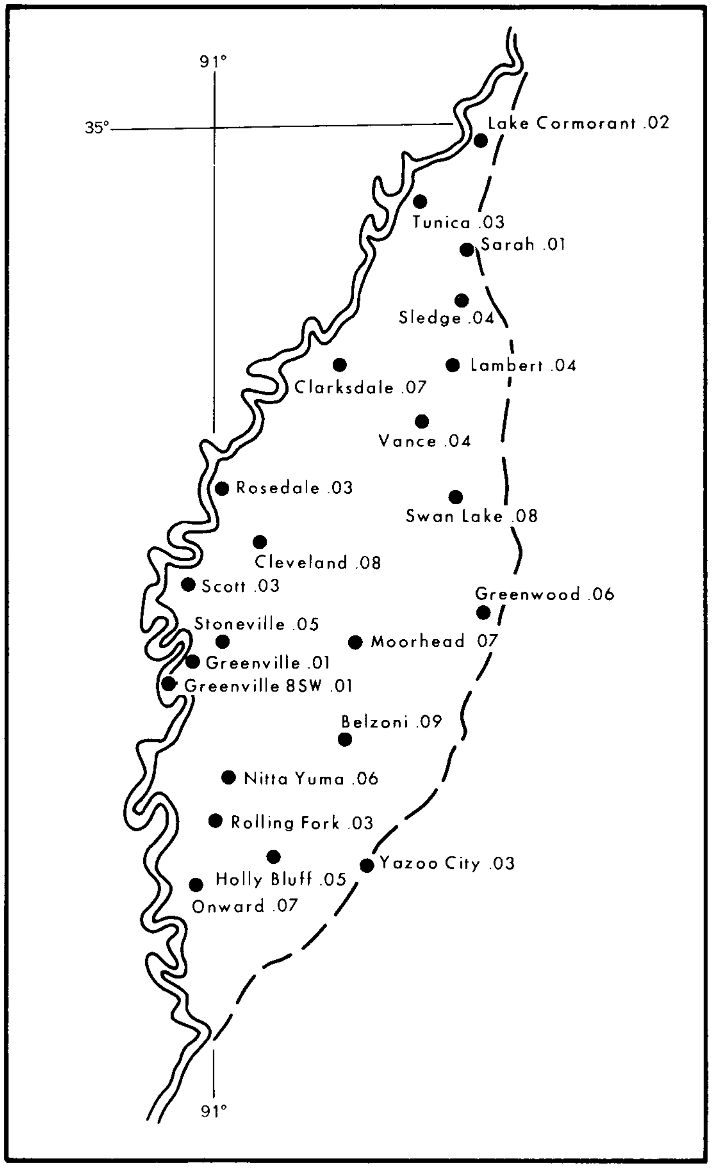

FIGURE 2.—Stations in the Mississippi Delta for which daily rainfall reports were available and the associated weighting function.

4. DATA

The Mississippi Delta is a rich agricultural plain in northwestern Mississippi between the Mississippi River on the west and the bluffs along the Yazoo River on the east (fig. 1). Rainfall observations taken at approximately 7:00 a.m. local time for about 20 stations in the Delta were available. These observations and a weighting function determined from the relative size of the geographical areas represented by the stations were used to estimate the areal coverage of rainfall occurrence. Each of these areas, represented by the station within it, was

Table 1.—*Scale used to assign weather code values. The lowest value derived from the hourly and special observations between the hours of 0000 and 0600* GMT *inclusive was used. The two levels of classification were used separately. For example, if precipitation occurred at any time during the 6 hours, the lowest ceiling height that was observed during that period was used in assigning the weather code value within that category.*

Classification	No clouds reported								≤0.5 clouds reported								>0.5 clouds, No precipitation											>0.5 clouds, Precipitation					
	$T-T_d$ (° F.)								$T-T_d$ (° F.)								Ceiling height (hundreds of ft.)											Ceiling height (hundreds of ft.)					
	≥24	20-23	16-19	12-15	8-11	4-7	2-3	0-1	≥24	20-23	16-19	12-15	8-11	4-7	2-3	0-1	≥100	90	80	70	60	50	40	30	20	10	0	≥50	40	30	20	10	0
Weather code value	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

determined such that all points within it were closer to that station than to any other station. Figure 2 shows the locations of the stations for which reports were available in 1959 and the associated weighting function. These 24-hr. areal coverage estimates were used as the predictand data.

For example, the predictand value for the 24-hr. period ending at 0700 EST, June 1, 1959 (assigned to the date June 1) was 0.50, which is the sum of all the weights of all stations which reported a trace or more of rainfall during that period.

