

WEATHER BUREAU
Systems Development Office
Techniques Development Laboratory
Silver Spring, Md.

October 1969

An Operational Method For Objectively Forecasting Probability of Precipitation

HARRY R. GLAHN AND DALE A. LOWRY



Technical Memorandum WBTM TDL 27

U.S. DEPARTMENT OF COMMERCE / ENVIRONMENTAL SCIENCE SERVICES ADMINISTRATION

ESSA TECHNICAL MEMORANDA

Weather Bureau, Techniques Development Laboratory Series

The primary purpose of the Techniques Development Laboratory of the Office of Systems Development is to translate increases in basic knowledge in meteorology and allied disciplines into improved operating techniques and procedures. To achieve this goal, TDL conducts and sponsors applied research and development aimed at improvement of diagnostic and prognostic methods for producing weather information. The laboratory carries out studies both for the general improvement of prediction methodology used in the National Meteorological Service System and for more effective utilization of weather forecasts by the ultimate user.

ESSA Technical Memoranda in the Weather Bureau Techniques Development Laboratory series facilitate rapid distribution of material which may be preliminary in nature and which may be published formally elsewhere at a later date. The first five papers in the TDL series are part of the former Weather Bureau Technical Notes series.

Papers listed below are available from the Clearinghouse for Federal Scientific and Technical Information, U.S. Department of Commerce, Sills Bldg., 5285 Port Royal Road, Springfield, Va. 22151. Price: \$3.00 hard copy; \$0.65 microfiche. Order by accession number shown in parentheses at end of each entry.

- TN 10 TDL 1 Objective Prediction of Daily Surface Temperature. William H. Klein, Curtis W. Crockett, and Carlos R. Dunn, October 1965. (PB-168 590)
- TN 11 TDL 2 Hurricane Cindy Galveston Bay Tides. N. A. Pore, A. T. Angelo, and J. G. Taylor, September 1965. (PB-168 608)
- TN 29 TDL 3 Atmospheric Effects on Re-Entry Vehicle Dispersions. Karl R. Johannessen, December 1965. (PB-169 381)
- TN 45 TDL 4 A Synoptic Climatology of Winter Precipitation from 700-mb. Lows for the Intermountain Areas of the West. D. L. Jorgensen, W. H. Klein, and A. F. Korte, May 1966. (PB-170 635)
- TN 47 TDL 5 Hemispheric Specification of Sea Level Pressure from Numerical 700-mb. Height Forecasts. William H. Klein and Billy M. Lewis, June 1966. (PB-173 091)
- WBTM TDL 6 A Fortran Program for the Calculation of Hourly Values of Astronomical Tide and Time and Height of High and Low Water. N. A. Pore and R. A. Cummings, January 1967. (PB-174 660)
- WBTM TDL 7 Numerical Experiments Leading to the Design of Optimum Global Meteorological Networks. M. A. Alaka and F. Lewis, February 1967. (PB-174 497)
- WBTM TDL 8 An Experiment in the Use of the Balance Equation in the Tropics. M. A. Alaka, D. T. Rubsam, and G. E. Fisher, March 1967. (PB-174 501)
- WBTM TDL 9 A Survey of Studies of Aerological Network Requirements. M. A. Alaka,
 May 1967. (PB-174 984)
- WBTM TDL 10 Objective Determination of Sea Level Pressure from Upper Level Heights.
 William Klein, Frank Lewis, and John Stackpole, May 1967. (PB-179 949)
- WBTM TDL 11 Short Range, Subsynoptic Surface Weather Prediction. H. R. Glahn and D. A. Lowry, July 1967. (PB-175 772)
- WBTM TDL 12 Charts Giving Station Precipitation in the Plateau States from 700-Mb.
 Lows During Winter. D. L. Jorgensen, A. F. Korte, and J. A. Bunce, Jr.,
 October 1967. (PB-176 742)
- WBTM TDL 13 Interim Report on Sea and Swell Forecasting. N. A. Pore and W. S. Richardson, December 1967. (PB-177 038)
- WBTM TDL 14 Meteorological Analysis of 1964-65 ICAO Turbulence Data. DeVer Colson, September 1968. (PB-180 268)
- WBTM TDL 15 Prediction of Temperature and Dew Point by Three-Dimensional Trajectories. Ronald M. Reap, September 1968. (PB-180 727)

(Continued on inside back cover)

U.S. DEPARTMENT OF COMMERCE Environmental Science Services Administration Weather Bureau

ESSA Technical Memorandum WBTM TDL 27

AN OPERATIONAL METHOD FOR OBJECTIVELY FORECASTING PROBABILITY OF PRECIPITATION

Harry R. Glahn and Dale A. Lowry



Systems Development Office Techniques Development Laboratory

> Silver Spring, Md. October 1969

UDC 551.509.314:551.509.324(73)

551.5

Meteorology

.509

Forecasting

. 314

.314 Objective forecasting

(73)

United States

CONTENTS

	Page
Abstract	1
Introduction	1
The Screening Regression Procedure	3
Development of Equations	5
Seasonal Regression Equations	6
Summer 1968	7
Winter 1968-69	8
Summer 1969	10
Operational Transmissions	12
Accuracy of the POP Forecasts	16
Future Plans	22
Conclusions	22
Acknowledgments	22
References	23

AN OPERATIONAL METHOD FOR OBJECTIVELY FORECASTING PROBABILITY OF PRECIPITATION

Harry R. Glahn and Dale A. Lowry

ABSTRACT

An operational method for objectively producing forecasts of probability of precipitation (PoP) is discussed. The screening regression technique is used to relate precipitation observations to variables forecast by two existing numerical models, the TDL Subsynoptic Advection Model (SAM) and the NMC Primitive Equation (PE) Model. Verification figures for a 1-year period are presented for the eastern United States. These figures indicate that the objective forecasts were slightly more accurate than the official forecasts issued to the public by Weather Bureau local offices.

INTRODUCTION

A method has been developed that objectively produces forecasts of probability of precipitation (PoP). The method is based on a procedure called Model Output Statistics (MOS), which statistically relates precipitation observations to variables forecast by existing numerical models. Two numerical models have been used in this particular application—the Subsynoptic Advection Model (SAM) developed by the Techniques Development Laboratory (TDL) [10] and the Primitive Equation (PE) Model developed by the National Meteorological Center (NMC) [24].

The PoP forecasts are produced as an integral part of the twice-daily operational SAM computer program [11, 9, 10]. This program was developed to use the latest observations and numerical forecasts available in order to produce objective forecasts of PoP at the time they are needed by the local Weather Bureau offices for issuance to the public. The area covered and the SAM numerical grid are shown in figure 1. The evolution of the SAM project, from the testing of the model through the current operational program, is described in the Weather Bureau Technical Procedures Bulletin Series [3, 4, 5, 6, 7]. One of the unique features of SAM, in addition to the data time and grid length considerations, is the specification of initial layer moisture by using only surface observations [10, 15]. This feature allows the moisture field to be defined on a small scale (fig. 1) and at times other than 0000 and 1200 GMT. The latest surface data used in SAM are 0700 GMT for the "today" forecast and 1900 GMT for the "tonight" forecast.

Statistical relationships were derived by using the screening regression technique. A brief description of this procedure follows.

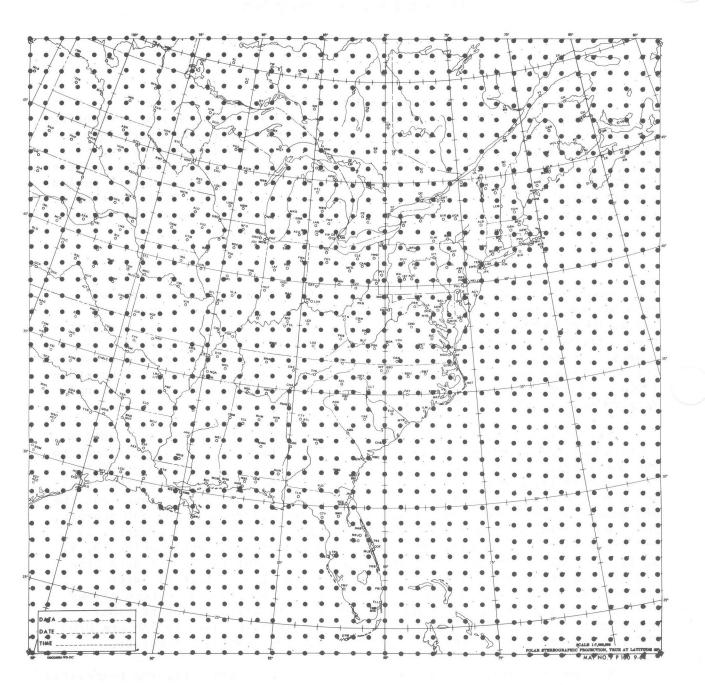


Figure 1.--The SAM 39 x 40 grid. The gridlength is exactly ½ that used for the NMC synoptic scale products. This is approximately 50 miles and not much different from the average spacing of hourly reporting stations in the eastern and central United States.

