

Forecasting

SIMULATED STRATIFICATION FOR PREDICTION OF PRECIPITATION TYPE

by Harry R. Glahn

Techniques Development Laboratory,
Office of Systems Development,
National Weather Service, NOAA,
Silver Spring, Maryland 20910

ABSTRACT

The updating of MOS forecasts of the probability of precipitation type (freezing, frozen, or liquid) with a recent observation is used as an example in using indicator variables to simulate stratification of the developmental sample. In this example, inclusion of the observation in a primitive manner in the updating regression equations does not make full use of the new information. It is proposed that simulated stratification is useful for updating guidance forecasts, such as MOS, for projections of a few hours, and that the simulation may be easier from both developmental and operational aspects than actually stratifying the sample and producing regression relationships for each stratum.

1. INTRODUCTION

The Local AFOS MOS Program (LAMP) under development in the Techniques Development Laboratory (TDL) will produce an objective prediction system that can be run on a minicomputer and provide updated MOS guidance Glahn (2) and Glahn and Unger (3). Inputs to the system will be the most recent station observations and analyses of those observations; output from simple, advective numerical models initialized with observations, output from NMC's primary guidance model, radar data when available, and the centrally-produced MOS forecasts.

As a part of the LAMP effort, Bocchieri and Forst (4), (5) performed a series of experiments in predicting precipitation type. They used data at 0800 and at 1300 GMT and made 2-, 5-, 8-, and 11-h predictions. The central MOS forecasts which correspond to the 0800 (1300) GMT start time can be considered to be 7- (12-), 10- (15-), 13- (18-), and 16- (21-) h projections, since 0300 GMT observations are used as the latest input to them. Bocchieri and Forst were surprised to find that the centralized MOS 7- and 12-h forecasts were better than 2-h regression forecasts produced from observations alone when evaluated on independent data. Also, the inclusion in regression equations of LAMP model output and observations did not improve on the MOS forecasts alone as much as expected.

In attempting to see why these results were so disappointing, Bocchieri and Forst stratified the developmental sample according to whether or not precipitation was occurring at initial time—that is, whether or not the observation included

precipitation. By so doing, and developing separate equations for each of the two subsamples, results were better, and in particular, the observations alone now produced better 2-h forecasts than the MOS 7- and 12-h forecasts.

This paper follows up on these experiments and uses indicator variables to simulate stratification—thus, "simulated stratification."

2. PREDICTAND DEFINITION

Following Bocchieri and Forst (4), observations of precipitation type were divided into three mutually exclusive categories: freezing rain or drizzle, snow or ice pellets (called frozen precipitation), and rain and mixed types (called liquid precipitation). Only cases in which precipitation occurred at the forecast valid time were included in the developmental sample; therefore, the precipitation type forecasts are conditional on the event that precipitation occurs. In this paper, 2-h predictions of precipitation type are made starting from 0800 GMT.

3. PREDICTOR DEFINITION

MOS forecasts are available and are used here as the probabilities of the three predictand categories. That is, we have a MOS conditional probability forecast of each of the three predictand categories valid at the same time as the predictand—1000 GMT. The only other variable used in this study is the precipitation observation—whether or not precipitation was occurring at 0800 GMT and if it were, the type, defined in the same way as the predictand.

4. MODEL

a. Multiple regression

A statistical model that has been used over and over in meteorology in developing objective prediction systems is multiple linear regression. Non-linear relationships between the predictand and the predictors can be, to some degree, accounted for by transformations of the predictors. The predictors can be binary (zero or one variables) either because they occur that way (e.g., precipitation occurring or not) or they can be created from a quasi-continuous variable by breaking it into categories and giving a (new) dummy variable the value of one (zero) when the value of the original variable is (is not) in that category.

The predictand can also be binary; Miller (6) has termed this use of the regression model

Regression Estimation of Event Probabilities (REEP). In this case, the model estimates the relative frequency of the binary predictand for realistic combinations of values of the predictor variables.

b. Simulated Stratification

Developers many times believe that they will get better overall results (i.e., a better relationship between the predictand, or predictands, and the predictors) if they stratify their data sample. That is, they believe that the predictand-predictor relationships are fundamentally different in subsets of the sample—so different that they cannot be adequately accounted for by parameterization of the subset characteristics. This is a very appealing concept, but it doesn't always work out in practice because the subdivision of the sample may cause overfitting, and while the results with stratification will be better on the developmental data, they may not be as good on independent test data. However, there are situations where stratification, or something very close to it, should be a definite advantage. The example in this paper is such a situation.

Stratification will produce a set of equations for a single predictand like the following:

$$\hat{Y}_1 = a_{10} + a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n \quad (1)$$

$$\hat{Y}_2 = a_{20} + a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n$$

⋮

$$\hat{Y}_g = a_{g0} + a_{g1}X_1 + a_{g2}X_2 + \dots + a_{gn}X_n$$

where there are g exhaustive and mutually exclusive groups or strata, and the same n predictors are in each equation. (Actually, not all n predictors need be in each equation—some of the a 's can be zero.) Note that these groups are determined by data available at the time the forecast is to be made.

