5.2 PRE-PROCESSING OF ATMOSPHERIC FORCING FOR ENSEMBLE STREAMFLOW PREDICTION

John Schaake*, Sanja Perica, Mary Mullusky, Julie Demargne, Edwin Welles and Limin Wu
Hydrology Laboratory, Office of Hydrologic Development
National Weather Service, NOAA, Silver Spring, Maryland

1. NCEP GLOBAL ENSEMBLE FORECASTS

Atmospheric ensemble forecasts contain biases that must be removed before they are used as input to hydrological models. Also the spread of the adjusted ensembles underestimates the true uncertainty. Correcting these limitations of weather and climate ensemble forecasts is essential to produce skillfull and reliable ensemble streamflow forecasts. Results of a study of alternative approaches to re-scale NCEP global ensemble forecasts and to compensate for under-estimation of ensemble spread are presented. These approaches are being used to support the NWS Advanced Hydrologic Prediction Services (AHPS)

The NCEP ensemble truncates the resolution of the nominal MRF and AVN runs (T126 truncation, ~100 km) to T62 (~200 km) at a lead times of 7 and 3 days at 00Z and 12Z, respectively. At 00Z there is also a "control" totally T62 run. In addition to this control forecast, 10 forecasts with T62 resolution are run from 00Z starting from slightly perturbed initial conditions. At 12Z four additional forecasts are generated from perturbed initial analyses. Hence, there is a total of 17 individual global predictions generated daily. All forecasts are run to 16 days with the latest version of the EMC MRF global model.

Precipitation forecasts for 12 and 24 hr durations have been archived by EMC on a 2.5 degree grid since 1977. The analyses in this study used ensemble forecast data for the 3 year period 1997-1999. Only 12Z-12Z 24hr precipitation forecasts were analyzed. Values of observed 24 hr precipitation for each 2.5 degree grid element were estimated using observations from the NCDC Coop network and the SNOTEL network. Only 2.5 degree grid elements where there were at least 10 precipitation gages were used in the analysis.

---

* Corresponding author address: John Schaake, NWS/NOAA, Office of Hydrologic Development, OHD12, 1325 East_West Highway, Silver Spring,MD 20910; e-mail: john.schaake@noaa.gov

2. ALTERNATIVE ANALYSIS STRATEGIES

Figure 1 illustrates the bias that exists in the global ensemble forecasts for mean daily precipitation in the 3 month period centered on July. The upper panel in Figure 1 gives the average 24 hr forecast precipitation (mm) for forecasts for day - 1. The lower panel gives the corresponding observed value. Areas in black did not have at least 10 gage observations for each day. The inset graph compares the observed and forecast values for each pixel. Clearly, the model has some large biases that vary over the model domain. The large over-prediction in summer in the south east is well known.

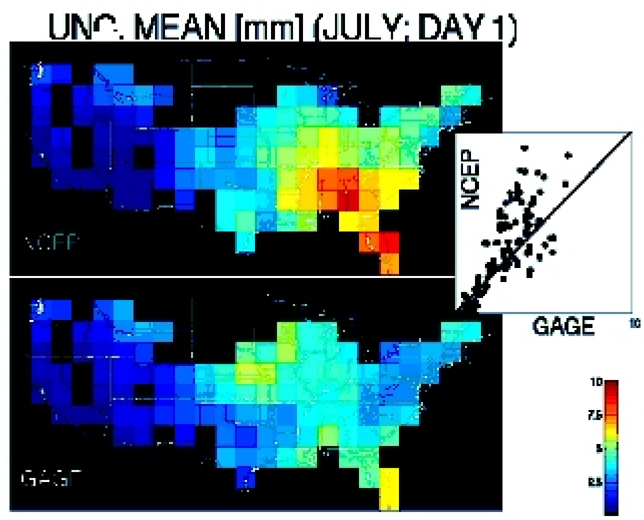

Figure 1. - Example Bias in NCEP Global Ensemble 24 hr Precipitation Forecasts for July.

