

STATISTICAL QUANTITATIVE PRECIPITATION FORECASTS  
BASED ON THE MEDIUM RANGE FORECAST MODEL

[Mark A. Shirey](#) and [Mary C. Erickson](#)

Techniques Development Laboratory  
Office of Systems Development  
National Weather Service, NOAA  
Camp Springs, Maryland

**1. INTRODUCTION**

Accurate Probabilistic Quantitative Precipitation Forecast (PQPF) guidance has long been one of the operational meteorologist's most important and desired tools, given PQPF's direct relationship to flash flooding, river flooding, and snowfall accumulation. In fact, the development of improved PQPF's has recently been targeted as a top research priority of the National Weather Service (NWS) and the United States Weather Research Program. The Techniques Development Laboratory (TDL) currently provides statistical PQPF guidance based on output from the Nested Grid Model (NGM; Hoke et al. 1989). This system was developed by using the Model Output Statistics (MOS) technique (Glahn and Lowry 1972) and, like other MOS systems, has been shown to provide valuable information to the forecaster (Antolik 1995). In addition, Probability of Precipitation (PoP, specifically defined as the probability of 0.01 inches or more of liquid-equivalent precipitation) guidance is available based on output from the Medium Range Forecast (MRF) and Aviation (AVN) runs of the Global Spectral Model (Kanamitsu 1989). With the modernization of the NWS, there is a call for increased lead times in forecasts, hence an elevated demand for more detailed products in the medium-range. To help meet this need, we are beginning to develop and test statistical PoP/QPF guidance by applying the MOS technique to output from the MRF.

In this paper, we discuss our development thus far on the new MRF-based MOS PoP/QPF system. In particular, we describe the definition of the proposed weather elements, the statistical techniques used to develop test forecast equations, and the procedures for verification. Preliminary results from test forecasts generated on independent data are also presented. PoP forecasts from current test equations are compared against the operational MRF-based MOS PoPs, and PQPF forecasts are compared against a forecast based on climate.

**2. PROPOSED MRF MOS PoP/QPF SYSTEM**

The new MRF MOS PoP/QPF system will include three weather elements: PoP, PQPF, and categorical QPF. Guidance for all three weather elements will encompass 12- and 24-h forecast periods, ending at both 0000 and 1200 UTC. Therefore, both 12- and 24-h forecasts will be available every 12 hours (Fig. 1). The PoP will be available for projections valid 24 to 192 hours after 0000 UTC, while the more detailed categorical QPF and PQPF will be provided out 156 hours (if skill is still attainable). When fully developed, guidance will be available for approximately 1000 stations, encompassing all 50 U.S. states as well as Puerto Rico.

Projection (hrs)	24	36	48	60	72	84	96	108	120	132	144	156	168	180	192	
24-h fcst period ending at 00z			X		X		X		X		X		P		P	X=PoP/QPF
24-h fcst period ending at 12z		X		X		X		X		X		X		P		P=PoP only
12-h fcst period	X	X	X	X	X	X	X	X	X	X	X	X	P	P	P	

**Figure 1. Forecast projections for the new MRF-based PoP/QPF product.**

Probabilistic forecasts will be produced for several cutoffs, including the accumulation of 0.01, 0.10, 0.25, 0.50, 1.00, and 2.00 inches of precipitation. A forecast of the accumulation of 3.00 inches of precipitation will also be available for 24-h periods. It is important to note that the probabilities produced at these cutoffs will be cumulative. For instance, a report of 0.65 inches of precipitation will qualify as an occurrence of four separate events (i.e., accumulation 0.01 inches, 0.10 inches, 0.25 inches, and 0.50 inches). In other words, a PQPF value can be defined as the probability that precipitation accumulation will equal or exceed its corresponding cutoff value.

