6.4 EXPERIMENTAL MOS PRECIPITATION TYPE GUIDANCE FROM THE ECMWF MODEL

Phillip E. Shafer * and David E. Rudack
Meteorological Development Laboratory
Office of Science and Technology
National Weather Service, NOAA
Silver Spring, MD

1. INTRODUCTION

Forecast skill of the European Centre for Medium-Range Weather Forecasts (ECMWF) model is widely recognized as being strongly competitive with, and often exceeding, that of other global models (e.g., Hamill 2012). Although very skillful, ECMWF forecasts do contain systematic bias, and the model does not directly provide forecasts for some sensible weather elements such as probability of precipitation or precipitation type. To enhance the usefulness of the ECMWF output to NWS forecasters, the Experimental Meteorological Development Laboratory (MDL) has recently developed an experimental suite of Model Output Statistics (MOS) guidance from the ECMWF model. The MOS technique (Glahn and Lowry 1972) has been employed by MDL to post-process numerical model output for several decades. The initial development effort for ECMWF MOS focused on the following elements: temperature, dewpoint, maximum and minimum temperature, wind speed and direction, sky cover, and probability of precipitation (this effort is described in detail in Rudack et al. 2014).

The experimental suite of ECMWF MOS has now been enhanced with the addition of probabilistic and best category precipitation type guidance. A three-category short-range and four-category extended-range precipitation type system has been developed at stations over the contiguous U.S. (CONUS) and Alaska, for the 0000 and 1200 UTC ECMWF model cycles. For the short-range system (through 84 hours in advance), equations for the conditional probability of freezing, frozen, and liquid precipitation types were developed for projections every 3 hours valid on the hour (hereafter referred to as PoPT03), while the extended-range product includes a fourth rain/snow mix category and forecasts are valid for 12-hour periods through 192 hours (hereafter referred to as PoPT12). Best category forecasts for the short-range and extended-range products are produced by applying statistically-derived thresholds to the probability forecasts.

This paper describes the development of the new ECMWF MOS precipitation type system and its performance when compared to climatology and GFS MOS precipitation type guidance. Section 2 of this paper gives an overview of the development procedure. Example forecast products are shown in Section 3. Verification scores are presented in Section 4. Finally, a summary of this paper is given in Section 5.

2. EQUATION DEVELOPMENT

The procedure described here for developing the ECMWF MOS precipitation type system follows very closely the approach used for the most recent GFS MOS development (see Shafer 2010 for a detailed description of the GFS MOS precipitation type system).

2.1 Observations

MOS precipitation type guidance was developed from present weather observations at METAR sites. Observations were examined for nearly 2000 stations in the CONUS and Alaska, for the period April 2008 through March 2013. This five-year period corresponds to the sample of ECMWF model data that was available for this development (Rudack et al. 2014). Only sites that report present weather reliably and have a sufficiently long record of observations were used in the development. After eliminating part-time sites, sites that have stopped reporting, and sites that never (or rarely) report precipitation, roughly 1450 stations remained. Of these, 1321 are in the CONUS and 49 are in Alaska. As in previous developments, some reliable Canadian stations (totaling 64) in close proximity to the CONUS and Alaska were used to supplement the sample. Due to the lack of freezing and frozen cases in Hawaii and Puerto Rico, these sites were not used in the

*Corresponding author address:
Phillip E. Shafer, 1325 East-West Highway, Station 10434, Silver Spring, MD 20910-3283; e-mail: PhilShafer@noaa.gov
development and precipitation type guidance is not available for them.

2.2 Definition of PoPT03 predictand

For the short-range precipitation type predictand, the present weather observations valid every 3 hours on the hour (i.e., at 0000, 0300, 0600,... 2100 UTC) were classified into one of three mutually exclusive categories: freezing, frozen, or liquid. A separate “null” category was used for cases when no precipitation of any type occurred, or when the exact type could not be determined. All “null” cases were treated as missing and not included in the development. Thus, only precipitation cases of discernable type comprised the developmental sample. Table 1 lists the present weather observations that were assigned to each precipitation type category. As in previous developments, ice pellets were included with the freezing category and any mixture of liquid precipitation with snow was classified as liquid. Freezing events are exceptionally rare, comprising only around 1.5% of all cool season precipitation cases over the CONUS, and about 0.5% of precipitation cases over Alaska. This makes correctly forecasting freezing precipitation events very difficult.