THE SCREENING REGRESSION PROCEDURE

Multiple linear regression relates one variable Y, called the dependent variable or predictand, to k other variables X, called the independent variables or predictors. The result is an equation which can be used for estimating the predictand as a linear combination of the predictors:

$$\hat{Y} = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_k X_k$$

The carat indicates an estimate, and the a 's are the regression constant and coefficients. The a 's are determined such that the sum of the squares of the estimation errors is a minimum on the developmental (or dependent) sample of size n, i.e.,

$$\sum_{j=1}^{n} \left(y_{j} - \hat{y}_{j} \right)^{2} = Minimum$$

A measure of the goodness of the equation for estimating Y is the reduction of variance RV, where

$$RV = \frac{\frac{1}{n} \sum_{j=1}^{n} (y_{j} - \bar{y})^{2} - \frac{1}{n} \sum_{j=1}^{n} (y_{j} - \hat{y}_{j})^{2}}{\frac{1}{n} \sum_{j=1}^{n} (y_{j} - \bar{y})^{2}}$$

This is the fractional part of the variation of Y about its mean Y, measured by the variance

$$\sigma_y^2 = \frac{1}{N} \sum_{j=1}^N (\gamma_j - \overline{\gamma})^2$$

that is "explained" by the regression equation. RV is the square of the multiple correlation coefficient, i.e.,

$$RV = R_{Y,X_1,X_2,\ldots,X_k}^2$$

It is clear from the above equations that decreasing the sum of squares of the estimation errors is tantamount to increasing the reduction of variance RV and to decreasing the root mean square error (or standard error of estimate), where

RMSE =
$$\left[\frac{1}{n}\sum_{j=1}^{n}\left(\frac{1}{j}-\hat{y}_{j}\right)^{2}\right]^{\frac{1}{2}}$$

Many times it is not known which or how many predictors to include in a regression equation. Even though the predictand may be correlated with hundreds of variables, a regression equation containing only a very few of them usually explains nearly as much of the variance as an equation containing many. This is due to the high intercorrelations among the variables. Also, if many predictors are included, the predictand may be estimated extremely well in the dependent data sample, but the equation may be showing not only the real physical relationships but also the chance relationships in the dependent data that will not be present in other samples. Therefore, the equation with many terms may perform more poorly on independent data than the one with fewer terms.

A technique for selecting predictors to include in an equation, called screening regression, was used in meteorology as early as 1944 by Bryan [1]. Since being popularized by Miller [18], it has had many applications in meteorology. (For instance, see Klein, Lewis, and Crockett [13] and Pore [21]; other applications are discussed by Glahn [8].) Actually, several variations of this general technique have been used. The one explained below, sometimes called the forward stepwise method, is perhaps the simplest.

The first step in the procedure is to select the variable which correlates most highly (in either a positive or negative sense) with the predictand. This is the variable which explains a greater fraction of the predictand variance than any other of those available. Then, the next variable selected is the one that together with the first increases the reduction of variance the most. Selection can continue in this way until some specified cutoff criterion is met. Usually the cutoff criterion is some function of the additional reduction of variance afforded by the next best predictor. A discussion of the screening technique and the necessary matrix operations is given by Efroymson [2].

Screening regression, as a mathematical technique, can be used no matter what the joint distribution of the predictand and predictors. (However, this distribution is important in the application of significance tests and the interpretation of results.) In fact, any or all of the variables involved can be binary (take on only one of two possible values, 0 and 1).

If a predictand can assume only one of two states, it can still be estimated by regression by giving it the value of zero for the first state and one for the second state. The estimate provided by the regression equation can then be considered as the probability of the second state for the particular combination of predictor values on which the estimate was calculated. (See Mook [19] and Lund [16] for early uses of this particular application.)

If a predictand can assume only one of several, say q, states, it can be transformed into q binary predictands and each treated as discussed above. Miller [17] used this technique for q > 2 when all predictors were binary and called it regression estimation of event probabilities (REEP).

Even though the individual estimates are not bounded by zero and one, their sum over all categories is always unity, provided exactly the same predictors are included in each of the q regression equations. Screening algorithms can also be specified for this application of regression.

The equations for estimating probabilities can include continuous as well as binary variables. It is also worthy of note that minimizing the RMSE is the same as minimizing the P-Score defined by Panofsky and Brier [20] and generally used today in PoP verification.

The assumptions underlying tests of significance are usually far from actuality when the screening procedure is used. Therefore, even though screening stopping rules may be based on some such test, no exact (or perhaps even approximate) level of significance can be attached to them. In practice, experience and subjective judgment are probably as good or better than objective rules in deciding which equation (how many predictors) to use in operation. For instance, if one is interested in explaining the variance of maximum temperature, then using several more predictors to reduce the RMSE from 5.00 degrees to 4.95 degrees is certainly questionable from a practical point of view. Other considerations may play a more important role than the slightly lower RMSE in deciding on the final equation.

DEVELOPMENT OF EQUATIONS

Our work in PoP forecasting to date has been restricted to a generalized operator approach. This amounts to a pooling of the developmental data for all stations and means that the same regression equation is used for all stations for a given period. An application of the generalized operator approach is contained in the work of Russo, Enger, and Merriman [23]. This approach can be used when the various station equations are almost identical or must be used when the dependent data sample is small. We consider the data sample still too small to derive dependable individual station equations. When it becomes possible to develop single station equations, a noticeable improvement in the accuracy of the forecasts for some stations should result.

Also, forecasts for nighttime have been produced by equations developed on daytime data. On October 1, 1969 we will have available both daytime and nighttime regression equations.

The variety of possible predictors used in the screening regression technique includes not only PE and SAM forecasts, but also initial observed data (0700 or 1900 GMT) and climatology. Initial data include weather, clouds, ceiling, visibility, temperature, dew point, sea-level pressure, and wind components at the surface. The initial saturation deficit (the moisture parameter in SAM) was also considered.

The saturation deficit is a term defined by Younkin, La Rue, and Sanders [25] as

$$S_d = h_5 - S_T$$

where h_5 is the 1000-500 mb thickness and S_T the saturation thickness. The saturation thickness, for our purposes, is that thickness between 1000 mb and 500 mb for which precipitation will occur for a given amount of moisture between those levels. The saturation deficit is the moisture parameter used in the SLYH precipitation model [25] that was operational for several years at NMC. The SLYH method was later adapted for on-station use by Kulawiec [14].

Climatology variables consisted of the relative frequency of precipitation during 6- and 12-hour periods for each month [12]. Another variable, the location of the observation point, could be considered to be in this category also.

Variables from the PE model output included forecast fields of the following at 6-hour intervals: mean relative humidity in the column from the surface to between 500 mb and 400 mb; quantitative precipitation; vertical velocities at 1000 mb, 850 mb and 500 mb; temperatures at 1000 mb and 500 mb; and wind components at 500 mb.

Forecasts from SAM were hourly and 3-hourly saturation deficits (see pp 8 and 16 of reference [10] for explanations of these terms), sea-level pressures, 1000 mb geostrophic wind components, and indicators of terrain induced vertical velocities for selected hours during the forecast period.

The screening procedure outlined above correlated the predictors with the observed 6- and 12-hour precipitation (an indicator that precipitation had occurred) during the season of interest (October-March for winter and April-September for summer). If precipitation was observed during the period, the predictand was given the value of one; otherwise it was given the value of zero.

The screening technique usually chooses a number of significant parameters for each regression equation. Those selected were the 3-hourly saturation deficits and sea-level pressure forecasts during the forecast period from SAM, and the 6-hourly mean relative humidity and quantitative precipitation forecasts from the PE model. Invariably, one of the SAM saturation deficit forecasts was chosen first. This variable was available in a continuous form and in a binary form. Interestingly enough, the binary form was always the one selected for any period in any season. This indicates that non-linear relationships exist that can be explained by the binary form that can not be explained by the continuous form. Use of initial data and climatology resulted in very little increase in the multiple correlation coefficient over that for the forecast variables alone.

SEASONAL REGRESSION EQUATIONS

One year of operations is covered by this article, July 1, 1968, to June 30, 1969. Portions of three seasons are involved--summer 1968, winter 1968-69, and summer 1969. Each season has its own set of regression equations. A set consists of three equations, one each for: the first 6-hour period, 1200-1800 GMT (0000-0600 GMT); the second 6-hour period, 1800-2400 GMT (0600-1200 GMT); and the 12-hour period, 1200-2400 GMT (0000-1200 GMT).

