

# Calibrating forecasts of heavy precipitation in river catchments

**AMARILLA MÁTRAI** (General Directorate of Water Management, Hungary),  
**ISTVÁN IHÁSZ** (Hungarian Meteorological Service)

Ensemble forecasts of severe weather can provide valuable information on the range of possible scenarios and the likelihood of their occurrence. However, to make sure ensemble forecasts are reliable they need to be well calibrated. We have used a re-forecast-based method called quantile mapping to calibrate ECMWF ensemble forecasts (ENS) of precipitation. High-quality forecasts of heavy precipitation can assist hydrologists in their decision-making. We have therefore investigated re-forecast-based ensemble calibration for 120 extreme events in the catchments of the rivers Danube and Tisza in the period from 2008 to 2013. Although there are limitations when applying the method to extreme events, we found the calibration to be useful for the case of the extreme floods that occurred in May and June 2013 along the Danube.

## Comparing model and observed climates

ECMWF has regularly provided ensemble re-forecasts since March 2008 (Hagedorn, 2008; Gneiting, 2014). Ensemble re-forecasts are generated by using the current model version to produce forecasts for previous years within a time window starting on the current date. Today 11-member 46-day re-forecasts are operationally generated for the last 20 years every Monday and Thursday. In the period investigated (2008–2013), five-member re-forecasts were available once a week (on Thursdays).

Ensemble calibration (Box A) can bring valuable improvements if there is a significant difference between the probability distributions of model and observed climates (Ihász *et al.*, 2010). Significance was investigated with two-sample Kolmogorov-Smirnov tests. A stable

## Calibration method

A

Calibration is the statistical adjustment of a forecast to improve its quality. In the approach presented here, to perform a calibration the following data are needed, averaged over each river catchment area:

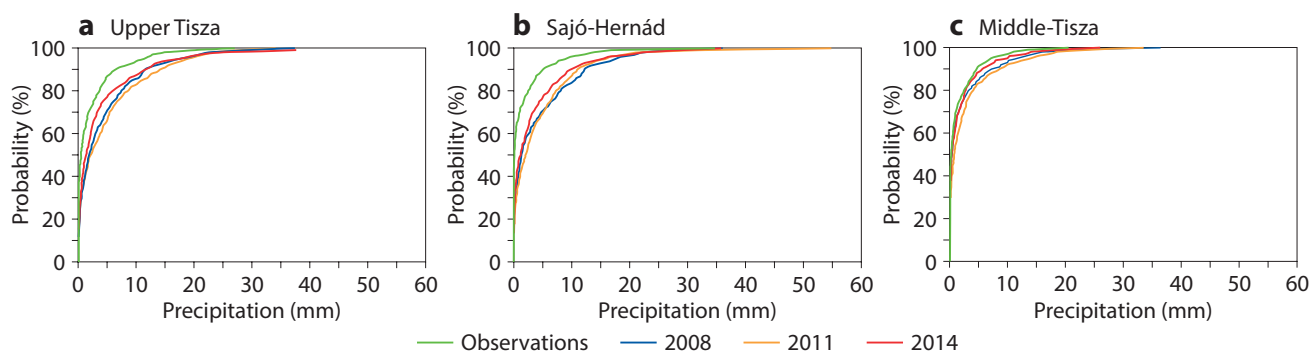
- Model climate: the model climate cumulative distribution function (CDF) derived from re-forecasts.
- Observed climate: the CDF derived from observed data.
- Ensemble forecast: the CDF of the current ENS forecast.

To calibrate the ensemble forecast, it was adjusted by the difference between the observed climate and the model climate. A greater difference between the climates requires a greater adjustment of the ensemble forecast. If the observed climate and the model climate are close, the required adjustment is small.

It is important for all climate CDFs to cover the same period of time. If the period under consideration is too short, it may not include any extreme events. The result of the calibration is an adjusted ENS CDF. Forecasters can compare this with the uncalibrated, raw CDF to help them decide whether or not to adjust the precipitation forecast.

model climate can be produced by using re-forecasts from five consecutive weeks centred on the current date. A model climate was produced for each week of every year in the selected period (2008–2013).

Differences between the probability distributions of model and observed climates are liable to change as a result of changes to the model (Figure 1). For example, the horizontal resolution of ENS was 50 km between 2006 and 2010 and 32 km (up to day 10) between 2010 and 2016. The vertical resolution was 62 levels between 2006 and 2013 and it has been 91 since 2013.



