EWTD 76-13
02358

U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
SYSTEMS DEVELOPMENT OFFICE
TECHNIQUES DEVELOPMENT LABORATORY

TDL Office Note 76-6

Testing the LFM for PoP Forecasting - Summer Season

David B. Gilhousen

May 1976

TESTING THE LFM FOR PoP FORECASTING - SUMMER SEASON

David B. Gilhousen

INTRODUCTION

Until recently, objective probability of precipitation (PoP) forecasts used operationally by the National Weather Service have been based on only primitive equation (PE) and trajectory (TJ) model output. During the winter of 1975-76, predictors from the National Meteorological Center's limited area fine mesh model (LFM) were added to PE and TJ predictors for winter PoP forecasts out to thirty-six hours. Glahn and Bocchieri (1976) showed that the addition of LFM predictors to PE and TJ predictors increased wintertime forecast accuracy. However, the same study showed the opposite result for the April to September summer season. The small sample of days available to develop the summer equations (72 as opposed to 149 for winter development) may have been the cause for the inconsistent results.

This paper describes a continuation of their experiment for the summer season; a larger developmental data sample is now available. There were several objectives of this latest study:

(1) We wanted to determine the accuracy of PoP forecasts based on LFM data alone relative to the operational PE and TJ PoPs.

(2) We wanted to know if PE and TJ variables combined with the LFM data will provide a substantially better PoP than will LFM predictors alone.

(3) We wanted to determine if we can reduce the occasionally large inconsistencies between a PoP forecast for a 12-hr period and the two 6-hr PoP forecasts within that 12-hr period. The PoP forecasts are inconsistent when either one of the 6-hr PoPs is greater than the 12-hr PoP or the sum of the 6-hr PoPs is less than the 12-hr PoP. This can happen because the three equations are derived independently of each other.

This study was limited to regressions using 0000 GMT model output as predictors and the occurrence of .01 inch of precipitation between 1200-2400 GMT as the predictand. Surface observations were not used as predictors.

DEVELOPMENTAL PROCEDURE

The various regression equations that were derived and tested are discussed in three groups--the operational equations, small sample equations, and small sample equations derived together in one screening run. The first two groups are identified in Table 1 which shows verification statistics to be discussed later. In each case, we developed generalized operator equations by region, each limited to 12 terms. The equations used operationally during the 1975 summer were derived on a 570-day sample from the summers of 1971 through 1974. Derivation of these equations is described in a Technical Procedures Bulletin

(National Weather Service, 1975a) and definition of the 26 regions and station locations are shown in Figure 1.

The small sample equations were developed on a 281-day 1973 and 1974 summer sample of 0000 GMT LFM model output. The data sample dropped to 211 days when output from all three models was combined. If equations had been developed in each of the 26 regions used for the operational system, the data sample in several of the regions might have been prohibitively small. Therefore, we decided to decrease the number of regions. When the equation set contained LFM variables, we developed equations in the 18 regions shown in Figure 2. The regions were determined subjectively from an analysis of the relative frequency of precipitation when the LFM precipitation amount forecast (12-24 hr period) was $\geq$ .01 inch. When the equation set did not contain LFM variables, we derived equations in the 18 regions shown in Figure 3. These regions were determined from an analysis of the relative frequency of precipitation when the 24 hr forecast of PE mean relative humidity was $\geq$ 60%.

The 85 predictors screened in developing the small sample equations included both binary and continuous predictors. The PE and TJ predictors screened are listed in Table 2. The LFM predictors screened were the same ones screened in the operational winter equations (National Weather Service, 1975b) and are listed in Table 3. The binary limits were slightly different from the ones screened in the winter equations. Since we kept the total number of predictors screened in any regression to 85 to conserve computer storage, addition of new predictors necessitated omitting a few old ones. Elimination of specific predictors was based on their relatively poor performance when included in previous runs.

To see if we could reduce inconsistencies between 6-hr and 12-hr PoP forecasts, we developed two other sets of equations. The predictands were 12-18, 18-24, and 12-24 hr occurrence of precipitation and the predictors were composed of only LFM output. Equations for each projection were derived separately (as before) and also together in one screening run. When they were derived together, each equation had the same predictors selected. The screening regression procedure selected the predictor which had the highest correlation coefficient with any one of the predictands and then introduced that predictor into all three of the equations. Equations derived separately do not necessarily contain the same predictors.