Summer 1968

The generalized equations are based on developmental data from 100 stations for 72 days from April through October 1967. A total of 81 predictors were tested. None of these were PE model forecasts. Those chosen by the screening regression program are shown in table 1. The binary predictors enter into the equations in the form of a series of yes-no questions. A "yes" answer results in that term being included as a contribution while a "no" answer results in the term being dropped. The "contribution" due to a particular binary predictor X, being one is the regression coefficient a, and the constant, which is always included as a contribution, is a o

Table 1. Summer 1968 regression equations for PoP forecasting. S_d is 3-hourly saturation deficit in meters and SLP is sea-level pressure in mb.

a. 6-hour PoP, 1200-1800 GMT (0000-0600 GMT)

	Predictor	Contribution to PoP (%)	Cumulative Reduction of Variance
1)	Constant	1.02	
2)	SAM $S_d \leq 30$ at 18Z (06Z)	14.06	.2243
3)	SAM $S_d \leq -7.5$ at 15Z (03Z)	18.62	.2609
4)	SAM $S_d \leq 105$ at 21Z (09Z)	8.95	. 2805
5)	SAM $S_d \le 45$ at 15Z (03Z)	9.74	. 2864
6)	SAM $S_d \leq -15$ at 15Z (03Z)	12.67	.2920
7)	SAM SLP \leq 1010 at 15Z (03Z)	12.32	.3001

(Probability Range is 1% to 77%)

b. 6-hour PoP, 1800-2400 GMT (0600-1200 GMT)

	Predictor	Contribution to PoP (%)	Cumulative Reduction of Variance
1)	Constant	1.01	
2)	SAM $S_d \leq 60$ at 21Z (09Z)	9.34	.1856
3)	SAM $S_d \le -7.5$ at 21Z (09Z)	16.29	.2305
4)	SAM $S_{d} \le -22.5$ at 21Z (09Z)	26.19	.2417
5)	SAM $S_{d} \le 90$ at 24Z (12Z)	10.59	.2528
6)	SAM $S_d \le 30$ at 21Z (09Z)	8.22	.2564
7)	SAM SLP \leq 1015 at 21Z (09Z)	7.63	.2657

(Probability Range is 1% to 79%)

c. 12-hour PoP, 1200-2400 GMT (0000-1200 GMT)

	Predictor	Contribution to PoP (%)	Cumulative Reduction of Variance
1)	Constant	0.40	
2)	SAM $S_d \leq 45$ at $18Z$ (06Z)	20.90	.2877
3)	SAM $S_d \le -7.5$ at 18Z (06Z)	11.44	.3211
4)	SAM $S_d \leq 90$ at 24Z (12Z)	14.05	. 3460
5)	SAM $S_d \leq -15$ at 15Z (03Z)	17.19	.3526
6)	SAM $S_d \leq 15$ at 21Z (09Z)	11.72	.3572
7)	SAM $S_d \leq 210$ at 24Z (12Z)	4.87	.3596
8)	SAM SLP \leq 1010 at 21Z (09Z)	17.75	. 3744

(Probability Range is 0% to 98%)

Winter 1968-69

The generalized equations are based on developmental data for 80 stations for 139 days from October 1967 through March 1968. A total of 99 predictors were examined in the final regression screening. These include various forecasts produced by both SAM and the PE model. Those chosen by the program are shown in table 2. Again, the binary predictors enter into the equations in the form of a series of yes-no questions.

Table 2. Winter 1968-69 regression equations for PoP forecasting. S is 3-hourly saturation deficit in meters, SLP is sea-level pressure, and $^{\rm d}$ PE precipitation amounts are in inches.

a. 6-hour PoP, 1200-1800 GMT (0000-0600 GMT)

	Predictor	Contribution to PoP (%)	Cumulative Reduction of Variance
1)	Constant	34.68	
2)	SAM $S_d \leq 0$ at 15Z (03Z)	15.85	.3388
3)	PE 12-hr precipitation \leq .10 at 24Z (12Z)	-13.71	.3999
4)	SAM $S_d \leq -7$ at 15Z (03Z)	13.54	.4155
5)	SAM $S_d \leq 60$ at $18Z$ (06Z)	6.65	.4254
6)	SAM SLP \leq 1010 at 18Z (06Z)	6.06	.4305
7)	PE 6-hr precipitation \leq .20 at 18Z (06Z)	-12.12	.4343
8)	PE mean relative humidity \leq 70% at 18Z (06Z)	- 3.48	.4377
9)	SAM $S_d \leq -15$ at 15Z (03Z)	8.74	.4390
10)	PE 12-hr precipitation = 0 at 24Z (12Z)	- 5.02	.4402
11)	SAM $S_d \leq 105$ at 15Z (03Z)	3.85	.4414
	(Probability Range is 0% to	80%)	

(Probability Range is 0% to 89%)

Ъ.	6-hour PoP, 1800-2400 GMT (0600-1200 GMT)		
r 1. (a.	Predictor	Contribution to PoP (%)	Cumulative Reduction of Variance
1)	Constant	41.25	
2)	SAM $S_d \le 15$ at 21Z (09Z)	8.83	.2750
3)	PE 12-hr precipitation \leq .10 at 24Z (12Z)	- 5.55	.3312
4)	SAM $S_d \leq -15$ at 21Z (09Z)	18.45	. 3495
5)	PE mean relative humidity \leq 80% at 24Z (12Z)	- 4.13	. 3655
6)	SAM $S_d \le 75$ at 21Z (09Z)	9.30	.3720
7)	PE 12-hr precipitation \leq .40 at 24Z (12Z)	-10.20	.3778
8)	PE mean relative humidity \leq 90% at 24Z (12Z)	- 7.60	.3801
9)	SAM $S_d \leq -7$ at 15Z (03Z)	7.32	.3823
10)	PE mean relative humidity \leq 70% at 24Z (12Z)	- 4.78	. 3835
11)	PE 12-hr precipitation \leq .20 at 24Z (12Z)	- 7.44	. 3845
	(Probability Range is 2% to 85	%)	
C.	12-hour PoP, 1200-2400 GMT (0000-1200 GMT)		
С.	12-hour PoP, 1200-2400 GMT (0000-1200 GMT) Predictor	Contribution to PoP (%)	Cumulative Reduction of Variance
c.		COLOR COLORS COLORS	Reduction
	Predictor	to PoP (%)	Reduction
1)	Predictor	to PoP (%) 42.67	Reduction of Variance
1)	Predictor Constant SAM $S_d \leq 0$ at 18Z (06Z)	to PoP (%) 42.67 10.73	Reduction of Variance
1) 2) 3)	Predictor Constant SAM $S_d \leq 0$ at 18Z (06Z) PE 12-hr precipitation \leq .05 at 24Z (12Z)	to PoP (%) 42.67 10.73 - 7.68	Reduction of Variance .3606 .4334
1) 2) 3) 4)	Predictor Constant SAM $S_d \leq 0$ at 18Z (06Z) PE 12-hr precipitation \leq .05 at 24Z (12Z) SAM $S_d \leq 75$ at 21Z (09Z)	to PoP (%) 42.67 10.73 - 7.68 10.19	Reduction of Variance .3606 .4334 .4529
1) 2) 3) 4) 5)	Predictor Constant SAM $S_d \leq 0$ at 18Z (06Z) PE 12-hr precipitation \leq .05 at 24Z (12Z) SAM $S_d \leq 75$ at 21Z (09Z) SAM $S_d \leq -5$ at 15Z (03Z)	to PoP (%) 42.67 10.73 - 7.68 10.19 12.68	Reduction of Variance .3606 .4334 .4529 .4669
1) 2) 3) 4) 5) 6)	Predictor Constant SAM $S_d \le 0$ at $18Z$ (06Z) PE 12-hr precipitation \le .05 at 24Z (12Z) SAM $S_d \le 75$ at 21Z (09Z) SAM $S_d \le -5$ at 15Z (03Z) PE mean relative humidity $\le 70\%$ at $18Z$ (06Z)	to PoP (%) 42.67 10.73 - 7.68 10.19 12.68 - 6.33	Reduction of Variance .3606 .4334 .4529 .4669 .4761 .4817
1) 2) 3) 4) 5) 6) 7)	Predictor Constant SAM S $_{\rm d} \le 0$ at 18Z (06Z) PE 12-hr precipitation \le .05 at 24Z (12Z) SAM S $_{\rm d} \le 75$ at 21Z (09Z) SAM S $_{\rm d} \le -5$ at 15Z (03Z) PE mean relative humidity $\le 70\%$ at 18Z (06Z) PE 12-hr precipitation \le .20 at 24Z (12Z)	to PoP (%) 42.67 10.73 - 7.68 10.19 12.68 - 6.33 -12.97	Reduction of Variance .3606 .4334 .4529 .4669 .4761 .4817 .4863
1) 2) 3) 4) 5) 6) 7) 8)	Predictor Constant SAM $S_d \leq 0$ at 18Z (06Z) PE 12-hr precipitation \leq .05 at 24Z (12Z) SAM $S_d \leq 75$ at 21Z (09Z) SAM $S_d \leq -5$ at 15Z (03Z) PE mean relative humidity $\leq 70\%$ at 18Z (06Z) PE 12-hr precipitation \leq .20 at 24Z (12Z) SAM SLP \leq 1015 at 18Z (06Z)	to PoP (%) 42.67 10.73 - 7.68 10.19 12.68 - 6.33 -12.97 6.44 8.63	Reduction of Variance .3606 .4334 .4529 .4669 .4761 .4817 .4863 .4889
1) 2) 3) 4) 5) 6) 7) 8) 9)	Predictor Constant SAM $S_d \leq 0$ at 18Z (06Z) PE 12-hr precipitation \leq .05 at 24Z (12Z) SAM $S_d \leq 75$ at 21Z (09Z) SAM $S_d \leq -5$ at 15Z (03Z) PE mean relative humidity $\leq 70\%$ at 18Z (06Z) PE 12-hr precipitation \leq .20 at 24Z (12Z) SAM $S_d \leq 45$ at 15Z (03Z)	to PoP (%) 42.67 10.73 - 7.68 10.19 12.68 - 6.33 -12.97 6.44 8.63	Reduction of Variance .3606 .4334 .4529 .4669 .4761 .4817 .4863 .4889 .4912
1) 2) 3) 4) 5) 6) 7) 8) 9) 10) 11)	Predictor Constant SAM $S_d \leq 0$ at $18Z$ (06Z) PE 12-hr precipitation \leq .05 at 24Z (12Z) SAM $S_d \leq 75$ at 21Z (09Z) SAM $S_d \leq -5$ at 15Z (03Z) PE mean relative humidity $\leq 70\%$ at $18Z$ (06Z) PE 12-hr precipitation \leq .20 at 24Z (12Z) SAM $S_d \leq 45$ at $15Z$ (03Z) PE mean relative humidity $\leq 90\%$ at 24Z (12Z) PE mean relative humidity $\leq 90\%$ at 24Z (12Z)	to PoP (%) 42.67 10.73 - 7.68 10.19 12.68 - 6.33 -12.97 6.44 8.63 - 7.36	Reduction of Variance .3606 .4334 .4529 .4669 .4761 .4817 .4863 .4889 .4912