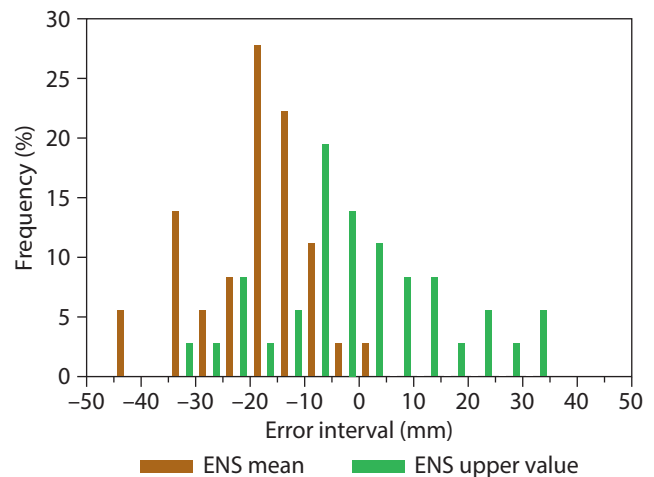
**Figure 1** Cumulative distribution functions for 20-year model climates for 24-hour precipitation based on 78-hour re-forecasts over a five-week period centred on the end of May, using ECMWF model versions operational in 2008, 2011 and 2014, and for the observed climate for (a) a mountainous catchment area (Upper-Tisza), (b) a mixed catchment area (Sajó-Hernád), and (c) a flat catchment area (Middle-Tisza). The cumulative distribution functions show the probability that the amount of 24-hour precipitation will not exceed a given threshold.

To investigate the differences between model and observed climates, we compared the observed climate with consecutive model climates for 20 individual catchment areas of the Danube and Tisza rivers. The catchments were divided into three catchment types: flat, mountainous and mixed. Model climates for different years were also compared to capture the impact of changes to the model.

The following general conclusions can be drawn:

- There tend to be considerable differences among the model climates for the same catchment depending on the model version used.
- The model climates based on the model versions operational in 2011 and 2008 are closer to each other than those based on 2014 and 2011.
- The differences between the model and observed climates are relatively small for small to moderate amounts of precipitation in flat regions. As a result, there is generally no need for calibration in these cases. This is especially true for the 2014 model climate.
- In the case of mountainous or mixed catchments and generally in the case of heavy or extreme precipitation, the differences are larger, so calibration is beneficial.
- The smallest differences between the model climate and the observed climate can be seen in the climate based on the model version operational in 2014.

Seasonal and annual similarities and differences were examined by applying the Kolmogorov-Smirnov test to model climates. Model climates based on the model versions operational in 2008 and 2014 were considered in order to capture the influence of model developments. A similar investigation was carried out for the observed and the 2014 model climate to discern the strengths and weaknesses of the model and to support decision-making in situations when there is a risk of flooding. Results show that larger differences usually appear in summer due to more intense convection. The largest differences between the model and observed climates for 2014 appear in spring and summer. The largest differences between model climates (2008 and 2014) were found in summer. This highlights the positive impact of model development on convective precipitation forecasts.

