JP1.5 USING SKILL SCORES TO ASSIST ASSESSMENT OF THE "MAN-MOS-MET MIX" IN PROBABILITY OF PRECIPITATION (POP) FORECASTING

Preston W. Leftwich, Jr.*
NOAA, National Weather Service
Central Region Headquarters
Scientific Services Division
Kansas City, Missouri

1. INTRODUCTION

Probability of Precipitation (PoP) is an integral part of public forecasts issued by the National Weather Service (NWS). Model Output Statistics (MOS) [Glahn and Lowry 1972] have provided primary objective guidance for PoP forecasts. Evolving operational numerical models have continued to provide improved guidance to forecasters. These tools are utilized along with meteorological diagnosis/analysis to support forecast decisions by forecasters. As such, the overall PoP forecast process is a mix of forecaster (Manual) decision, MOS guidance, and other Meteorology derived from various sources (e.g., direct model output). That is, the forecast is a "Man-MOS-Met Mix." Analysis of each component of this mix is important in efforts to improve PoP forecasts provided to the public. Additionally, clarification of roles of these components will be valuable in establishment of a modernized forecast process executed via the Interactive Forecast Preparation System (IFPS).

Performance measures are used to track trends in forecast accuracy. For PoP forecasts, a Brier score, defined by the NWS as one half the P-score of Brier (1950) is the commonly used statistic. Also, a MOS-based skill score (Wilks 1995) serves as a performance measure via comparisons of PoPs issued in forecasts from local NWS offices to those provided by operational MOS guidance, currently that from the operational Nested Grid Model (NGM). Further, particular focus is placed on situations when MOS PoPs and local (LCL) PoPs differ by at least 20%.

Efforts are underway to (1) assist NWS forecasters in improving PoPs issued at Weather Forecast Offices (WFOs) within the Central Region (CR) and (2) provide basic direction for localization of the IFPS for future PoP forecasts.

* *Corresponding author address:* Preston W. Leftwich, Jr., NOAA/NWS, Room 1836, Federal Building, 601 East 12th Street, Kansas City, Missouri 64106; e-mail: Preston.Leftwich@noaa.gov.

In support of these efforts, a procedure has been developed to utilize PoP skill scores to identify cases targeted for review of both the meteorological situation and the related forecast process carried out at the WFO. First, skill scores for forecast situations when MOS guidance and LCL PoPs differed by at least 20% were acquired. Then, these skill scores were stratified by annual, seasonal and quarterly periods of the fiscal year to identify periods during which LCL PoPs within the CR reflected either increased or decreased skill relative to MOS. Reviews of the meteorological situation and the related forecast process for these cases identify sources of error or improvement that can provide a foundation for future forecasts.

2. BACKGROUND

Statistical guidance on predictions of numerous sensible weather elements is provided via MOS (Glahn and Lowry 1972). Over the years, such statistical guidance has been associated with several operational numerical models. Currently, MOS is developed from the NGM (Su 1993), Aviation Model (AVN) [Dallavalle and Erickson 2000] and the Medium Range Forecast Model (MRF) [Erickson and Dallavalle 2000].

Current performance measures reflect percent improvement relative to NGM MOS. In development of NGM MOS, linear regression techniques were used to select predictors of PoP from NGM forecasts, station geographical variables and trigonometric functions for the day of the year. No observed data were used. Forecast equations exist for 6-h and 12-h periods for projections out to 60 h. Details of the development procedure, predictors, and routine messages are given by Su (1993). For demonstration of the analysis procedure in the current study, 24-h forecasts of 12-h PoPs are used.

Currently, NGM MOS is considered a standard for measurement of skill, and forms the basis for a skill score, SS_{MOS}, given by

$$SS_{MOS} = [(BS_{MOS} - BS_{LCL}) / BS_{MOS}] \times 100 \qquad (1)$$

where BS_{MOS} and BS_{LCL} are Brier scores for MOS and LCL PoP forecasts, respectively. This statistic expresses the percent improvement over MOS guidance provided by local forecasts. A positive/negative value indicates local forecasts have produced a lower/higher Brier score for the sample of forecasts considered.