From these probabilities, a categorical QPF will be computed. This process will be similar to the one used currently in producing the NGM QPF guidance, and details concerning the process may be found in Antolik (1998). The following QPF categories will be available:

No measurable precipitation  
0.01 - 0.09 inches  
0.10 - 0.24 inches  
0.25 - 0.49 inches  
0.50 - 0.99 inches  
1.00 - 1.99 inches  
2.00 inches or more (12-h forecasts)  
2.00 - 2.99 inches (24-h forecasts)  
3.00 inches or more (24-h forecasts)

This new system will increase the amount of information available to the forecaster in the medium-range. In the current MRF MOS system, PoP is the only precipitation forecast created, and it is available for approximately 200 stations. The current 24-h PoP ends at 0000 UTC only, and is determined objectively by combining the two relevant 12-h PoPs in the period, whereas all PoP/QPF weather elements covering a 24-h period in the new system will be developed from 24-h equations. For more information on the entire new MRF MOS system, please refer to Erickson and Carroll (1999).

### **3. CURRENT DEVELOPMENTAL TECHNIQUES**

#### **3.1 Equation Development**

The MOS technique is the statistical procedure used in this development. Least squares linear regression is used to derive statistical equations that best explain the relationship between the observed QPF amount (predictand) and various forecast variables (predictors). To decrease the chance of over-fitting in the testing done thus far, a maximum of 10 predictors was allowed in each equation set. A predictor was not chosen unless it added at least 0.5% to the total reduction of variance. Equations for PoP and PQPF were simultaneously developed in order to enhance the meteorological consistency among the forecasts produced by the equations. The final PoP/QPF system will have two sets of equations to account for seasonal change; warm season equations will be developed on April-September data, and cool season on October-March data.

Two different data sets were available for TDL's use in this development, spanning a period from January 1992 to March 1999. The first contained MRF model data from 1992 to 1996 and was obtained from NCEP's reanalysis project (Kalnay et al. 1996). Forecast data from 12-192 hours, valid every 12 hours, were available every 5<sup>th</sup> day (making it equivalent to one full year of daily archived data), and were archived on a 190.5-km (at 60N) grid. The second data set contained 1997 to 1999 MRF model data which were saved daily on a 95.25-km (at 60N) grid.

In general, developing with a large sample size helps produce robust equations which perform well on independent data. Because of this, we wanted to develop using both data sets simultaneously, but were not sure how the different characteristics (especially the different grid spacing) would affect results. After performing many tests, we found that applying to each data set spatial filters which resulted in the same effective smoothing provided the best results. During development, the predictor variables were smoothed, then interpolated to station locations through bi-linear interpolation. In the future, we may investigate applying additional smoothers to the model fields at the longer projections.

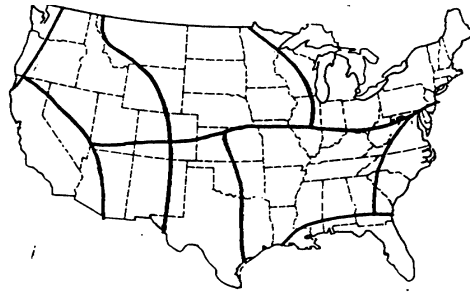
#### **3.2 Predictors**

The potential predictors offered to the regression were MRF forecasts, station climatic variables, and station geographic variables. The MRF predictors included relative humidities, precipitation amount, u- and v- wind components, moisture convergence, vertical velocities, k index, and relative vorticity. A number of these were presented in grid-binary form (Jenseniuss 1992) which resulted in predictor values ranging from 0 to 1. This technique provides a smoother transition, both spatially and temporally, between the extremes of the predictor. Various time averages were also applied to the predictors to capture the nature of the predictors over the 24-h period.

The climatic relative frequencies included in the development process were produced by using 18 years of precipitation observations (1980-1998). Climatic normals of all the cutoffs were calculated for each of the 12 months, and offered as predictors in the regression. The relative frequencies help add climatic value to the regression, hence are especially needed at the later projections. The relative frequencies also provide individual station information when regional equations are developed and used.