The predictor data collected for study were of 14 types: the 0600 GMT values of sea level pressure, 3-hr. pressure change, surface temperature, and surface dew point for 91 United States stations; the 0000 GMT observations of temperature, relative humidity, and height at each of the levels 850 mb., 700 mb., and 500 mb. for 63 United States stations; and a "weather code" parameter derived from observations of precipitation, cloud height and amount, and temperature-dew point spread over a 6-hr. period (see table 1) for 61 stations covering the area of the United States east of the Rocky Mountains, south of 40° N and west of 88° W.

These data collected for June, July, and August, 1957 and 1958 composed the 184-case developmental sample. Also, these same parameters were collected for 1959 and were used for testing purposes.

5. VERIFICATION OF OBJECTIVELY MADE FORECASTS

The objective development techniques were applied to the data in a variety of ways in an attempt to learn something about their behavior on this type of problem. Some of the questions studied and the results are presented below.

A. LINEAR INFORMATION FURNISHED BY DIFFERENT TYPES OF PREDICTORS

Screening regression was carried out on each type of reported predictor separately. The stations selected for each type of predictor and the reductions of variance on the 1957–58 developmental data and on the 1959 test data are presented in figure 3. For these comparisons predictors were furnished the screening technique from Area 2 (see fig. 4). It seems apparent that, except for the sea level pressure, surface temperature, and surface dew point, the predictors as they were used here show little evidence of predictive information.

B. EFFECT OF SIZE OF AREA USED IN PREDICTOR SELECTION

Regression screening was performed on the sea level pressures from Area 1 so that the effect of including possible predictors from an area larger than Area 2 could be assessed. The reductions of variance on dependent and test data are shown in table 2.

Although Area 1 includes stations outside the boundaries of Area 2 and is made up of 62 stations, Area 2 is composed of a more dense network of stations and includes a total of 54 stations. The first three predictors picked were the same in both instances; the fourth differed only slightly in geographical location, and Concordia, the fourth picked from Area 2, was not included as an Area 1 station. The highest verification value reached with test data was with eight Area 2 predictors (RV=0.238), while Area 1 predictors gave a somewhat lower maximum value (RV=0.187). The latter reduction of variance was attained with five predictors all of which were from the smaller geographical region of Area 2. This suggests that

Table 2.—*Stations picked by screening sea level pressures and the reductions of variance (RV) on dependent and test data for Areas 1 and 2*

Order of selection	AREA 1 Station	RV dependent data	RV test data	AREA 2 Station	RV dependent data	RV test data
1	Shreveport	0.078	0.005	Shreveport	0.078	0.005
2	North Platte	.216	−.072	North Platte	.216	−.072
3	Montgomery	.325	.129	Montgomery	.325	.129
4	Topeka	.349	.159	Concordia	.361	.181
5	Little Rock	.375	.187	Fort Worth	.380	.193
6	Flint	.389	.159	Midland	.396	.195
7	Nashville	.404	.170	Billings	.405	.208
8	Green Bay	.423	.164	Memphis	.415	.238
9	Fort Worth	.438	.174	Tupelo	.433	.171
10	Midland	.451	.170	Farmington	.455	.093
11	Santa Maria	.459	.156	Nashville	.465	.080
12	Charleston	.463	.109	Galveston	.482	.116
13	Hatteras	.476	.084	Lake Charles	.492	.099
14	Norfolk	.488	.018	Mobile	.502	.135
15	Omaha	.496	--------	Burrwood	.513	.194
16				Jackson	.525	.146
17				Columbia	.530	.141
18				Broadus	.533	.139
19				Springfield	.535	.138