If the sample had not been stratified, the one equation with the same n predictors would have been simply

$$\hat{Y} = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n. \quad (2)$$

One crude attempt to embellish this equation without stratification is to add $g-1$ binary predictors, each signifying membership or non-membership in a particular group, i.e.,

$$\hat{Y} = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n +$$

$$a_{n+1}X_{n+1} + \dots + a_{n+g-1}X_{n+g-1}.$$

Eq. (3) may be an improvement over Eq. (2). However, it's obvious that the relationship between the predictand and the predictors for the different groups differs by only a constant. For example, this equation applied to group 1 is

$$\hat{Y}_{(1)} = a_0 + a_{n+1} + a_1X_1 + a_2X_2 + \dots + a_nX_n \quad (4)$$

since

$$X_{n+1} = 1 \text{ and } X_{n+i}, i = 2, 3, \dots, g-1, = 0.$$

There are many situations where this treatment does not allow enough flexibility. One such situation exists when we have an estimate of the probability of an event and want to improve upon that probability by using a recent observation. Suppose that the MOS forecasts of the probability of precipitation (PoP) for a 12-h period are available and the forecaster now has an observation taken 7 hours after the MOS forecast was made and only 2 hours before the start of the 12-h period. One might expect that the conditional relative frequency of precipitation (a new, better estimate of PoP) as a function of the MOS PoP would be different for those cases when precipitation was and was not observed 2 hours before the start of the period, and that the two frequencies would not just differ by a constant. The (linear) relationships might be something like those shown in Fig. 1.

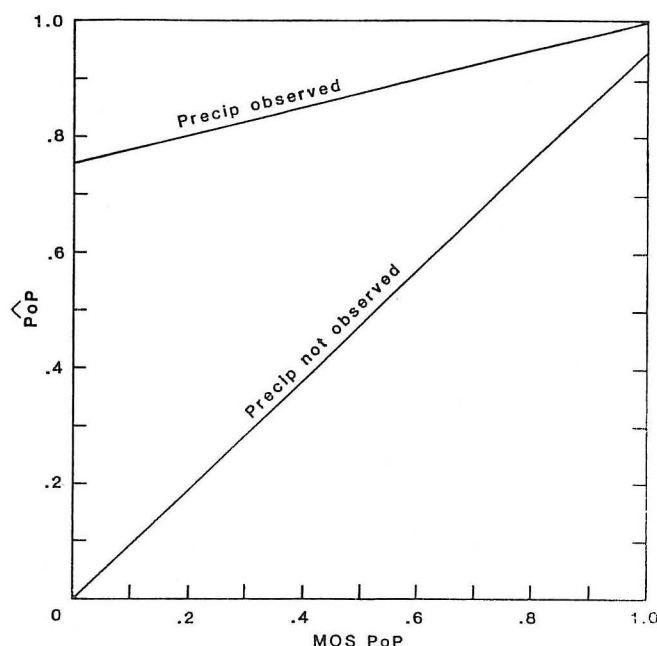


Fig. 1. Hypothetical relationships between the relative frequency of precipitation in a 12-h period and the MOS PoP when precipitation was and was not observed 2 hours before the start of the period.

The two relationships in Fig. 1 could be derived from data stratified on the observation of precipitation and no precipitation. Both relationships can also be represented by a single equation

$$\hat{Y} = a + bX_1 + cX_2 + dX_1X_2 \quad (5)$$

where X_1 is the MOS PoP and X_2 is a binary indicator variable taking the value of zero when precipitation is not observed and the value one when precipitation is observed. Then, for the observed-precipitation cases, the equation becomes

$$\hat{Y} = (a+c) + (b+d) X_1 \quad (6)$$

and for the no-precipitation cases the equation becomes

$$\hat{Y} = a + bX_1. \quad (7)$$

Note that these equations can be represented by the two lines in Fig. 1. The single equation (Eq. (5)) can be derived by least-squares regression on the total sample by using the binary variable X_2 , and the results will be identical to the results where a different relationship is derived by least squares on the two subsamples.

In the application to precipitation type presented in this paper, there are three binary predictands, so there will be three equations for each specific application of the technique.

5. APPLICATION OF THE MODEL

a. Data Sample

Data from 30 stations in a five-state area around the Washington, D.C., Weather Service Forecast Office were available for five winter seasons, defined as October through March, for the years 1977-78 through 1981-82. For those stations for which MOS forecasts were not available, forecasts were made by interpolation from stations for which forecasts were available (see Bocchieri and Forst, (5) for more details.) These MOS forecasts were not necessarily those transmitted in real time; they were generated from quality-controlled data with the probability of precipitation type (PoPT) system now in use (see Bocchieri and Maglaras (7)). Also, since MOS forecasts are not made for the 1000 GMT valid time, linear interpolation in time was made for each station from forecasts valid at 0600 and 1200 GMT. The observations were taken from TDL's hourly data archive.

