Ensemble Post-Processing over Complex Terrain Using High-Resolution Anomalies

Reto Stauffer[12], Nikolaus Umlauf[1], Jakob W. Messner[12], Georg J. Mayr[2], Achim Zeileis[1]

[1] Department of Statistics, Univ. of Innsbruck, Austria
[2] Institute of Atmospheric and Cryospheric Sciences, Univ. of Innsbruck, Austria

E-mail for correspondence: `reto.stauffer@uibk.ac.at`

Abstract: Probabilistic weather forecasts computed by numerically solving physical equations describing atmospheric processes have systematic errors, particularly over complex terrain. Statistical post-processing is often applied to alleviate these errors. We will present a novel fully scalable full-distributional post-processing method for precipitation, using high-resolution local anomalies to account for the high spatial variability. The application of the new method to the central Alps improves the skill of forecasts for both the probability of occurrence and the amount of precipitation.

Keywords: GAM; BAMLSS; Censored Normal Distribution; Precipitation; Climatology and Weather.

1 Introduction & Data

In mountainous regions, large amounts of precipitation can lead to severe floods and land slides during spring and summer, and to dangerous avalanche conditions during winter. An accurate and reliable knowledge about the expected precipitation can therefore be crucial for strategical planning and to raise awareness among the public.
Precipitation forecasts are typically provided by numerical weather prediction (NWP) models using physical prognostic equations. Ensemble prediction systems (EPS) provide several independent weather forecasts based on slightly different initial conditions to depict the forecast uncertainty. A crucial limitation of these forecasts is the horizontal resolution. Therefore several approaches to correct the NWP forecasts for unresolved features and systematic errors are available, known as post-processing methods.

We present a novel spatial post-processing method for precipitation over complex terrain using high-resolution spatial climatologies as background information, and apply it to Tyrol, Austria. Due to the local alpine topography, observations vary strongly across the domain, increasing the complexity for spatial modelling. The new approach uses high-resolution climatologies to remove local features from (i) the observations and also from (ii) EPS forecasts. The remaining short-term derivations can be used to create high-resolution spatial corrected EPS forecasts.
We use an ensemble consisting of 50 forecasts computed by the EPS of the European Center for Medium-Range Weather Forecasts (ECMWF). The horizontal mesh of the current model is roughly 32 *km* (see Figure 1, left). Approximately 2200 single days (2010–2015) are used, including 90 grid points from the EPS model covering the area of interest. Precipitation observations at 117 stations cover the period 1971–2013 and constitute roughly 1.5 million unique observations.

2 Censored Spatio-Temporal Anomaly Model

Ensemble model output statistics (EMOS; Gneiting et al., 2005) model the statistical relationship between past observations and the corresponding EPS forecasts. As the EPS provides 50 individual forecasts, the corrections can be accounted to both, the expected mean, and the uncertainty of the EPS, typically represented by the EPS standard deviation.

$$y \sim \mathcal{N}\big(\mu, \sigma\big) \quad \text{with: } \mu = \beta_0 + \beta_1 m(\text{eps}),\ \sigma = \gamma_0 + \gamma_1 s(\text{eps}) \tag{1}$$

Gneiting et al. (2005) proposed that the response y is assumed to follow a normal distribution with location μ represented by a linear function of the EPS mean (m(eps)), and standard deviation σ represented by a linear function of the EPS standard deviation (s(eps)).
However, for the application of high-resolution precipitation post-processing on a daily time scale, two major problems arise. Daily sums of observed precipitation are *no longer normally distributed*, as they contain a large fraction of zero-observations (dry days), and the observations show a *large variability across the area of interest* – especially over complex terrain like e.g., the Alps.
To account for the distribution of the observations, the conditional response distribution in Equation 1 has to be modified first. Messner et al. (2014) showed that the response distribution of precipitation can be seen as left-censored normal, as precipitation is physically limited to 0 *mm*.
Furthermore, a way has to be found to include the information of all available stations within the area of interest, but to account for the different location and season dependent characteristics across the domain at the same time. Therefore we are using the concept of local standardised anomalies, based on high-resolution precipitation climatologies. Both, the observations

and all 50 individual EPS forecast members, will therefore be standardised using:

$$y^* = \frac{y - \mu_{y,clim}}{\sigma_{y,clim}} \tag{2}$$

While the climatological location $\mu_{y,clim}$ and scale $\sigma_{y,clim}$ represent the long-term spatio-temporal patterns in both, the observations and the individual EPS forecast members respectively (y), anomalies are the short-term deviations from the underlying climatology. By removing location and season dependent characteristics, the observations and the EPS forecasts can be brought to a compareable level, what will be called "standardised anomalies", denoted by superscript "*".
We are using a Bayesian framework estimating generalized additive models for the climatologies (R package `bamlss`, Umlauf et al., 2016) to estimate heteroscedastic spatio-temporal climatologies of the observations, and the EPS forecasts. Therefore, similar assumptions to Equation 1 will are used, replacing the linear predictors for μ, and $\log(\sigma)$ by (Stauffer et al., 2015):

$$\beta_0 + \beta_1 \text{alt} + s(\text{yday}) + s(\text{lon}, \text{lat}) + s(\text{yday}, \text{lon}, \text{lat}) \tag{3}$$