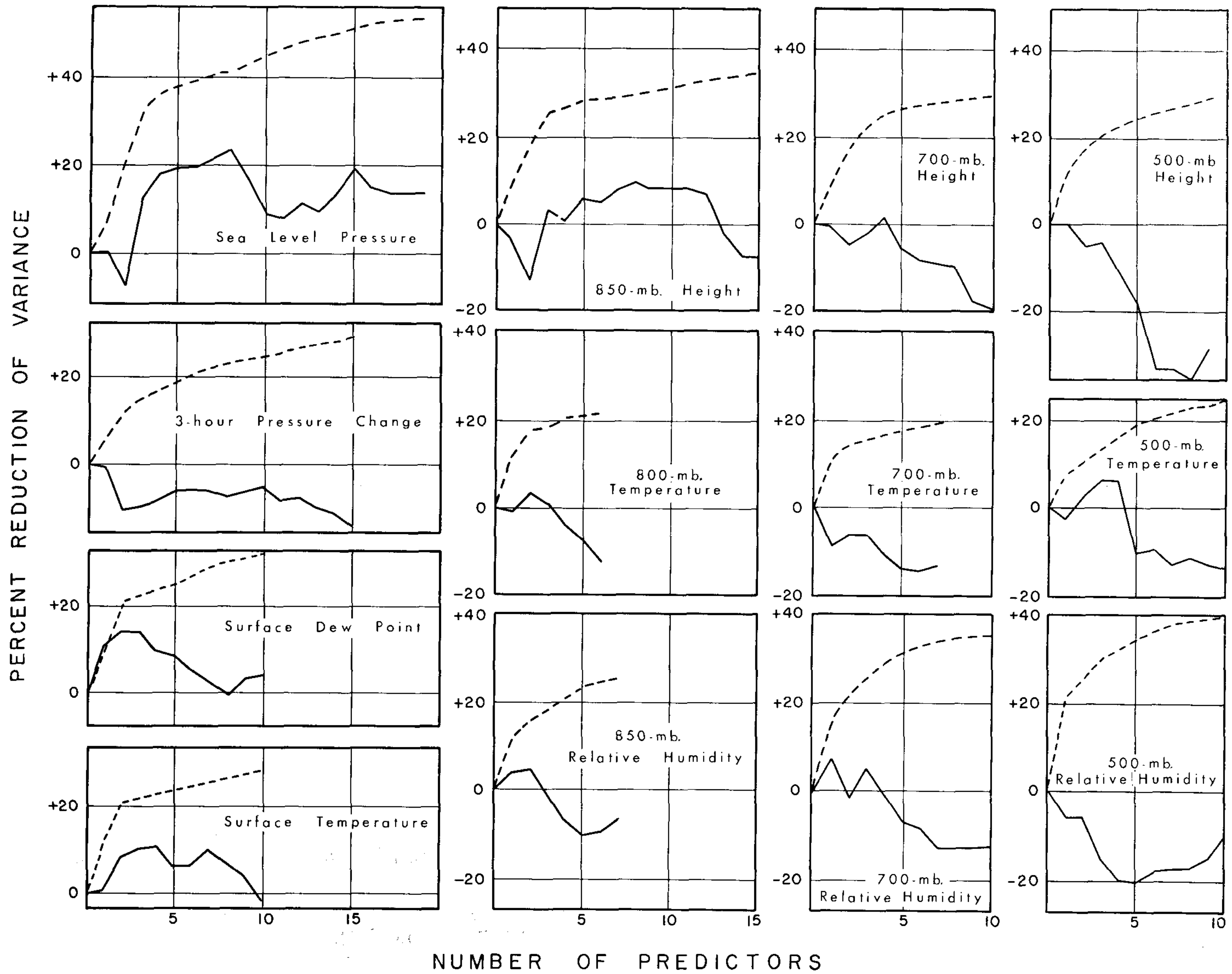

FIGURE 3.—Percent reductions of variance for dependent (dashed line) and test (solid line) data for the indicated types of predictors.

a network of stations in a limited region around the predictand area furnishes more predictive information than does a more sparse network in a larger geographical area.

Regression equations were also generated for Area 1 and Area 2 with orthogonal functions of the sea level pressures used as predictors. Verification results for these equations are shown in table 3. For a particular equation, the number of constants determined by the regression analysis is one more than the number of included functions. However, each equation can be put in terms of all variables which describe the orthogonal functions, 62 in the case of Area 1 and 54 in the case of Area 2.

TABLE 3.—*Orthogonal functions of sea level pressures picked by screening, the percent of predictor variance explained by each function, and the reductions of variance of the predictand on dependent and test data for Area 1 and 2 predictors*

	AREA 1				AREA 2			
Order of selection	Function number	Predictor explained variance	RV dependent data	RV test data	Function number	Predictor explained variance	RV dependent data	RV test data
1	4	0.058	0.082	−0.021	2	0.153	0.155	−0.041
2	8	.017	.136	−.158	9	.006	.213	.046
3	3	.095	.192	−.079	6	.021	.267	.102
4	9	.016	.240	.038	11	.004	.312	.215
5	14	.005	.270	.077	8	.009	.334	.220
6	15	.004	.293	.085	3	.086	.342	.212
7					13	.003		
8					5	.025		
9					1	.603	.360	.202
10					12	.004		
11					10	.004		
12					7	.010	.367	.199

C. COMPARISON OF SCREENING AND PRINCIPAL COMPONENT ANALYSIS WITH COMBINATIONS OF TYPES OF PREDICTORS

In order to obtain a prediction equation which involved several types of data, the first few predictors picked in individual previous screenings of all 14 types of data from Area 1, except the weather code, were screened by pairs. These same 73 variables that were screened were also put into regression equations with the principal component technique. The results are shown in table 4.

