Comprehensive study of the calibrated EPS products

István Ihász, Máté Mile and Zoltán Üveges Hungarian Meteorological Service, Budapest, Hungary

1. Introduction

Calibration of ensemble forecasts is a new challenge in meteorology (Hamill and Hagedorn, 2007), in most cases different techniques are used for different types of meteorological variables depending on their distribution function. Even though a number of derived ensemble products (e.g. forecast probabilities) are provided directly by ECMWF, it encourages its members and co-operating states to make local developments. Therefore while reforecast products have been available for calibration purposes (Hagedorn, 2008) since 2008, calibration itself is not made by ECMWF, calibration methods are expected to be developed locally for local needs. We have successfully applied a simple calibration method for several meteorological variables such as 2m temperature, daily precipitation, 10m wind speed, which have completely different distribution functions. Second part of this paper shows the method and verification results. This method was successively introduced operationally in the second part of 2009 at the HMS.

2. Method of the calibration

Several ways can be followed to make ensemble calibration, the most simple method is the so called bias correction, when systematic model error is eliminated. Our chosen calibration method is based on empirical distribution functions, when forecasted ensemble distribution is compared to the distribution of the model climate and observed climate. The quality of the model climate is an important point of this procedure.

The classical and simplest method is to create a model climate from several years of ensemble forecasts of the past. It is important to note, there are always some differences between past and recent behaviour of the model due to model developments.

In recent years to avoid this problem a new method was started at the largest centres (NCEP, ECMWF), the operationally generated reforecast datasets. A reforecast dataset is a collection of forecasts with start and prediction dates from the past, usually going back for a considerable number of years (Hamill and Whitaker, 2007, Hagedorn, 2008). In November 2006 daily operational forecasts were extended from 10 to 15 days by the introduction of the VarEPS system (Buizza, et al., 2006). By the middle of March 2008 VarEPS was merged with the former monthly forecast. At the same time a new reforecast method was introduced, i.e. a 5 ensemble member reforecast datasets operationally produced weekly for the last 18 years (Hagedorn, 2008). Model climate distributions can be generated from 5 consecutive weekly reforecast datasets coming from ECMWF's Meteorological Archival and Retrieval System (MARS).

Hagedorn (2008) proved the usefulness of calibration for 2m temperature and encourage the member and co-operating states to develop and apply calibration procedures for their local needs. Hagedorn used non-homogeneous Gaussian regression method based (Hamil and Whitaker, 2007) on daily value of archived observation and model. Verification results showed that the success of calibration is partly derived from simple bias correction, but additional improvements could be achieved by ensemble calibration. 2m temperature ensemble forecast was calibrated for 250 European stations and a verification measure, the so called continuous rank probability score (CRPS) showed that calibrated EPS forecasts for day 4 are as good as uncalibrated forecasts for day 1. The disadvantage of this method, is that it can be used only for meteorological variables which have Gaussian distribution. Another type of calibration method can be applied for non Gaussian meteorological variables, like precipitation (Hamill et al., 2007).

In early 2007 HMS planed to develop calibration algorithm, test and verify the system. A quite simple method was chosen which does not need the a priori knowledge of cumulative density functions of meteorological variables and does not need daily values of archived observation of ensemble forecasts. As it was presented at the end of 2007 Webster (2007) successfully used similar method for calibration of daily precipitation for the water basin of river Ganges and Brahmaputra. In the same time our first verification results were presented and published for 2m temperature (Ihász, 2007). After promising results several meteorological variables, like 2m minimum and maximum temperature, 2m temperature at 00 and 12 UTC, 24 hour accumulated precipitation and 00 and 12 UTC wind speed was chosen to be calibrated. Ten reliable Hungarian meteorological stations were selected for our investigations in both periods (1971-2000 and 1990-2007. In first part of our work a calibration method was developed, model climate was created from the +48 h reforecast control model for the period of 1971-2000 (Ihász, 2007, Ihász and Mile, 2008). Due to the fact model climate was generated from +48 h control model forecast calibration could be aimed between +24 and +120 hours (Mile, 2008).