Data from all stations within the five-state area were combined to produce one set of equations valid for all stations in the area. This was necessary because the number of cases (at most, one per day per station) would be too small to develop a reliable set of equations for each station separately or even for just a few stations combined. As noted previously, only those occasions when precipitation was occurring at the predictand time were used. Also, freezing precipitation is quite rare, as will be seen later, and data must be aggregated over several months and locations to acquire a sample adequate to define a relationship involving that variable.

b. Benchmark No. 1 - Observations not used as predictors

Purely as a benchmark against which to test the simulated stratification, we derived the following equations:

¹ Actually, X_2 is not limited to zero and one—any two (non-equal) values can be used.

$$\hat{P}_Z = 0.00 + 0.81 (\text{MOS } P_Z) - 0.02 (\text{MOS } P_F) \quad (8)$$

$$(\text{RV} = 0.250)$$

$$\hat{P}_F = 0.00 - 0.04 (\text{MOS } P_Z) + 1.02 (\text{MOS } P_F) \quad (9)$$

$$(\text{RV} = 0.813)$$

$$\hat{P}_L = 1.00 - 0.77 (\text{MOS } P_Z) - 1.00 (\text{MOS } P_F) \quad (10)$$

$$(\text{RV} = 0.786)$$

where P_Z , P_F , and P_L represent the (conditional) probability of freezing, frozen, and liquid precipitation at 1000 GMT, respectively. The reductions of variance (RV) by each equation are also shown.

The MOS P_L is not included as a predictor because it is redundant with MOS P_Z and MOS P_F . $\text{MOS } P_Z + \text{MOS } P_F + \text{MOS } P_L = 1$, and if all three are included, there is an infinite number of solutions, and the cross-product matrix cannot be inverted. We can also note that the sum of the constants = 1.00, and the sum of the coefficients of a particular predictor = 0.00; thus, the sum of the forecast probabilities = 1.0. This is a characteristic of the regression model when the predictands represent exhaustive and mutually exclusive groups and the same predictors are in each equation.

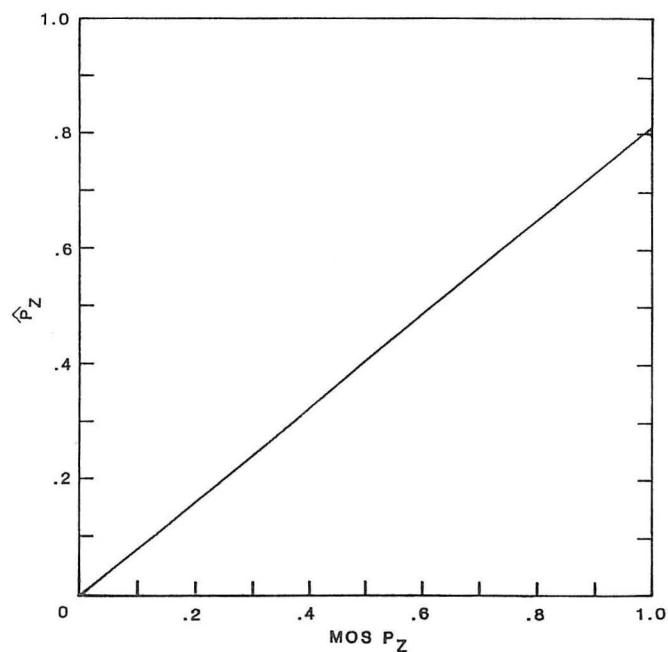


Fig. 2. Equation 8 when MOS $P_F = 0.0$. When MOS $P_F = 0.5$, the result is a parallel line only 0.01 units below the one shown. That is, for a given MOS P_Z , P_Z is decreased by 0.01 as MOS P_F goes from 0.0 to 0.5. Since MOS $P_F + \text{MOS } P_Z$ will not generally exceed 1.00, when MOS $P_F = 0.5$ only the range of values 0.0 to 0.5 for MOS P_Z are appropriate.

Equations (8) and (9) are represented in Figs. 2 and 3, respectively. It can be seen from Fig. 3 that the MOS P_F is very reliable as nearly as can

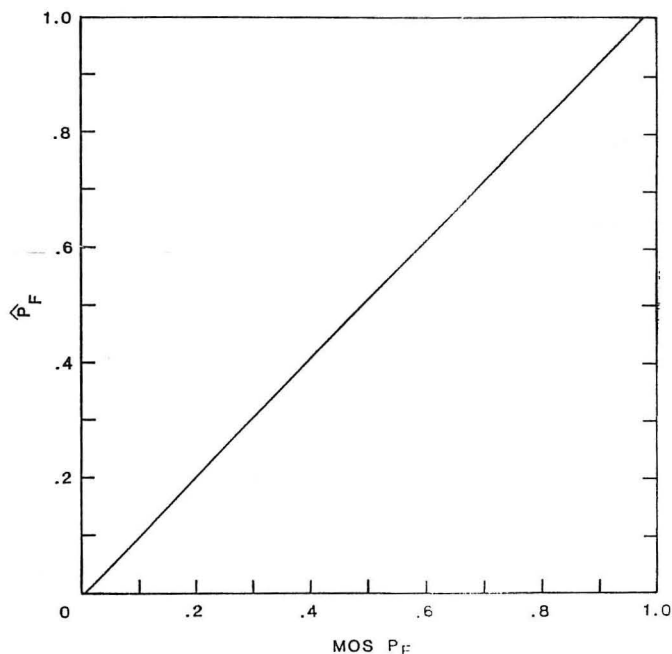


Fig. 3. Equation 9 when $\text{MOS } P_Z = 0.0$. When $\text{MOS } P_Z = 0.5$, the result is a parallel line only 0.02 units below the one shown.

be ascertained from a linear fit to the data. Usually $\text{MOS } P_Z$ is low, and $\hat{P}_F \approx \text{MOS } P_F$. The reliability of the $\text{MOS } P_Z$ is not as good on this sample as is the reliability of the $\text{MOS } P_F$, but is respectable (Fig. 2). Because the relative frequency of freezing precipitation is low, this result could vary considerably from sample to sample (there were only 91 cases of freezing precipitation in this sample).

