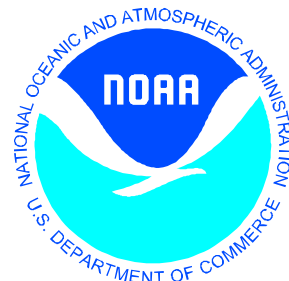National Weather Service
Office of Meteorology

Technical Procedures Bulletin

The Technical Procedures Bulletin was written by Rebecca L. Allen. It describes the precipitation type guidance for 12-h periods ending at 0000 and 1200 UTC, up to 192 hours in advance which is based on the Model Output Statistics.

Paul Hirschberg
Chief, Science Plans Branch

MRF-Based MOS Precipitation Type Guidance for the United States
by Rebecca L. Allen

1. INTRODUCTION

Since October 4, 2000, the National Weather Service has been disseminating precipitation type guidance for 12-h periods ending at 0000 and 1200 UTC, up to 192 hours in advance. The forecasts are based on the application of the Model Output Statistics (MOS) (Glahn and Lowry 1972) technique to output from the Medium Range Forecast (MRF) run of NCEP's Global Spectral Model (Kanamitsu 1989). Forecast equations predict the conditional probabilities of freezing precipitation, snow, rain, and rain mixed with snow. These probabilities are then used to determine a categorical forecast of precipitation type. The guidance is available for approximately 1000 sites in the contiguous United States and Alaska, and is issued from September 1 through May 31 as part of the MRF MEX forecast bulletin (Erickson and Dallavalle 2000). This Technical Procedures Bulletin summarizes the development, testing, and dissemination of the MRF MOS precipitation type forecasts.

2. DEVELOPMENT

The MOS approach statistically relates observed predictand data to predictor data such as forecasts from dynamical models, surface observations, and geoclimatic information. In applying the MOS technique to precipitation type, multiple linear regression was used to develop the statistical equations.

a. Predictand

The predictands used in this development were obtained from the hourly METAR observations of present weather. Each 12-h period ending at 0000 and 1200 UTC was classified into one of five categories: freezing precipitation, snow, rain, rain mixed with snow, or the null category, according to the present weather reported in the hourly observations. The thirteen individual present weather reports from the 12-h period are used to determine one precipitation type categorization for the entire period. The null category included those periods where no precipitation of any kind occurred, or cases where we were unable to determine the exact type of precipitation. Freezing precipitation was comprised of freezing rain, freezing drizzle, ice pellets, or any precipitation in combination with any of these three events. Snow was defined as pure snow or snow grains, while rain was comprised of pure rain or drizzle. The rain/snow mixed category (RS) consisted of those periods where both rain and snow occurred. For any 12-h period to be considered a valid case, a station had to report at a minimum of seven of the possible 13 hours during the period, and three of those reports must have been precipitation. For the null case where none of the four types occurred, the predictands were set to missing. Every valid case in the sample was characterized by four mutually exclusive binary predictands, each taking a value of 1 or 0 for freezing/no freezing, snow/no snow, RS/no RS, and rain/no rain respectively. For example, a snow event would be characterized as a non-freezing/snow/non-RS/non-rain event. A separate linear regression equation was developed for each predictand to forecast the conditional

probability of that precipitation type occurring. The probability equations (and subsequent forecasts) are conditional upon precipitation occurring since only precipitation cases were used to develop the forecast equations.

Note that for any cases in the mixed categories - freezing precipitation or rain/snow mixed - one cannot tell whether a transition between precipitation types occurred during the 12-h period, or whether the mixture occurred at a particular hour. For example, rain and ice pellets reported for three separate hours would be categorized as a freezing precipitation type case; 2 hours of rain followed by 1 hour of ice pellets would also be categorized the same.