2.3 Definition of PoPT12 predictand

The extended-range precipitation type predictand is defined somewhat differently from the short-range. Each 12-h period ending at 0000 and 1200 UTC was classified into one of four mutually-exclusive categories: freezing, frozen, liquid, or rain/snow mix. Again, a “null” category was used to define cases in which no precipitation occurred or when the type of precipitation could not be determined, and all such cases were treated as missing and not included in the development. The 13 individual reports spanning the 12-h period were used to determine one precipitation type for the period. In order to be considered a valid case, an observation must have been available for at least seven of the possible 13 hours, and of these at least three must have been a report of precipitation (Allen 2001). The freezing, frozen, and liquid category definitions are the same as those for PoPT03, except here a fourth rain/snow mix category is included which consists of those periods in which both liquid and frozen types were observed (Table 2). This category can represent periods when precipitation is transitioning from rain to snow or vice versa. If any observation during the 12-h period was a report of freezing precipitation, then the category assigned to that period was freezing.

2.4 Gridded geoclimatic predictors

Gridded monthly conditional relative frequencies of freezing, frozen, and liquid precipitation were used as predictors in the regression analysis. The relative frequencies, valid for 12-h periods centered on each 3-hourly forecast valid time, were originally calculated at each METAR site from 10 years of predictand data and then analyzed to high resolution grids over the CONUS and Alaska as part of the most recent GFS MOS development (see Shafer 2010 for more details on the analysis procedure). The relative frequencies provide station-specific information. Stations may have similar model forecasts but often experience vastly different weather due to localized effects. Figure 1 shows example plots of the conditional relative frequency of frozen precipitation over the CONUS in January (Fig. 1a), and liquid precipitation over Alaska in April (Fig. 1b).

Additional geoclimatic information was incorporated through the use of logit 50% (or equal-probability) values. As part of the most recent GFS MOS development, the 50% values were calculated at each METAR site for several parameters that are good discriminators of precipitation type. These include: 2-m temperature, 850-hPa temperature, 1000-850 hPa thickness, 1000-500 hPa thickness, and freezing level (see Shafer 2010 for a detailed description of how the 50% values were derived). The model forecast of 2-m temperature, for example, is then “transformed” by subtracting the 50% value to form a new predictor that helps to account for climatological differences between stations (Shafer 2010). As with the relative frequencies, the 50% values are available in gridded form over the CONUS and Alaska. An example gridded 50% value for 1000-850 hPa thickness valid at 0000 UTC is shown in Figs. 2a and 2b, for the CONUS and Alaska, respectively.

Having the relative frequencies and 50% values in gridded form has two main advantages:

1. They help capture localized effects that may not be well-resolved on the model scale. This allows data for all stations (for which relative frequencies and 50% values are available) to be pooled together into one large region for development, while still retaining station specificity in the equations. The ability to combine data in-
to larger samples is critical when forecasting rare events such as freezing precipitation, when the number of cases in the sample is often limited.

2. The gridded constants can be used to obtain geoclimatic information at any desired point by interpolating values directly from the grid, thus allowing for forecasts to be made even at stations that were not in the developmental sample.

2.5 Predictors offered to regression analysis

Several ECMWF model-derived predictors were offered to the regression analysis. These include, for example, various thicknesses, temperature and wet-bulb temperature at various levels, temperature advection, and a predictor based on the vertical profile of wet-bulb temperature, called the “Z-R predictor”. The latter 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 melting of the frozen precipitation (Erickson 1992, Allen and Erickson 2001). Geoclimatic predictors offered to the regression include the aforementioned logit transforms (see Section 2.4), the monthly relative frequencies of freezing, frozen, and liquid precipitation, and the sine and cosine day of the year. For early lead times (i.e. through 18 hours for PoPT03 and through 36 hours for PoPT12), surface observations of temperature, dewpoint, and precipitation type valid at 3 hours past the model cycle time were offered as predictors.