Four different analysis strategies to deal with the effects of ensemble bias are considered in this study. These strategies are:

- (i) Use the raw ensemble forecasts with no adjustment
- (ii) Adjust each ensemble value so that the mean ensemble value is the same as the observed mean
- (iii) Adjust each ensemble value so that the cumulative distribution of

ensemble values is the same as the cumulative distribution of corresponding observations

(iv) Use the joint distribution of ensemble mean values and observed values for each forecast day to re-construct an ensemble of precipitation forecasts.

Analysis strategy (iv) is being tested by several NWS River Forecast Centers as an intial approach to construct ensemble forecasts from single-value precipitation forecasts produced by the Hydrometeorologic Prediction Center (HPC) (Herr, et al, 2002). In this analysis, the ensemble mean value from the global ensemble forecast is used essentially as a surrogate for the HPC single value forecast. As a result this analysis not only illustrates some of the science issues that need to be considered for hydrologic application of NCEP ensemble forecasts but it also illustrates how the current RFC operational ensemble procedures compare with potential applications of global ensemble forecasts.

The effect of applying these analysis strategies to individual ensemble member values on the cumulative distribution of adjusted 24 hr forecast precipitation amounts is illustrated in Figure 2 and compared to the cumulative distribution of observed values. The data in Figure 2 are for July for a grid element in the southeast where there is a large bias as illustrated in Figure 1.

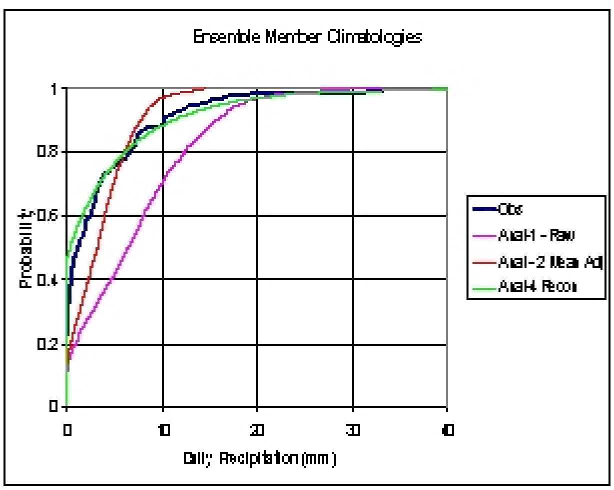

Figure 2 - Cumulative Distribution of Adjusted Global Ensemble Members Compared to the Cumulative Distribution of Observations for Day-1 24 hr Precipitation Forecasts in July in the Southeast U.S.

The observed distribution is shown in blue. The raw ensemble distribution, analysis (i), is shown in violet. It appears far to the right of the observed distribution as a result of the high bias in the ensemble forecasts. The distribution for analysis (ii) is shown in red. It has the same shape as the distribution of the raw members but is shifted to the left of the raw distribution and is closer to the observations. The distribution for adjustment (iii) is identical to the distribution of observations in blue and therefore cannot be plotted as a separate curve. The distribution for adjustment (iv) is in green and is very close to the distribution of observations in blue.

3. VERIFICATION STATISTICS

Each of the four analysis strategies was applied to the ensemble forecasts for 8 representative grid points throughout the U.S. for forecast lead times of 1 to 5 days and for forecasts in January and July. A number of different verification statistics were computed and averaged over the 8 representative grid points. Some of these results are presented below.

Figure 3 presents the average bias for July. There are four bars for each forecast lead time, one corresponding, from left to right, to each of the analysis strategies. Figure 3 shows that the bias in the raw ensemble data are effectively removed by each of the analysis strategies ii, iii and iv.