(Probability Range is 2% to 100%)

Summer 1969

The generalized equations are based on developmental data for 99 stations for 129 days from July through September 1967 and April through September 1968. A total of 99 predictors were examined in the final regression screening. These include various forecasts produced by both SAM and the PE model. Those chosen by the program are shown in table 3. As in tables 1 and 2, the binary predictors enter into the equations in the form of a series of yes-no questions.

Table 3. Summer 1969 regression equations for PoP forecasting. S_d is 3-hourly saturation deficit in meters, SLP is sea-level pressure, $\overset{ ext{d}}{\text{and}}$ PE precipitation amounts are in inches.

a.	6-hour	PoP,	1200-1800	GMT	(0000-0600	GMT)
----	--------	------	-----------	------------	------------	------

a.	6-hour PoP, 1200-1800 GMT (0000-0600 GMT)		
	Predictor	Contribution to PoP (%)	Cumulative Reduction of Variance
1)	Constant	24.87	
2)	SAM $S_d \le 30$ at 18Z (06Z)	7.89	.2173
3)	PE 6-hr precipitation \leq .05 at 18Z (06Z)	- 8.64	. 2537
4)	SAM $S_d \leq -7$ at 15Z (03Z)	15.80	. 2751
5)	PE mean relative humidity \leq 75% at 18Z (06Z)	- 6.99	.2886
6)	PE mean relative humidity \leq 95% at 24Z (12Z)	- 8.10	.2927
7)	SAM $S_d \leq 45$ at 21Z (09Z)	8.39	.2959
8)	SAM $S_d \leq 15$ at 15Z (03Z)	9.54	.2990
9)	SAM $S_d \leq -15$ at 15Z (03Z)	12.49	.3012
10)	SAM SLP \leq 1010 at 18Z (06Z)	4.84	.3027
	(Probability Range is 1% to	84%)	
b.	6-hour PoP, 1800-2400 GMT (0600-1200 GMT)		C
	Predictor	Contribution to PoP (%)	Cumulative Reduction of Variance
1)	Constant	26.73	
2)	SAM $S_d \le 60$ at 24Z (12Z)	11.51	. 1449
3)	PE mean relative humidity \leq 85% at 24Z (12Z)	-10.54	.1855
4)	SAM $S_d \leq -15$ at 24Z (12Z)	22.12	.1993
5)	SAM $S_d \le 120$ at 24Z (12Z)	6.52	.2081
6)	SAM SLP \leq 1015 at 18Z (06Z)	5.20	.2135
7)	PE 12-hr precipitation \leq .20 at 24Z (12Z)	- 9.03	.2169
8)	PE mean relative humidity \leq 70 at 24Z (12Z)	- 6.49	.2203

(Probability Range is 1% to 72%)

c. 12-hour PoP, 1200-2400 GMT (0000-1200 GMT)

	Predictor	Contribution to PoP (%)	Cumulative Reduction of Variance
1)	Constant	31.09	
2)	SAM $S_d \leq 45$ at 21Z (09Z)	10.76	.2233
3)	PE mean relative humidity \leq 75% at 24Z (12Z)	- 5.02	.2780
4)	PE 6-hr precipitation \leq .05 at 18Z (06Z)	- 7.66	. 2954
5)	SAM SLP \leq 1015 at 18Z (06Z)	4.81	.3051
6)	SAM $S_d \leq -5$ at 18Z (06Z)	12.55	.3133
7)	SAM $S_d \le 105$ at 18Z (06Z)	7.38	.3203
8)	PE mean relative humidity \leq 90% at 24Z (12Z)	-10.60	.3252
9)	PE mean relative humidity \leq 70% at 24Z (12Z)	- 8.11	.3271
10)	SAM $S_d \le 45$ at 18Z (06Z)	9.09	.3290
11)	SAM SLP \leq 1010 at 18Z (06Z)	6.87	.3305
12)	SAM $S_d \le -15$ at 24Z (12Z)	10.43	.3319
	SAM SLP \leq 1020 at 18Z (06Z)	3.80	.3332
	(P1-1-1-1 P	0 7713	

(Probability Range is 0% to 97%)

The multiple correlation coefficient with each added predictor is given by the square root of the corresponding cumulative reduction of variance.

Since a separate equation is used for each of the three forecast periods, the 12-hour PoP may not be consistent with the two 6-hour PoPs covering the same 12-hour period. When these inconsistencies occur in the set of probabilities forecast at a grid point, the following adjustments are made:

- a. If either or both of the 6-hour PoPs exceed the 12-hour PoP, either or both are reduced to the value of the 12-hour PoP.
- b. If the 12-hour PoP is larger than the sum of the two 6-hour PoPs, the 12-hour PoP is reduced by one-half the difference and each of the 6-hour PoPs is increased by one-quarter of the difference.

The ranges shown in tables 1, 2, and 3 are unadjusted. Therefore, it is possible for the final forecast to fall outside the range shown.

When only SAM predictors were available in summer 1968 (see table 1), the saturation deficits were far more important than the sea-level pressures. In the later months, when the various forecast fields from the PE model were also available, several of these fields were chosen but none as a first predictor. It was found that for the winter 1968-69 equations, the SAM predictors alone gave about 3 percent greater reduction of variance than the PE predictors alone. Also, it was found that by allowing both SAM and PE predictors to be available, about 3 percent greater reduction of variance was shown than for the

SAM alone. These figures were obtained on dependent data but quite likely independent data would indicate about the same thing. Thus the decision was made to include the PE forecast information along with the SAM predictions.

Both the mean relative humidities and the precipitation amounts from the PE model were chosen frequently. It is interesting that several PE vertical velocity fields were screened but not once was any of them chosen. This indicates that the vertical velocity predictions, in their present form, offer no help in forecasting PoP after the other fields are known.

In general, the screening regression program had available as possible predictors fields valid before, during, and after the predictand verification period. It is encouraging that most of the fields chosen were forecasts for the period of interest rather than before or after the period. This indicates certain favorable features concerning the timing of forecasts produced by the numerical models.

Seasonal variations are apparent from tables 1, 2, and 3. The reduction of variance figures are higher (better) in winter than in summer. This is to be expected because of the convective nature of the summer precipitation.

OPERATIONAL TRANSMISSIONS

At the present time, two scheduled SAM runs are made each day, one to cover the "today" period and one to cover the "tonight" period. Information from the most recent PE operational run is available on a history tape during each SAM run.

Two methods of transmitting the SAM/PE PoP forecasts from NMC to the users are being used. These are the Service "C" teletypewriter service and the FOFAX facsimile circuit. Service "C" has a fairly large user coverage while FOFAX is rather limited.

A portion of a typical "today" period teletypewriter bulletin is shown in figure 2. Transmission times for the 79-station SAM bulletin are 0839 and 2039 GMT daily. A station list with call letters is given by Glahn, Lowry, and Hollenbaugh [10]. PoP forecasts contained in the bulletin (last three groups for each station) include a 12-hour prediction (three digits), the first 6-hour prediction (two digits), and the last 6-hour prediction (two digits). The values shown are the adjusted regression estimates rounded to the nearest percent. As an example (see fig. 2), the PoP forecast for Caribou, Maine, is 21 percent for the 12-hour "today" period. The PoP prediction for the first 6 hours (1200-1800 GMT) of the same period is 5 percent and the forecast is 18 percent for the second 6 hours (1800-2400 GMT).