2.6 Probability equations

ECMWF MOS precipitation type guidance was developed for the cool season, defined as the period 9/1 - 5/31 over the CONUS and 9/1 – 6/15 over Alaska. Roughly five cool seasons of ECMWF forecast output and present weather observations were available for the development (April 2008 through March 2013). As with previous MOS precipitation type developments, multiple linear regression was used to derive the equations. This method, called “Regression Estimation of Event Probabilities” (REEP), relates the binary predictands (see Sections 2.2 and 2.3) to a linear combination of predictor variables, using a forward stepwise selection procedure (Miller 1964). The equations for all predictands (three categories for PoPT03 and four categories for PoPT12) were developed simultaneously; that is, the equations contain the same predictor variables but have different regression coefficients. This insures that the probability forecasts are consistent and sum to 100%. The predictors most often selected were the logit transforms (1000-850 hPa thickness was most influential), the Z-R predictor, the conditional relative frequencies, 2-m wet bulb temperature, 850 hPa temperature, and observed precipitation type. To account for possible missing observations, a secondary (i.e. backup) set of equations was developed without observed predictors. When no observation is available for a particular station, the secondary equation is used.

In order to develop stable forecast equations, all stations within the CONUS were combined into one large region – known as a “generalized operator” approach. This technique is necessary because cases of freezing precipitation do not occur frequently enough (during the 5-yr training sample) at individual stations to obtain a stable single-station equation. Also, in many parts of the CONUS, even snow is a rare event. Pooling data into one or more regions in this way helps to increase the number of freezing and frozen cases in the sample, and the resulting equation is applicable to all stations within the region. Testing showed that two regions (a coastal region and an interior region) worked best for Alaska.

2.7 Post processing

The probability forecasts are first normalized by truncating any negative probabilities to zero and then dividing each by the sum of the positive probabilities to get the normalized probability (i.e. probabilities which sum to 100%). Next, a conditional best category forecast is generated by applying statistically-derived thresholds to the normalized probabilities. Here, the thresholds were chosen which maximized the threat score on the dependent sample, while constraining the bias to a reasonable range (0.98 and 1.02).

3. EXPERIMENTAL PRODUCTS

Equations for the conditional probability of freezing, frozen, and liquid precipitation types (PoPT03) were developed for projections every 3 hours from 6 to 84 hours in advance for the 0000 and 1200 UTC cycles, and are used to populate the experimental short-range ECMWF MOS text bulletin (example shown in Fig. 3a). Equations for the extended-range system, which includes a fourth rain/snow-mix category, are valid for 12-h periods from 24 to 192 hours in advance, and are
used to populate the experimental extended-range ECMWF MOS message (example shown in Fig. 3b). Each bulletin includes all of the probabilities except for the rain category (which can easily be deduced by summing the remaining probabilities and subtracting from 100). The conditional best category is included below the probabilities in the row indicated by “TYP.”

To support possible future generation of a gridded ECMWF MOS suite for the National Digital Guidance Database (NDGD), the PoPT03 system was developed out to 192 hours for the 0000 and 1200 UTC cycles. The gridded guidance is produced by evaluating the PoPT03 equations directly at each NDGD grid point over the CONUS and Alaska (at 2.5 km and 3 km resolution, respectively). This is possible because the equations are generalized operators (i.e. are applicable to all points within the region), and any required geoclimatic predictors can easily be interpolated from the gridded constant datasets (Section 2.4). An example gridded ECMWF MOS forecast for the conditional probability of frozen precipitation over the CONUS is shown in Fig. 4.

4. VERIFICATION

To assess the skill of the new ECMWF MOS precipitation type guidance, verification scores were calculated for an independent sample and compared to climatology and the operational GFS MOS precipitation type system. For rare events such as freezing precipitation, it is desirable to have as large a verification sample as possible in order to minimize the effects of sampling variability on the results. Similar to the procedure used for the most recent GFS development (Shafer 2010), this was accomplished using a “k-fold” cross-validation approach. Each season was withheld as an independent sample and equations were developed from the remaining four seasons. The procedure was repeated five times, each time withholding a different season for testing. Scores were then calculated for the aggregate of all five independent tests, which has the effect of smoothing out any sampling variability from year to year. The guidance was evaluated for the same set of stations used in the development. The results presented in this section (and shown in Figs. 5-9) are for the 0000 UTC cycle and are aggregated for all CONUS stations and separately for all Alaska stations.