The test data verifications of the screening equations again reflect the fact that most of the types of data do not determine stable coefficients on this sample of data. The inclusion of the sea level pressure for Fort Worth, Montgomery, and Concordia seems to be the main reason for the positive reductions of variance on test data.

Although the reductions of variance are not high, the principal component technique seemed to give a more stable relationship than did the screening.

D. THE STABILITY OF THE WEATHER CODE AS A PREDICTOR

The weather code was the only derived predictor used; all of the others were directly reported parameters. To test the usefulness of the weather code alone the first 10 stations picked from the 61-station network by screening were tested. Table 5 indicates the results. Although satisfactory reductions of variance were found on the dependent data, the equations were not useful on the test data.

TABLE 4.—*Results of screening by pairs and principal component techniques applied to 73 variables picked on previous screenings from all available types of data except weather code*

SCREENING BY PAIRS				PRINCIPAL COMPONENT TECHNIQUE				
Order of selection	Predictor identification	RV dependent data	RV test data	Order of selection	Function number	Predictor explained variance	RV dependent data	RV test data
1,2	Shreveport 500-mb. R.H. Dodge City 850-mb. R.H.	0.288	−0.021	1 2	3 2	0.096 .134	0.274 .353	0.122 .124
3,4	Brownsville 700-mb. R.H. Fort Worth S.L. pressure	.397	−.011	3 4	1 5	.227 .042	.410 .443	.156 .152
5,6	Montgomery S.L. pressure Concordia S.L. pressure	.514	.101	5	35	.005	.472	.147
7,8	Fort Smith 3-hr. pressure change Portland 3-hr. pressure change	.568	.078					
9,10	Montgomery 850-mb. height Nashville Sfc. dew point	.598	.123					
11,12	New York Sfc. temperature Midland 3-hr. pressure change	.633	.114					
13,14	Little Rock 850-mb. temperature Oklahoma City S.L. pressure change	.655	.046					
15	Jackson 850-mb. R.H.	.667	−.001					

TABLE 5.—*Weather Code predictors and factors chosen by screening and the reductions of variance on dependent and test data*

	REGRESSION SCREENING			PRINCIPAL COMPONENT TECHNIQUE			
Order of selection	Predictor chosen	RV dependent data	RV test date	Factor number	Predictor explained variance	RV dependent data	RV test data
1	Pine Bluff, Ark	0.253	0.050	1	0.271	0.343	−0.035
2	Oklahoma City, Okla	.328	−.014	4	.056	.373	.038
3	Walnut Ridge, Ark	.359	−.032	42	.003	.392	.015
4	Burrwood, La	.384	−.027	6	.032	.402	.008
5	Garden City, Kans	.406	.009	38	.004	.411	
6	Monroe, La	.418	−.012	33	.005	.418	
7	Baton Rouge, La	.427	.008	26	.007	.425	
8	Greenwood, Miss	.434	−.054	18	.010	.431	
9	McComb, Miss	.440	−.047	14	.013	.436	
10	College Station, Tex	.447	−.070	34	.005	.441	

In addition to the attempt to find a useful relationship between the predictand and the weather code alone, a test was made of the combined effects of weather code and other types of predictors. A screening was performed on 119 predictors which included the ones picked first in previous individual screening of each type of predictor from Area 2 except the weather code. Another screening was made on these same 119 predictors plus the 10 weather code predictors judged best by a previous screening of weather code alone. The results are shown in table 6.

The inclusion of the weather code in the list of possible predictors caused a marked change in the first 10 predictors picked. When weather code was omitted, no sea level pressure predictors were picked. With the inclusion of weather code only two of the first 10 picked were in-

TABLE 6.—*Predictors chosen and reductions of variance for equations produced by screening 119 predictors which did not include weather code (on the left) and the same 119 predictors plus 10 weather code predictors (on the right)*