After having quite promising verification results, in the second part of our investigation, we started to adapt our calibration system using the new type of reforecast products when the 5 member reforecast had became available in March 2008.

3. Comparison of model and observation climate distribution

Before having experiences on calibration it is useful to investigate and compare the monthly distributions of model and observed climates. This investigation was done for the middle of the months for several meteorological variables. Some typical and interesting results are shown on fig. 1 and 2).

3.1 2m temperature at 12 UTC

Fig. 1 shows the distribution of 2m temperature at 12 UTC in Pécs and Szombathely for April and July. It can be seen in April that the model overestimates temperature especially in Pécs. In July the model slightly overestimates the temperature in case if it is below average and it is quite correct if it is higher than average. We can easily see simple bias correction can not be applied successfully and realise the potential benefit of calibration using two distribution functions.

3.2 Daily precipitation

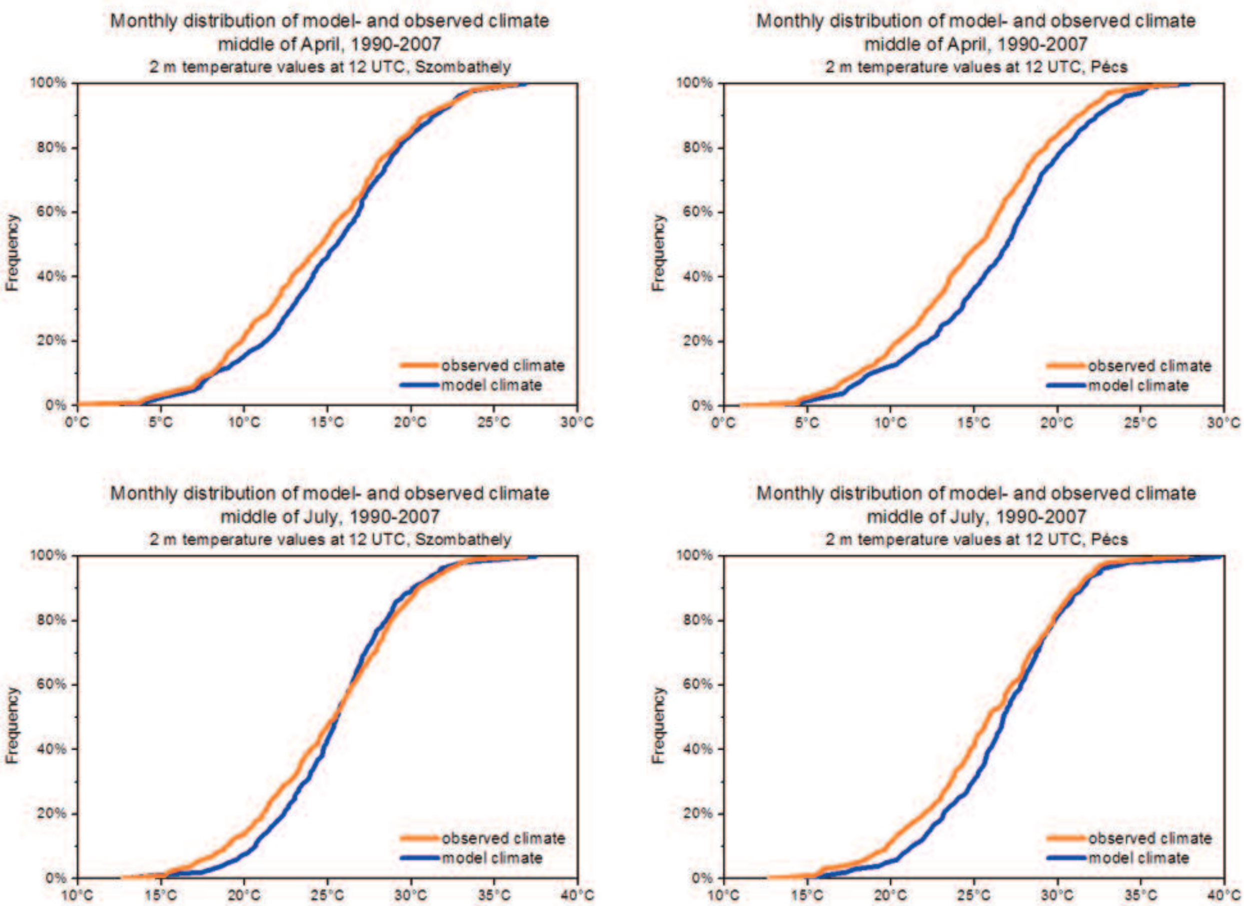

Fig. 1 Monthly distribution of 2 m temperature at 12 UTC: blue – model climate: orange – observed climate: (Szombathely and Pécs: April /upper row/ and July /lower row/)

Fig. 2 shows the distribution of daily precipitation in Miskolc and Nagykanizsa in April and July. It can be seen, that in all cases the rate of small and moderate precipitation is overestimated meanwhile the rate of daily precipitation exceeding 10 mm / day is underestimated.