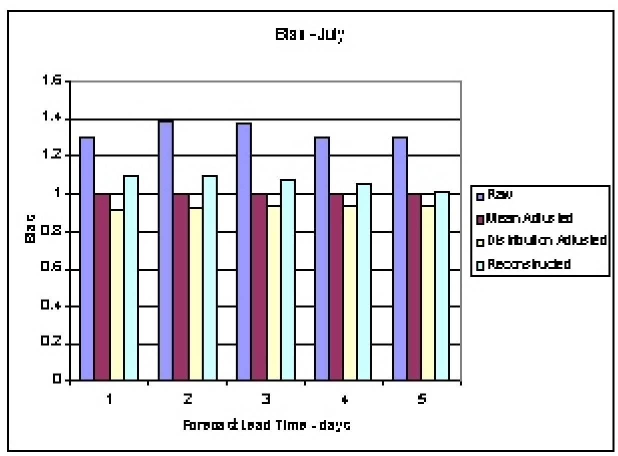

Figure 3 - Average Bias for July 24 hr Ensemble Precipitation Forecasts

Figures 4a and 4b present the average Nash-Sutcliffe efficiency for ensemble mean forecasts for January and July. The Nash-Sutcliffe efficiency is essentially a skill score for the ensemble mean. Figure 4a shows that the efficiencies in the raw ensemble data are negative, which means they are not as good as simply predicting the climatological mean to occur every day. But the bias adjustments effectively included in each of the analysis strategies ii, iii and iv corrects the problem for January, and the efficiencies are seen to be positive and to decrease with lead time. Figure 4b shows average efficiencies for July. As in Figure 4a, the effects of bias can be corrected but the average efficiency improves as the analysis strategy changes from (ii) to (iv). Nevertheless, the average efficiencies in July are less than in January. But analysis strategy (iv) gives a positive skill score for the ensemble mean, even on day 5 in the warm season.

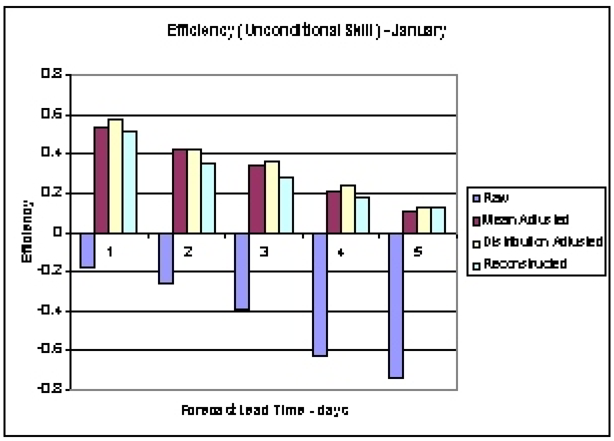

Figure 4a - Average Nash-Sutcliffe Efficiency for January 24 hr Ensemble Precipitation Forecasts

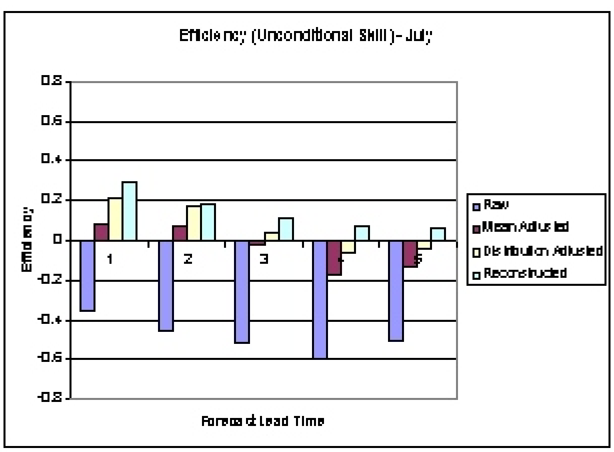

Figure 4b - Average Nash-Sutcliffe Efficiency for July 24 hr Ensemble Precipitation Forecasts

Figures 5a and 5b present the average Brier skill score for ensemble probability of precipitation forecasts for January and July. Figure 5a shows that comparable Brier skill scores for January are obtained by all of the analysis options, although options (ii) and (iii) are slightly better than the others. In July, option (iv) gives the best results and is the only option to give a positive value for day 5 forecasts.