c. Benchmark No. 2 - Observed precipitation types as binary predictors

This benchmark is the model many times used when a categorical variable such as precipitation type is included in a regression equation. The following equations were developed:

$$\begin{aligned} \hat{P}_Z &= 0.01 + 0.48 (\text{MOS } P_Z) + 0.00 (\text{MOS } P_F) \quad (11) \\ &+ 0.63 W_Z - 0.02 W_F - 0.01 W_L \\ &(\text{RV} = 0.488) \end{aligned}$$

$$\begin{aligned} \hat{P}_F &= 0.04 + 0.01 (\text{MOS } P_Z) + 0.79 (\text{MOS } P_F) \quad (12) \\ &- 0.18 W_Z + 0.23 W_F - 0.06 W_L \\ &(\text{RV} = 0.841) \end{aligned}$$

$$\begin{aligned} \hat{P}_L &= 0.95 - 0.49 (\text{MOS } P_Z) - 0.79 (\text{MOS } P_F) \quad (13) \\ &- 0.45 W_Z - 0.21 W_F + 0.07 W_L \\ &(\text{RV} = 0.821) \end{aligned}$$

where W_Z , W_F , and W_L take the value of one when precipitation at 0800 GMT is observed as freezing, frozen, or liquid, respectively, and zero otherwise.

All three variables, W_Z , W_F , and W_L , can be in the equation because the no precipitation cases complete the set. Note that, again, the three constants sum to unity and the sum of the coefficients for each of the other variables is zero. The reductions of variance have improved over those for Eqs. (8), (9), and (10) as indeed they must on this developmental sample, the one for P_Z most substantially. In fact, the RV for P_Z nearly doubled.

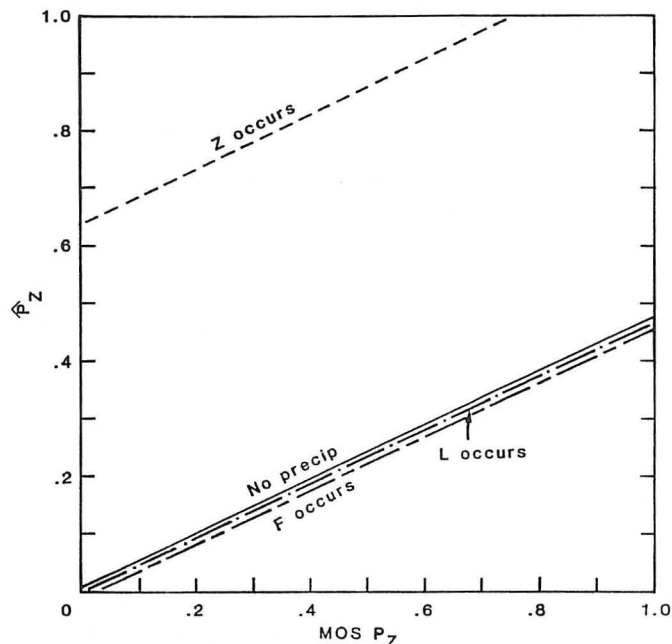


Fig. 4. Equation 11 when $\text{MOS } P_F = 0.0$. When $\text{MOS } P_F = 0.5$, the relations are essentially unchanged.

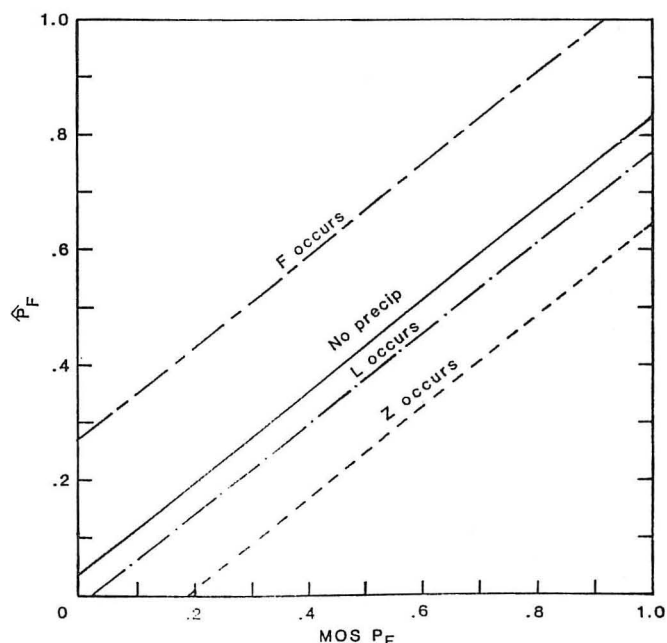


Fig. 5. Equation 12 when $\text{MOS } P_Z = 0.0$. When $\text{MOS } P_Z = 0.5$, the result is that each line is displaced downward by only about 0.005 units; the range of the $\text{MOS } P_F$ is then 0.0 to 0.5.