The bulletins also contain predictions of saturation deficit and 1000-mb geostrophic winds. These are direct output from SAM and are explained in detail in another Technical Memorandum [10].

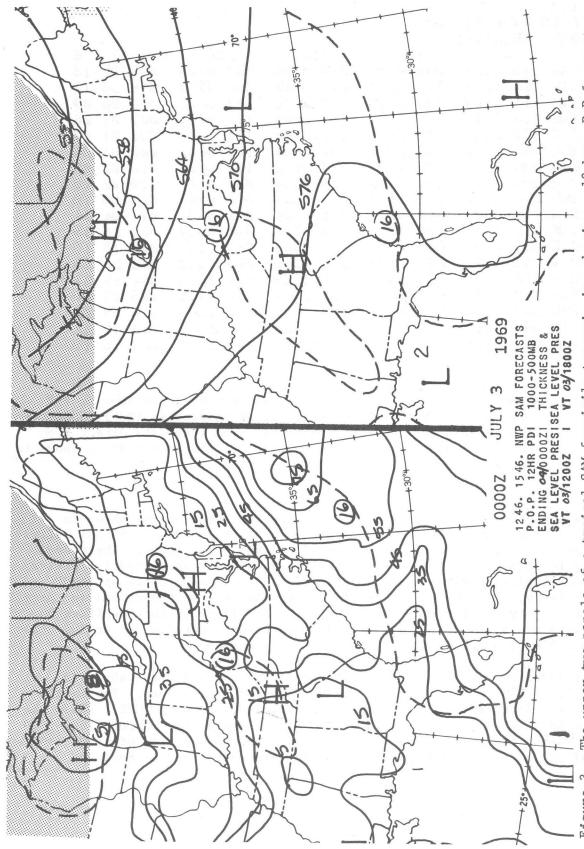
ZCZC F OUS KWBC Ø3Ø8ØØ SAM F ORECASIS

DUIT	1 011	TOUDI D										
	1:	57	18	BZ	2	1Z	Ø	ØZ	P OP	12 6	5 6	
CAR	049	0204	058	0506	Ø68	0508	080	0509	Ø21	05	18	
BIV	26 Ø	3105	Ø5 7	3604	Ø59	04 05	066	0507	018	08	13	
PWM	113	3207	116	3205	117	35 Ø3	117	03 04	010	02	09	
BOS	116	0106	124	0204	127	0303	130	05 03	007	Ø1	06	
PVD	102	9496	129	0406	116	04 05	121	0504	008	ØI	07	
BDL	105	0304	107	0203	107	0203	107	04 03	พีพี5	01	104	
L GA	0 84	0512	Ø 96	0409	108	04 06	115	04 05	010	02	08	
AL B		3003		3303	Ø85	0204	Ø82	0406	010	02	Ø8	
BGM	064	2705	050	3103	04 1	0203	043	05 05	021	09	18	
SYR	056	2703	046	3402	046	03 04	055	0505	017	04	14	
BUF	010	3002	014	0302	Ø27	0704	044	0807	031	26	19	
CLE	031	1806	019	1706	015	1306	Ø17	1108	030	15	19	
DTW	009	1408	017	1109	028	1010	04 1	1010	032	26	20	
BIL	013	10 10	027	0910	043	1009	Ø5 7	1208	034	25	23	
FWA	010	2123	014	1002	023	Ø 9 Ø 5	031	1106	Ø3 6	27	24	
IND	030	2507	025	24 04	020	1903	-02	1606	036	17	24	
CV G	048	2511	033	2411	020	2409	-02	2207	233	15	22	
SDF	Ø63	2411	045	2312	-00	2212	-06	2111	Ø27	10	19	

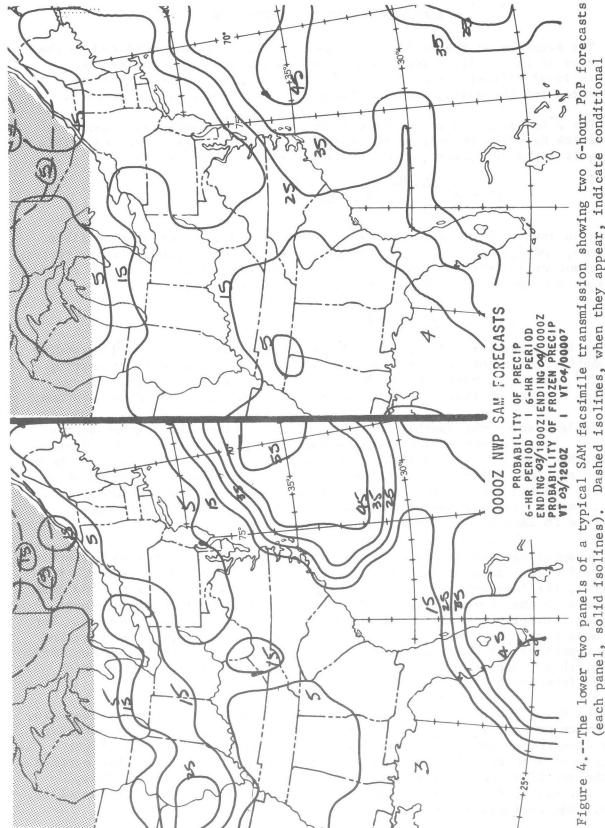
Figure 2.--A portion of a typical SAM teletypewriter bulletin showing PoP forecasts as derived from Model Output Statistics (MOS). Saturation deficit and 1000-mb geostrophic wind forecasts direct from SAM are also shown. This example is for the same case as shown in figures 3 and 4.

The upper two panels of the SAM facsimile transmission are shown in figure 3. These transmissions are sent on FOFAX at 0849 and 2045 GMT daily. Only the solid isolines on the left panel are PoP predictions. These are 12-hour forecasts and should agree with those values transmitted for cities on the SAM teletypewriter bulletin (see fig. 2). In case of an apparent inconsistency, it is better to accept the teletypewriter report unless there has been an obvious transmission error. The specific isolines depicted on PoP charts are 5, 15, 25,, 75, 85, 95 percent. Details concerning other forecast fields shown in figure 3 are given elsewhere [10].

The lower two panels of the SAM facsimile transmission are shown in figure 4. Again, only the solid isolines are PoP predictions. The left panel shows forecasts for the first 6-hour period (1200-1800 GMT), while the right panel includes the predictions for the second 6-hour period (1800-2400 GMT). Dashed isolines, when they appear, represent forecasts of conditional probability of frozen precipitation [6] and are not discussed here.



3.--The upper two panels of a typical SAM facsimile transmission showing a 12-hour PoP forecast on the left panel (solid isolines). Other forecasts direct from SAM are sea-level pressure and 1000-500-mb thickness. This example, from July 3, 1969, is referred to as a 0000 GMT run because that was the time of the initial upper-air input data. Figure



This example is for the same case as shown in figures (each panel, solid isolines). Dashed isolines, when they appear, indicate conditional probability of frozen precipitation. 2 and 3.

ACCURACY OF THE POP FORECASTS

The accuracy of the machine produced PoP forecasts was compared to the accuracy of the official forecasts issued to the public by the Weather Bureau local offices. Verification figures covering a 1-year span are presented for the eastern United States for the 12-hour "today" period.

The Brier score is in common use throughout the Weather Bureau as the yardstick by which to measure the goodness of probability forecasts. This score is $\frac{1}{2}$ the "P" score defined by Panofsky and Brier [20] as

$$P = \frac{1}{N} \sum_{j=1}^{r} \sum_{i=1}^{N} (F_{ij} - E_{ij})^{2}$$

where on each of N occasions an event can occur in only one of r possible classes, and f_{i1} , f_{i2} ,, f_{ir} represent the forecast probabilities that the event will occur in classes 1, 2,, r, respectively. If the r classes are chosen to be mutually exclusive and exhaustive,

$$\sum_{j=1}^{r} f_{ij} = 1$$

for each and every $i=1, 2, \ldots, N$. E takes the value 1 or 0 according to whether the event occurred in class j or not. For perfect forecasting the "P" score will have a value of zero and for the worst possible forecasting a value of two. The Brier score, then, has a range of zero to one.

Table 4 shows the monthly breakdown of Brier scores by Weather Bureau Region. Five Eastern Region stations were used (Washington, D. C., Boston, Mass., New York, N. Y., Raleigh, N. C., and Cleveland, Ohio); four Southern Region stations (Miami, Fla., Atlanta, Ga., New Orleans, La., and Memphis, Tenn.); and two Central Region stations (St. Louis, Mo., and Chicago, Ill.). The 12-month averages showed the SAM/PE forecasts to be slightly more accurate in each of the three regions than the local forecasts.