### 3.3 Regions

The stations were combined into regions with specific equations developed for each region since single station development for precipitation variables does not provide enough events for adequate and stable equations. Stations with similar relative frequencies and climatology were grouped together, resulting in the 10 regions shown (the same regions currently in the operational guidance) in Fig. 2. In an effort to increase sample size (especially for the higher amount categories) we may test in the future equations developed for larger regions.



**Figure 2. Regions used for development of the warm season PoP/PQPF equations.**

## 4. PRELIMINARY TEST RESULTS

Since the time and computer resources needed to evaluate forecasts at over 1000 stations is significant, we decided to develop test equations for a representative smaller number of stations. This allowed us to develop several sets of equations in a short period of time, and permitted us to test the impact of different smoothings and predictor combinations. The following is a snapshot of the work to date, with more information to be provided at the conference.

### 4.1 Test Development and Verification

Our test sample consisted of data from roughly 250 stations that representatively covered the contiguous U.S. Warm season (April - September) equations were developed for the 24-h PoP and PQPF's 0.25 and 1.00 inches at a sampling of the proposed final projections. Equations valid for periods ending at both 0000 UTC (48-h, 72-h, 96-h, 120-h, and 144-h projections) and 1200 UTC (108-h, 132-h, 156-h, and 180-h projections) were tested. During the equation development, the 1st through 15th of each month in 1998 was held out to be used as independent data.

These test equations were then used to produce forecasts on independent data. These forecasts were verified against forecasts based on climate (observed relative frequencies) and the operational MRF MOS (PoP only). Brier scores (mean squared errors) were calculated for all three systems and compared. Note that the Brier score ranges from 0 to 1 with lower values indicating more accurate forecasts. Reliability tables were also used to help verify the forecasts. A reliability table compares the mean forecast in any given forecast interval (e.g., 25-35%), with the observed relative frequency of precipitation that actually occurred in that same interval. In large samples, the mean forecast will usually be close to the middle of the forecast interval (near 30% in the above example). The closer the observed relative frequency is to the mean forecast, the more reliable the forecast.

### 4.2 PoP Results

As one can see from Fig. 3, both the new MOS PoP and operational MOS PoP were skillful (compared to climate) in terms of the Brier score for all projections tested. While the new PoP was more accurate than the operational at early projections (48-h and 72-h), the new PoP failed to improve at the later projections (96-h, 120-h and 144-h). One explanation for this difference may be related to the reliability of the two systems. While both systems were generally reliable, the new PoP forecasts (Fig. 4a) showed much greater variability in its forecasts than that of the operational (Fig. 4b). A large difference between the two PoPs in the number of cases forecast can be seen at both ends of the probability spectrum. While the operational MOS has 433 forecasts below 5% and 462 above 55%, the new MOS contains 829 below 5% and 904 above 55%. Although the MOS technique causes forecasts to trend towards climate with increasing projection, these results show that the operational MOS trends toward climate more strongly than the new MOS. Since it is difficult to reliably deviate from the climatic mean at longer-range projections, on this independent sample the operational PoP verifies better. One can see from the reliability figure that the new PoP shows a slight dry bias at the low end and a wet bias at the high end, and, thus, seems over-confident on this sample.

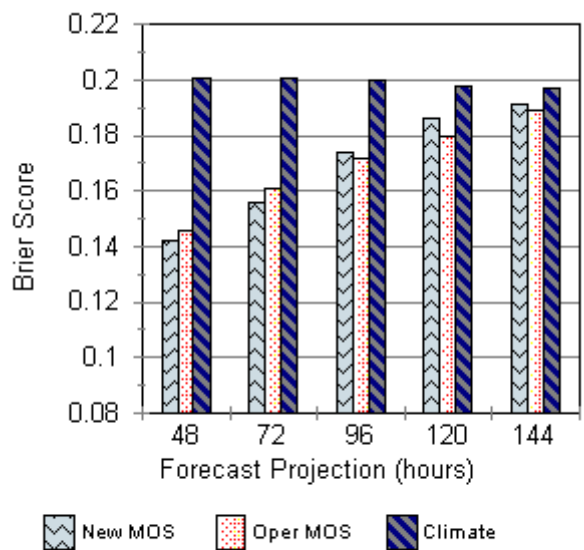


Figure 3. Brier scores for the new MOS PoP, the operational MOS PoP, and climate for the 1998 warm season test.