Equations (11) and (12) are represented in Figs. 4 and 5, respectively. Each of the lines in Fig. 4 is parallel to the others; the same is true for Fig. 5. The slopes of the lines in Fig. 4 (Fig. 5) are not the same as the slope of the line in Fig. 2 (Fig. 3).

d. Simulated stratification on precipitation event

This model treats the precipitation cases differently than no-precipitation cases. The following equations were developed:

$$\begin{aligned}\hat{P}_Z &= 0.00 + 0.85 (\text{MOS } P_Z) - 0.01 (\text{MOS } P_F) \quad (14) \\ &+ 0.73 W_Z - 0.01 W_F + 0.00 W_L \\ &- 0.63 (\text{MOS } P_Z) W_P + 0.02 (\text{MOS } P_F) W_P \\ &(\text{RV} = 0.519)\end{aligned}$$

$$\begin{aligned}\hat{P}_F &= -0.01 - 0.03 (\text{MOS } P_Z) + 1.00 (\text{MOS } P_F) \quad (15) \\ &- 0.01 W_Z + 0.65 W_F + 0.01 W_L \\ &- 0.08 (\text{MOS } P_Z) W_P - 0.62 (\text{MOS } P_F) W_P \\ &(\text{RV} = 0.867)\end{aligned}$$

$$\begin{aligned}\hat{P}_L &= 1.01 - 0.82 (\text{MOS } P_Z) - 0.99 (\text{MOS } P_F) \quad (16) \\ &- 0.72 W_Z - 0.64 W_F - 0.01 W_L \\ &+ 0.71 (\text{MOS } P_Z) W_P + 0.60 (\text{MOS } P_F) W_P \\ &(\text{RV} = 0.848)\end{aligned}$$

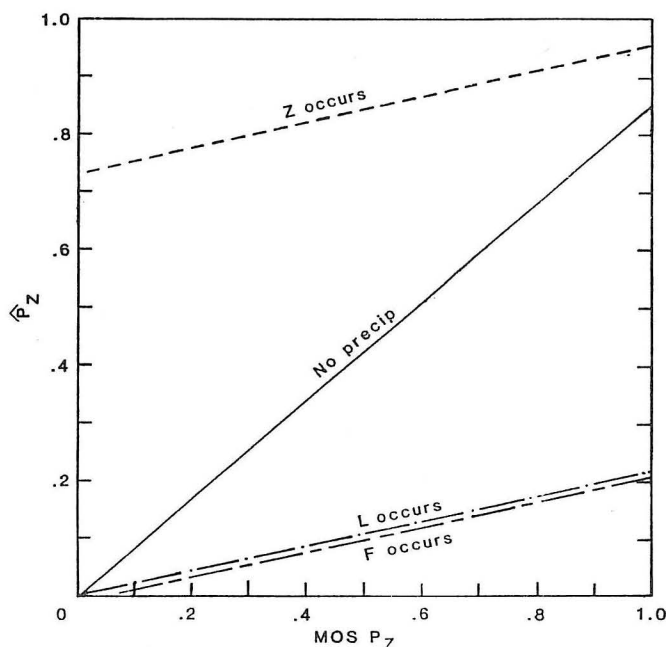


Fig. 6. Equation 14 when $\text{MOS } P_F = 0.0$. When $\text{MOS } P_F = 0.5$, the result is that no-precip line is displaced downward by only about 0.005 units and the other lines are displaced upward by about 0.005 units; the range of the $\text{MOS } P_Z$ is then 0.0 to 0.5.

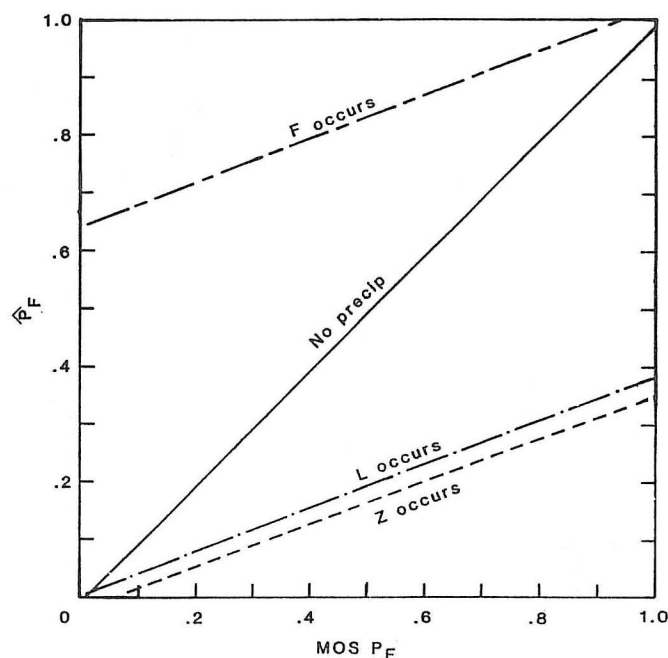


Fig. 7. Equation 15 when $\text{MOS } P_Z = 0.0$. When $\text{MOS } P_Z = 0.5$, the result is that the no-precip line is displaced downward by only about 0.015 units and the other lines are displaced downward by about 0.055 units; the range of the $\text{MOS } P_F$ is then 0.0 to 0.5.

where W_P takes the value of one when precipitation is observed at 0800 GMT, and zero otherwise. (Note that $W_P = W_Z + W_F + W_L$.) The slopes of the no-precipitation lines in Figs. 6 and 7, which represent Eqs. (14) and (15), respectively, are different than the slopes of the other lines in the same figure. The inclusion of the two additional terms has given much more flexibility in fitting the data, and the reductions of variance have increased about 3% for each of P_Z , P_F , and P_L .