	SCREENING WITHOUT WEATHER CODE		SCREENING WITH WEATHER CODE		
Order of selection	Predictor chosen	RV dependent data	Predictor chosen	RV dependent data	RV test data
1	Shreveport 500-mb. R.H.	0.215	Pine Bluff, Ark. Wea. code	0.253	0.048
2	Dodge City 850-mb. R.H.	.288	Shreveport 500-mb. R.H.	.321	.039
3	Nashville Surface dew point	.336	Garden City, Kans. Wea. code	.368	−.014
4	Little Rock 500-mb. height	.377	Columbia 500-mb. R.H.	.402	−.010
5	Brownsville 700-mb. R.H.	.421	North Platte 3-hr. pressure change	.432	−.041
6	Little Rock 850-mb. temperature	.446	Little Rock 3-hr. pressure change	.460	−.016
7	Fort Smith 3-hr. pressure change	.471	Brownsville 700-mb. R.H.	.478	−.027
8	St. Cloud 3-hr. pressure change	.497	Fort Worth S.L. pressure	.507	−.000
9	Dodge City 500-mb. R.H.	.514	Concordia S.L. pressure	.544	.076
10	Midland 3-hr. pressure change	.526	Montgomery S.L. pressure	.581	.129

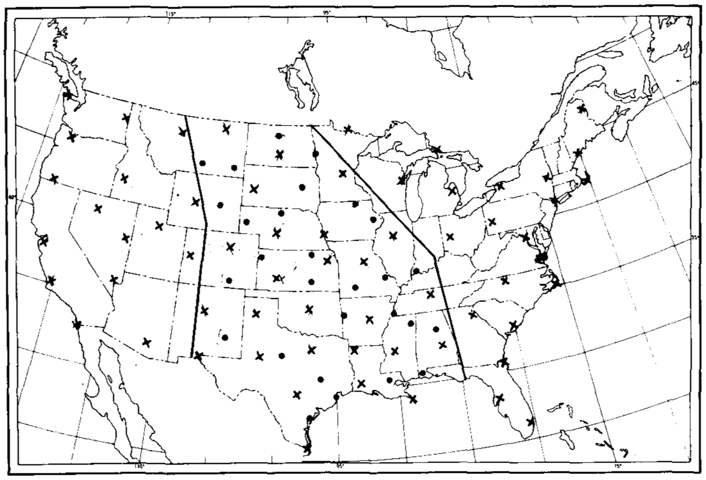

FIGURE 4.—Stations from which data were used. Crosses indicate the 62 stations in Area 1. All 54 stations (crosses and dots) in the central United States between the indicated boundaries are included in Area 2.

cluded in the first 10 picked with the screening without weather code. Again, the only promising verifications on test data were those which included the Fort Worth, Montgomery, and Concordia sea level pressures.

E. COMPARISON OF THREE VARIATIONS OF SCREENING

It was found by screening a group of 146 predictors composed of sea level pressure, 3-hr. pressure change, and surface temperature at selected stations in Area 2 that the first 10 predictors selected were not as efficient in explaining the variance on dependent data as the first 10 predictors selected from the subset of sea level pressures alone. Also, reductions of variance on test data were more promising when sea level pressures alone were used.

Two more screenings were performed on the 146 variables. One was screening by pairs and the other was screening singly but after the first predictor was designated to be the first sea level pressure selected when the sea level pressure subset was screened. The latter was accomplished by use of a provision in the computer program whereby a list of predictors could be "forced" before the screening started. The results of these three comparative runs are shown in table 7; the results of screening the subset of Area 2 sea level pressures are shown in table 2.

The first few predictors selected by the screening singly method differed markedly from the first few selected by the other two screening variations. The first two selected in the former case were surface temperatures, while the first eight in screening by pairs and the first eleven in screening with the first predictor forced did not include a surface temperature.

F. SPECIFICATION VERSUS PREDICTION

The question of whether it is better to develop aids for forecasting small-scale weather elements based on large-scale forecast parameters or to develop the aids based on only observed data has not been settled. Forecast aids are sometimes developed from parameters observed at the time for which the forecast is to be made and at application

TABLE 7.—*Predictors selected and reductions of variance for dependent and test data for three different variations of screening.*