Seasonal variations are apparent from table 4 but can be seen better in figure 5 where the monthly scores for all regions combined are shown. In general, the forecasts are shown to be more accurate in the winter and less accurate in the summer. This is due to the relative difficulty of forecasting convective precipitation which occurs predominately in summer. For the most part, the local and the SAM/PE scores are in phase (i.e. rise and fall together). Exceptions were the months of December 1968 and March 1969. Again, this in-phase relationship is tied to the difficulty factor which varies from month to month.

The Brier score alone does not show the reliability of the forecasts in different probability ranges. Table 5 shows the number of SAM/PE forecasts in a given range and the number of times precipitation occurred. Eight additional stations were added to the previously listed 11 for this portion of the verification. The cities are Norfolk, Va., Jacksonville, Fla., Jackson, Miss., Albany, N. Y., Charleston, S. C., Buffalo, N. Y., Nashville, Tenn., and Louisville, Ky.

Objective forecasts for July, August, and September used only SAM predictors. Later forecasts were based on SAM and PE predictors. Five Eastern Region, four Southern Region, and two Central Region stations were used. The total number of forecasts was 3682. Brier scores for one year of PoP forecasts for the 12-hour "today" period (1200-2400 GMT). Table 4.

1968 July 1.258 .1248 1.05AL SAN/PE 1.06AL SAN/PE 1.06AL SAN/PE 1.06AL SAN/PE 1.058 1.1258 1.1248 1.157 1.1429 1.1749 1.1942 1.1942 1.1953 1.155 1.1258 1.1248 1.1050 1.1077 1.1574 1.1608 1.1323 1.125 1.1258 1.1054 1.1056 1.1144 1.1068 1.1140 1.0997 1.101 1.1056 1.1054 1.1064 1.1064 1.1069 1.1064		THE THE THE PERSON OF THE	EASTERN REGION	REGION	CENTRAL REGION	REGION	SOUTHERN	REGION	ALL R	ALL REGIONS
July .1258 .1248 .1157 .1429 .1749 .1942 .1419 August .1216 .1073 .1090 .1007 .1574 .1608 .1323 September .0817 .0878 .1026 .1114 .1208 .1140 .0997 November .0831 .0784 .0487 .0615 .1062 .0890 .0852 November .0912 .1004 .1279 .0910 .1084 .0793 .0852 January .0963 .0776 .1243 .0916 .1055 .1016 .1047 March .0770 .0939 .0574 .0673 .0731 .0779 .0968 April .0941 .0811 .1419 .1263 .0903 .0915 .0564 .0673 .0915 .0928 June .1428 .1346 .1833 .1499 .1263 .1405 .1186 .1186 .1186 .1053 .1058 .1186 .1186 <			LOCAL	SAM/PE	LOCAL	SAM/PE	LOCAL	SAM/PE	LOCAL	SAM/PE
August .1216 .1073 .1090 .1007 .1574 .1608 .1323 September .0817 .0878 .1026 .1114 .1208 .1140 .0997 November .0912 .1004 .1279 .0910 .1084 .0852 .0852 November .1101 .1221 .0945 .1099 .0679 .0853 .1041 January .0963 .0776 .1243 .0916 .1055 .1016 .1047 February .0963 .0776 .0878 .0694 .0921 .0685 .0867 April .0937 .0687 .0673 .0731 .0779 .0928 May .0941 .0811 .1419 .1279 .1263 .0915 .0928 Average .1000 .0973 .1048 .0989 .1124 .1145 .1145	1968	July	.1258	. 1248	,1157	. 1429	.1749	.1942	.1419	.1534
September .0817 .0878 .1026 .1114 .1208 .1140 .0997 October .0831 .0784 .0487 .0615 .1062 .0890 .0852 November .0912 .1004 .1279 .0910 .1084 .0793 .1041 December .1101 .1221 .0945 .1099 .0679 .0553 .1041 January .0963 .0776 .1243 .0916 .1055 .1016 .1047 March .0770 .0939 .0504 .0673 .0731 .0779 .0777 April .0941 .0811 .1419 .1279 .1263 .0915 .0779 June .1428 .1346 .1583 .1499 .1256 .1405 .1394 Average .1000 .0973 .1048 .0989 .1124 .1056 .1405 .1124 .1125 .1405 .1125 .1405 .1125 .1125 .1124 .1125		August	.1216	.1073	.1090	.1007	.1574	.1608	.1323	.1256
October .0831 .0784 .0487 .0615 .1062 .0890 .0852 November .0912 .1004 .1279 .0910 .1084 .0793 .1041 December .1101 .1221 .0945 .1099 .0679 .0553 .1041 January .0963 .0776 .1243 .0916 .1055 .1016 .1047 March .0770 .0939 .0504 .0673 .0731 .0779 .0867 April .0937 .0687 .0659 .0632 .0903 .0915 .0928 April .0941 .1419 .1279 .1263 .0915 .0928 June .1428 .1583 .1499 .1256 .1405 .1394 Average .1000 .0973 .0978 .0989 .1124 .1056 .1053		September	.0817	.0878	.1026	.1114	.1208	.1140	. 0997	.1016
November .0912 .1004 .1279 .0910 .1084 .0793 .1041 December .1101 .1221 .0945 .1099 .0679 .0553 .0919 January .0963 .0776 .1243 .0916 .1055 .1016 .1047 February .0820 .0913 .0878 .0694 .0921 .0685 .0867 April .0937 .0687 .0673 .0731 .0779 .0928 April .0941 .0811 .1419 .1279 .1263 .0915 .0928 June .1428 .1346 .1583 .1499 .1256 .1405 .1394 Average .1000 .0973 .1048 .0989 .1124 .1056 .1053		October	.0831		. 0487	.0615	.1062	0680°	.0852	.0792
December .1101 .1221 .0945 .1099 .0679 .0553 .0919 January .0963 .0776 .1243 .0916 .1055 .1016 .1047 March .0870 .0878 .0694 .0921 .0685 .0867 April .0770 .0939 .0504 .0673 .0731 .0779 .0707 May .0941 .0811 .1419 .1279 .1263 .0941 .1145 June .1428 .1346 .1583 .1499 .1256 .1405 .1394 Average .1000 .0973 .1048 .0989 .1124 .1056 .1053		November	.0912	,1004	.1279	.0910	.1084	.0793	.1041	.0910
January .0963 .0776 .1243 .0916 .1055 .1016 .1047 February .0820 .0913 .0878 .0694 .0921 .0685 .0867 March .0770 .0939 .0504 .0673 .0731 .0779 .0707 April .0937 .0687 .0959 .0632 .0903 .0915 .0928 June .1428 .1346 .1583 .1499 .1256 .1405 .1145 Average .1000 .0973 .1048 .0989 .1124 .1056 .1053		December	.1101	.1221	.0945	. 1099	6290°	.0553	.0919	9960°
1ry .0820 .0878 .0694 .0921 .0685 .0867 .0770 .0939 .0504 .0673 .0779 .0707 .0937 .0687 .0632 .0903 .0915 .0928 .0941 .0811 .1419 .1279 .1263 .0941 .1145 .1428 .1346 .1583 .1499 .1256 .1405 .1394 .1000 .0973 .1048 .0989 .1124 .1056 .1053	1969	January	.0963	.0776	.1243	,0916	.1055	.1016	.1047	. 0889
.0770 .0939 .0504 .0673 .0779 .0779 .0707 .0937 .0687 .0959 .0632 .0903 .0915 .0928 .0941 .0811 .1419 .1279 .1263 .0941 .1145 .1428 .1346 .1583 .1499 .1256 .1405 .1394 .1000 .0973 .1048 .0989 .1124 .1056 .1053		February	.0820	.0913	. 0878	,0694	.0921	. 0685	.0867	0620.
.0937 .0687 .0632 .0903 .0915 .0928 .0941 .0811 .1419 .1279 .1263 .0941 .1145 .1428 .1346 .1583 .1499 .1256 .1405 .1394 .1000 .0973 .1048 .0989 .1124 .1056 .1053		March	.0770	6860°	,0504	.0673	.0731	. 0779	.0707	.0832
e .1419 .1279 .1263 .0941 .1145 .1145 .1263 .0941 .1145 .1394 .1000 .0973 .1048 .0989 .1124 .1056 .1053		April	. 0937	. 0687	6960°	.0632	. 0903	. 0915	. 0928	0920.
.1428 .1346 .1583 .1499 .1256 .1405 .1394 .1000 .0973 .1048 .0989 .1124 .1056 .1053		May	.0941	. 0811	.1419	,1279	.1263	.0941	.1145	7760.
.1000 .0973 .1048 .0989 .1124 .1056 .1053		June	. 1428	.1346	.1583	. 1499	.1256	.1405	.1394	.1395
		Average	.1000	. 0973	.1048	6860°	.1124	.1056	.1053	,1006