e. Simulated stratification on precipitation and precipitation type

This model simulates stratification on the precipitation event, as observed at 0800 GMT, and on precipitation type when precipitation is observed. The following equations were developed:

$$\begin{aligned}\hat{P}_Z &= 0.00 + 0.85 (\text{MOS } P_Z) - 0.01 (\text{MOS } P_F) \quad (17) \\ &+ 0.80 W_Z + 0.03 W_F - 0.00 W_L \\ &- 0.72 (\text{MOS } P_Z) W_Z - 0.20 (\text{MOS } P_F) W_Z \\ &- 0.75 (\text{MOS } P_Z) W_F - 0.02 (\text{MOS } P_F) W_F \\ &- 0.56 (\text{MOS } P_Z) W_L + 0.05 (\text{MOS } P_F) W_L \\ &(\text{RV} = 0.522)\end{aligned}$$

$$\begin{aligned}\hat{P}_F = & -0.01 - 0.03 (\text{MOS } P_Z) + 1.00 (\text{MOS } P_F) \quad (18) \\ & + 0.01 W_Z + 0.79 W_F + 0.00 W_L \\ & - 0.06 (\text{MOS } P_Z) W_Z - 0.71 (\text{MOS } P_F) W_Z \\ & - 0.57 (\text{MOS } P_Z) W_F - 0.77 (\text{MOS } P_F) W_F \\ & + 0.12 (\text{MOS } P_Z) W_L - 0.58 (\text{MOS } P_F) W_L \\ & (\text{RV} = 0.869)\end{aligned}$$

$$\begin{aligned}\hat{P}_L = & 1.01 - 0.82 (\text{MOS } P_Z) - 0.99 (\text{MOS } P_F) \quad (19) \\ & - 0.81 W_Z - 0.82 W_F - 0.00 W_L \\ & + 0.78 (\text{MOS } P_Z) W_Z + 0.91 (\text{MOS } P_F) W_Z \\ & + 1.32 (\text{MOS } P_Z) W_F + 0.79 (\text{MOS } P_F) W_F \\ & + 0.44 (\text{MOS } P_Z) W_L + 0.53 (\text{MOS } P_F) W_L \\ & (\text{RV} = 0.852)\end{aligned}$$

None of the lines in Figs. 8 and 9, representing Eqs. (17) and (18), respectively, is constrained to have the same slope as any other line. The reductions of variance have increased only slightly over the previous model. This is consistent with the fact that Fig. 8 (Fig. 9) is not greatly different from Fig. 6 (Fig. 7). The added flexibility that this stratification on precipitation event and type gives over stratification only on the precipitation event is that the "F occurs," "Z occurs," and "L occurs," lines can have slopes that differ from each other. Evidently, this added flexibility is not greatly

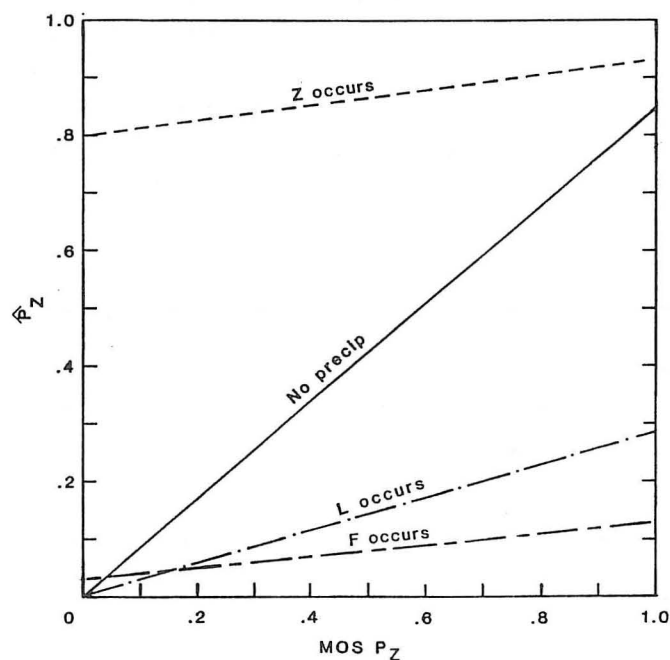


Fig. 8. Equation 17 when $\text{MOS } P_F = 0.0$. When $\text{MOS } P_F = 0.5$, the result is that the no precip, F occurs, L occurs, and Z occurs lines are displaced by -0.005 , -0.015 , 0.02 , and -0.105 units, respectively; the range of the $\text{MOS } P_Z$ is then 0.0 to 0.5.