	SCREENING SINGLY			SCREENING BY PAIRS			FIRST PREDICTOR FORCED		
Order of selection	Predictor chosen	RV dependent data	RV test data	Predictor chosen	RV dependent data	RV test data	Predictor chosen	RV dependent data	RV test data
1	Dodge City Sfc. temp.	0.123	0.014	Fort Worth S.L. pressure	0.262	0.041	Shreveport S.L. pressure	0.078	0.005
2	Nashville Sfc. temp.	.203	.087	Concordia S.L. pressure			North Platte S.L. pressure	.216	−.072
3	Little Rock 3-hr. pres. change	.248	.160	Montgomery S.L. pressure	.407	.133	Montgomery S.L. pressure	.325	.129
4	North Platte 3-hr. pres. change	.287	.121	Fort Smith 3-hr. pres. change			Little Rock 3-hr. pres. change	.380	.176
5	Galveston S.L. pressure	.313	.163	Shreveport 3-hr. pres. change	.470	.241	Concordia S.L. pressure	.411	.212
6	Mason City S.L. pressure	.358	.121	Memphis S.L. pressure			Fort Worth S.L. pressure	.435	.231
7	Burrwood S.L. pressure	.385	.155	Farmington S.L. pressure	.518	.116	Shreveport 3-hr. pres. change	.460	.250
8	Fort Smith 3-hr. pres. change	.408	.123	North Platte 3-hr. pres. change			Fort Smith 3-hr. pres. change	.476	.215
9	Shreveport S.L. pressure	.428	.111	San Antonio Sfc. temp.	.547	.095	North Platte 3-hr. pres. change	.495	.167
10	Montgomery S.L. pressure	.446	.116	Tupelo S.L. pressure			Omaha 3-hr. pres. change	.511	.100
11	Mobile S.L. pressure	.470	.141	Minot 3-hr. pres. change	.571	.079	Minot 3-hr. pres. change	.524	.123
12	Nashville S.L. pressure	.491	.124	Omaha 3-hr. pres. change			San Antonio Sfc. temp.	.535	.144
13	Farmington S.L. pressure	.511	.086	Galveston S.L. pressure	.593	.051	Mobile 3-hr. pres. change	.544	.159
14	Lake Charles S.L. pressure	.526	.086	Nashville S.L. pressure			Nashville S.L. pressure	.552	.178
15	Shreveport 3-hr. pres. change	.539	.102				Farmington S.L. pressure	.568	.125

TABLE 8.—*Stations picked for specification by screening sea level pressures and 500-mb. heights and the reduction of variance on dependent and test data for areas 1 and 2*

Number of predictor	500-MB. HEIGHTS						SEA LEVEL PRESSURES					
	Area 1			Area 2			Area 1			Area 2		
	Station	RV dependent data	RV test data	Station	RV dependent data	RV test data	Station	RV dependent data	RV test data	Station	RV dependent data	RV test data
1	Little Rock	0.203	0.081	Little Rock	0.203	0.081	Jackson	0.085	0.004	Jackson	0.085	0.004
2	Jacksonville	.292	.216	Montgomery	.262	.151	Dodge City	.209	.006	Dodge City	.209	.006
3	Boise	.313	.219	Brownsville	.284	.129	Burrwood	.287	.087	Mobile	.288	.173
4	Greensboro	.339	.301	Columbia	.307	.177	Tuscon	.325	.069	Albuquerque	.317	.149
5	Brownsville	.366	.283	Jackson	.318	.179	Peoria	.356	—.026	Peoria	.349	.072
6	Santa Maria	.382	.292	El Paso	.330	.146	Shreveport	.374	.098	Memphis	.370	.075
7	Tampa	.390	.288	Topeka	.339	.165	Hatteras	.385	.165	Amarillo	.387	.032
8	Columbia	.396	.303	North Platte	.348	.149	Nashville	.397	.160	Billings	.398	—.015
9	Nashville	.406	.299	Burrwood	.352	.160	St. Cloud	.407	.187	St. Cloud	.409	.021
10	Oklahoma City	.417	.284	Oklahoma City	.356	.161	Lander	.416	.153	Mason City	.418	.016

time the appropriate forecast values are used for dependent parameters.

Regression equations were developed with sea level pressure and 500-mb. height data used to specify the rainfall. The 0600 GMT pressure observations were taken after about three-fourths of the forecast period had passed and the 0000 GMT height observations were taken mid-way during the period. The test verifications are shown in table 8.

Area 2 was chosen originally as being desirable from a forecast point of view and is probably too far west for good specification. It is somewhat surprising to note that the sea level pressures as used in this study seem to be better predictors than specifiers. However, for 500-mb. heights the opposite is true.

6. VERIFICATION OF CONTROL FORECASTS

In order to compare the results of the objective techniques with the level of capability that now exists for this type of forecast, the official forecasts issued at 5:00 a.m. at the Weather Bureau Airport Station, Jackson, Miss. for the Delta were verified for the two summers of 1959 and 1960. Also, as another control, each of three meteorologists independently made an experimental probability-of-rain forecast for the Delta for each day in June, July, and August, 1959 and 1960. These forecasts were made at 7:30 a.m. CDT with the use of only the *Daily Weather Map*, published by the U.S. Weather Bureau, dated that same day. The arithmetic mean of the three experimental forecasts was used as an "average" daily experimental forecast.

The official forecasts were issued for two consecutive 12-hour periods instead of a 24-hr. period and for each period a range of values was used. For example, the forecast for the first 12-hr. period may have read "60 to 80 percent of the area will have rain," and for the second period, "0 to 20 percent of the area will have rain."