4.1 Performance relative to climatology

The score most often used to assess the accuracy of multi-category probability forecasts is the Brier p-score. The p-score is essentially the mean squared error for the probability forecasts summed over each of the nominal binary events to which the probabilities relate (Wilks 2006). Lower p-scores are better. P-scores were calculated for the three-category PoPT03 system and for a reference climatology forecast. Here, climatology is simply the conditional relative frequency of freezing, frozen, and liquid precipitation. Fig. 5 shows the percent improvement in p-score over the reference climatology forecasts for the ECMWF MOS. Skill scores range from a 65-70% improvement over climatology in the early projections to between 45-55% at 72 hours. Forecast skill for the extended-range projections is positive through 192 hours. Recall that only projections through 72 hours are used for the short-range ECMWF MOS text bulletins, while projections through 192 hours are used to generate gridded guidance for NDGD. Relative to climatology, forecasts for the CONUS (blue curve) are more skillful than forecasts for Alaska (red curve). This is most likely due to the reduced sample size resulting from a much smaller number of stations being available over Alaska, and the less frequent occurrence of freezing precipitation events.

4.2 Comparison to GFS MOS

For comparison with operational GFS MOS guidance, p-scores were calculated for the ECMWF MOS and the GFS MOS for the 2010-2011, 2011-2012, and 2012-2013 cool seasons (the most recent GFS precipitation type system was implemented in 2009). Scores for PoPT03 are shown in Figs. 6a and 6b, for the CONUS and Alaska, respectively, for the 0000 UTC cycle. The ECMWF MOS guidance is more skillful than the GFS through all forecast projections. For many projections, ECMWF MOS forecasts have similar or lower p-scores than GFS forecasts valid up to 24 hours earlier. The difference in skill is not as pronounced over Alaska (Fig. 6b), but even here the ECMWF MOS is better through all projections. A similar comparison is shown in Fig. 7 for the four-category extended-range (PoPT12) system. Here too p-scores are better for the ECMWF MOS than for the GFS MOS, with gains of 12-24 hours in skill seen for most projections. The differences in skill seen here are similar to those reported for other MOS elements (Rudack et al. 2014).
The best category forecasts for PoPT03 and PoPT12 were verified by computing the Heidke Skill Score (HSS). The HSS is the proportion of correct forecasts (computed from a contingency table) relative to the proportion correct that would be achieved by random forecasts that are independent of the observations (Wilks 2006). Scores for the ECMWF MOS and GFS MOS are shown in Figs. 8 and 9, for the PoPT03 and PoPT12 systems, respectively. Over the CONUS (Figs. 8a and 9a), skill relative to the GFS seems to improve slightly with increasing projection, while over Alaska (Figs. 8b and 9b) the opposite appears to be true. This is likely due to differences in how the GFS and ECMWF handle synoptic systems at higher latitudes.

4.3 Reliability

Aside from forecast accuracy, it is important that probability forecasts be well-calibrated and reliable. That is, an event with a forecast probability of 50% should occur roughly 50% of the time over a large sample of forecasts. The reliability diagram is one method to visually assess this calibration (Wilks 2006). Forecasts are deemed reliable when the average forecast probability and observed relative frequency of the event are roughly the same in each probability bin. Reliability diagrams for the three-category PoPT03 system are shown in Fig. 10 for the 72-h and 192-h projections. Curves for freezing, frozen, and liquid are plotted (bins with fewer than 100 cases are omitted). The ECMWF MOS forecasts are generally reliable, with slight underforecasting for the freezing category at the 72-h projection. Freezing probabilities at 192 hours rarely exceed 10%. Model temperature profiles used as predictors at extended projections usually are not able to distinguish freezing from frozen precipitation, and the ECMWF model is no exception.

5. SUMMARY

A new short-range and extended-range ECMWF MOS precipitation type system has been developed at stations for the 0000 and 1200 UTC model cycles. Probabilistic and best category precipitation type guidance is now available in the experimental ECMWF MOS short-range and extended-range text bulletins, and also available experimentally in gridded format.1 A cross-validation was performed for the 0000 UTC cycle to compare the skill of the new ECMWF MOS precipitation type guidance to that of climatology and the operational GFS MOS. The skill of the ECMWF MOS is superior to the GFS MOS for all projections through 192 hours, and in some cases ECMWF MOS precipitation type forecasts have similar or better skill than GFS MOS forecasts valid up to 24 hours earlier.