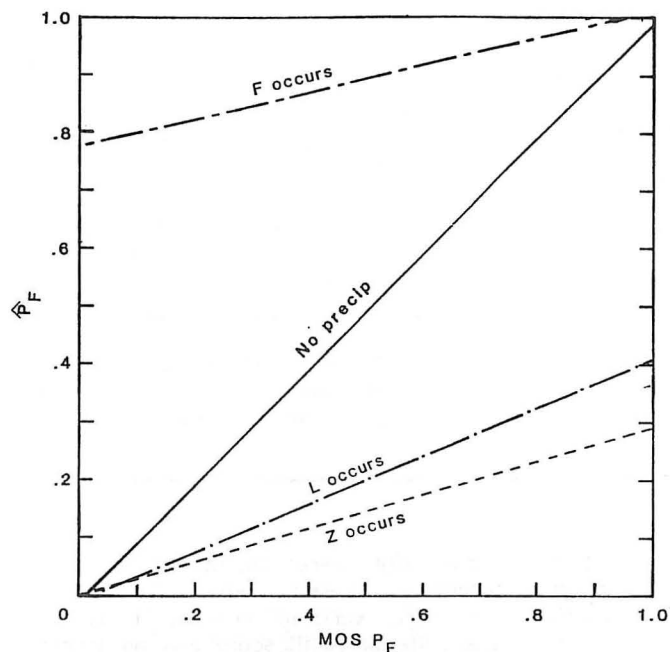


Fig. 9. Equation 18 when $\text{MOS } P_Z = 0.0$. When $\text{MOS } P_Z = 0.5$, the result is that the no precip, F occurs, L occurs, and Z occurs lines are displaced by -0.015 , -0.30 , 0.045 , and -0.045 units, respectively; the range of the $\text{MOS } P_F$ is then 0.0 to 0.5.

important. Note that the no-precipitation lines in Figs. 6 and 8 (7 and 9) are identical, as they must be.

f. Discussion of results

Table 1 gives the reduction of variance and mean-square errors (MSE) for each of the four model applications to the same dependent data sample discussed above — two benchmarks and two simulated stratifications. Each model improved upon the lower order ones below it, as it must on this developmental data. The most notable improvement is that for P_Z in benchmark No. 2 over benchmark No. 1; the introduction of the observations reduced the MSE by $.007$ or $.007/.021 = 33$ percent. The addition of observations (Benchmark No. 2) reduced the MSE by $.006$ or $.006/.040 = 15$ percent for the frozen category and by $.008$ or $.008/.048 = 17$ percent for the liquid category. The simulated stratification models further reduced the MSE by $.006 = .006/.034 = 18$ percent for the frozen category and $.006/.040 = 15$ percent for the liquid category. Prediction of the freezing category was not helped much by stratification.

To see how each of the four systems of equations discussed above would perform on independent data, they were rederived on each combination of 4 years of data and tested on the fifth. That is, five sets of equations were derived for each of the four systems, one set derived on years 1, 2, 3, and 4, another on years 1, 2, 3, and 5, etc. Forecasts were made for each of these five sets for the year omitted in the development. Then the 5 years of forecasts were verified, all 5 years being "independent" data. All samples were matching; the cases in any year were the same for development and testing.

Predictand		Equation Nos.											
Precip. Type	No. of Occurrences in Sample	8-10			11-13			14-16			17-19		
		No. of Terms	RV	MSE	No. of Terms	RV	MSE	No. of Terms	RV	MSE	No. of Terms	RV	MSE
Freezing	91	2	0.250	.021	5	0.488	.014	7	0.519	.014	11	0.522	.013
Frozen	953	2	0.813	.040	5	0.841	.034	7	0.867	.028	11	0.869	.028
Liquid	2108	2	0.786	.048	5	0.821	.040	7	0.848	.034	11	0.852	.033

Table 1. Statistics for the four models discussed. RV is reduction of variance and MSE is mean square error. The number of terms given for each equation is in addition to the constant.

P-scores (Brier (8)) were calculated. Also, categorical forecasts were made from the probabilities and were verified on the basis of percent correct and Heidke skill score and on threat score of the freezing category.² Improvement over the MOS forecasts was calculated for each score. Categorical forecasts were made from the MOS probabilities by using the thresholds used in the operational system for the Washington, D.C. area for 12-h forecasts valid at 1200 GMT—0.28 for the freezing category and 0.40 for the frozen category. Categorical forecasts are made by first comparing \hat{P}_Z with the freezing threshold; if \hat{P}_Z is the larger, a forecast of freezing precipitation is made. If \hat{P}_Z does not exceed the threshold, \hat{P}_F is compared to the frozen threshold; if \hat{P}_F is the larger, a forecast of frozen precipitation is made. Otherwise, the forecast is for liquid precipitation. Some initial testing indicated that these thresholds should be reduced for the regression models used in this study to give biases of each category near unity. The thresholds chosen, without additional tuning, were 0.25 and 0.37 for the freezing and frozen categories, respectively.

Table 2 shows the results on the 5 years of data. The bias of a category is the number of

forecasts of that category divided by the number of observations of that category. A bias of 1.0, or perhaps slightly greater for the rather rare freezing category, is appropriate. It is obvious that Benchmark No. 1, which is really just a recalibration of the MOS probability forecasts, offers no improvement over MOS even in terms of the P score which does not involve a transformation to categorical forecasts.

²Heidke skill score, defined by Panofsky (1958), $= (H-E)/(T-E)$ where H = the number of correct forecasts, T = the total number of forecasts, and E = the expected number of correct forecasts based on the marginal totals of the contingency table. The threat score, defined as early as 1884 by Gilbert (9), which he called the "ratio of verification," by W. C. Palmer and R. A. Allen in an unpublished manuscript in 1949, and by Donaldson, et al. (10) as the Critical Success Index, is $H/(F+O-H)$ where H = the number of correct forecasts, F = the number of forecasts of a (rare) event, and O = the number of observations of that event.