There are a few caveats concerning the METAR present weather observations. First, ASOS sites often have trouble distinguishing between light rain and light snow, and, therefore, report unknown precipitation (UP). Because one cannot be sure of the exact precipitation type, these cases of UP were not included in the equation development. Reports of hail and squalls were also left out of the development. Finally, in our data, when a thundershower was reported, the precipitation type was unknown. Therefore, if a thundershower was reported along with any other report of precipitation at that hour, that other reported type took precedence in determining the precipitation type. Otherwise, the thundershower was assumed to be a rain shower.

b. Predictors

Predictors offered in the regression process included MRF model data, geoclimatic information, and, for some projections, surface observations. The MRF model fields included temperature, wet-bulb temperature, temperature advection, wind components, relative humidity, vertical velocity, and relative vorticity; these model variables were included at levels of 500, 700, 850, 925, and 1000 mb. Some surface variables were offered, including 2-m temperature and wet-bulb temperature, and 10-m wind speed. Thicknesses were also offered for many layers including 1000-700 mb, 1000-850 mb, 1000-925 mb, 925-850 mb, 850-700 mb, and 700-500 mb. For most predictors, model fields valid at the beginning and end of the 12-h forecast period were offered, along with the average of the field over the 12 hours.

Several predictors derived from model fields were made available to the regression. These included the pressure (mb) of the freezing level, as well as a predictor based on the vertical profile of wet-bulb temperature forecast by the model. This "zr" predictor identifies cases where freezing precipitation is likely to occur based on the presence of a sufficiently cold surface layer, and a warm layer aloft that will allow for the melting of frozen precipitation. Finally, several predictors were offered that give the probability of snow occurring at an individual station based solely on the forecast value of a particular model field (1000-850 mb thickness, 850-mb temperature, or 2-m temperature).

Geoclimatic predictors included the sine and cosine of the day of the year, which help infer variations within the season, and the conditional monthly relative frequencies of freezing, frozen, and liquid precipitation. These relative frequencies were computed from 2 years of data, and are valid for the 12-h period centered on the forecast projection. The freezing, frozen, and liquid categorization was done according to the method used in the AVN MOS precipitation type

development (Allen and Erickson 2001). The relative frequencies provide specific information about individual stations, some of which might have similar model forecasts, but regularly experience vastly different weather due to local effects.

Surface observations taken at 0600 UTC were also offered to the regression for the 24- and 36-h projections. Surface temperature, dew-point temperature, and the average of these two fields were included as predictors. In addition, the present weather reported at 0600 UTC was included in the form of three binary predictors: freezing precipitation/no freezing precipitation, frozen/no frozen, and liquid/no liquid.

Predictors were offered to the regression in a continuous, point-binary, or grid-binary (Jensenius 1992) form. To compute a point-binary variable, the original gridded predictor is first interpolated to the specific station and then compared to the appropriate binary cutoff. The resulting value of the predictor is either 0 or 1, according to whether the predictor is less than, or greater than/equal to the breakpoint, respectively. Conversely, to compute a grid-binary variable, the binary cutoff is applied at the model gridpoints, and then that field of 0's and 1's is smoothed and interpolated to specific stations. The resulting variable can have any value between 0 and 1. This technique provides a smoother transition, both spatially and temporally, between the extremes of the predictor than does the point-binary approach. In this development, the model fields were offered as continuous and grid-binary variables, and the surface observations were offered as point-binaries.

The most frequently chosen predictors included the relative frequency of freezing precipitation, the zr predictor, the 2-m temperatures, grid binaries of various thicknesses and temperatures, and four predictors that indicated the probability of snow at individual sites. These later variables were based on temperatures or thicknesses. The surface observations of temperature, average of temperature and dew point, and present weather (freezing/no freezing) were also chosen quite frequently in the equations for the 24- and 36-h projections.

c. Regions

Since freezing rain, ice pellets, and snow are relatively rare events in some parts of the country, stations were combined into geographic regions in order to develop stable forecast relationships. These forecast equations are then applied to any station within a region. This also allows forecasts to be produced for stations that were not included in the developmental sample, but that are located within a region for which an equation was developed. Figures 1 and 2 show the four regions used in the contiguous U.S. and the two regions used in Alaska, respectively. These regions were developed after considering both climatic and geographic similarity. Precipitation type guidance was not developed for stations in Puerto Rico, Hawaii, southern Florida, and much of California. Consequently, MRF MOS precipitation type guidance is not available for these sites.