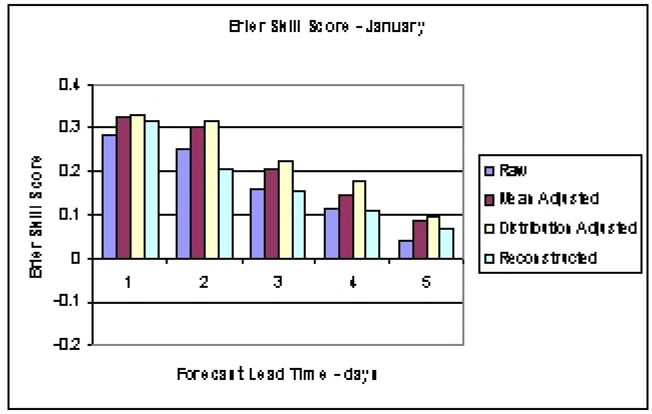

Figure 5a - Average Brier Skill Score for January 24 hr Ensemble Precipitation Forecasts

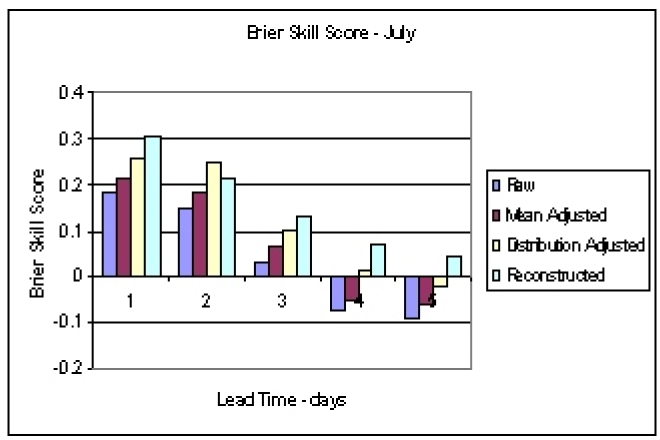

Figure 5b - Average Brier Skill Score for July 24 hr Ensemble Precipitation Forecasts

4. ENSEMBLE SPREAD

Much of the difference in verification statistics is a result of differences in representation of forecast uncertainty by the different approaches to adjusting the raw ensemble forecasts. These differences can be studied by considering differences in the spread of ensemble values. One way to analyze ensemble spread is through the distribution of forecast exceedence probabilities of the observed positive precipitation amounts. If the forecast exceedence reliability were reliable, the distribution of exceedence probabilities corresponding to the observed values would be uniform. The fraction of exceedence probabilities in different probability quantiles is known as a Talagrand diagram. The Talagrand diagram for the southeast grid point in July is presented in Figure 6. If probability forecasts from the ensemble were reliable, each bar in Figure 6 would have a value of 0.25. The raw ensemble forecasts tended to occur in just one of the 4 quartiles, a result of bias in the ensemble members. All of the other approaches remove this bias. But neither the second nor third approach that remove bias can correct the problem associated with the ensemble spread.

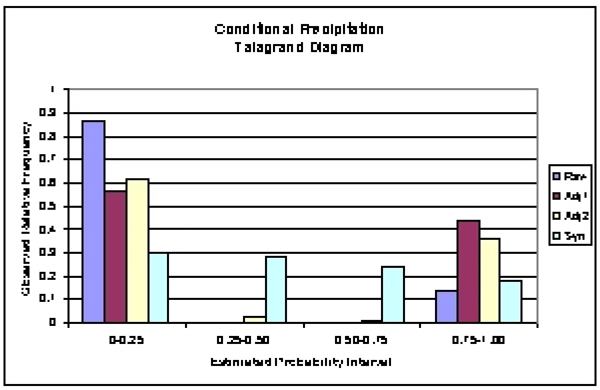

Figure 6 - Talagrand Diagram for Day - 1 24 hr Precipitation Forecasts for the Southeast U.S. for July

Because the spread is underestimated in Figure 6, observed precipitation values do not lie within the main part of the forecast probability distribution. As a result they appear in the first and fourth quartiles for analysis approaches (ii) and (iii). Analysis approach (iv) did a much better job of compensating for the underestimate of the ensemble spread. Almost equal fractions of the observations occurred in each of the quartiles.