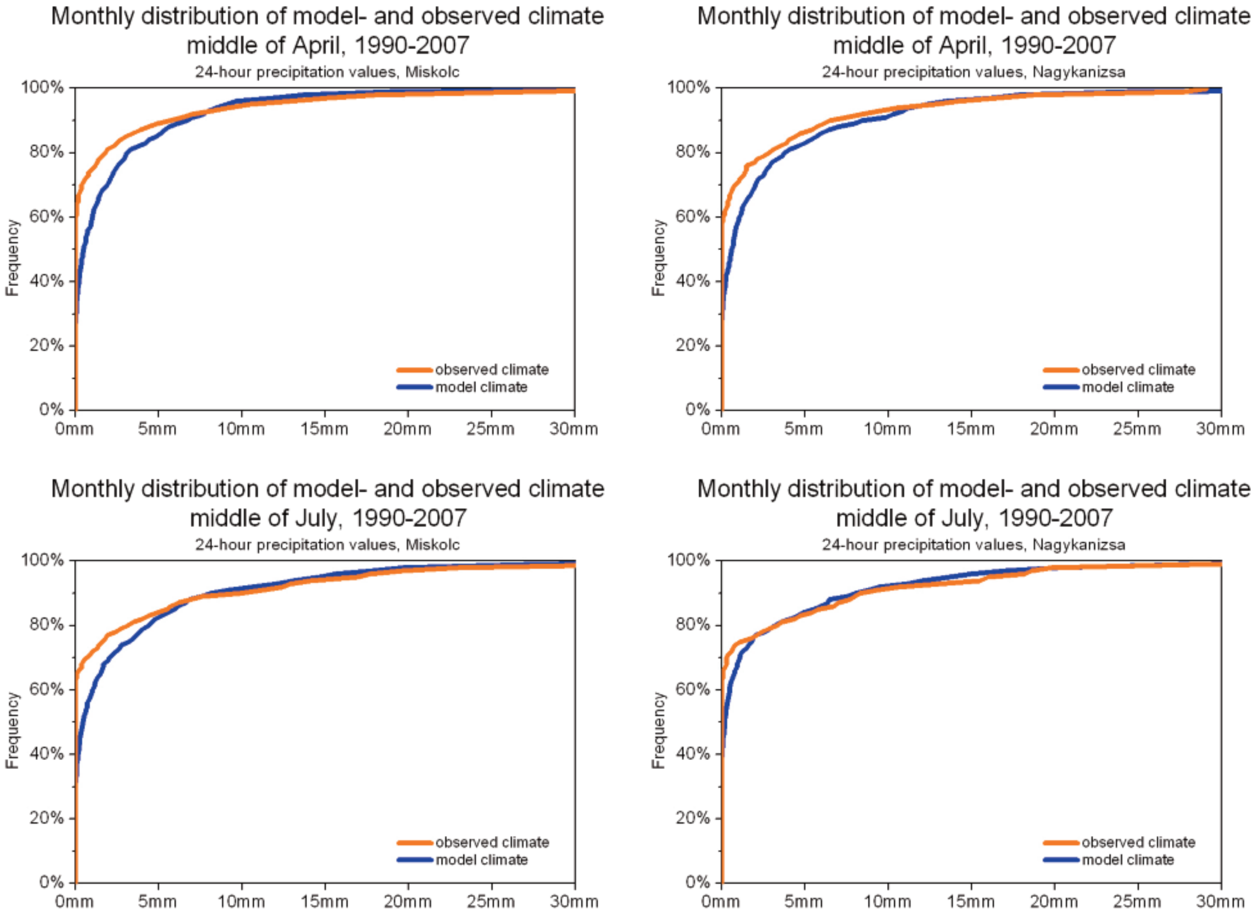

Fig. 2 Monthly distribution of daily precipitation: blue – model climate: orange – observed climate: (Miskolc and Nagykanizsa: April /upper row/ and July /lower row/)

4. Comprehensive verification of uncalibrated and calibrated ensemble forecasts

As it was mentioned, at the ECMWF once a week a 5 member reforecast has been made for the latest 18 years since 13 March 2008. We succeeded a quite stable model climate generated from the superset of 5 consecutive reforecasts centred around the current week. Success of calibration is highly depends on quality of model climate, so it can be useful to compare the former and current model climate. As we can remember the generation method of the model climates are quite different from each other and it is important to note that the new reforecast dataset covers the period of last 18 years (1991-2007) and the old one covers thirty years (1971-2000) so some changes in climate can be reflected in model climate as well. Similarities and differences between old and new model climates were thoroughly investigated (Üveges, 2008). Expect autumn new model climate is 0.5 -1.5 degrees warmer in the Carpathian basin. After having comparison of the model climate it is important to note it is necessary to chose the same time interval for model and observation climate anyway successful calibration could not be achieved if climate changes.

