Calibration of Aladin EPS Precipitation Forecasts by Logistic Regression

Lovro Kalin[1], Alexander Kann[2]

[1] Meteorological and Hydrological Service (DHMZ), Croatia, kalin@cirus.dhz.hr
[2] Zentralanstalt für Meteorologie und Geodynamik (ZAMG), Austria, alexander.kann@zamg.ac.at

1 – Introduction

For calibration of precipitation, various techniques are proposed in the literature (Bayesian model averaging, analogue method etc.) Based on paper by Hamill et al. (2008), we applied a logistic regression to a larger set of Aladin EPS forecasts. First attempt in 2008 – with a very small dataset - gave significant improvement only for small precipitation thresholds: www.rclace.eu/File/Predictability/2008/report.doc

2 – Method

Logistic regression is represented by

$$P(O>T)=1.0-\frac{1.0}{1.0+\exp\left\{\beta_0+\beta_1\left(\bar{x}^f\right)+\beta_2\left(\sigma^f\right)\right\}}.$$

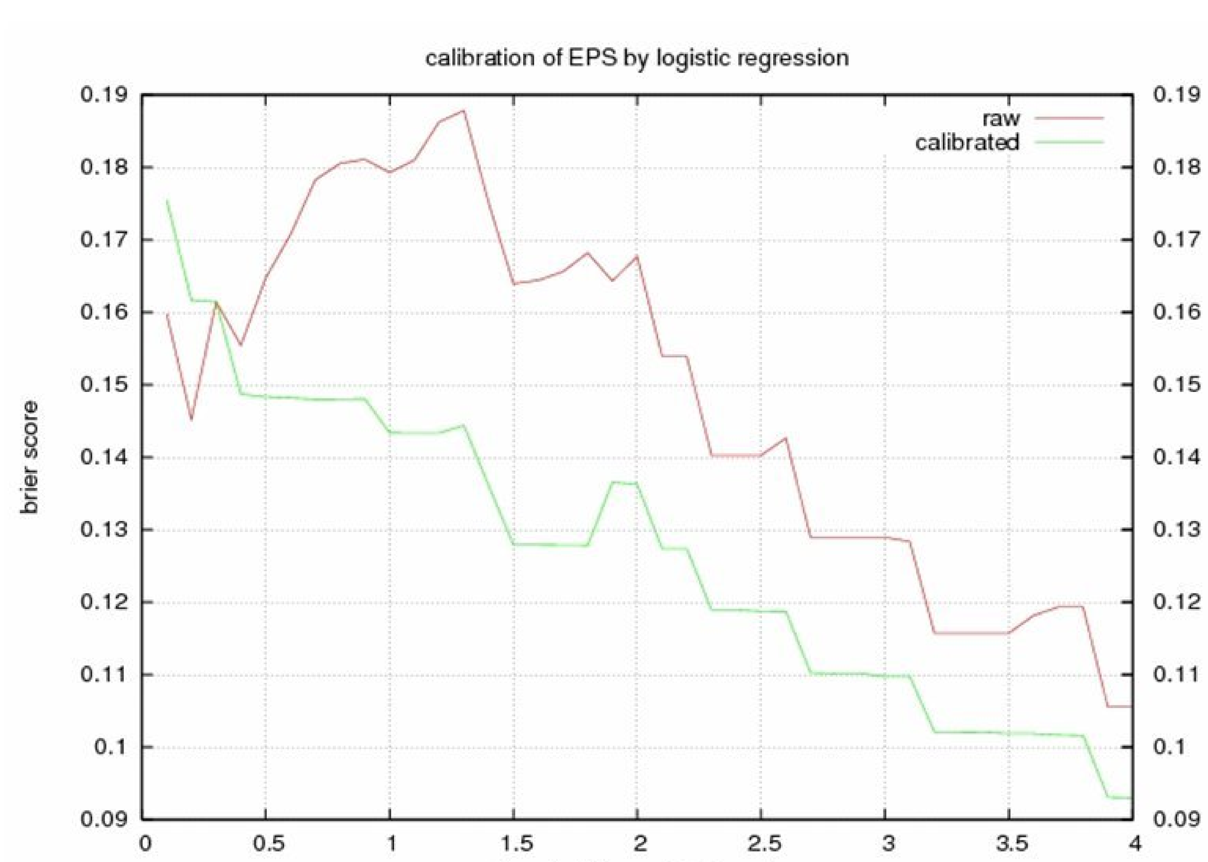
where P iz probability that precipitation O (predictand) will reach certain threshold T. In this case, the predictors are:

x - ensemble mean
σ - spread of ensemble (RMSE)

