Multivariate Statistical Post-Processing of Ensemble Forecasts of Precipitation and Temperature over four River Basins in California

Thomas M. Hamill and Michael Scheuerer

NOAA/ESRL, Physical Sciences Division

97th AMS Annual Meeting
Univariate post-processing: temperature

Hydrological streamflow forecast models strongly rely on predictions of temperature and precipitation amounts as inputs (’forcings’).

Ensemble forecasts provide such forcings and give an idea about the associated forecast uncertainty.

Typically, statistical post-processing is required to re-calibrate the ensemble forecasts and produce reliable, probabilistic guidance.

To post-process ensemble temperature forecasts, we use a variant of the regression-type approach proposed by Scheuerer and Büermann (2014).
To post-process ensemble precipitation forecasts, we use the approach proposed by Scheuerer and Hamill (2015), modeling precipitation amounts by censored, shifted gamma distributions.

This method also accounts for an increase of forecast uncertainty with the expected amount of precipitation.

Both methods yield reliable, probabilistic forecasts at each forecast lead time and each sub-basin.
Serial dependence of temperature forecast trajectories

By calculating certain quantiles, the predictive forecast distributions can be turned back into an ensemble (of any desired size).

Univariate post-processing, however, does not provide any information about serial dependence, i.e. we don’t know how to connect the ensemble forecasts at different lead times.
Serial dependence of precipitation forecast trajectories

Serial dependence also affects accumulated precipitation amounts:
Spatial dependence of precipitation forecast trajectories

Hydrologists need to know not only the intensity of rainfall, but whether or not that intense rainfall is expected at several locations simultaneously.

No problem

a problem when marginal & joint probs. are not well forecast
Modeling spatio-temporal dependence: the Schaake Shuffle

Idea of the Schaake Shuffle:

1. Draw time series of observed temp. & precip. at this date in the past
2. Determine the rank order at each location and lead time
3. Apply these to the univariate distributions created by post-processing
Modeling spatio-temporal dependence: the Schaake Shuffle

Idea of the Schaake Shuffle:

1. use historical observation trajectories as 'dependence template'
2. re-map the values of this historical ensemble to the values of the (univariate) quantile forecasts while retaining its rank order
Modeling spatio-temporal dependence: the Schaake Shuffle

Spatio-temporal precipitation trajectories are obtained in the same way:
Modeling spatio-temporal dependence: the Schaake Shuffle

Spatio-temporal precipitation trajectories are obtained in the same way:
Main idea for improving the Schaake Shuffle

The Schaake Shuffle (Clark et al., 2004) has been used very successfully since it has been proposed. Yet, there are a few concerns:

- the re-mapping of ensemble values can be quite substantial. Are the re-mapped trajectories still realistic?
- in particular: is it adequate to assume that spatial/ temporal correlation is state-independent, i.e. for example the same for low and high levels of precipitation?

To address these concerns we propose an algorithm that selects historical trajectories whose marginal distributions already resemble the distributions of the forecast ensemble, thus reducing the required amount of re-mapping.

Temperature and precipitation are considered simultaneously, but since precipitation is more important and more complex it will be treated in a more sophisticated way.
Improving the Schaake Shuffle: step 1

Starting from a much increased set of historical dates/trajectories, we apply a simple subsetting criterion based on temperatures and discard all trajectories with too many values outside the 99.9% predictions intervals of the forecast distributions.
Improving the Schaake Shuffle: step 1

Starting from a much increased set of historical dates/trajectories, we apply a simple subsetting criterion based on temperatures and discard all trajectories with too many values outside the 99.9% predictions intervals of the forecast distributions.
Improving the Schaake Shuffle: step 2

Further thinning is based on similarity of precip. marginal distributions.
Improving the Schaake Shuffle: step 2

Similarity is measured by the divergence of the marginal distributions

\[ d(F_m, G_n) = \int (F_m(x) - G_n(x))^2 \, dx \]

where \( F_m \) is the CDF of the \( m \)-member predictive ensemble and \( G_n \) is the CDF of the \( n \) historical trajectories at a particular lead time and location.

A subset of those \( n \) trajectories will be chosen iteratively based on the mean divergence over all lead times and all locations in the basin:

- for each trajectory \( i \), we calculate the divergence \( d(F_m, G_{n,-i}) \) between \( F_m \) and the CDF \( G_{n,-i} \) obtained after omitting that trajectory.
- trajectories whose omission reduces the divergence from \( F_m \) most will be discarded, all others retained for the next iteration

This process will be repeated until only \( m \) trajectories remain.
Improving the Schaake Shuffle: step 2

The original 553 trajectories are successively reduced to 60:
Improving the Schaake Shuffle: step 2

The original 553 trajectories are successively reduced to 60:
Improving the Schaake Shuffle: step 2

Now, less re-mapping is required to match the predictive ensemble:
Application to four river basins in Northern California

In the following we consider

- 25 sub-basins in 4 different basins
- verification period: 1985 to 2010
- 6-h precipitation accumulations & 6-h average surface temperatures
- forecasts lead times up to 15 days

First, the GEFS ensemble (11 members) forecasts are calibrated with the univariate techniques outlined above, using NOAA’s 2nd generation reforecast data set.