d. Developmental Sample

The developmental sample consisted of precipitation type observations and predictor data for 669

stations in the contiguous U.S. and 36 sites in Alaska. Although forecasts are produced operationally for roughly 1000 stations, only those that report present weather reliably were used to develop the equations. The final equations were derived by using approximately 4 years of data from two different data sets. The first set consisted of MRF model data from the cool seasons of 1997-98, 1998-99, and 1999-2000. For developmental purposes, the cool season is defined as September 16 through May 15 for the contiguous U.S., and September 1 through May 31 for Alaska. The second set of data consisted of MRF model data from the National Centers for Environmental Prediction's (NCEP) reanalysis project (Kalnay et al. 1996). Forecast data valid every 12 hours from 12 to 192 hours were used every 5th day for 1992 through 1997, thus resulting in the equivalent of one full year of data.

e. Equation Characteristics

In the precipitation type development, the equations to predict each of the four conditional probabilities for a specific projection were developed simultaneously. As a result, all four equations contain the same terms, but the coefficients vary among the predictands. The simultaneous development and the nature of the mutually exclusive binary predictands insure the forecasts sum to 100%.

The multiple linear regression routine that produces the forecast equations uses a forward selection method where predictors are added to the equation until a specified stopping criterion is reached. In the precipitation type development, the regression procedure stopped when 10 terms had been added to the equation, or when none of the remaining terms reduced the variance by an additional 0.5%.

Both primary and secondary equations were developed for the 24- and 36-h projections. The primary equations included surface observations as predictors, and are used in the operational setting when observations are available. When no observations are available for a particular station, the secondary equations are used. In the development of these secondary equations, surface observations were not offered as potential predictors.

f. Determining Thresholds for Categorical Forecasts

Categorical forecasts, that is, a forecast of freezing precipitation (Z), snow (S), rain mixed with snow (RS), or rain (R), conditional upon precipitation occurring, are produced from the probabilistic forecasts. The probability forecasts are compared to three threshold probabilities in order to determine the categorical precipitation type forecast. These threshold probabilities are calculated from the developmental sample for each forecast projection and region. For precipitation type, we chose thresholds that maximized the threat score on the dependent sample, while also maintaining a bias between 0.98 and 1.02.

3. POST PROCESSING

Once a day, at 0000 UTC, the conditional probability of precipitation type equations are evaluated by using the appropriate MRF model predictors, recent surface observations, and geoclimatic

predictors. Once the probability forecasts have been generated, they are normalized. Any probabilities less than zero are set to zero, and the remaining probabilities are divided by the sum of all the positive probabilities to obtain the normalized probability. In other words, all of the precipitation type probability forecasts will be greater than or equal to zero and less than or equal to one, and the four probabilities will sum to 100%. Next, the categorical forecasts are generated by comparing the normalized forecast probabilities to the thresholds. In making an operational forecast, the following procedure is used to choose the categorical forecast. To begin, the freezing precipitation probability is compared to the first threshold probability. If the freezing precipitation probability is greater than the threshold value, then freezing precipitation (Z) is chosen as the categorical forecast. If not, the freezing probability and snow probability are added together and compared to the next threshold value. If this threshold value is exceeded, then snow (S) is chosen as the categorical forecast. If that threshold value is not exceeded, the rain/snow mix probability is added to the freezing and snow sum, and that new sum is compared to the third threshold. If the third threshold is exceeded, rain/snow mix (RS) is chosen as the categorical forecast. If none of the three threshold values is exceeded, rain (R) is chosen as the categorical forecast.