6. ACKNOWLEDGMENTS

The authors wish to thank Tamarah Curtis for her assistance with processing the ECMWF model data. We also wish to thank Bob Glahn, David Ruth and Kathryn Gilbert for reviewing this manuscript.

7. REFERENCES


---

1 ECMWF data is provided to the National Weather Service for internal use only. ECMWF model data is considered proprietary and confidential. The ECMWF MOS guidance derived from ECMWF data is restricted for internal use only.


Table 1. Definitions of MOS short-range precipitation type categories.

<table>
<thead>
<tr>
<th>Freezing</th>
<th>Frozen</th>
<th>Liquid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freezing rain (FZRA)</td>
<td>Snow (SN)</td>
<td>Drizzle (DZ)</td>
</tr>
<tr>
<td>Freezing drizzle (FZDZ)</td>
<td>Snow showers (SHSN)</td>
<td>Rain/drizzle (RADZ)</td>
</tr>
<tr>
<td>Ice pellets (PL)</td>
<td>Snow grains (SG)</td>
<td>Rain (RA)</td>
</tr>
<tr>
<td>Any precipitation in combination with any of the above.</td>
<td></td>
<td>Rain shower (SHRA)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thunderstorm (TSRA)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any mixture of liquid precipitation with snow.</td>
</tr>
</tbody>
</table>

Table 2. Definitions of MOS extended-range precipitation type categories.

<table>
<thead>
<tr>
<th>Freezing</th>
<th>Frozen</th>
<th>Liquid</th>
<th>Rain/Snow Mix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freezing rain (FZRA)</td>
<td>Snow (SN)</td>
<td>Drizzle (DZ)</td>
<td>12-h periods in which both rain and snow were observed.</td>
</tr>
<tr>
<td>Freezing drizzle (FZDZ)</td>
<td>Snow showers (SHSN)</td>
<td>Rain/drizzle (RADZ)</td>
<td></td>
</tr>
<tr>
<td>Ice pellets (PL)</td>
<td>Snow grains (SG)</td>
<td>Rain (RA)</td>
<td></td>
</tr>
<tr>
<td>Any precipitation in combination with any of the above.</td>
<td></td>
<td>Rain shower (SHRA)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thunderstorm (TSRA)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any mixture of liquid precipitation with snow.</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Example gridded precipitation type relative frequencies for (a) snow over the CONUS and (b) liquid over Alaska.
Figure 2. Analyzed 1000-850 hPa thickness 50% value (meters) for (a) the CONUS and (b) Alaska, valid at 0000 UTC.
Figure 3. Example (a) short-range and (b) extended-range ECMWF MOS text bulletins. The precipitation type portions of each bulletin are highlighted by the blue rectangles.
Figure 4. Example gridded ECMWF MOS forecast for the conditional probability of frozen precipitation over the CONUS.
Figure 5. P-score percent improvement over climatology for the short-range ECMWF MOS precipitation type system. CONUS scores (blue curve) and Alaska scores (red curve) are shown.
Figure 6. P-scores for the short-range ECMWF MOS precipitation type system over (a) the CONUS and (b) Alaska, for the 0000 UTC cycle. The ECMWF MOS (blue curve) and GFS MOS (red curve) are shown.
Figure 7. P-scores for the extended-range ECMWF MOS precipitation type system over (a) the CONUS and (b) Alaska, for the 0000 UTC cycle. The ECMWF MOS (blue curve) and GFS MOS (red curve) are shown.
Figure 8. Heidke skill scores for the short-range ECMWF MOS precipitation type system over (a) the CONUS and (b) Alaska, for the 0000 UTC cycle. The ECMWF MOS (blue curve) and GFS MOS (red curve) are shown.
Figure 9. Heidke skill scores for the extended-range ECMWF MOS precipitation type system over (a) the CONUS and (b) Alaska, for the 0000 UTC cycle. The ECMWF MOS (blue curve) and GFS MOS (red curve) are shown.
Figure 10. Reliability diagrams for the short-range ECMWF MOS precipitation type system over the CONUS for (a) the 72-h projection and (b) 192-h projection. Only bins with 100 or more cases are plotted.