A wide range of methods is available for ensemble verification (Persson and Grazzini, 2007). During our investigation Talagrand diagram, ROC diagram, Brier score and calculation of outliers were used.

Verification of the ensemble mean was done for 2m temperature at 12 UTC. It can be seen on fig. 3 that the calibration significantly improved the quality of the forecast for all stations. Large improvement could be found in Miskolc, Nyíregyháza, Pécs and Siófok. In Békéscsaba, Budapest, Szeged and Szombathely small changes were found. The largest changes could be seen in Siófok, where the meteorological station is located very close to Lake Balaton, so especially in summer during daytime, the lake could strongly influence the air temperature around the lake. Verification was done for the full range of the VarEPS between +36 and +348 hours, it can be seen calibration is effective throughout the full time range.

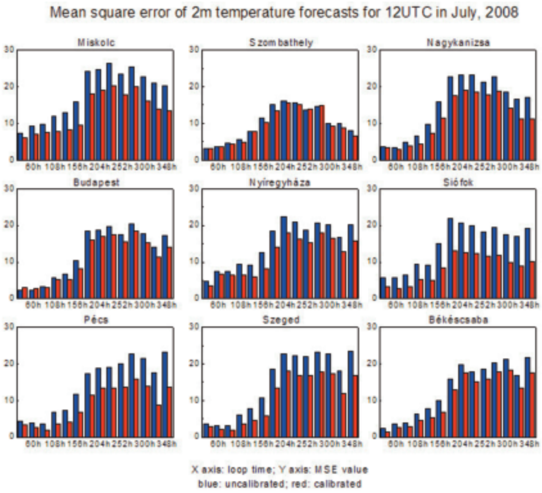

Fig. 3. Mean square error of uncalibrated (blue boxes) and calibrated (red boxes) 2m temperature at 12 UTC between +36 and +348 hours with +24 h resolution.