4. OPERATIONAL PRODUCTS

MRF MOS precipitation type forecasts are available in the new MRF MOS forecast bulletin distributed under the WMO headers FEPA20, FEUS21 - 26, and FEAK37 - 39, and the AWIPS product identifier MEX. Technical Procedures Bulletin No. 460 (Erickson and Dallavalle 2000) describes the complete MEX MOS message. Probabilistic and categorical precipitation type forecasts are available every 12 hours from 24 to 192 hours in the alphanumeric message. The precipitation type guidance is also available in a binary format (BUFR) message. A sample of the precipitation type portion of the alphanumeric forecast bulletin is shown in Fig. 3. The PZP line contains the conditional probability of freezing precipitation, the PSN line contains the conditional probability of snow, and the PRS line contains the conditional probability of a mixture of rain and snow. The categorical precipitation type is shown in the TYP line: Z indicates freezing precipitation, S indicates snow, RS indicates rain/snow mix, and R indicates rain. For users outside the National Weather Service, the guidance is available through NOAAPORT, the Family of Services, and specific military communication circuits.

The MRF MOS forecast message is produced for 1060 sites in the contiguous U.S. and Puerto Rico. Of these 1060 sites, there is no precipitation type guidance for sites in Puerto Rico, Hawaii, southern Florida, and much of California (see Fig. 1). Precipitation type forecasts are issued for the remainder of the sites from September 1 through May 31. Therefore, the forecast bulletins for some sites may never contain the PZP, PSN, PRS and TYP lines, and all the bulletins will be missing these lines from June 1 through August 31.

5. VERIFICATION

During the development of the new MRF precipitation type guidance, verification was done to ensure that the forecasts were skillful. Test equations were developed from a portion of the developmental data; the probabilistic and categorical forecasts produced from these equations on

an independent sample were then verified. The data from the 1999-2000 cool season was set aside as the independent sample. Forecasts for all 705 developmental stations in the contiguous U.S. and Alaska were verified. In addition, a system of forecast equations was developed by using only relative frequencies and observations as predictors. This system represented a blend of climate and persistence, and served as a basis of comparison. Forecasts produced by these equations on the independent sample were also verified, and the results for both systems are presented in Figs. 4 and 5. Note that the precipitation type forecasts can only be verified for cases where precipitation occurred.

Figure 4 shows the P-scores for the MRF MOS and the climate/persistence probabilistic precipitation type forecasts for the projections from 24 through 192 hours. The values for P-score can range from 0 to 2 where a smaller score represents a more accurate forecast. The P-scores show that the MRF MOS does indeed exhibit measurable skill at all forecast projections.

Figure 5 shows the verification of the categorical forecasts for both the MRF MOS and the climate/persistence system. The Heidke Skill Score (HSS) can range from 0 to 1, with a 1 denoting a perfect forecast. Like the probabilistic forecasts, the MRF MOS categorical forecasts are more skillful than the climate/persistence forecasts at all projections.

Figs. 4 and 5 also show that the primary MRF MOS forecasts are more skillful than the secondary MRF MOS forecasts at the two projections where both are available. Despite this lower accuracy, these secondary forecasts are preferable to a missing forecast in the event that the observational predictors are not available.

6. OPERATIONAL CONSIDERATIONS

The forecaster must remember that the probabilities of precipitation type are conditional, that is, given that precipitation occurs, these are the probabilities that precipitation will fall in the specified form during a 12-h period. Therefore, one should use this guidance in conjunction with the probability of precipitation (PoP) forecasts. If the conditions are not right for precipitation to occur, the precipitation type forecast is virtually meaningless.

The MOS technique can account for some systematic model biases, as well as the reduced skill of the model with increasing projection. Unfortunately, the MOS guidance is not able to overcome a bad model forecast; the guidance will reflect the main patterns in the MRF model output. Also, the MOS equations are tuned to the specific sample on which they are developed. For example, if the developmental sample represented a relatively dry period, the equations might not perform as well during a significantly wet period. Additionally, future changes to the MRF model that affect the model forecasts and biases may affect the performance of the MOS forecasts.