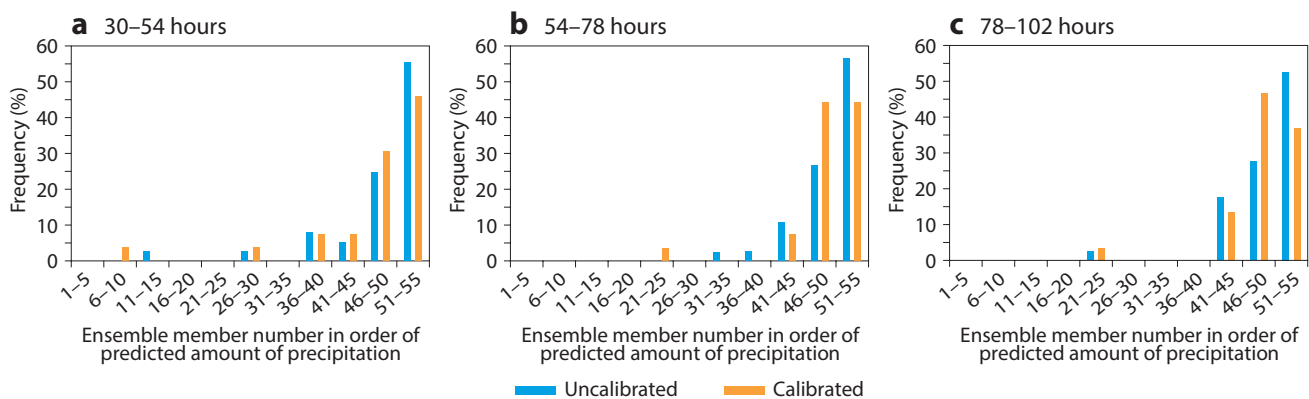


**Figure 2** Error distribution of uncalibrated ensemble forecasts of 24-hour precipitation 30 to 54 hours ahead, for 120 cases of extreme precipitation in the period from 2008 to 2013. The chart shows the frequency in per cent for the ensemble mean (brown) and for the ensemble member predicting the largest amount of precipitation (green).

**Verification of 120 extreme events**

Figure 2 shows the error distribution of uncalibrated ensemble forecasts for 120 extreme 24-hour precipitation events in the upper Danube area in the period from 2008 to 2013. It can be seen that the ensemble mean tends to underestimate the amount of precipitation in these cases. The ENS member predicting the largest amount of precipitation under and over-estimates the observed precipitation amount in approximately the same proportion.

For the verification of ENS forecasts the Talagrand diagram is widely used. This type of diagram shows how often observations match different parts of an ensemble forecast distribution. To this end, the ensemble forecast distribution is divided into bins of equal size by ensemble member number, for example going from low predicted amounts of precipitation to high predicted amounts of precipitation. In a reliable ensemble forecast, the frequency of observations in each bin will be the same as each part of the ensemble forecast distribution is equally likely.



**Figure 3** Talagrand diagrams for calibrated and uncalibrated ensemble forecasts of 24-hour precipitation for 120 cases of extreme precipitation in the period from 2008 to 2013, for lead times of (a) 30 to 54 hours, (b) 54 to 78 hours and (c) 78 to 102 hours

Figure 3 shows such diagrams for uncalibrated and calibrated ensemble forecasts and different lead times for 120 cases of extreme precipitation in the upper Danube area in the period from 2008 to 2013. The distribution is slightly more even in the case of calibrated forecasts, which means that extreme events are less likely to be outliers in the ensemble forecast distribution. The calibration method thus improved forecasts of extreme precipitation.

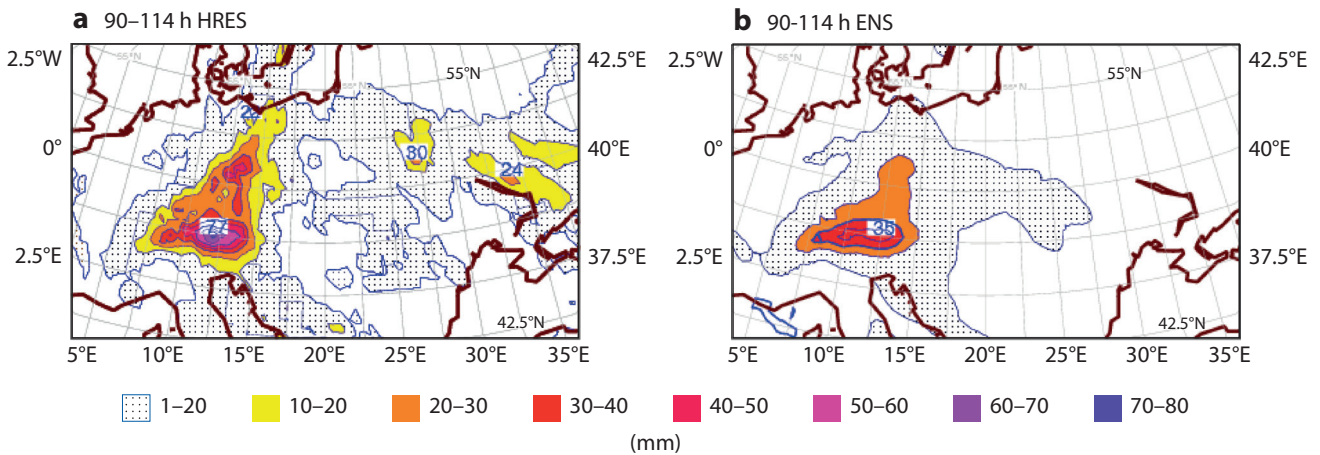
**Danube flood May–June 2013**

In late May and early June 2013, due to intense cyclonic activity over a few days in the area of the Alps, a severe flood event caused massive damage in the upper Danube region. In Hungary the water level exceeded the previous record level reached in 2002 in most parts of the river except the southern part, near the Hungarian-Serbian border. The flooding was caused by extreme precipitation that fell over the course of four days in the three upper catchments of the Danube. The largest amount of daily precipitation was recorded on 2 June 2013: an average amount of 34.6 mm/24 h in the upper Danube region,