VERIFICATION PROCEDURE AND RESULTS

Table 1 shows the verification results on a 137-case independent data sample from the summer of 1975. The scores are the P-score as defined by Brier (1950). The paired t-tests were made in the following manner. Forecasts and observations at the 233 U.S. stations shown in Figure 1 were combined to compute a P-score on each of the days (cases) in the independent data sample. These 137 P-scores were then "paired" with scores associated with another set of equations. An equivalent way of thinking about this is to create a new variable, the differences of the P-scores for each day. We assumed that the cases were independent (which is not strictly true) and that the differences were normally distributed (a plot of the data showed this to be true), so we conducted a t-test whose null hypothesis is that the mean difference is zero. An asterisk in Table 1 indicates that the mean difference was significantly different from zero at the 1 percent level.

From the scores shown in Table 1, we observed that PoP forecasts from equations with LFM predictors developed on the two year sample were significantly better than the operational forecasts. Examining the results on a regional basis, we see that the Southern region is the only one where LFM forecasts were of little additional value. Looking at the day to day performance of both sets of equations, we found that the LFM equations did not yield uniform improvement throughout the independent data sample. Rather, sizable improved P-scores on about twenty-five days were largely responsible for the statistical significance.

Also from Table 1, we found that equations with only LFM predictors gave no worse (perhaps slightly better) PoP forecasts than equations with predictors from all three models. The only exception to this result is in the Eastern region where the equations with all three models performed better than the ones with LFM predictors only.

In drawing conclusions from Table 1 about the relative effectiveness of PE, TJ, and LFM predictors, it should be remembered that the differences in development sample size and in the regions for which generalized operator equations were developed, as described above, may have accounted for part of the differences in score. We believe these are small effects compared to the effect of differing numerical models.

Table 4 shows the independent verification and a consistency check of LFM PoP forecasts whose equations were derived separately versus those derived together. P-scores for all 233 stations were used to see if deriving the equations together deteriorated the quality of the forecasts. The percentage of forecasts that were inconsistent was tabulated to provide a measure of success for this approach. Since the PoP forecasts were not rounded for this consistency check, it is likely that a large percentage of minor inconsistencies would not appear after the forecasts are rounded for teletype transmission. Therefore, we also tabulated the percentage of forecasts that were inconsistent by at least 5%, since these are more likely to appear inconsistent after rounding.

The results from this experiment show that equations derived together performed as well as the equations derived separately for two of the three verification periods, and only marginally less on the third. Also, equations derived together did reduce the number of inconsistencies, especially the number of major inconsistencies.

CONCLUSIONS AND PLANS

Based on results from Table 1, early guidance PoP1 equations with only LFM predictors will be developed for use in the summer of 1976. Also, we will not derive equations with predictors from all the models for first period (12-24 hr) forecast guidance. Instead, we will simply retransmit the early guidance based on LFM model output. However, second period (24-36 hr) forecast guidance will be based on predictors from all three models since we haven't archived LFM variables beyond 24 hours until very recently.

Based on results from Table 4, whenever we need to develop equations for two 6-hr periods within a 12-hr period we will use a screening package that derives all three equations together.

REFERENCES

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. Monthly Weather Review, 78, 1-3.

Glahn, H. R. and J. R. Bocchieri, 1976: Testing the limited area fine mesh model for probability of precipitation forecasting. Monthly Weather Review, 104, 127-132.

National Weather Service, 1975a: Operational probability of precipitation forecasts based on model output statistics--No. 10. Technical Procedures Bulletin No. 136, 6 pp.

National Weather Service, 1975b: Operational probability of precipitation forecasts based on model output statistics--No. 11, Technical Procedures Bulletin No. 147, 11 pp.

Table 1. Verification on 137-case independent data sample of forecasts from summer 1975 from various regression equations. The comparisons and t-tests were based on data from all stations. An asterisk indicates significance at the 1% level.

Type of Equations	Predictors			P-Scores by NWS Regions					Comparisons and t-tests	
	PE	TJ	LFM	Eastern	Southern	Central	Western	All Stations	Compared to OPNL PE+TJ	Compared to small Sample PE+TJ+LFM
Operational	X	X		.2654	.2872	.2415	.1575	.2390	(comparison set)	
Small Sample	X			.2725	.2893	.2472	.1615	.2437	Worse*	
	X	X		.2708	.2897	.2462	.1609	.2429	Worse*	
	X	X	X	.2524	.2879	.2285	.1552	.2316	Better*	(Comparison Set)
	X		X	.2569	.2869	.2289	.1549	.2325	Better*	Worse
			X	.2576	.2872	.2271	.1506	.2312	Better*	Better

Table 2. The PE and TJ predictors that were screened in developing small sample equations.