The user may notice the absence of the probability of rain forecasts in the alphanumeric message. Remember that the four probability equations are developed simultaneously and, therefore, the probabilities produced from the equations sum to 100%. In the interest of space, the probability of rain was omitted, but can be obtained by subtracting the sum of the PZP, PSN and PRS from 100%.

7. REFERENCES

Allen, R. L., and M. C. Erickson, 2001: AVN-based MOS precipitation type guidance for the United States. NWS Technical Procedures Bulletin No. 476, NOAA, U.S. Dept. of Commerce, 10 pp.

Erickson, M. C., and J. P. Dallavalle, 2000: MRF-based MOS guidance - The alphanumeric messages. NWS Technical Procedures Bulletin No. 460, NOAA, U.S. Dept. of Commerce, 9 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.

Jensenius, J. S., Jr., 1992: The use of grid-binary variables as predictors for statistical weather forecasting. Preprints Twelfth 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. Ropelewski, 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. and Forecasting, 4, 335 - 342.
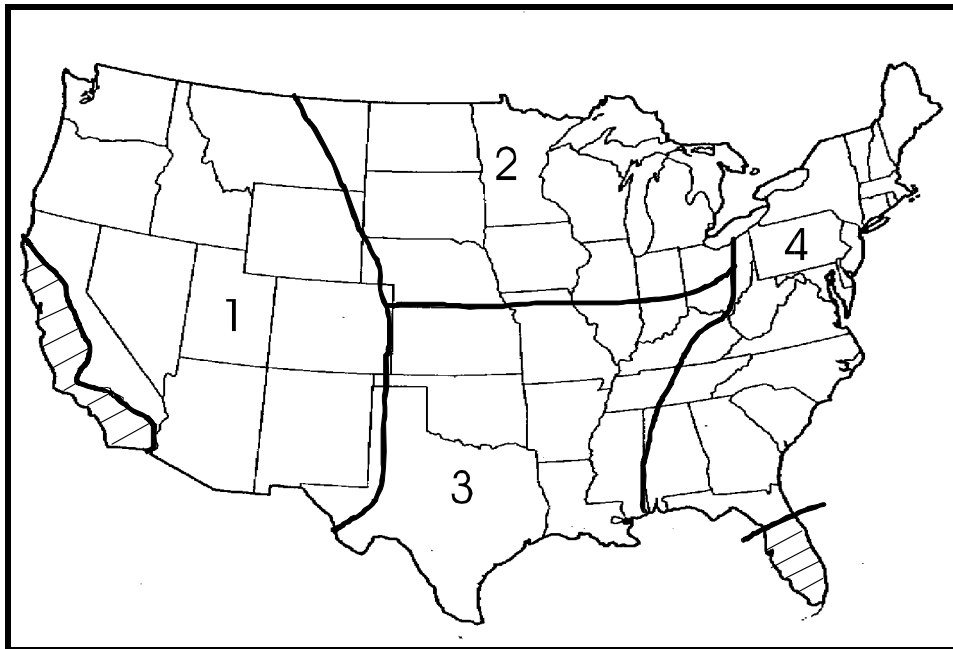
Figure 3. Regions used in the MRF precipitation type development for the contiguous U.S. Precipitation type forecasts are not produced for stations in the hatched areas.
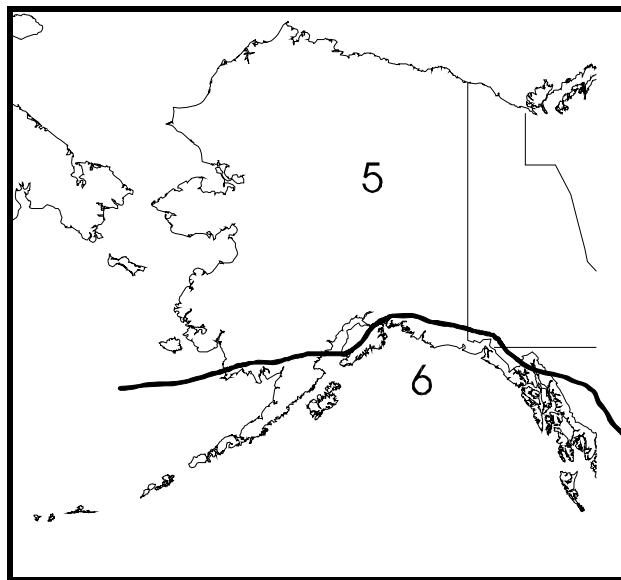