3. USE OF PERFORMANCE MEASURES

Skill scores, as described in Section 2, are computed routinely for PoP forecasts issued by the NWS. This study will use information for PoP forecasts issued by field offices within the Central Region of the NWS during fiscal year 2000 (FY00), i.e., October 1999 through September 2000. The goal is to identify individual forecast situations where LCL PoP and MOS PoP differ in such a way that the forecaster made a clear decision to have a forecast PoP different from MOS guidance. These cases will then be candidates for detailed review of both the meteorological situation and the forecast process. Because the Skill Scores used as performance measures address cases when the two PoPs differed by at least 20%, these statistics provide an appropriate starting point. All further discussions will address cases having PoPs differing by at least 20%.

Stratification of PoP forecast performance measures for FY00 was used to isolate periods of greatest/least percent improvement over MOS. These periods were (1) the entire year (October 1999 - September 2000), (2) the cool season, H1, (October 1999 - March 2000), (3) the warm season, H2, (April 2000 - September 2000) and (4) four quarters (October - December, January - March, April - June, July - September) denoted as Q1,Q2, Q3 and Q4, respectively. As seen in Fig. 1, least improvement over MOS was evident during Q2, while the greatest improvement was evident during Q4. As a demonstration of the procedure for isolating cases for further review, focus is placed on Q2.

Next, skill scores from Q2 were stratified by the 0000 UTC and 1200 UTC model forecast cycles (Fig. 2). MOS PoPs showed higher skill than LCL PoPs for the 0000 UTC cycle, while LCL PoPs showed some improvement over the MOS guidance for the 1200 UTC cycle. All local (LCL)

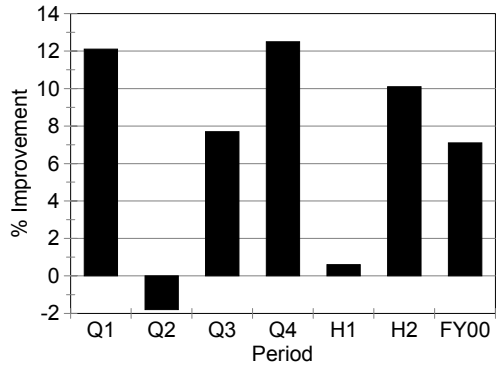

Fig. 1. MOS-based skill scores for 24-h forecast PoPs during various periods of FY00. See text for description of periods.

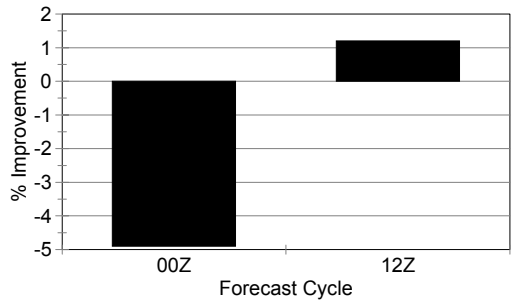

Fig. 2. MOS-based skill scores for 24-h forecast PoPs associated with the 0000 UTC and 1200 UTC cycles during Q2 of FY00.

and MOS PoP forecasts associated with these forecast cycles during Q2 were collected.

At this point a variety of directions can be taken. For example, WFOs having the most/least skill for either of these forecast cycles can be identified. However, such distinction is not the purpose of this study. An example of how a review of the elements behind the PoP issued by a WFO can be used to sort the "Man-MOS-Mix" is presented.

In the afternoon public forecast package of 11 February 2000 the WFO at Pleasant Hill, Missouri, issued a PoP of 70% for the period 0000 to 1200 UTC 12 February. The corresponding NGM MOS PoP, a 24-h forecast of a 12-h PoP from 1200 UTC, was 20%. This case provides an example for which the local forecaster made a clear decision to use a PoP different from NGM MOS guidance.