Then, we apply the standard Schaake Shuffle (StSS) and the minimum divergence Schaake Shuffle (MDSS) described above to generate calibrated 60-member forecast trajectories for the two weather variables.
Application to four river basins in Northern California

To compare the performance of StSS and MDSS we

- use leave-one-year-out cross validation
- aggregate the temperature trajectories to a univariate quantity by considering 3-day mean temperatures, averaged over all stations within each basin
- aggregate the precipitation trajectories to a univariate quantity by considering 3-day accumulated precipitation, averaged over all stations within each basin
- calculate the continuous ranked probability skill score (CRPSS)

\[
crps(F, y) = \int (F(x) - 1(x \geq y))^2 \, dx
\]

of the StSS and MDSS forecasts for these aggregated quantities.
While the improvement of mean areal temperature (MAT) forecasts is rather limited, mean areal precipitation (MAP) forecasts at shorter lead times are noticeably better if MDSS is used instead of StSS.
Assessing the representation of spatial MAP gradients

We take a closer look at how well the two approaches reconstruct the spatial structure of precipitation fields. A useful tool for that is the variogram score (Scheuerer and Hamill, 2015):

\[
S_{VG}(F, y) = \sum_{i,j=1}^{d} \left( \frac{|y_i - y_j|^{0.5} - E_F |X_i - X_j|^{0.5}}{2} \right),
\]

where \( F \) is the \( d \)-variate forecast distribution, \( y \) is the \( d \)-variate observation vector (here: observed MAP at \( d \) subbasins).

The variogram score measures if spatial differences in MAP are adequately represented by the post-processed forecast ensemble.

To facilitate interpretation, we convert the variogram score into a skill score with respect to climatology (\( \rightsquigarrow \) VSS).
Average VSS over all four basins

Variogram skill scores for MAP (Jan)

Variogram skill scores for MAP (Apr)

Variogram skill scores for MAP (Oct)

The improvement of the spatial structure of MAP forecasts constructed via MDSS compared to StSS is small but consistent over all lead times, all seasons, and all basins.
Streamflow forecasts

Does the improvement in MAP forecasts translate into an improvement of streamflow predictions?

To answer this question we

- use the meteorological forcings constructed above as inputs to NOAA’s Community Hydrologic Prediction System (CHPS)
- generate streamflow hindcasts for the two headwater basins UKAC1 (Ukiah, Russian River) and LAMC1 (Coyote Dam, Lake Mendocino)
- during the period from 1 January 1985 to 15 September 2010
- consider 1-day average flow forecasts at UKAC1 and 3-day average flow at LAMC1
- verify against observed streamflow, using the CRPSS
At both locations, streamflow forecasts are significantly improved; this confirms the importance of modeling serial and spatial dependence of the forcings for hydrological models, and shows that the MDSS has the potential to improve upon the current gold standard StSS.
CRPSS for streamflow forecasts

At both locations, streamflow forecasts are significantly improved; this confirms the importance of modeling serial and spatial dependence of the forcings for hydrological models, and shows that the MDSS has the potential to improve upon the current gold standard StSS.

Thanks for listening!

Scheuerer, M. and L. Büermann

Scheuerer, M. and T. M. Hamill

Scheuerer, M. and T. M. Hamill

Scheuerer, M., T. M. Hamill, B. Whitin, M. He, and A. Henkel
While the improvement of mean areal temperature (MAT) forecasts is rather limited, mean areal precipitation (MAP) forecasts at shorter lead times are noticeably better if MDSS is used instead of StSS.
While the improvement of mean areal temperature (MAT) forecasts is rather limited, mean areal precipitation (MAP) forecasts at shorter lead times are noticeably better if MDSS is used instead of StSS.
While the improvement of mean areal temperature (MAT) forecasts is rather limited, mean areal precipitation (MAP) forecasts at shorter lead times are noticeably better if MDSS is used instead of StSS.
While the improvement of mean areal temperature (MAT) forecasts is rather limited, mean areal precipitation (MAP) forecasts at shorter lead times are noticeably better if MDSS is used instead of StSS.
The improvement of the spatial structure of MAP forecasts constructed via MDSS compared to StSS is small but consistent over all lead times, all seasons, and all basins.
The improvement of the spatial structure of MAP forecasts constructed via MDSS compared to StSS is small but consistent over all lead times, all seasons, and all basins.
Supplemental slide: VSS for American River Basin

The improvement of the spatial structure of MAP forecasts constructed via MDSS compared to StSS is small but consistent over all lead times, all seasons, and all basins.
The improvement of the spatial structure of MAP forecasts constructed via MDSS compared to StSS is small but consistent over all lead times, all seasons, and all basins.