~~Dr. Glahn~~
file

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

TDL Office Note 79-8

IMPROVED PREDICTION OF LIQUID PRECIPITATION TYPE

David B. Gilhousen

March 1979

IMPROVED PREDICTION OF LIQUID PRECIPITATION TYPE

David B. Gilhousen

1. INTRODUCTION

For nearly 10 years, the Techniques Development Laboratory (TDL) has been experimenting with producing public weather forecasts in worded form by computer (Glahn, 1970). With the advent of objective forecasts of all the weather elements contained in public forecasts, daily computer worded forecasts (CWF'S) (Glahn, 1978) have become a reality. One of the objective forecast products that serve as input to the CWF program is the conditional probability of liquid precipitation type. When liquid precipitation is likely, the CWF uses these forecasts to decide whether "showers", "rain", or "drizzle" would best describe the event. Admittedly, this is not the most ponderous decision that the CWF will have to make. However, if the CWF'S forecasts of liquid precipitation type are often inaccurate or inappropriate, forecasters will change many of the CWF-generated phrases. This action would defeat the CWF's time-saving benefits.

Forecasts of the conditional probability of liquid precipitation type have been produced by the MOS (Model Output Statistics) technique for several years (Carter, 1974). Last year, meteorologists at TDL noted several cases where these forecasts performed poorly. For example, they observed several cases where drizzle was inaccurately forecast at Miami, a location where drizzle is an extremely rare event. They also discovered three reasons why the performance might be poor. First, the original equations were derived for one region covering the entire conterminous United States. This was necessary because of the small sample size when the data were stratified to include only those cases where measurable liquid precipitation occurred. No attempt was made to include the effect of the local climatologies of showers, rain, or drizzle. Second, different model output is now being used by our equations than when the equations were developed. The equations were developed using predictors from the Primitive Equation (PE) Model (Shuman and Hovermale, 1968) and the Trajectory (TJ) Model (Reap, 1972). They are now being applied to output from the Limited-area Fine Mesh (LFM) model (National Weather Service, 1977) and its associated trajectory model (National Weather Service, 1978). Finally, the data sample for the summer equations consisted of only one summer season (April through September of 1973), and the winter data sample consisted of only two seasons (October through March of 1972-73 and 1973-74). In order to overcome these limitations, I decided to rederive the equations. The new equations would be based on the LFM model output, be regionalized, and be derived from a larger data sample.

2. DEVELOPMENT

The first step in the development process was determining the regions. This was done semi-objectively from two analyses. These analyses were the relative frequency of drizzle and showers given that precipitation occurred. Separate analyses were performed for both the winter and summer seasons. The data on which the analyses were based were 233 station relative frequencies for 0600 and 1800 GMT, for both winter and summer seasons. Figures 1 and 2 show the regions which we determined for the winter and summer seasons, respectively.

In Figure 1, note the small number of regions in the Interior U.S. Snow is the most common form of precipitation during the winter months at these locations, and more stations must be grouped together to form an adequate sample of liquid precipitation cases. Across the southern tier of states and near both oceans, liquid precipitation is more common than frozen, and a greater number of regions could be determined to reflect changes in event relative frequency. To illustrate these changes, let's consider two stations at roughly the same latitude, Tallahassee, Fla. and Austin, Texas. At Tallahassee, the relative frequency of drizzle is 7% and that of showers is 27%. At Austin, the relative frequencies are 35% and 14%, respectively. Since the central and western Gulf is a favorable area for cyclogenesis, Texas is more likely to observe drizzle and rain that falls from stratiform clouds. From this it follows that we should not group data from Tallahassee and Austin together.

After the regions were determined, generalized operator equations were developed for each region. For test purposes, we chose to concentrate on making 18-h, winter season forecasts from the 0000 GMT model run. We had four seasons (1972-73 through 1975-76) of developmental data. Again, the sample size was reduced since only those cases where measurable liquid precipitation occurred were included in the sample. The sample size ranged from 2184 cases in region 6 to 183 cases in region 9.