Figure 4. Regions used in the MRF precipitation type development for Alaska.

```
KGTF   MRF MOS GUIDANCE   10/04/2000   0000 UTC
FHR   24| 36  48| 60  72| 84  96|108 120|132 144|156 168|180 192
WED   04| THU 05| FRI 06| SAT 07| SUN 08| MON 09| TUE 10| WED 11 CLIMO

PZP    3|  1   3|  0   1|  0   0| 25  13|  1   1|  1   0|  1   1
PSN   71| 82  88| 98  95| 88  61| 29  35| 20  23| 22  23| 25  39
 PRS   27| 17   9|  2   5|  6  21| 30  16| 17  11| 12  11| 15  10
  TYP    S|  S   S|  S   S|  S   S|  Z   Z|  R   R|  R   R|  R  RS
```

Figure 3. Precipitation type portion of sample 0000 UTC MRF MOS message
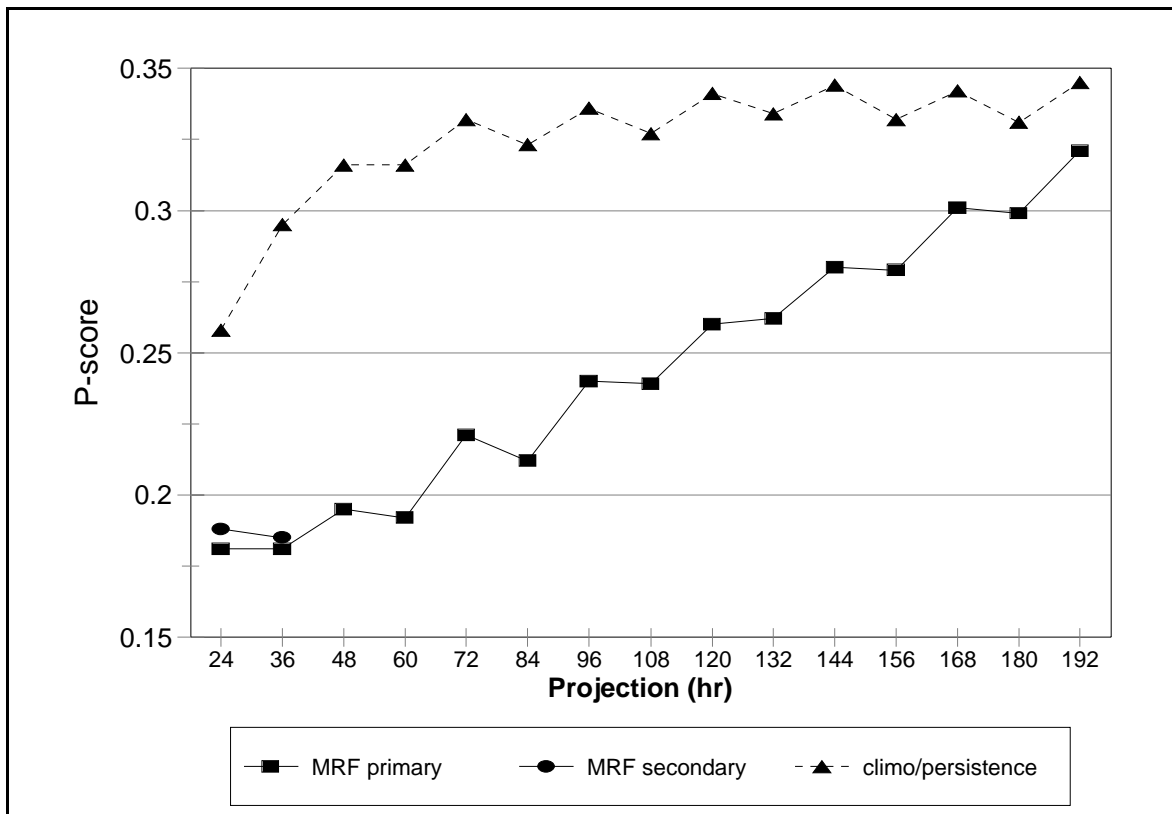