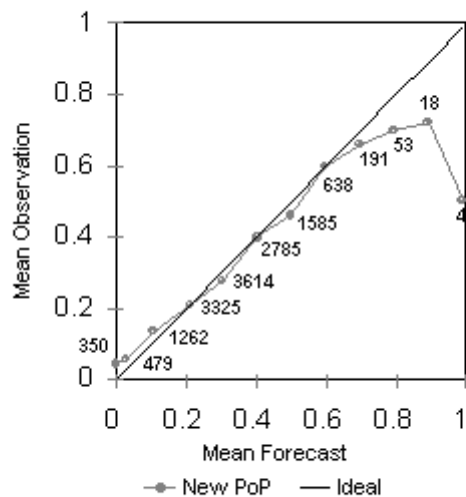


Figure 4a. Reliability of the 24-h new PoP forecasts valid at 144 hours after 0000 UTC.

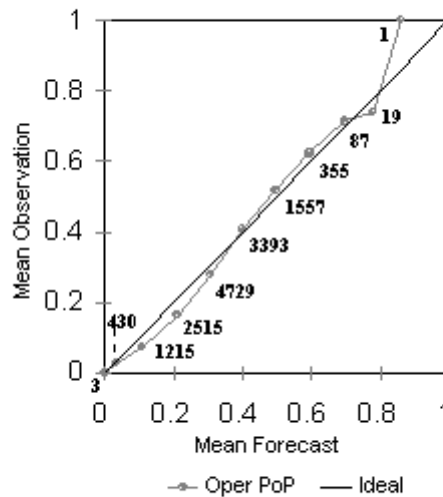


Figure 4b. Reliability of the 24-h operational PoP forecasts valid at 144 hours after 0000 UTC.

Forecasters often say that higher variability in a forecast system is more helpful, providing an increased amount of information and a stronger signal of when to deviate from the mean. Ideally, the best solution is one that weighs the significance of both the mean squared error and increased variability, balancing out the pros and cons of each. While we can not yet explain these differences in variability, we suspect there are several contributing factors. For one, there have been three major changes to the MRF over the course of the current developmental period. Furthermore, the two systems were developed on very different seasons (1988-1994 for the operational vs. 1992-1998 for the new). Finally, the new MOS was developed from MRF data on much finer grids (95.25 km and 190.5 km) than the operational MOS (381 km). Statistically speaking, we do not think that the increased variability is due to overfitting the dependent sample since most equation sets average four to five predictors, and the number of predictors offered was comparable to past developments. We will continue to investigate the cause of the increased variability.

#### 4.3 PQPF Results

Probabilities of accumulating 0.25 and 1.00 inches of precipitation were also verified. The PQPF for 0.25 inches (Fig. 5) performed favorably (especially in the first 3 periods), improving over climate in six out of the seven projections tested. The PQPF for 1.00 inch (Fig. 6), however, did not verify as well, narrowly improving over climate only at 48-h and 72-h projections. Since there are fewer cases of heavier (and rarer) events in the dependent sample and since observed precipitation amounts of an inch or more in the summer are usually related to convection, we may not have the tools yet to forecast mesoscale convection very accurately in the medium-range. Cool season results, which will be presented at the conference, should yield better verification scores given that heavy precipitation is more correlated to large-scale synoptic features.