To combine these two 12-hr. ranges into a single number forecast for the 24-hr. period (a consumer in the Delta desiring a 24-hr. probability forecast would have to do this) the midpoint of the "extreme range" was used, where the extreme range was defined as that range specified by the smallest and the largest numbers, consistent with the two 12-hr. forecasts, that could have been forecast for the 24-hr. period. Thus, in the preceding example, the extreme range would be from 60 to 100 percent and the midpoint would be 80 percent. Verification of the two 12-hr. forecasts separately was not possible since only 24-hr. rainfall observations were available at most of the stations in the Delta.

The verification statistics are shown in table 9. In 1960 there were only 40 days on which all three experimental forecasts were available. The average for each of the 91 days was determined by only two forecasts for the other 51 days. Since for some purposes it might be desirable to omit a trace of rain as a rain occurrence, the 1960 forecasts were also verified with a trace being counted as no rain in the areal coverage determination.

The economic utility of the official and experimental forecasts was computed as described by Thompson and Brier [5] and the graphs of the saving in dollars per dollar potential loss versus the cost-over-loss ratio are shown in figure 5.

7. DISCUSSION AND COMMENTS

In the verifications of the objectively made forecasts the reduction of variance has been computed. A more realistic statistic is the reduction of error defined by Lorenz [4]. However, in this case it does not matter as the mean areal coverage for each of the two samples was 0.33. The 18-year climatological value available when the forecasts were made was 0.26.

It is difficult to form a definite conclusion as to why many of the prediction equations did not yield acceptable forecasts for the test period. It is possible that the data sample was too small to be treated as it was. The indications are that the geographical area from which variables are selected for use in screening or principal component analysis should be picked rather carefully. The inclusion of predictors outside Area 2 did not improve test results.

It is probable that some method which attempts to consider advection parameters and parameters derived

TABLE 9.—*Verification data for forecasts of the probability of rainfall in the Mississippi Delta for 1959 and 1960. Traces were counted as a rain occurrence except where noted*

	Number of forecasts	Mean observed areal coverage	Mean of forecasts	Reduction of error based on climatological probability of 0.26	Comments
Forecaster A, 1959	92	0.335	0.305	0.214	
Forecaster B, 1959	92	.335	.224	.143	
Forecaster C, 1959	92	.335	.336	.163	
Official, 1959	92	.335	.195	.121	
Average of A, B, C, 1959	92	.335	.288	.263	
Persistence	92	.335	.332	−.225	
Forecaster A, 1960	40	.320	.261	.422	Days when all forecasters A, B, and C made forecasts.
Forecaster B, 1960	40	.320	.269	.408	
Forecaster C, 1960	40	.320	.267	.185	
Official, 1960	40	.320	.166	.049	
Official, 1960	91	.264	.144	.164	
Average of A, B, C, 1960	91	.264	.229	.296	
Official, 1960	91	.227	.144	.226	Traces not counted as rain.
Average of A, B, C, 1960	91	.227	.229	.313	