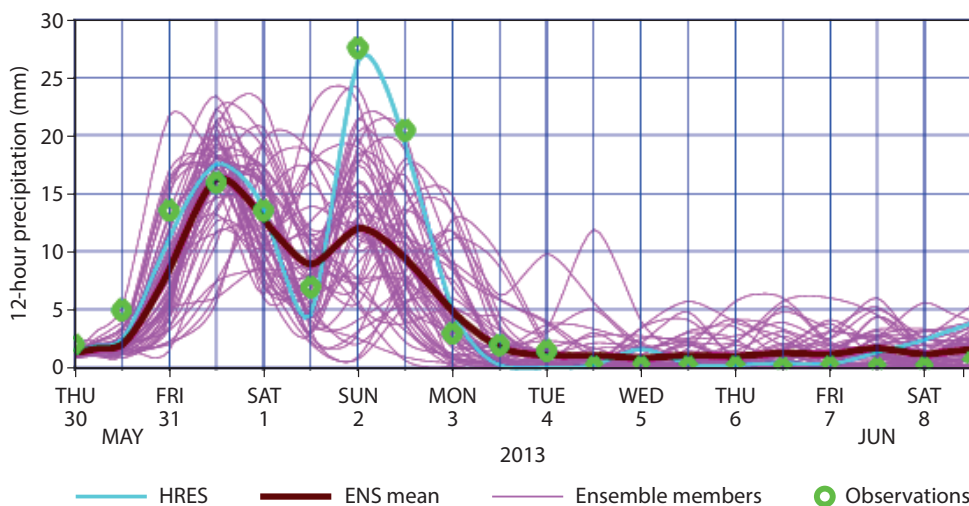
48.2 mm/24 h in the Inn region, and 53.1 mm/24 h in the Traun-Enns region.

Figure 4 shows ECMWF’s 90-hour high-resolution forecast (HRES) and ensemble forecast (ENS) of 24-hour precipitation starting at 12 UTC on 30 May 2013. It can be seen that the area of intense precipitation was well predicted. However, the HRES over-predicted and the ENS mean under-predicted the daily precipitation amount by about 10–20 mm throughout the period. It is important to note that the position and the intensity of the extreme event were well predicted by both HRES and ENS several days ahead.

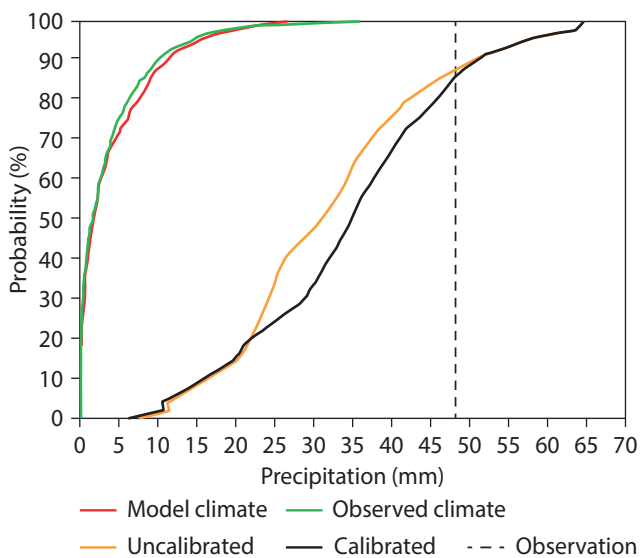
Figure 5 shows an ENS 12-hour precipitation plume and HRES 12-hour precipitation forecast for the upper Danube area starting at 00 UTC on 29 May 2013. The forecast predicts intense precipitation between days 2 and 5, and this is when heavy rain was indeed observed. The ENS prediction comes with a large spread and the ENS mean is much lower than the HRES on day 4. However, the ENS spread decreased in subsequent, shorter-term forecasts.