Talagrand diagrams provide useful information about the relationship of ensemble spread and observations. In case of uncalibrated forecasts for Miskolc and Pécs, (Fig.4) a high amount of observed values can be found close to the left end of the horizontal axis representing the spread of the current ensemble forecast. That means ensemble system makes a constant overestimation of temperature, the observed values are lower than most of the EPS members. However our calibration process can correct this error, the distribution of observations is much flatter in case of calibrated forecasts.

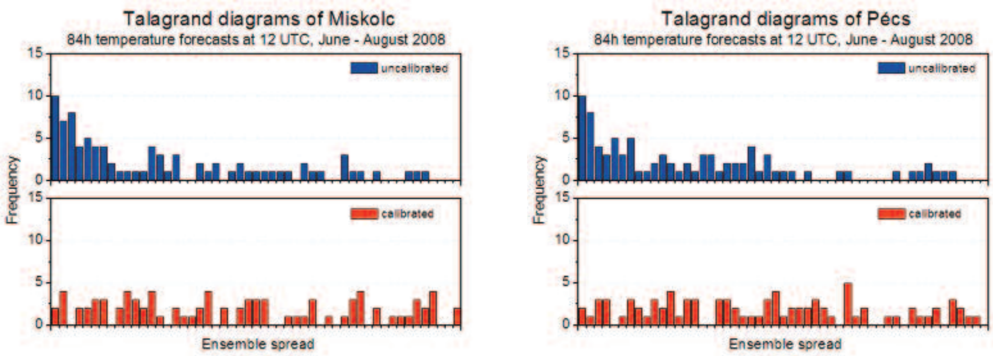

Fig. 4 Talagrand diagrams of uncalibrated (blue boxes) and calibrated (red boxes) 2m temperature at 12 UTC: Miskolc and Pécs (+84 h forecast) /June - August 2008/

ROC diagram can be a good tool for investigation of success of calibration of forecasts exceeding a predefined limit, such as temperature exceeding 30 degrees. On fig. 5 ROC diagram is shown for Pécs in July at timestep +162 - +258 hours.

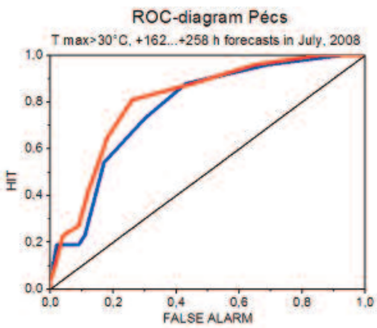

Fig. 5 ROC diagram of uncalibrated (blue line) and calibrated (red line) temperature at 12 UTC when forecast is above 30 degrees Pécs (+162-258 h timestep) July 2008

As we know precipitation is much more variable in space and time than temperature. Precipitation is a non-continuous meteorological variable that, on two out of three days is not observed in most part of the country. So making a successful forecast and verification is quite challenging, especially true in summer where convection can be quite intensive. Fig. 6 shows ROC diagram valid for daily precipitation, it can be seen, that calibration significantly reduces the rate of false alarms so the quality of the precipitation forecast is significantly improved.

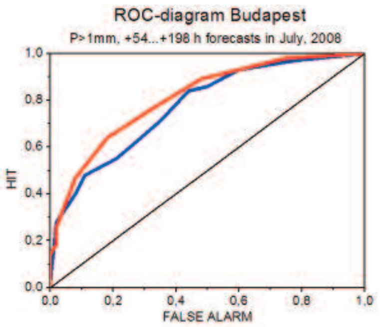

Fig 6 ROC diagram of uncalibrated (blue line) and calibrated (red line) 24 h precipitation when forecast is above 1 mm Budapest (+54-198 h timestep)

In the light of the encouraging verification results of calibration, the extension of the number of stations to 70, and the start of operational service is planned.