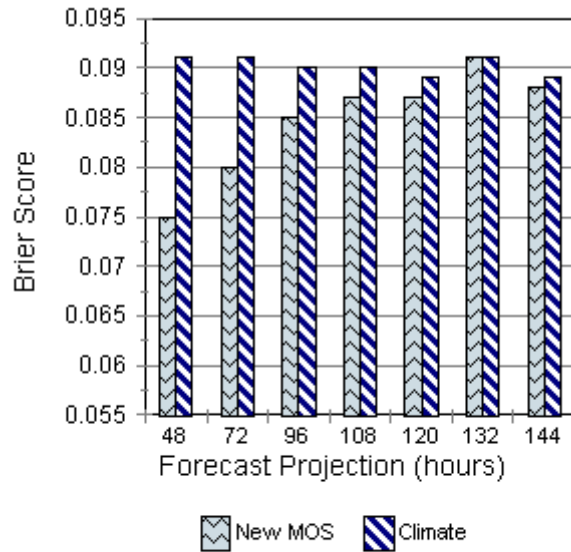


Figure 5. Brier scores for MOS and climate forecasts of  $\geq 0.25$  inches for the warm season test.

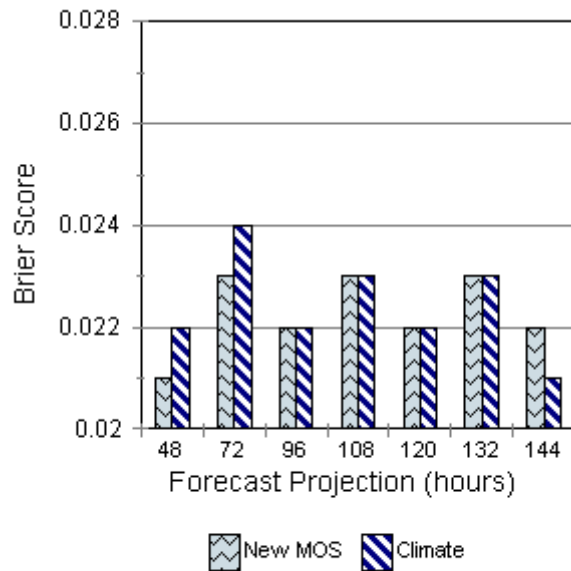


Figure 6. Brier scores for MOS and climate forecasts of  $\geq 1.00$  inches for the warm season test.

## 5. CONCLUSION

In this paper, we have discussed the proposed new MRF MOS QPF system, and provided preliminary test results. We've shown that some success can be obtained forecasting QPF in the mediumrange, but how much and at which projections are yet to be determined. Over the next 6 months, we will resolve these issues, conduct additional tests on independent data, and implement new forecast equations. At that time, further documentation of our development efforts and verification results will be provided to forecasters.

## 6. REFERENCES

Antolik, M. S., 1995: NGM-based quantitative precipitation forecast guidance: Performance tests and practical applications. Preprints 14<sup>th</sup> Conference on Weather Analysis and Forecasting, Dallas, Amer. Meteor. Soc., 182-187.

\_\_\_\_\_, 1998: NGM-based statistical quantitative precipitation forecast guidance for the contiguous United States and Alaska. NWS Technical Procedures Bulletin No. 451, National Oceanic and Atmospheric Administration, U.S. Department of Commerce.

Erickson, M. C. and K. L. Carroll, 1999: Updated MRF-based MOS guidance: Another step in the evolution of objective medium-range forecasts. Preprints 17<sup>th</sup> Conference on Weather Analysis and Forecasting, Denver, Amer. Meteor. Soc., (this volume, 5.1).

Glahn, H. R., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. J. Appl. Meteor., 11, 1203-1211.

Hoke, J. E., N. A. Phillips, G. J. DiMego, J. J. Tucillo, and J. G. Sela, 1989: The regional analysis and forecast system of the National Meteorological Center. Wea. Forecasting., 4, 323-334.

Jensenius, J. S., Jr., 1992: The use of grid-binary variables as predictors for statistical weather forecasting. Preprints 12<sup>th</sup> Conference on Probability and Statistics in the Atmospheric Sciences, Toronto, Amer. Meteor. Soc., 225-230.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Roelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-year re-analysis project. Bull. Amer. Meteor. Soc., 77, 437-471.

Kanamitsu, M., 1989: Description of the NMC global data assimilation and forecast system. Wea. Forecasting., 4, 335-342.