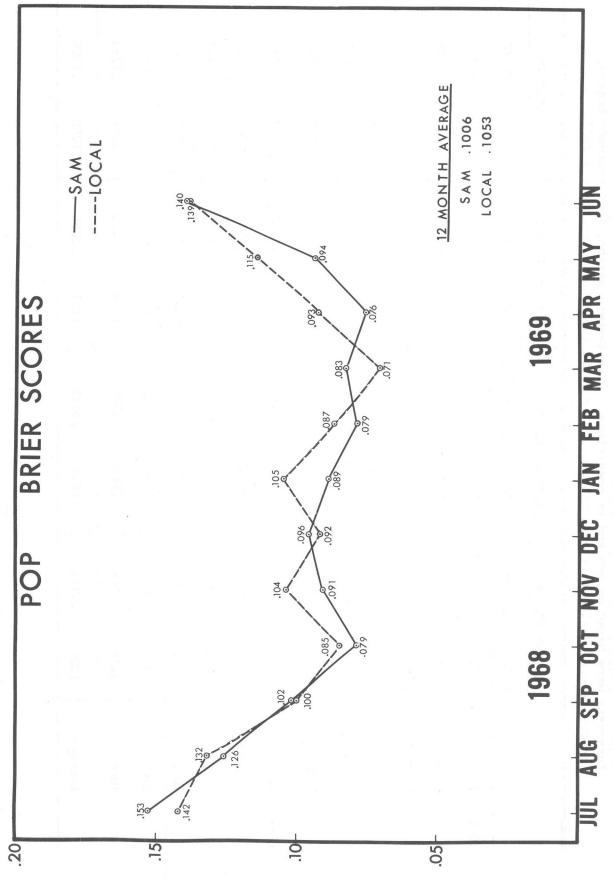


Figure 5.--Comparative verification of SAM/PE and local forecasts for 12-hour "today" period. number of forecasts was 3682, representing 11 stations.

Distribution of SAM/PE PoP forecasts covering the 12-hour "today" period (1200-2400 GMT). "A" is the number of forecasts in a given range; "B" is the number of times precipitation occurred. The total number of forecasts was 6361 from 19 stations for a 1-year period (July 1968-June 1969). Table 5.

00.7														19
00.1 Z 4 Z 06.	B		1	2	17	36	41	29	32	31	15	98 01	7	205
	A	0	0	n	18	37	94	33	34	32	16	0	n	222
e8. Z q Z 08.	B	1	I	4	4	4	10	16	12		22	4	m	98
	A	0	0	4	4	2	11	20	17	6	26	4	7	105
er. z 4 z 0r.	B	-	7	2	10	19	16	13	16		9	9	6	
02.	A	2	2	9						3 11	-	3 16		134
69. =					13	26	22	20	19	13	18	18	12	171
69. Z 4 Z 09.	B	2	-	9	11	9	4	10	10	4	14	15	14	97
	A	3	9	7	14	6	6	16	15	00	15	18	24	141
ez. z q z oz.	B	18	6	22	14	16	17	9	6	00	16	6	13	.57
	A	27	26	32	16	25	29	24	20	15	19	16	29	278 1
64. Z q Z 04.	B	29	26	15	00	6	7	7	0	6	9	12	18	149
	A	57	61	28	14	19	18	24	20	21	16	21	44	343 1
ee. z 4 z 0e.	B	2	0	0	9	12	00	00	7	13	7	H	13	87
	A	3	3	1	16	26	20	21	21	25	19	26	39	220
es. z q z os.	В	0	0	2	00	00	12	2	6	11	0	21	20	105
	A	4	2	9	37	27	25	28	34	26	28	99	69	342 1
61. Z 4 Z 01.	B	55	32	19	15	00	00	_	4	00	11	12	18	197
60:-	A	183	171	91	64	40	36	39	32	64	57	98	131	796
60. Z 4 Z 20.	B	29	16	00	7	6	5	2	7	6	n	2	00	106
	А	251	177	220	80	92	47	54	65	83	89	09	111	1300
40. Z 4 Z 00.	В	0	0	0	9	7	13	7	2	7	7	7	2	54 1
	A	38	30	28	263	242	269	234	255	288	230	265	103	2275
	MONTH	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June	TOTAL 2

A close inspection of table 5 reveals that some ranges were rarely forecast in a certain season. For example, during the summer of 1968 there were relatively few forecasts in the .00 to .04, the .20 to .39, and the .60 to 1.00 ranges. Correspondingly, there were relatively many forecasts in the .05 to .19 and .40 to .59 ranges. This is probably due to the binary nature of the predictors and the inclusion of a fairly small number of them in the equation. Other equations do not exhibit this same pattern of favoring certain forecast ranges. Table 5 shows clearly the large number of forecasts being made in the very low ranges.

The values from table 5 were used to compute the observed relative frequencies shown in table 6. These figures give an indication as to the reliability of the forecasts. Good reliability is indicated if the figure being examined falls within the range shown at the top of the table. Considerable variation exists from month to month and from range to range. However, the totals show rather good reliability.

Table 6. Observed relative frequency (RF) in percent for given ranges of the SAM/PE PoP forecasts. Values shown are computed using the figures

from table 5 where RF = B/A. .39 .09 . 59 .04 29 69 89 7 VI 2 R 2 D 4 R D 2 2 R VI 05 00 10 40 80 MONTH RF .51 July .00 .12 .30 .00 .67 .67 .67 ,50 .00 .09 .00 August .19 .00 .43 .35 .33 1.00 September .00 .04 .21 .33 .00 .54 .69 .86 .83 1.00 .67 October .94 .02 .08 .31 .22 .38 .79 .57 .88 ,77 1.00 November .03 .12 .20 .30 .46 .47 .97 . 64 .67 .73 .80 .39 .48 .73 December .05 .11 .22 .40 . 59 .44 .91 .89 .88 January .01 . 04 .18 .18 . 38 .29 .25 .65 .80 .63 February .02 .08 .13 .26 .33 .15 .45 .84 .70 .94 .67 March .02 .11 .16 .42 .52 .43 .53 .50 .85 .78 .97 April .04 .19 .32 .37 .38 .84 . 94 .01 .93 .89 .85 .42 . 14 May .03 .08 .38 .57 .56 .83 .89 1.00 .29 . 14 June .05 .07 .33 .41 .45 .58 .75 .60 .67 .78 TOTAL .02 .08 .20 .31 .40 .43 .56 .69 .82 .92

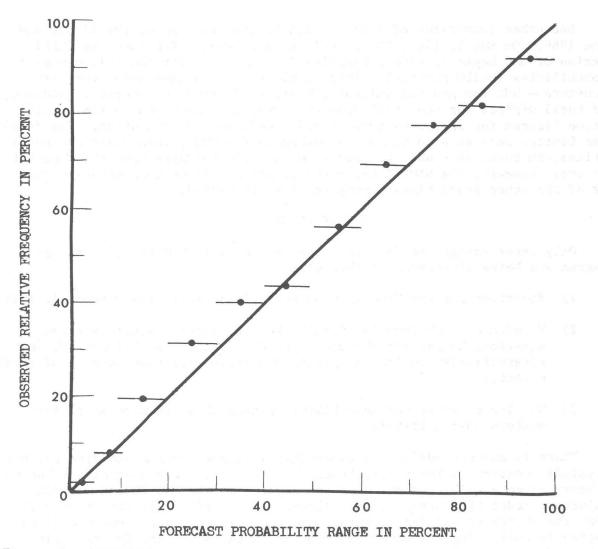


Figure 6.--Reliability of SAM/PE forecasts of PoP for 12-hour "today" period. The number of forecasts was 6361, representing 19 stations over a 1-year span. Each horizontal line represents the forecast interval over which the corresponding observed relative frequency was calculated.

Figure 6 gives a graphical presentation of the totals. The diagonal represents perfect reliability. If the short horizontal lines (which, together, represent the full forecast probability range) intersect or touch the diagonal, this indicates favorable reliability. If the forecasts within a given range are assumed to be centered about the mid-point (shown as dots on fig. 6) and if the ranges are weighted according to the number of forecasts they contain, the forecasts were, on the average, 1.6 percent too high. This is considered to be insignificant, especially in view of the fact that Weather Bureau offices normally round probability forecasts to the nearest 10 percent. SAM/PE probability forecasts are made to the nearest percent and there is no indication that this is hurting the forecasts. The slight overforecasting may not hold true for future seasons since new equations will be derived from different samples.

One other comparison of interest can be made for the months of May and June 1969. On May 1, 1969, the Quantitative Precipitation Forecast (QPF) Section of NMC began issuing subjective PoP forecasts for the full range of probabilities (0-100 percent). This completed the three-echelon forecast structure — NMC for general guidance, Forecast Centers for specific guidance, and Local Offices for the final product. Over the two month period, verification figures for this structure showed a desirable distribution. The Forecast Centers were able to improve somewhat on the NMC guidance and the Local Offices, in turn, were able to improve upon their guidance from the Forecast Centers. However, the SAM/PE forecasts turned out to be more accurate than any of the other predictions during the 2-month period.

FUTURE PLANS

Only three changes and/or additions in the basic SAM/PE PoP forecast program are being considered at this time.