Table 1 contains the various forecast fields from the LFM model and surface observations that were screened as potential predictors. The three equations for drizzle, rain, and showers contained the same predictors with different regression coefficients, since it is desirable that for any given station, the three probability forecasts add to unity. Six-, eight-, ten-, and twelve-predictor equations were produced for test purposes. Table 2 shows the most important predictors based on their frequency and order of selection in the 10-term equations. The predictors selected are well related to the atmospheric processes that produce the different precipitation forms. Notice that Layer 2 (720-490 mb) relative humidity and boundary layer relative humidity are ranked higher than mean relative humidity. This suggests that the vertical distribution of humidity is more important than the overall total humidity in the column. The regression coefficients for the Layer 2 relative humidity predictor were always negative for the drizzle equation. This result agrees with Mason and Howorth (1952) who showed that drizzle is the most common form of precipitation when the cloud bases are below 2,000 feet and cloud thicknesses less than 6,000 feet.

Table 2 also shows that the temperature difference between 1000 and 850 mb is an important predictor while the two stability indicies are infrequently selected. The two stability indicies, the K Index and

Total Total Index, are based on 850-, 700-, and 500-mb temperatures and dew points. Therefore, it would appear that the temperature structure in the lower 6,000 feet of the atmosphere is more important than the temperature structure above 6,000 feet in determining precipitation type. This result agrees with Hughes (1977) who found a low-level inversion below 850 mb in all cases where drizzle was observed. Also, winter showers are often associated with low level instability below a subsidence inversion (Anderson et. al., 1969).

3. VERIFICATION

We compared forecasts from the new equations with those produced by the current operational equations on winter 1976-77 data. For this comparison, data from all 233 stations were grouped together to give us 2085 precipitation cases. To verify the probability forecasts, we calculated the P-score as defined by Brier (1950) for the three predictand categories. We transformed the probability forecasts to categorical forecasts by choosing the category with the highest probability. From the categorical forecasts, we computed percent correct, the skill score[1], and the bias[2] of the shower category.

Table 3 shows these verification scores for the various forecast systems. Of the new LFM equations, the eight-term equations produced forecasts with the lowest P-score and the highest percent correct and skill score. The system with ten-term equations was a close second. We concluded from this that eight terms is about the proper number to use in these equations. When we compared forecasts from the eight predictor equation with the operational PE and TJ forecasts, the new system produced considerably better scores. The improvement in P-score was 22.2% while the percent correct improved by 8.7%. Table 3 also shows that both systems underforecast showers. This may not be a serious deficiency since we have no idea what bias forecasters exhibit.

We realize that a more valid approach would have been to conduct two separate tests. One test would determine the number of terms that the new equations should have. On a different data set, the other test would compare forecasts from the new equations selected in the first test with forecasts from the old equations. We decided not to take this approach because of the limited amount of data. In view of the clear superiority of the new equations, this approach would probably not have yielded any different results.

---

[1] The skill score used throughout this paper is the Heidke skill score (Panofsky and Brier, 1965).

[2] The bias is the number of forecasts of an event divided by the number of observed events.

4. OPERATIONAL ASPECTS

Probability of liquid precipitation type equations were derived for 18-, 30-, and 42-h projections from both 0000 and 1200 GMT model runs. The developmental sample for the 30-h projection consisted of data from the winter seasons of 1975-76, 1976-77, and 1977-78. The sample for the 42-h projection consisted of data from only the last two winter seasons. The regions for the 18-h and 30-h projections were the same ones used in the experiment. However, because of the small amount of data for the 42-h projection, data in regions 9 and 10 were grouped together. Since surface observations were never important predictors in any of our equations, operational equations were derived without them.

Summer season equations were derived in a similar manner. Figure 2 shows the regions that were determined after we examined the relative frequencies of drizzle and showers during the summer. The developmental sample for the 18-h projection consisted of data from the summers of 1973 through 1977. The sample for the 30-h projection consisted of data from the summers of 1975 through 1977 while the 42-h sample consisted of data from the last two summers. No verification was conducted for the summer season forecasts.

On an operational basis, these forecasts are available only in the CWF matrix. There, the conditional probabilities of showers, drizzle, and rain are given in the rows labeled "R SHR (L)", "DRZL (L)", and "RAIN (L)" respectively.

ACKNOWLEDGEMENTS

I'm very thankful to George Lessons of TDL for programming support.