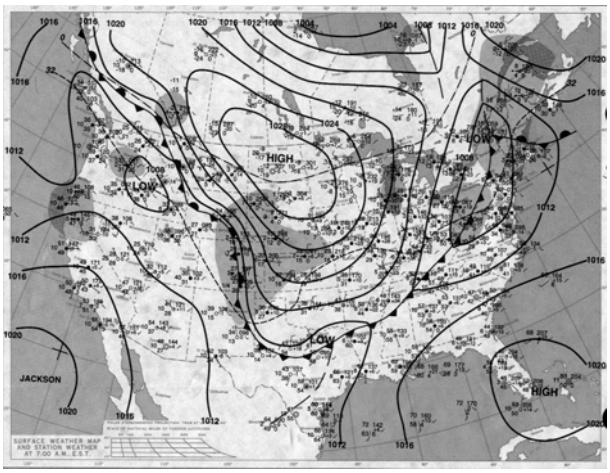

Fig. 3. Sea level pressure analysis from 1200 UTC 11 February 2000. (From *Daily Weather Map* series)

Figure 3 is the sea level pressure analysis for 1200 UTC 11 February 2000. A strong high pressure system is shown extending from the northern Plains across the central and southern Plains. Northwest Missouri has temperatures below freezing and northerly winds. Light precipitation is occurring to the northwest in northern and western Kansas and southern Nebraska.

In addition to the PoP provided via MOS from the NGM, a quantitative precipitation forecast (QPF) was available. Figure 4 is the 12-h QPF ending at 0000 UTC 12 February, the beginning of the verification period for precipitation occurrence associated with the PoPs of interest. Although light precipitation was predicted in north central Kansas, no precipitation was indicated in northwest Missouri. Figure 5 is the QPF from the Eta model (Black 1994) for the same 12-h period. A distinct area of precipitation extends from northern Kansas across northern Missouri. During the next 12-h period, which was associated with the PoPs of interest, no precipitation was predicted by either the NGM or Eta models in northwest Missouri. So, how were these guidance elements combined to produce the resulting PoP forecast?

Insight into the forecaster's decision process, or manual component of the forecast preparation, can be derived from the Area Forecast Discussion (AFD) written in association with the afternoon public forecasts (Fig. 6). An area of snow fit that predicted by the Eta model, as well as satellite and radar imagery, up to the time of the AFD. Although precipitation predicted by the Eta model (Fig. 5) ended in the Pleasant Hill area of responsibility by the period covered by the PoP

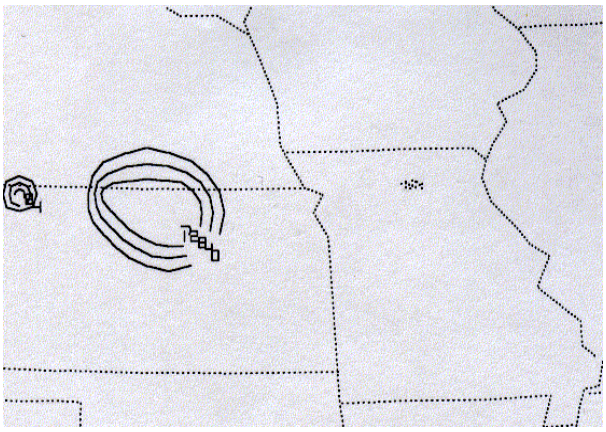

Fig. 4. Predicted precipitation amounts (.01 in) from the NGM for the 12-h period ending at 0000 UTC 12 February 2000.

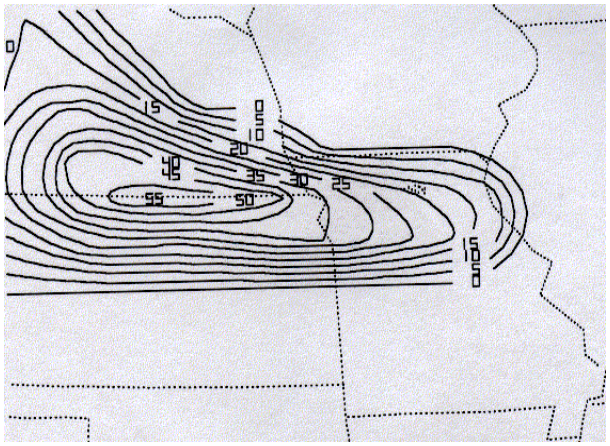

Fig. 5. Predicted precipitation amounts (.01 in) from the Eta model for 12-h period ending at 0000 UTC 12 February 2000.

of interest, the forecaster noted the precipitation was running a few hours behind the timing of the Eta numerical guidance. Also, the forecaster stated there was "plenty" of upward motion available to maintain snow into the evening. The forecaster thus extended a higher risk (a 70% PoP) into the next 12-h period.