- 1) Equations for the "tonight" period will be available October 1, 1969.
- 2) We plan to look into the feasibility of deriving single station equations to replace the generalized operators now being used, or, alternatively, to include predictors which will take account of local effects.
- 3) We plan to study the feasibility of extending the program to the western United States.

There is another addition planned that does not involve SAM directly but certainly pertains to PoP forecasting. We will extend PoP predictions out to 48 hours or more using the MOS approach applied to the NMC PE and the TDL Trajectory Model [22] outputs. The cities to be included in the study will cover the 50 states and Puerto Rico. The first year of the program will be devoted to collecting and error checking the data sample and in developing a number of computer programs necessary for processing it.

CONCLUSIONS

An operational program for objectively forecasting probability of precipitation with computers has been discussed and examined in some detail. The comparative verification figures for July 1968 - June 1969 indicate that the machine produced forecasts were slightly more accurate than the official forecasts issued to the public by Weather Bureau local offices.

ACKNOWLEDGMENTS

We wish to sincerely thank those other members of TDL who helped the SAM/PE PoP program in various ways, especially Mr. George Hollenbaugh, Mrs. Evelyn Boston, Mrs. Jackie Hughes, and Mrs. Anna Booth. We acknowledge the cooperation of NMC in making available PE history tapes and in implementation of the SAM products. We also wish to thank the Weather Bureau's Weather Analysis and Prediction Division (WXAP), the Regions, and the ESSA Computer Division for help in the various aspects of implementation and verification.

REFERENCES

- 1. Bryan, J. G., "Special Techniques in Multiple Regression," Unpublished Manuscript, 1944.
- 2. Efroymson, M. A., "Multiple Regression Analysis," Mathematical Methods for Digital Computers (Edited by A. Ralston and H. S. Wilf), John Wiley and Sons, Inc., New York, 1960, pp. 191-203.
- 3. ESSA Weather Bureau, "Implementation Test of a Numerical Subsynoptic Precipitation Prediction Model," <u>Technical Procedures Bulletin No. 6</u>, August 1967, 6 pp.
- 4. ESSA Weather Bureau, "Operational Forecasts with the Sub-synoptic Advection Model (SAM)," <u>Technical Procedures Bulletin No. 14</u>, June 1968, 19 pp.
- 5. ESSA Weather Bureau, "Operational Forecasts with the Sub-synoptic Advection Model (SAM)-No. 2," <u>Technical Procedures Bulletin No. 18</u>, September 1968, 4 pp.
- 6. ESSA Weather Bureau, "Operational Forecasts with the Sub-synoptic Advection Model (SAM)-No. 3," <u>Technical Procedures Bulletin No. 21</u>, February 1969, 12 pp.
- 7. ESSA Weather Bureau, "Operational Forecasts with the Sub-synoptic Advection Model (SAM)-No. 4," <u>Technical Procedures Bulletin No. 23</u>, April 1969, 5 pp.
- 8. Glahn, H. R., "Objective Weather Forecasting by Statistical Methods," The Statistician, Vol. 15, No. 2, 1965, pp. 111-142.
- 9. Glahn, H. R., G. W. Hollenbaugh, and D. A. Lowry, "An Operationally Oriented Objective Analysis Program," <u>ESSA Technical Memorandum WBTM TDL-22</u>, July 1969, 20 pp.
- 10. Glahn, H. R., D. A. Lowry, and G. W. Hollenbaugh, "An Operational Subsynoptic Advection Model," <u>ESSA Technical Memorandum WBTM TDL-23</u>, July 1969, 26 pp.
- 11. Hollenbaugh, G. W., H. R. Glahn, and D. A. Lowry, "Automatic Decoding of Hourly Weather Reports," <u>ESSA Technical Memorandum WBTM TDL-21</u>, July 1969, 27 pp.
- 12. Jorgensen, D. L., "Climatological Probabilities of Precipitation for the Conterminous United States," <u>ESSA Technical Report WB-5</u>, December 1967, 60 pp.
- 13. Klein, W. H., B. M. Lewis, and C. W. Crockett, "Objective Forecasts of Daily and Mean Surface Temperature," Monthly Weather Review, Vol. 90, No. 1, January 1962, pp. 11-17.

- 14. Kulawiec, M. H., "Local Cloud and Precipitation Forecast Method (SLYH)," ESSA Weather Bureau Technical Note 13-FCST-2, September 1965, 59 pp.
- 15. Lowry, D. A., and H. R. Glahn, "Relationships Between Integrated Atmospheric Moisture and Surface Weather," <u>Journal of Applied Meteorology</u>, Vol. 8, No. 5, October 1969, pp. 762-768.
- 16. Lund, I. A., "Estimating the Probability of a Future Event from Dichotomously Classified Predictors," <u>Bulletin of the American</u> <u>Meteorological Society</u>, Vol. 36, No. 7, September 1955, pp. 325-328.
- 17. Miller, R. G., "Regression Estimation of Event Probabilities," <u>Technical</u>
 <u>Report No. 1, Contract Cwb-10704</u>, The Travelers Research Center, Inc.,
 April 1964, 153 pp.
- 18. Miller, R. G., "The Screening Procedure" (in "Studies in Statistical Weather Prediction"), Final Report, Contract No. AF19 (604)-1590 (Edited by B. Shorr), The Travelers Research Center, Inc., Hartford, Conn., December 31, 1958, pp. 86-95.
- 19. Mook, C. P., "An Objective Method of Forecasting Thunderstorms for Washington, D. C. in May," Unpublished Manuscript, February 1948, 12 pp.
- 20. Panofsky, H. A., and G. W. Brier, <u>Some Applications of Statistics to Meteorology</u>, The Pennsylvania State University, University Park, **P**ennsylvania, 1958, 224 pp.
- 21. Pore, N. A., "The Relation of Wind and Pressure to Extratropical Storm Surges at Atlantic City," <u>Journal of Applied Meteorology</u>, Vol. 3, No. 2, April 1964, pp. 155-163.
- 22. Reap, R. M., "Prediction of Temperature and Dew Point by Three-Dimensional Trajectories," <u>ESSA Technical Memorandum WBTM TDL-15</u>, October 1968, 20 pp.
- 23. Russo, J. A., I. Enger, and G. T. Merriman, "A Statistical Approach to the 12-48-hr. Prediction of Precipitation Probability," <u>Final Report</u>, <u>Contract Cwb-11100</u>, The Travelers Research Center, Inc., Hartford, Conn., August 1966, 107 pp.
- 24. Shuman, F. G., and J. B. Hovermale, "An Operational Six-Layer Primitive Equation Model," <u>Journal of Applied Meteorology</u>, Vol. 7, No. 4, August 1968, pp. 525-547.
- 25. Younkin, R. J., J. A. LaRue, and F. Sanders, "The Objective Prediction of Clouds and Precipitation Using Vertically Integrated Moisture and Adiabatic Motions," <u>Journal of Applied Meteorology</u>, Vol. 4, No. 1, February 1965, pp. 3-17.

(Continued from inside front cover)

- WBTM TDL 16 Objective Visibility Forecasting Techniques Based on Surface and Tower Observations. Donald M. Gales, October 1968. (PB-180 479)
- WBTM TDL 17 Second Interim Report on Sea and Swell Forecasting. N. A. Pore and Lt. W. S. Richardson, USESSA, January 1969. (PB-182 273)
- WBTM TDL 18 Conditional Probabilities of Precipitation Amounts in the Conterminous United States. Donald L. Jorgensen, William H. Klein, and Charles F. Roberts, March 1969. (PB-183 144)
- WBTM TDL 19 An Operationally Oriented Small-Scale 500-Millibar Height Analysis.
 Harry R. Glahn and George W. Hollenbaugh, March 1969.
- WBTM TDL 20 A Comparison of Two Methods of Reducing Truncation Error. Robert J. Bermowitz, May 1969. (PB-184 741)
- WBTM TDL 21 Automatic Decoding of Hourly Weather Reports. George W. Hollenbaugh, Harry R. Glahn, and Dale A. Lowry, July 1969. (PB-185 806)
- WBTM TDL 22 An Operationally Oriented Objective Analysis Program. Harry R. Glahn, George W. Hollenbaugh, and Dale A. Lowry, July 1969. (PB-186 129)
- WBTM TDL 23 An Operational Subsynoptic Advection Model. Harry R. Glahn, Dale A. Lowry, and George W. Hollenbaugh, July 1969. (PB-186 389)
- WBTM TDL 24 A Lake Eric Storm Surge Forecasting Technique. William S. Richardson and N. Arthur Pore, August 1969. (PB-185 778)
- WBTM TDL 25 Charts Giving Station Precipitation in the Plateau States From 850- and 500-Millibar Lows During Winter. August F. Korte, Donald L. Jorgensen, and William H. Klein, September 1969. (PB-187 476)
- WBTM TDL 26 Computer Forecasts of Maximum and Minimum Surface Temperatures. William H. Klein, Frank Lewis, and George P. Casely, October 1969.