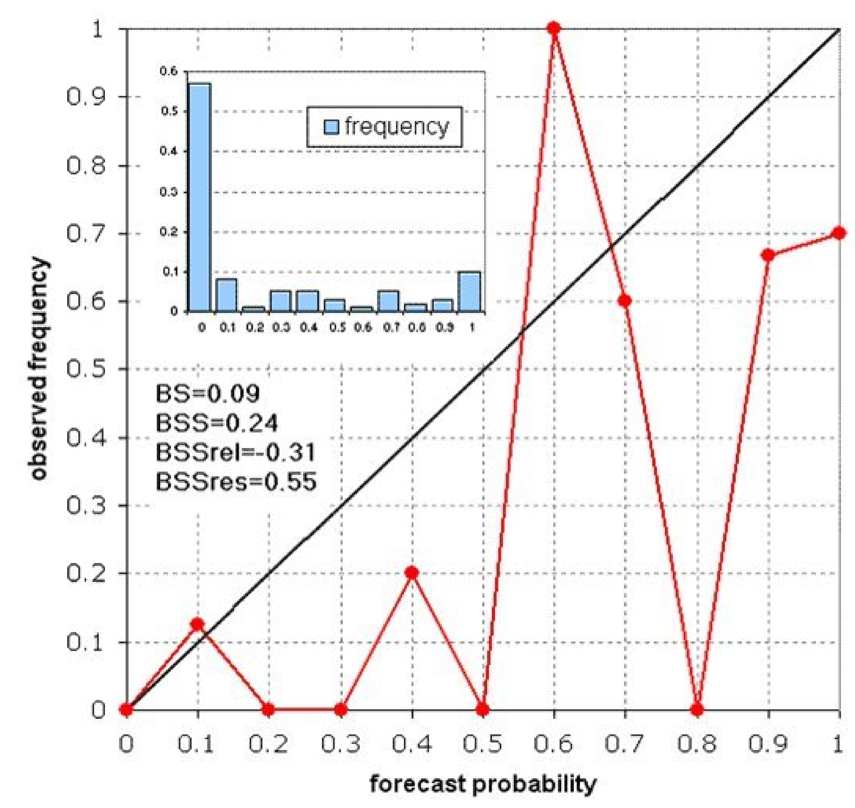
β0, β1 and β2 are coefficients of regression, that have to be obtained. We used method of least squares, fitting probability forecast (fraction of ensemble members exceeding O) to measurement (1 or 0).

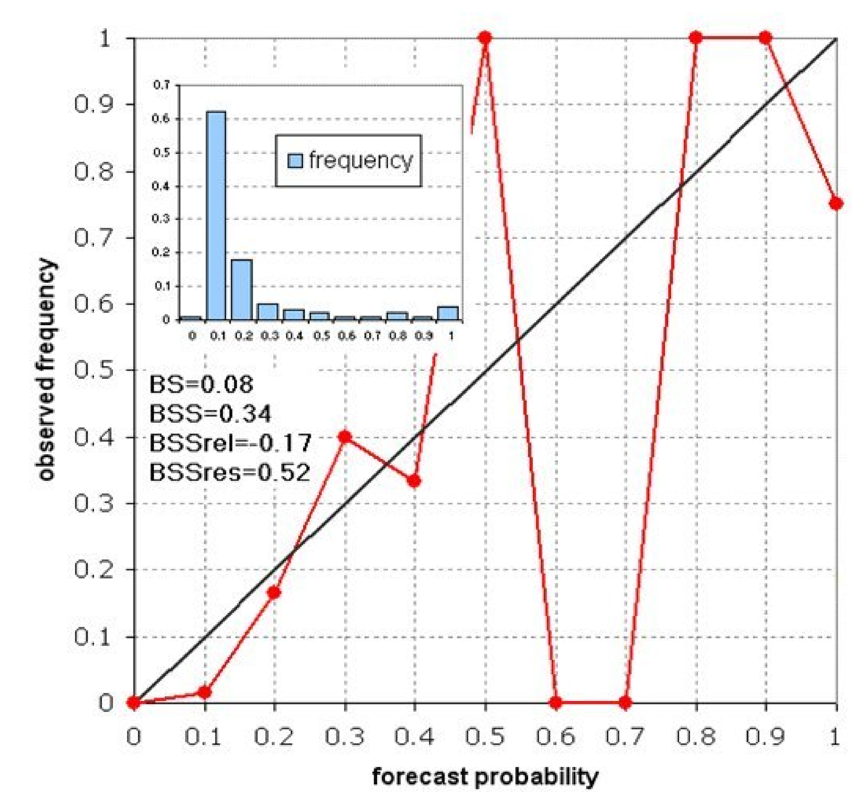
3 – Results

Testing sample was period from 2008 to 2009, for station Wien Hohe Warte (Austria). 24-hour precipitation forecasts have been verified, taking different rain thresholds.



Results for Brier score (above) show **significant positive impact for thresholds from 0.5 to 4 mm/24 hr.** For larger thresholds, the error of the calibrated forecast is similar or even higher than for raw forecast (not shown). This is probably due to insufficient training period, and also fact that in this range mean errors are very small. Also, raw forecast exhibits better skill for very small thresholds (less than 0.5 mm/24hr)





Reliability plots - for 1 mm threshold, for raw (upper left) and calibrated (lower left) forecasts - show that the distribution of calibrated probability forecasts (bar plot) is less U-shaped than the raw one. Therefore, calibration reduces sharpness of the forecasts, with lower frequency of 100% probability fc. However, forecasts up to 40%, that are most represented, are much better calibrated (red line closer to the diagonal).

Furthermore, decomposition of BS to reliability and resolution (Wilks, 2006), shows thats the improvement of the skill is in this case more influenced by improvement of the reliability, rather than improvement of the resolution. This confirms previous conclusion that calibrated forecasts are not much "sharper", but they are more accurate.

4 – Conclusions

Application of a larger dataset enabled considerably skillful results of calibration of the forecasting system. Comparison to previous results - obtained with smaller sample - show that the **increase of the training period has significant positive impact to the calibration process.**

Still, numerous problems exist, such as behavior of the system in dependence of forecast range, duration of event and particularly for threshold of precipitation amount.

References

Hamill, T. M., R. Hagedorn, and J. S. Whitaker, 2008: Probabilistic forecast calibration using ECMWF and GFS ensemble reforecasts. Part II: precipitation. Mon. Wea. Rev., **136**, 2620-2632

Wilks, D. S., 2006: Statistical Methods in the Atmospheric Sciences, 2nd Ed., Academic Press, 627 pp.