Figure 4. P-scores for the MRF MOS primary, MRF MOS secondary, and
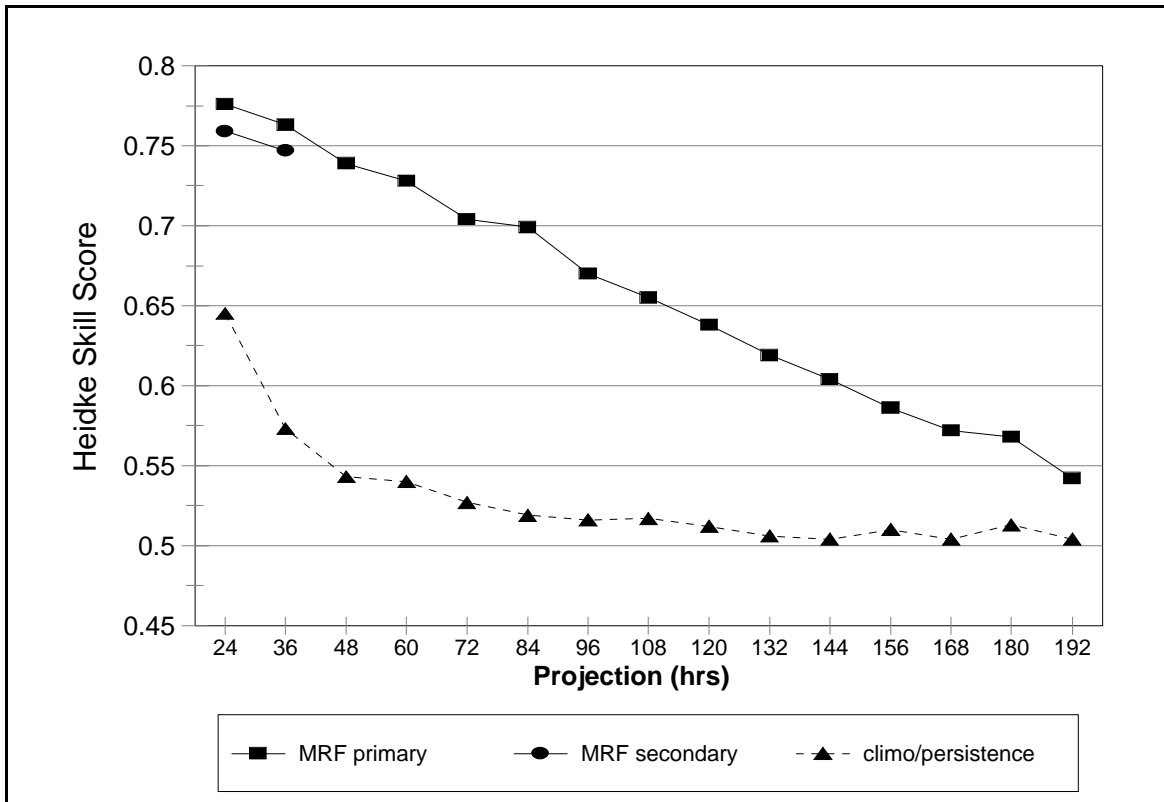climo/persistence precipitation type probabilistic forecasts.

Figure 5. Same as Fig. 4, except for Heidke skill scores for precipitation type categorical forecasts.