The error in the ensemble spread can be quantified as a function of the quartile fractions of observations used to construct the Talagrand Diagram in Figure 6, as follows

$$E_s = [(f_1 + f_4) - (f_2 + f_3)]^2$$

where f_i is the fraction of observations in the i-th quartile and E_s is the measure of spread error. Average rms values of E_s are presented in Figures 7a and 7b for January and July 24 hr precipitation forecasts. Results in Figure 7a for January shows that the spread error for analysis strategies (i), (ii) and (iii) becomes smaller as lead time increases. For each of these approaches the spread of the ensemble is basically governed by the raw ensemble, although the bias adjustments tend to have a slight beneficial affect on the spread when the bias is removed. The spread error for analysis strategy (iv) is not a function of the lead time and tends to have a minimum value for all lead times.

This effect is seen much more dramatically in Figure 7b for July where the spread error also becomes smaller with lead time for analysis strategies (I), (ii) and (iii) and is not a function of lead time for stragey (iv). But the improvement by using strategy (iv) over the other strategies is greater in July than in January.

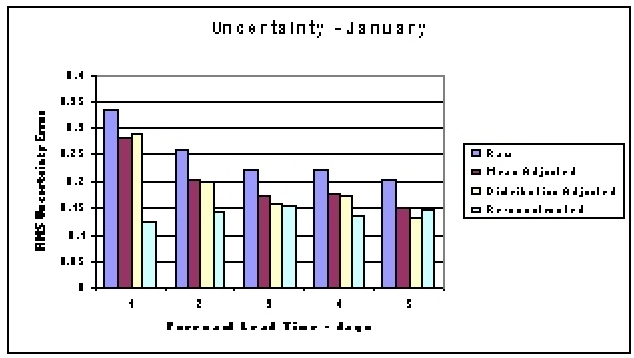

Figure 7a - Average Ensemble Spread Error for January

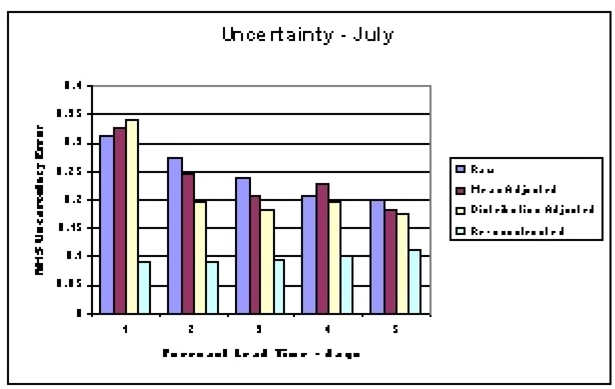

Figure 7b - Average Ensemble ?Spread Error for July

5. CONCLUSIONS

The raw precipitation ensemble forecasts from the GFS contain information that can be useful for hydrologic forecasting. But bias in the magnitude of the ensemble member values must be removed. This bias includes bias in the higher moments of the distribution of ensemble member values as well as in the mean value. These biases vary throughout the forecast domain but can be removed if there is a long enough historical archive of forecasts and corresponding observations.

Ensemble spread errors (underestimate of spread) tend to occur because the ensemble forecast system does not account for important sources of uncertainty. This occurs on average throughout the U.S. for all lead times and in all seasons. Analysis approach (iv), only using the ensemble mean together with a statistical approach to reconstruct the ensemble members, does not have a large spread error.

Joint relationships between ensemble member values at different locations in space and time will examined in future studies.

6. REFERENCES

Herr, H., E. Welles, M. Mullusky, L. Wu, J. Schaake, 2002, "Simplified Short Term Precipitation Ensemble Forecasts: Theory," Preprints of the Symposium on Observations, Data Assimilation and Probabilistic Prediction, 82nd AMS Annual Meeting, Orlando, Florida.