Model	Bias		P		Percent Correct		Heidke Skill		Threat Score	
	Freezing	Frozen	Score	Imp. over MOS	Score	Imp. over MOS	Score	Imp. over MOS	Score	Imp. over MOS
MOS	1.21	0.99	.109	-	92.2	-	0.832	-	0.314	-
Benchmark No. 1	1.10	1.01	.111	-0.013	92.1	-0.001	0.830	-0.002	0.299	-0.046
Benchmark No. 2	1.01	1.03	.088	0.189	94.2	0.021	0.871	0.051	0.441	0.406
Stratification on precipitation occurrence	1.09	0.99	.078	0.285	94.8	0.028	0.886	0.066	0.508	0.619
Stratification on precipitation occurrence and type	1.10	1.00	.078	0.285	94.6	0.026	0.883	0.061	0.504	0.606

Table 2. Independent data verification for the four models discussed. The scores are defined in the text.

Benchmark No. 2, the inclusion of the initial observation in a primitive way, did provide an improvement—considerable in terms of P score (19%), modest in terms of percent correct (2%) and skill score (5%), and large in terms of threat score of the freezing category (41%). It must be remembered that large improvements in percent correct and skill score are not possible, because they are already high for MOS, and even perfect forecasting would increase percent correct and skill score by only 8% and 17%, respectively; thus, the inclusion of the observation has improved upon MOS by 25 to 30% of the possible improvement.

Stratification on precipitation occurrence improved the P score (percent correct) (skill score) by an additional 10% (1%) (1%) over MOS. The improvement in threat score is notable—from 0.44 to 0.51, an additional improvement over MOS of 21%.

Stratification on precipitation type did not provide better results than stratification on precipitation occurrence. In fact, the additional fitting of the relationships to the developmental data by computing more coefficients tended to give slightly worse results on test data. This lack of improvement is in agreement with the similarity of Figs. 6 (7) and 8 (9).

6. SUMMARY AND CONCLUSIONS

Many times it is desirable to update a guidance forecast based on recent observations. This is a concept embodied in LAMP. An example has been shown in which the flexibility of stratification, or simulated stratification, produced better results than the inclusion of the observations in a more primitive way. The simulated stratification increased the improvement of the threat score of freezing precipitation over MOS from 0.41 to 0.62, a very worthwhile increase.

Overfitting can be a problem, and one must be careful to not include too many predictors in the statistical relationship. Also, objectively screening a large number of predictors and choosing the best may produce a relationship that is too heavily dependent on the specific characteristics of the developmental data set. The particular problem under investigation should be carefully analyzed, and a method used which is appropriate for that problem. Stratification on some condition may produce useful results, and it may be easier both in development and implementation to simulate that stratification rather than to actually stratify the sample and develop separate relationships.

NOTES AND REFERENCES

1. Dr. Harry R. (Bob) Glahn is Chief of the Techniques Development Laboratory. He was awarded a B.S. in Education degree from the Northeast Missouri State Teachers College, a M.S. degree from the Massachusetts Institute of Technology, and a Ph.D. from the Pennsylvania State University.
2. Glahn, H. R., 1980: Plans for the development of a Local AFOS MOS Program (LAMP). *Preprints Eighth Conference on Weather Forecasting and Analysis*, Denver, Amer. Meteor. Soc., 302-305.
3. _____, and D. A. Unger, 1986: A Local AFOS MOS Program (LAMP) and its application to wind prediction. *Mon. Wea. Rev.*, 114, (accepted for publication).
4. Bocchieri, and A. L. Forst, 1984a: Experiments in the use of the Local AFOS MOS Program (LAMP) for forecasting precipitation type in the Washington, D.C. WSFO forecast area. *TDL Office Note 84-1*, National Weather Service, NOAA, U.S. Department of Commerce, 14 pp.
5. _____, and _____, 1984b: Some results from the Local AFOS MOS Program (LAMP) for forecasting precipitation type in the Washington, D.C. WSFO forecast area. *Preprints Tenth Conference on Weather Forecasting and Analysis*, Clearwater Beach, FL, Amer. Meteor. Soc., 278-285.
6. Miller, R. G., 1964: Regression estimation of event probabilities. Technical Report No. 1, Contract Cwb-10704, The Travelers Research Center, Hartford, CT, 153 pp.
7. Bocchieri, J. R., and G. J. Maglaras, 1983: An improved operational system for forecasting precipitation type. *Mon. Wea. Rev.*, 111, 405-418.
8. Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, 78, 1-3.
9. Gilbert, G. K., 1884: Finley's Tornado predictions. *Amer. Meteor. Journal*, 1, 166-172.
10. Donaldson, R. S., R. M. Dyer, and M. J. Kraus, 1975: An objective evaluator of techniques for predicting severe weather events. *Preprints Ninth Conference on Severe Local Storms*, Norman Amer. Meteor. Soc., 321-326.

Join Us!

If you are interested and concerned about operational meteorology, join and participate in the National Weather Association. Annual dues are just \$20.00. Send your name, address and any particulars as to your occupation, affiliation and main meteorological interests to:

NATIONAL WEATHER ASSOCIATION
4400 STAMP ROAD, ROOM 404
TEMPLE HILLS, MD 20748

Name: _____

Address: _____

Dues enclosed (\$20.00 per year). _____ THANK YOU!