REFERENCES

Anderson, R. K., J. P. Ashman, F. Bitter, G. R. Farr, E. W., Ferguson, V. J. Oliver, A. H. Smith, 1969: Application of meteorological satellite data in analysis and forecasting. ESSA Technical Report NESC-51, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 182 pp.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. Mon. Wea. Rev., 78, 1-3.

Carter, G. M., 1974: Automated prediction of thunderstorms, drizzle, rain, and showers. TDL Office Note 74-7, National Weather Service, NOAA, U. S. Department of Commerce, 8 pp.

Glahn, H. R., 1970: Computer worded public weather forecasts. Bull. Amer. Meteor. Soc., 51, 1126-1131.

_____, 1978: Computer worded public weather forecasts. NOAA Technical Memorandum NWS TDL-67, National Oceanic and Atmospheric Administration, U. S. Department of Commerce, 24 pp.

Hughes, L. A., 1977: A program of chart analysis. NOAA Technical Memorandum NWS CR-63, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 56 pp.

Mason, B. J. and B. P. Howorth, 1952: Some characteristics of stratiform clouds over Northern Ireland in relation to their precipitation. Quart. J. Roy. Meteor. Soc., 78, 226-231.

National Weather Service, 1977: High resolution LFM (LFM-II). NWS Technical Procedures Bulletin No. 206, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 6 pp.

_____, 1978: The trajectory model. NWS Technical Procedures Bulletin No. 225, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 13 pp.

Panofsky, H. A. and G. W. Brier, 1965: Some applications of statistics to meteorology. Pennsylvania State University, University Park, Pa., 224 pp.

Reap, R. M., 1972: An operational three-dimensional trajectory model. J. Appl. Meteor., 11, 1193-1202.

Shuman, F. G., and J. B. Hovermale, 1968: An operational six-layer primitive equation model. J. Appl. Meteor., 7, 525-547.