The linear predictor includes a linear altitudinal, a cyclic seasonal ($s(\text{yday})$), a 2-D spatial ($s(\text{lon}, \text{lat})$), and a 3-D effect ($s(\text{yday}, \text{lon}, \text{lat})$) to account for changes in the seasonal pattern across the area of interest. Once the climatologies are known, the statistical relationship between standardised anomalies of the observations, and the standardised anomalies of the EPS forecasts can be modelled similar to the EMOS approach in Equation 1 using:

$$y^* \sim \mathcal{N}\big(\mu, \sigma\big) \text{ with: } \mu = \beta_0 + \beta_1 m(\text{eps}^*), \ \log(\sigma) = \gamma_0 + \gamma_1 \log(s(\text{eps}^*)) \tag{4}$$

As the standardised anomalies are no more location dependent, the prediction for any location within the area of interest can be made. This allows for a spatial correction of any future EPS forecast on an arbitrary fine horizontal resolution.

3 Summary

The novel approach for precipitation using anomalies provides an attractive and reliable new method for spatial ensemble post-processing. Once the climatologies are estimated, the computational costs are very low. Regarding the full probabilistic response, several quantities can be derived from one single model, like the expected amount of precipitation, quantiles, or probabilities. Figure 1 shows spatial sample prediction on a 800 m grid for a +30 h forecast, comparing the raw EPS mean (left) against the corrected forecasts (middle). In contrast to the EPS, several topographical features can be identified after the correction. Beside, probabilities for exceeding two different thresholds are plotted. First results have shown that the novel approach applied to the area of Tyrol, located in the Eastern Alps, increases the forecast skill for both, the probabilities of exceeding a certain threshold, and the amount of precipitation.

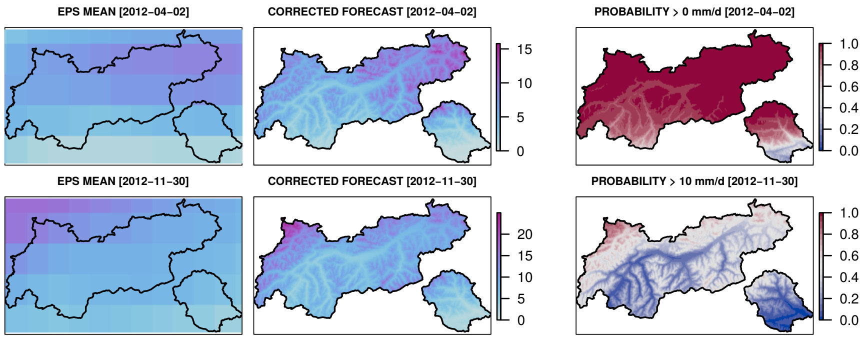

FIGURE 1. Sample predictions. Top: 2012-04-02, bottom: 2012-11-30. Left to right: raw uncorrected EPS forecast $[mm\ d^{-1}]$, corrected forecast $[mm\ d^{-1}]$, and probability of occurrence. Top $> 0\ mm\ d^{-1}$, bottom $> 10\ mm\ d^{-1}$. The color scale for the uncorrected and corrected forecast is identical for each individual day.

Acknowledgments: Ongoing project funded by the Austrian Science Fund (FWF): TRP 290-N26. The computational results presented have been achieved in part using the Vienna Scientific Cluster (VSC). Observation data set provided by the hydrographical service Tyrol (`ehyd.gv.at`).

References

Gneiting T., et. al. (2005): Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *Monthly Weather Review*, **133**, 1098–1118.

Messner J. W., et. al. (2014): Extending Extended Logistic Regression: Extended versus Separate versus Ordered versus Censored. *Monthly Weaather Reveview*, **142**, 3003–3014.

Stauffer, R., et. al. (2015). Spatio-temporal Censored Model of Precipitation Climatology. In: *Proceedings of the International Workshop of Statistical Modelling 2015*, Linz, Austria, 366–371.

Umlauf, N., et al. (2016): `bamlss`: Bayesian Additive Models for Location Scale and Shape (and Beyond). R *package version* `0.1-1` (https://R-Forge.R-project.org/R/?group_id=865).