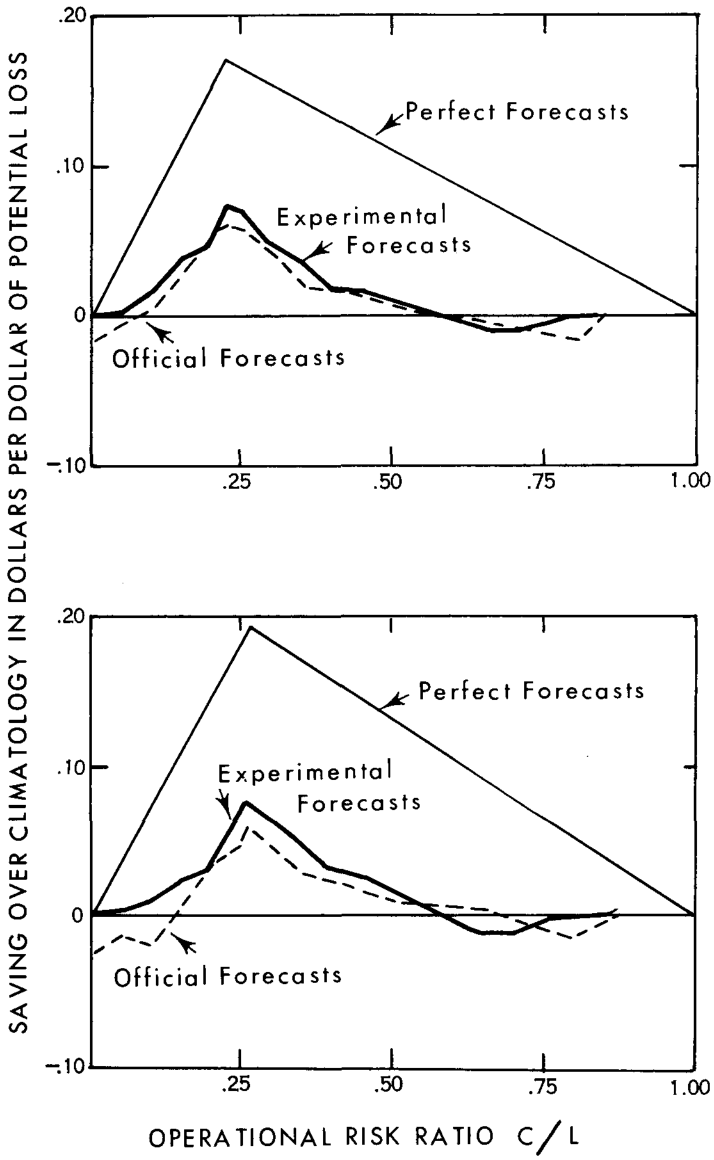

FIGURE 5.—Saving over climatology verification for forecasts of the probability of rainfall in the Mississippi Delta for June, July, and August, 1960. The arithmetic mean of three, or two if only two were available, independent experimental forecasts made by forecasters A, B, and C, and the official forecasts are verified with a trace counted as a rain occurrence.

from dynamic models such as vertical velocity as nonlinear operators would be more successful than the completely linear techniques described above. A nonlinear approach was tried but the data sample was too small to support this type of analysis and the results were inconclusive.

There seems to be some evidence to support the hypothesis that prediction equations determined through principal component analysis are more stable than those determined by screening, provided that the variables used are reasonable from a physical standpoint. In any predictor selection procedure it must be determined how many predictors to use in the application of the method. There is no good statistical significance test applicable to the screening technique. Experience is probably the best guide to follow in determining how many terms to retain in a regression equation. There is a decided advantage in using a technique that yields a forecast method such that the stability is not extremely sensitive to the number of terms retained. Table 4 indicates that the usefulness of the screening method depends quite heavily on the number of predictors used while the opposite is true of the principal component method.

Little evidence was found to suggest that screening by pairs was a better technique than screening singly. In many cases the variables picked by the two methods were very nearly the same, while the case presented in table 7 shows the opposite to be true. Miller's [6] statement that two prediction equations may contain mutually exclusive sets of predictors and still produce equally accurate forecasts is borne out here: the high redundancy of information in meteorological variables dictates this result.

It must be realized that all testing was done on the same data sample. Therefore, the many verifications performed are not independent. Much care must be taken in drawing conclusions from multiple verifications on a single small sample such as this. While sea level pressure has shown more promise as a predictor in this study, it may well be that another type of predictor would be more stable in future test samples.

ACKNOWLEDGMENT

Thanks are due the members of the Short Range Forecast Research Project who participated in this project and especially Mr. Roger Allen for his guidance and review of the manuscript.

REFERENCES

1. F. Sanders (editor), "Synoptic Application of Dynamical Concepts," *Final Report* on contract AF19(604)-1305, Massachusetts Institute of Technology, Feb. 1959.
2. F. Sanders, A. J. Wagner, and T. N. Carlson, "Specification of Cloudiness and Precipitation by Multi-Level Dynamical Models," *Scientific Report No. 1* on contract AF19(604)-5491, Massachusetts Institute of Technology, Sept. 1960.
3. A. Lubin and A. Summerfield, "A Square Root Method of Selecting a Minimum Set of Variables in Multiple Regression, Part I The Method," *Psychometrika*, vol. 16, No. 3, Sept. 1951, and "Part II A Worked Example," *Psychometrika*, vol. 16, No. 4, Dec. 1951.
4. E. N. Lorenz, "Empirical Orthogonal Functions and Statistical Weather Prediction," *Scientific Report No. 1* on contract AF19(604)-1566, Massachusetts Institute of Technology, Dec. 1956.
5. J. C. Thompson and G. W. Brier, "The Economic Utility of Weather Forecasts," *Monthly Weather Review*, vol. 83, No. 11, Nov. 1955, pp. 249-254.
6. R. G. Miller, "An Application of Multiple Discriminant Analysis to the Probabilistic Prediction of Meteorological Conditions Affecting Operational Decisions," *Technical Memorandum No. 4*, The Travelers Research Center, Inc., Hartford, Conn., Mar. 1961.