This case is only one of many that could have been selected from Q2 of FY00. Additional cases will be included in the conference presentation. Further, more complete reviews of cases are required to clearly define the relative contributions of the components of the "Man-MOS-Met Mix." Compilation of results from such reviews can lead to identification of "best practices."

AREA FORECAST DISCUSSION
NATIONAL WEATHER SERVICE
KANSAS CITY/PLEASANT HILL MO
300 PM CST FRI FEB 11 2000

YET ANOTHER VERY CHALLENGING FORECAST...THE TOUGHEST SO FAR THIS WEEK...

SNOW FOR THIS EVENING PANNING OUT MUCH AS EXPECTED...EXCEPT MAYBE BEHIND BY A FEW HOURS. RUC AND ETA QPF FIELDS FIT SATELLITE AND RADAR TRENDS VERY NICELY. UPPER LEVEL WAVE CONTINUES TO SHEAR OUT AND WEAKEN IN MOSTLY ZONAL FLOW...BUT BOTH RUC AND ETA SHOW PLENTY OF UPWARD FORCING IN LAYER DIV-Q PRODUCTS TO KEEP PCPN GENERATION GOING TROUGH EVENING...WILL PAINT A 1-3" STRIPE BASICALLY ALG AND N OF I-70 WITH LESSER AMTS EITHER SIDE. UPWARD FORCING END BY MIDNIGHT IN THE WEST AND 06-09Z TIME FRAME EAST. WILL TIME ACCORDINGLY.

Fig. 6. Excerpts from the Area Forecast Discussion (AFD) written at the Pleasant Hill, Missouri WFO discussing the local aspects of the forecast precipitation event of the evening of 11 February 2000.

4. PERSPECTIVE FOR IFPS

Operational use of the Interactive Forecast Preparation System (IFPS) brings new tools, procedures and end products to the forecaster. Although preparation of forecasts will be accomplished via new procedures, a fundamental "Man-MOS-Met Mix" will still exist for PoP, as well as other forecasts. With grids of forecast variables being constructed, this mix will be made in a more explicit manner than in the past. Results of best practices gleaned from case reviews as suggested in this study will enable a forecaster to more efficiently balance guidance and personal judgment in both space and time.

5. SUMMARY

An NGM MOS-based skill score for cases when MOS guidance and LCL PoPs differ by at least 20% is currently a primary NWS performance measure. Stratification of these skill scores for various periods allows identification of cases for which forecasters have made decisions leading to even larger differences in PoPs. Reviews of remote observations, numerical model output, objective statistical guidance and the AFDs written by forecasters assist assessment of the contribution of the various components of the forecast process. A better understanding of interactions, the "Man-MOS-Met Mix," of these components can lead to improved forecasts. In addition, this improved understanding can lead to better operational procedures for the IFPS.

6. ACKNOWLEDGMENTS

The author thanks Valery Dagostaro, Meteorological Development Laboratory, for providing NGM MOS PoPs, local PoPs and observed precipitation amounts for CR WFOs. Ted Funk and Steve Becka provided assistance with graphics. Pete Browning and Jeff Manion provided valuable discussions.

7. REFERENCES

Black, T.L., 1994: The new NMC mesoscale Eta model: Description and forecast examples. *Wea. Forecasting*, **9**, 265- 278.

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

Dallavalle, J.P., and M.C. Erickson, 2000: AVN-based MOS guidance - The alphanumeric messages. *NWS Technical Procedures Bulletin No. 463*, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce.12 pp.

Erickson, M.C., and J.P. Dallavalle, 2000: MRF-based MOS guidance - The alphanumeric messages. *NWS Technical Procedures Bulletin No. 460*, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce.12 pp.

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.

Su, J.C., 1993: NGM-based MOS guidance for the probability of precipitation (PoP). *NWS Technical Procedures Bulletin No. 409*, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce. 14pp.

Wilks, D.S., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction*. Academic Press, 467 pp.