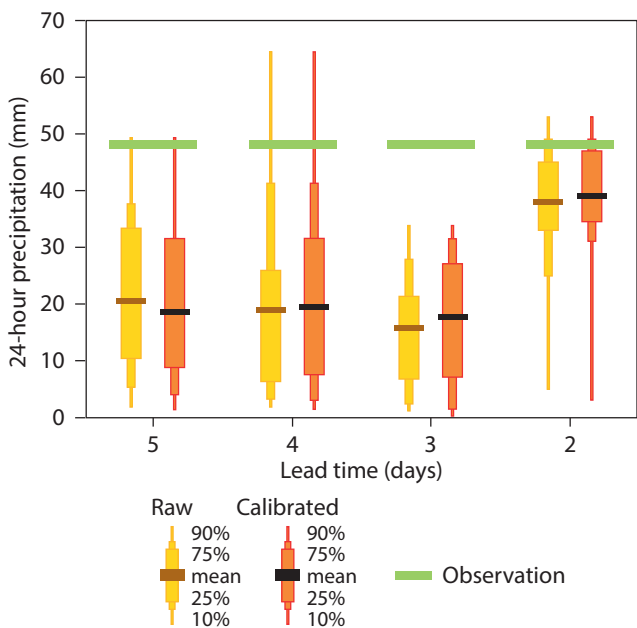
**Figure 4** Precipitation forecasts starting at 12 UTC on 30 May 2013 showing (a) the HRES 24-hour precipitation forecast 90 to 114 hours ahead and (b) the ENS mean 24-hour precipitation forecast 90 to 114 hours ahead.



**Figure 5** ENS 12-hour precipitation plume and HRES forecast for the upper Danube area starting at 12 UTC on 29 May 2013.



**Figure 6** Calibrated and uncalibrated two-day 24-hour precipitation forecasts starting at 00 UTC on 31 May 2013 for the Inn area, shown together with the observed climate and the model climate for that area at that time of year. The vertical dashed line shows the observed value.



**Figure 7** Calibrated and uncalibrated 24-hour precipitation forecasts valid for 06 UTC 1 June 2013 – 06 UTC 2 June 2013 for the Inn catchment area, initialised on four consecutive days starting from 06 UTC on 27 May. The horizontal line shows the observed value.

Figure 6 shows the effect of calibrating the two-day ENS for 24-hour precipitation in the Inn area starting from 00 UTC on 31 May 2013. In this case the observed climate and the model climate are fairly close together. For precipitation amounts up to about 22 mm, the model climate tends to be wetter than the observed climate, while beyond 22 mm it is drier. As a result, the calibration adjusts the ensemble forecast, which predicts a high probability of precipitation above 22 mm, towards even higher probabilities for large amounts of precipitation. However, there is no difference between the calibrated and the uncalibrated forecast beyond 50 mm because of a lack of re-forecast and observational data in that range.

Figure 7 shows how calibrated forecasts are shifted slightly towards higher precipitation values compared to uncalibrated ones in 4, 3 and 2-day forecasts of 24-hour precipitation in the Inn area valid for 06 UTC 1 June to 06 UTC 2 June 2013. It can be seen that the calibration moves the forecast slightly towards the observed value of 48.2 mm. Comparing the raw forecast with the calibrated forecast, forecasters can decide whether or not they need to modify the predicted amount of precipitation.

**Summary**

We have shown that ensemble precipitation forecasts can be improved using the calibration technique presented here. The observed and model climates were easy to produce from observational data and ECMWF re-forecasts. The model climate should be compared with the observed climate in each river catchment area separately because the differences in the climates can depend on differences in terrain. In our investigation we used regional averages in the calibration. However, in principle it would be better to apply the calibration to individual grid points since the forecasting model uses grid point data.

**FURTHER READING**

**Gneiting, T.**, 2014: Calibration of medium-range weather forecasts. *ECMWF Technical Memorandum No. 719*.

**Hagedorn R.**, 2008: Using the ECMWF reforecast dataset to calibrate EPS forecasts, *ECMWF Newsletter No. 117*, 8–12.

**Ihász, I., Z. Üveges, M. Mile & Cs. Németh**, 2010: Ensemble calibration of ECMWF’s medium-range forecasts. *Időjárás*, **114**, 275–286.