Field	Model	Smoothing (Points)	Time (Hours)
850-mb Height	PE	5	12,24
Precipitable Water	PE	5	18,30
Mean Rel. Humidity	TJ	5	24
Precipitation Amount	TJ	5	24
Total Totals Index	TJ	5	24
K Index	TJ	5	24
Boundary Layer U-Wind	PE	5	24
Boundary Laver V-Wind	PE	5	24
Mean Rel. Humidity	PE	5,9	24
Boundary Layer Humidity	PE	5,9	24
Second Layer Humidity	PE	5	24
850-mb Vertical Velocity	PE	5	24
650-mb Vertical Velocity	PE	5	24
Sine of Day of Year	-	-	-
Cosine of Day of Year	-	-	-

Table 3. The LFM predictors that were screened in developing small sample equations.

Field	Smoothing (Points)	Time (Hours)	Type[1]
Boundary Layer Rel. Humidity	1,5	12,18	B,C
Second Layer Rel. Humidity	1,5	12,18	B,C
Total Totals Index	5	12,24	B
Mean Rel. Humidity	1,5	18	B,C
Precipitable Water	5	18	B
Boundary Layer U Wind	5	18	B,C
Boundary Layer V Wind	5	18	B,C
700-mb Vertical Velocity	5	18	B
Precipitation Amount	1,5	18,24	B,C
Mean Rel. Humidity Trend[2]	1	18	B,C
Precipitable Water Trend[2]	1	18	B,C
Sine Day of Year	-	-	C
Cosine Day of Year	-	-	C

[1] B = binary form, C = continuous form

[2] 12-hour trend ending at time shown

Table 4. Independent verification and consistency check of LFM PoP forecasts whose equations were derived separately versus those derived together. Inconsistency definitions are given in the text.

	Derived Separately	Derived Together
12-hour period P-score	.2312	.2308
First 6-hour period P-score	.1407	.1407
Second 6-hour period P-score	.1937	.1946
Percent Inconsistencies	54.2	24.7
Percent Major Inconsistencies	10.1	.8