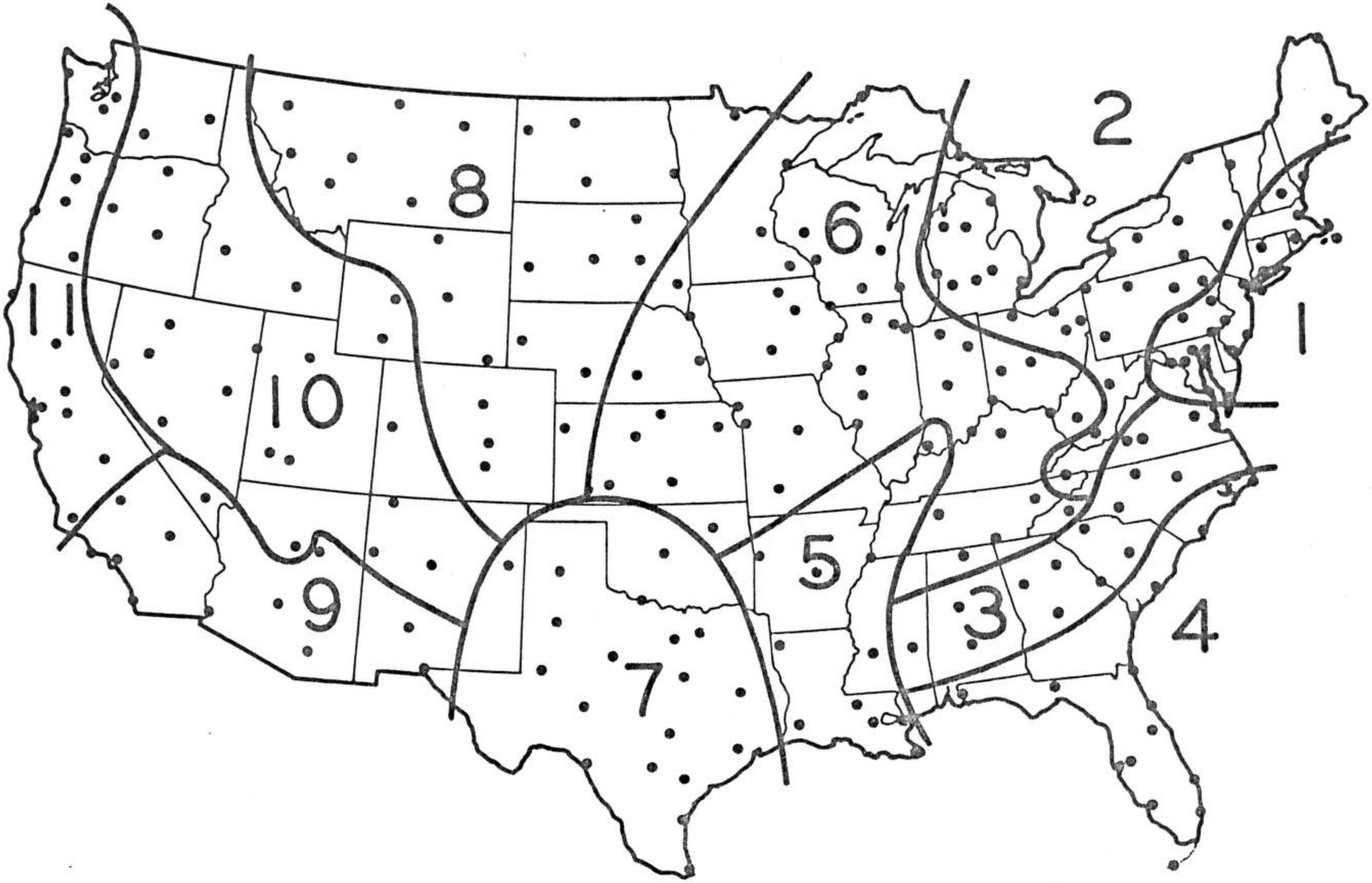

Figure 1. The 11 regions determined for the winter season.