5. Summary, conclusions

In this paper calibration post-processing method applied for ensemble forecasts were shown, which improve the forecast efficiency on regional and local scales. In the last two years a calibration method was developed at the HMS. This method was intensively investigated and verified for the two most important meteorological variables: 2m temperature and 24h precipitation. The calibration method was presented in medium, extended, and monthly ranges In the calibration two type of new reforecast datasets were used. The model climate based on 5 EPS member reforecasts was compared to the old one and investigated to examine the specialities of the Hungarian area. Verification of calibrated forecasts shows that in most cases the calibrated products are more skilful than the raw forecasts.

Talagrand diagrams indicated better ensemble spread, a flatter shape of diagram mainly and less over or underestimation in that stations located closely to a varied orography. The RMSE verification score was used side by side with Talagrand diagrams for 2m temperature and the ROC curve was computed for precipitation forecasts. The ROC diagram of precipitation is computed with low precipitation events, because the number of high precipitation events in our time period was not sufficient for a correct investigation. In weak precipitation situations the ensemble forecasts have higher false alarm rate and consequently lower skill. In case of calibrated forecasts the rate of false alarms is diminished. Calibration has been extended to 70 stations and grided calibration fields have been operationally generated for the forecasters.

Acknowledgements – The authors would like the express thanks for the support of the Hungarian National Office for Research and Technology (NKFP, grant No. 3A/051/2004). Authors are very grateful for László Hercsényi, Rita Hodossy and Zsuzsa Kövér for kindly providing the monthly observed distribution dataset by Software Development Division of Telecommunication and Informatics Department.

References

Buizza, R., Bidlot, J-R., Wedi, N., Fuentes, M., Hamrud, Holt, G., Palmer, T., Vitart, R., 2006: The ECMWF Variable Resolution Ensemble System (VAREPS), *ECMWF Newsletter*, **108**, 14-20

Hamill, T.M. and **R. Hagedorn**, 2007: Exploring ensemble forecast calibration issues using reforecast data sets Proceedings of the Eleventh ECMWF Workshop on Meteorological Operational Systems, Reading UK, 12-16 November 2007, 161-173.

Hamill, T.M. and **J.S. Whitaker**, 2007: Ensemble calibration of 500 hPa geopotential height and 850 hPa and 2-metre temperature using reforecasts. *Mon. Wea. Rev.*, **135**, 3273-3280

Hamill, T.M., Hagedorn, R. and **J.S. Whitaker**, 2007: Probalistic Forecast Calibration: Using ECMWF and GFS Ensemble Reforecasts. Part II: Precipitation. *Mon. Wea. Rev.*. **135**, 2620-2632

Hagedorn, R., 2008: Using the ECMWF reforecast dataset to calibrate EPS forecasts. *ECMWF Newsletter*, **117**, 8-13.

Ihász, I., 2007: Experiences using VarEPS products at the Hungarian Meteorological Service: Proceedings of the Eleventh ECMWF Workshop on Meteorological Operational Systems, Reading UK, 12-16 November 2007, 130-135

Ihász, I. and **Mile M.**, 2008: Calibration of ECMWF ensemble precipitation forecasts for hydrological purposes at the Hungarian Meteorological Service. Proceedings of the XXIV Conference of the Danubian Countries.2-4 June 2008, Bled, Slovenia

Mile, M., 2008: Subsequently calibration of ensemble forecasts of the ECMWF (in Hungarian). Master Thesis, Eötvös Loránd University, Budapest, Hungary

Persson, A., and **Grazzini, F.** 2007: User Guide to ECMWF forecasts products: *Meteorological Bulletin*, **3.2** (version 4.0), ECMWF Reading, UK.

Üveges, Z., 2009: Calibration of monthly forecasts of the ECMWF. (in Hungarian) Master thesis, Eötvös Loránd University, Budapest, Hungary.

Webster, P., 2007: Extended probabilistic hydrological forecasts of Ganges and Brahmaputra using ECMWF ensemble products. Proceedings of 11th Workshop on Meteorological Operational Systems, Reading UK, 12-16 November 2007, http://www.ecmwf.int/newsevents/meetings/workshops/2007/ensemble_prediction/