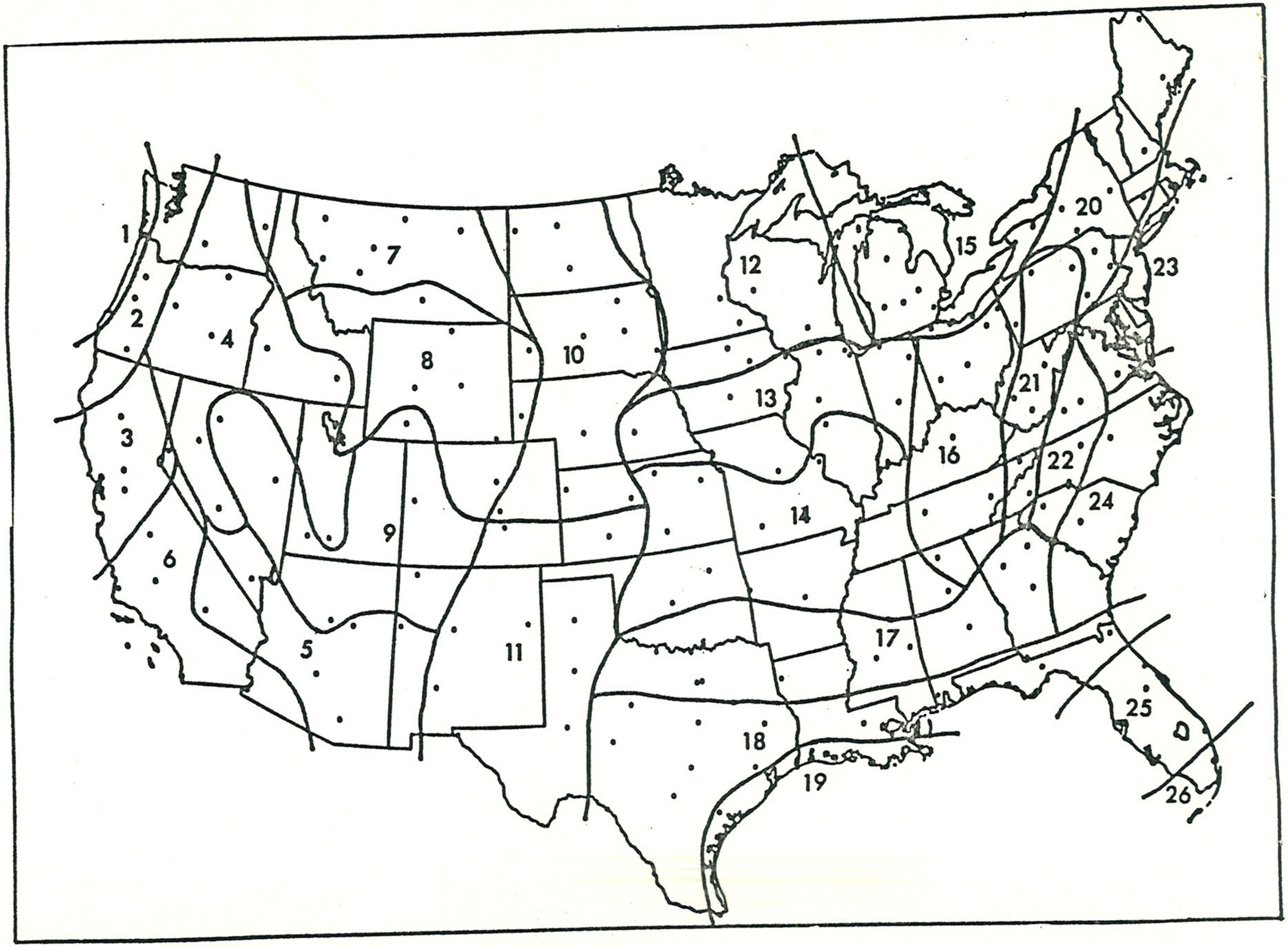

Figure 1. The 26 regions used for the operational PoP system during the 1975 summer season. The dots represent the 233 stations used in the developmental sample.

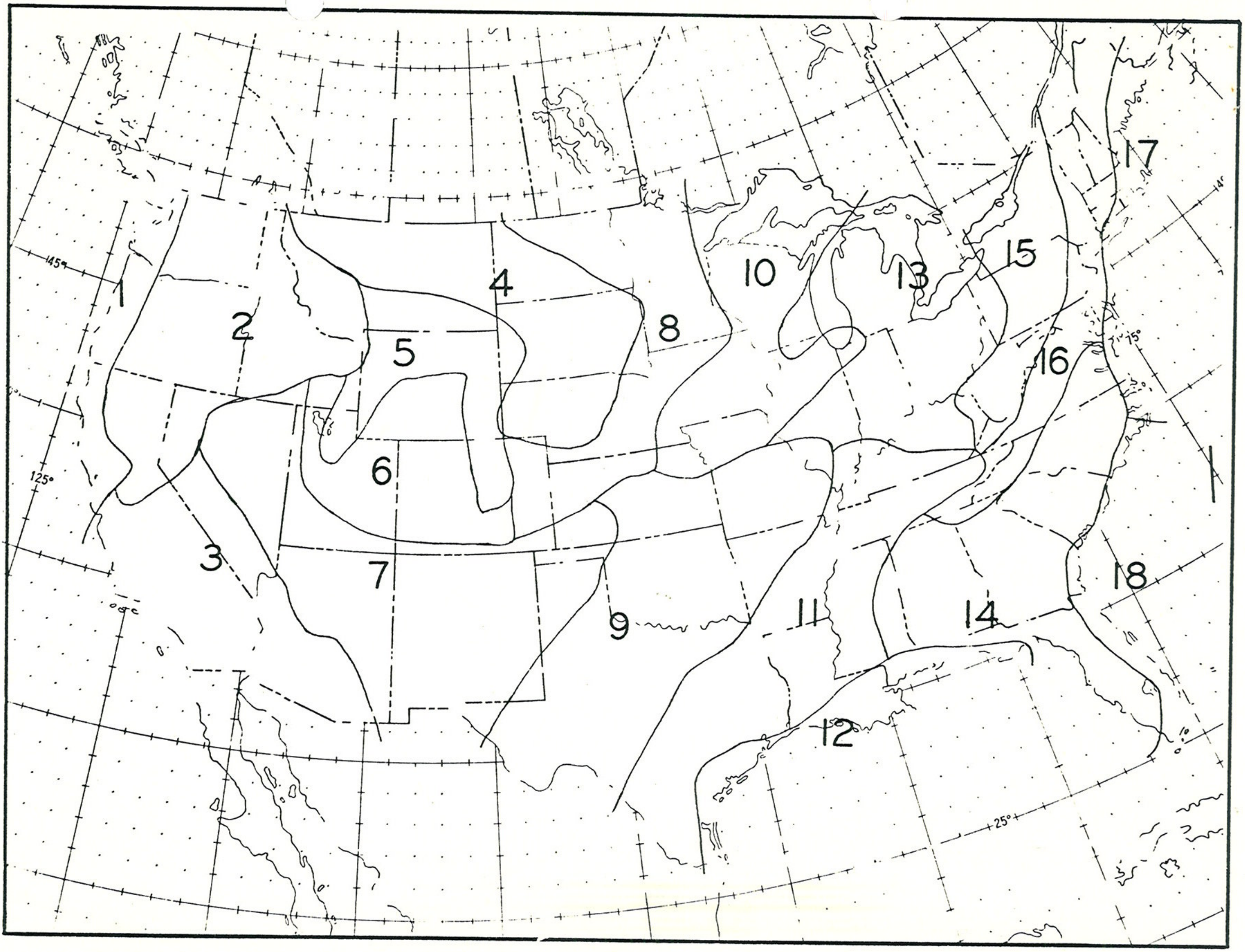

Figure 2. The 18 regions used to develop small sample PoP equations with LFM variables.