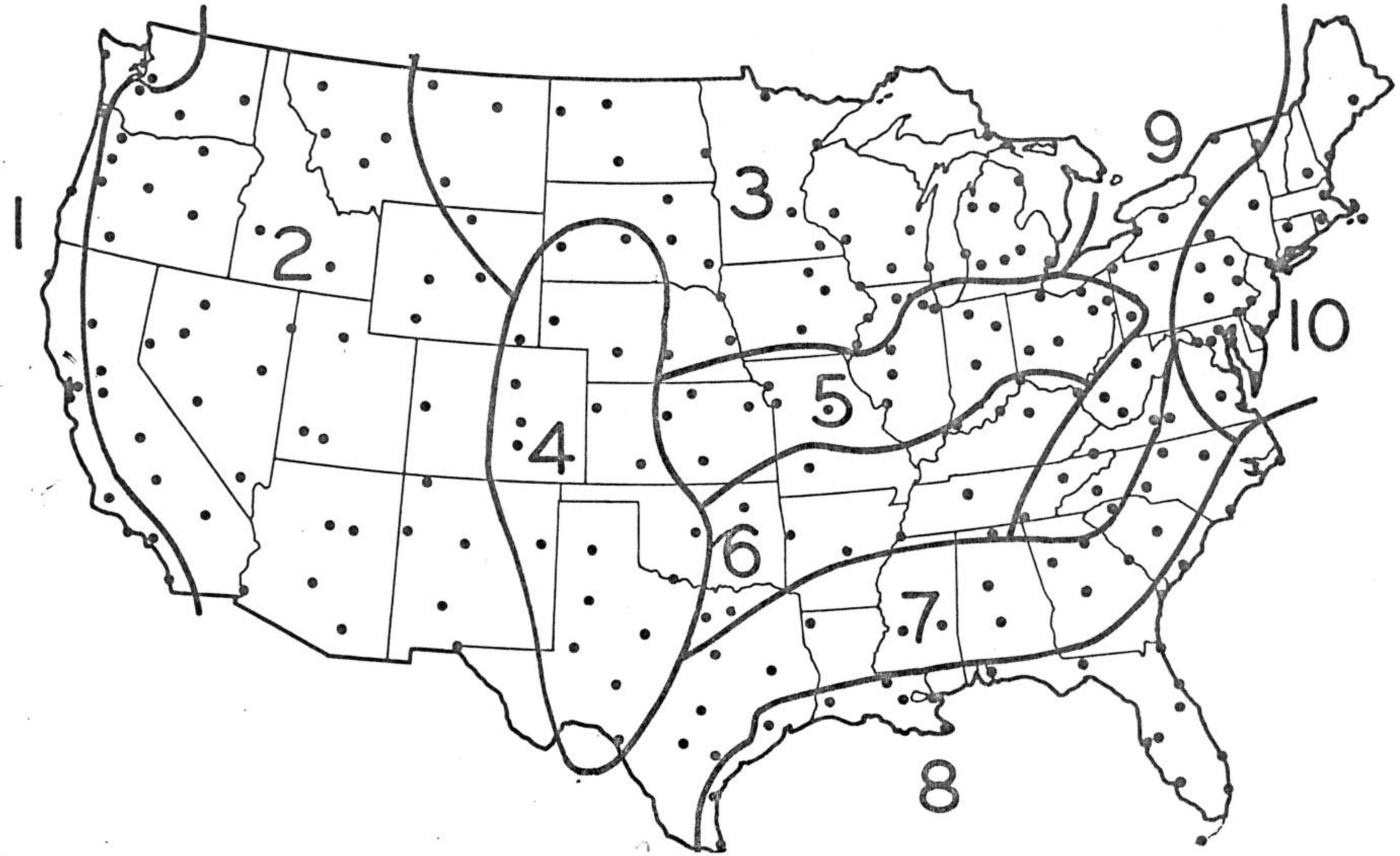

Figure 2. The 10 regions determined for the summer season.

Table 1. Surface observations and LFM output used as potential predictors for 18-h forecasts of the conditional probability of liquid precipitation type. Certain predictors have been space-smoothed by 5 or 9 points to eliminate small scale noise. The time is the number of hours after 0000 GMT or 1200 GMT that the surface observation is made or that the LFM forecast is valid.

Field	Smoothing (Points)	Time (Hours)	Form[1]
Observed Ceiling	--	3	B
Observed Liquid Precipitation Type	--	3	B
Observed U,V Wind	--	3	B
Sine Day of Year	--	--	C
Cosine Day of Year	--	--	C
Mean Rel. Humidity	1,5	12,18,24	B,C
Boundary Layer Rel. Humidity	1,5	12,18	B,C
Layer 2 Rel. Humidity	1,5	12,18,24	B,C
Precipitation Amount	1,5	12,18,24	B,C
850-mb Vertical Velocities	5	12,24	B,C
1000-mb Temperature	1	12,24	B,C
850-mb Temperature	1	12,24	B,C
(850-mb Temp.)-(1000-mb Temp.)	5	12,24	B,C
Boundary Layer Pot. Temperature	1,5	12,18,24	B,C
850-mb U,V Wind	5	12,18,24	B,C
Boundary Layer U,V,S Wind	5	12,18,24	B,C
Total Totals Index	5	12,24	B,C
K Index	5	12,24	B,C

[1] B=Binary , C=Continous

Table 2. The ten most important predictors in the 18-h liquid precipitation type equations for the winter season. The predictors were ranked on the basis of a weighted scoring system that emphasized both the frequency and order of selection of individual predictors. All predictors were from the LFM model.

Order	Predictor
1	Layer 2 Rel. Humidity
2	(850-mb Temp.) - (1000-mb Temp.)
3	Boundary Layer Humidity
4	Mean Rel. Humidity
5	Boundary Layer Pot. Temperature
6	1000-mb Temperature
7	Boundary Layer Wind Speed
8	Boundary Layer U Wind
9	Precipitation Amount
10	Cosine Day of Year

Table 3. Verification scores from several forecast systems of the probability of liquid precipitation type. These scores are for 18-h forecasts from the 0000 GMT model run. The verification was performed on 2085 precipitation cases during the period October 1976 - March 1977.

Forecast System	Verification Scores			
	Overall P-Score	Percent Correct	Skill Score	Bias-Shower-Category
Old PE and TJ Equations	.5525	.5986	.5299	.6900
New LFM Equations				
6 term	.4421	.6646	.5817	.6854
8 term	.4299	.6851	.6085	.6945
10 term	.4334	.6824	.6084	.6951
12 term	.4385	.6823	.6072	.6952