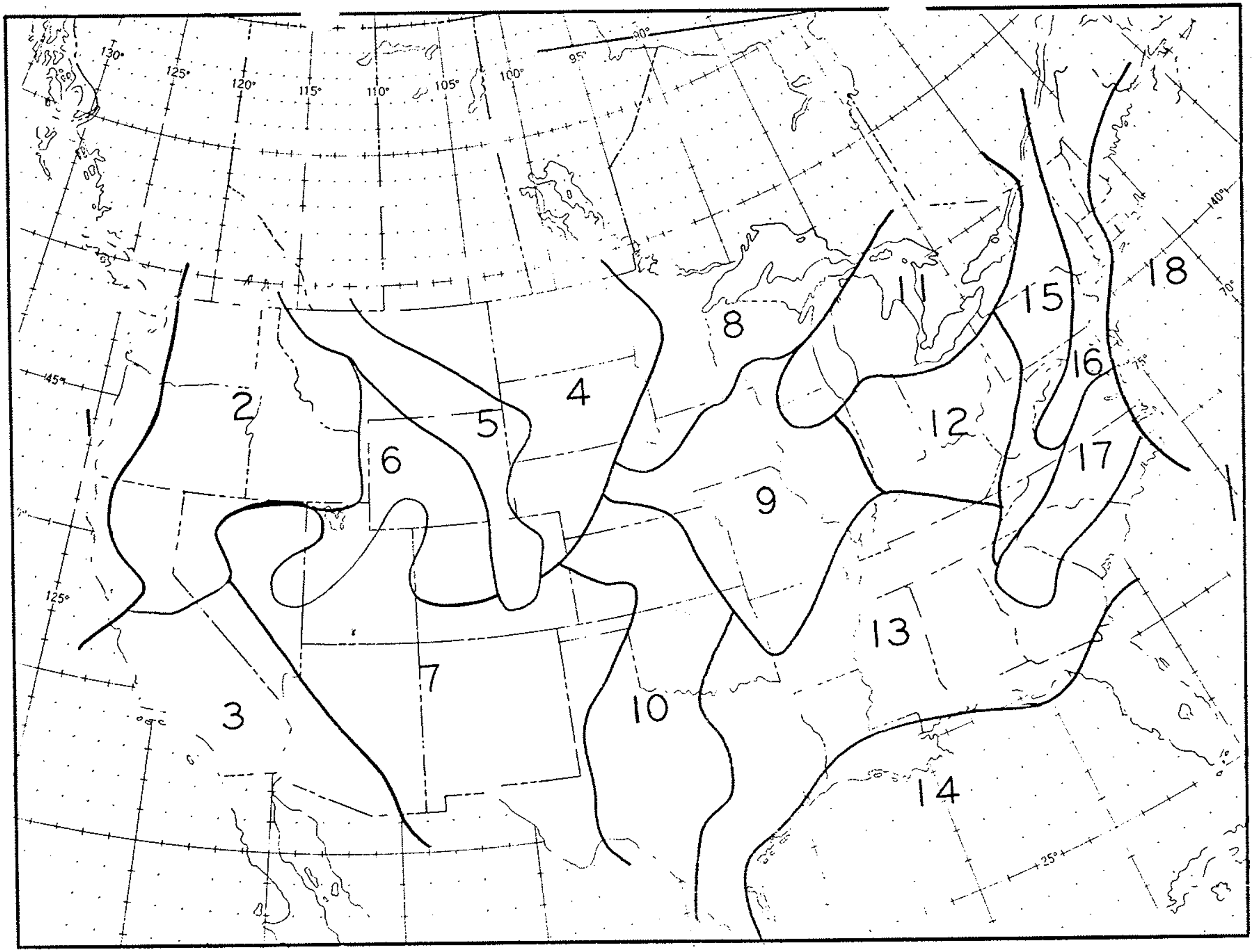

Figure 3. The 18 regions used to develop small sample PoP equations when no LFM variables are included.