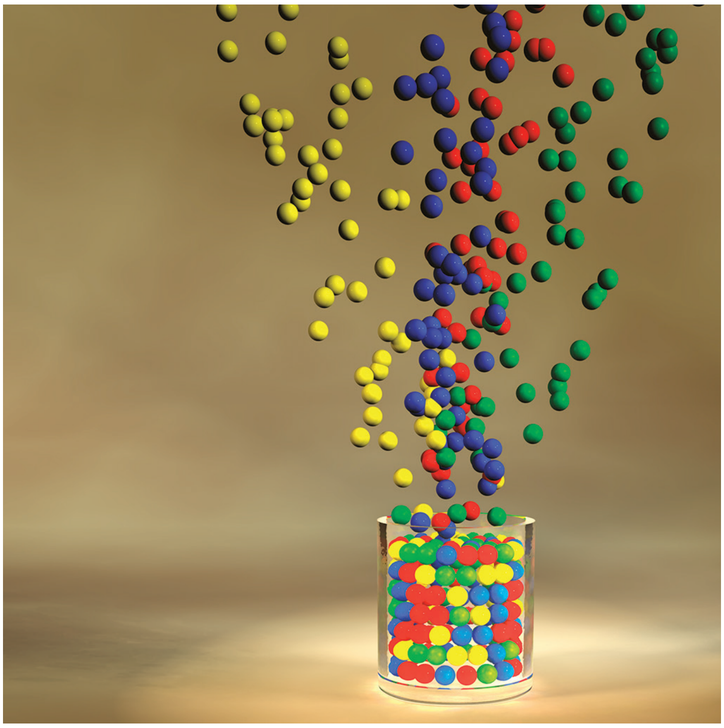
ECMWF Feature article

from Newsletter Number 116 – Summer 2008

METEOROLOGY

Predictability studies using TIGGE data



www.ecmwf.int/en/about/news-centre/media-resources

doi:10.21957/j9qbbbftjw

This article appeared in the Meteorology section of ECMWF Newsletter No. 116 – Summer 2008, pp. 16-20.

Predictability studies using TIGGE data

Roberto Buizza, Young-Youn Park, Martin Leutbecher, Florian Pappenberger

Operational medium-range ensemble prediction began in December 1992, when the National Centers for Environmental Prediction (NCEP) and ECMWF began producing global ensemble predictions as part of their operational products. In 1995, the Meteorological Service of Canada (MSC) implemented its ensemble prediction system. Following these examples, six other centres started running global ensemble prediction systems daily. At the beginning of 2008, ten meteorological centres are running a medium-range (i.e. with a forecast length of at least 7 days) global ensemble prediction system. The ensemble systems are based on a finite number of time integrations of a numerical weather prediction model. Most of them are designed to simulate the effect on forecast accuracy of initial uncertainties, with only few of them designed also to simulate the effect of model approximations.

TIGGE, the THORPEX Interactive Grand Global Ensemble, has among its objectives to achieve a deeper understanding of the contribution of initial and model uncertainties to forecast error, and to develop new methods of combining ensembles from different sources and of correcting systematic errors (biases and spread over- or under-estimation). More information about the concept of TIGGE and its objectives is given in the article by Philippe Bougeault on page 9 of this edition of the *ECMWF Newsletter.*

Here we summarise some preliminary conclusions from two studies using TIGGE data:

- ***Park et al.*** (2008), who presented some results on comparing and combining ensembles obtained using the TIGGE data available up to December 2007.
- ***Pappenberger et al.*** (2008), who illustrated how TIGGE weather forecasts can be used in hydrological ensemble prediction for a case of severe flooding that affected Romania in October 2007.

Assessing the strengths and weaknesses of single ensemble systems

For the first time, TIGGE gives access to forecasts generated by different ensemble systems but archived with the same format (GRIB2). Users can easily extract all available ensemble forecasts, and compare the performance of individual systems. Thereby, TIGGE helps to develop an understanding of the strengths and weaknesses of each single system. The TIGGE dataset can also be used to investigate the potential value of calibration methods, and of combining single ensembles into a multi-systems ensemble. An example is given in Figure 1 which uses ensembles from ECMWF, UK Met Office (UKMO), Japan Meteorological Administration (JMA) and China Meteorological Administration.

A comparison was made between the scores of the single ECMWF ensemble, with or without bias correction, with the scores of three bias-corrected combinations:

- Combined ECMWF and UKMO ensemble.
- Combined ECMWF, UKMO and JMA ensemble.
- Combined ECMWF, UKMO, JMA and CMA ensemble.

Note that the scores of the bias-corrected combined ECMWF, UKMO and JMA ensemble are almost identical to those of the bias-corrected combined ECMWF, UKMO, JMA and CMA ensemble. Consequently the results of the combined ECMWF, UKMO and JMA ensemble are not shown in Figure 1.

Results indicate that for the 500 hPa geopotential height over the northern hemisphere extra-tropics the scores of all these ensembles are very close (Figures 1(a) and 1(c)). It is found that all scores are slightly worse than the score of the single ECMWF ensemble without bias correction. By contrast, results indicate that for the 850 hPa geopotential height over the tropics (Figures 1(b) and 1(d)), adding the UKMO and JMA ensembles to the ECMWF ensemble improves the scores, but further adding the CMA ensemble does not bring any extra improvement. It is interesting that most of the improvements are achieved by combining the two ensembles characterized by the most similar performance characteristics (i.e. ECMWF and UKMO ensembles), while the addition of a third ensemble does not bring any extra improvement.

AFFILIATIONS
Roberto Buizza, Martin Leutbecher, Florian Pappenberger: ECMWF, Reading UK.
Young-Youn Park: ECMWF and Korea Meteorological Administration, Seoul, Republic of Korea.

 doi:10.21957/j9qbbbftjw

To further illustrate the impact of merging two ensembles with similar characteristics, Figure 2 compares the performance of the ECMWF ensemble without bias correction with three other ensembles:

- UKMO ensemble without bias correction.
- Combined ECMWF and UKMO ensemble with bias correction.
- Combined ECMWF and UKMO ensemble without bias correction.

For the 500 hPa geopotential height over the northern hemisphere extra-tropics (Figures 2(a) and 2(c)), results show that the single ECMWF ensemble performs better than the single UKMO ensemble. Also in the short-range the combined bias-corrected ECMWF and UKMO ensemble performs best, but in the long range the combined non-bias-corrected ECMWF and UKMO ensemble has the best performance. By contrast, for the 850 hPa temperature over the tropics (Figures 2(b) and 2(d)) it is the combined ECMWF and UKMO ensemble with bias-correction that outperforms the other systems.

In general, the preliminary results obtained by *Park et al.* (2008) indicate that, for the variables considered in their study, it is over the tropics where the performance of the ensembles differs most. Also it is over the tropics where the benefit of combining them is the highest. Overall, these preliminary results indicate that care must be taken when combining ensemble systems. It appears that the largest benefit of merging ensembles is for the prediction of, for example the 850 hPa temperature over the tropics, an area where the largest variability in the performance of the single ensembles have been detected. Refer to *Park et al.* (2008) for a more complete discussion of these and related issues.

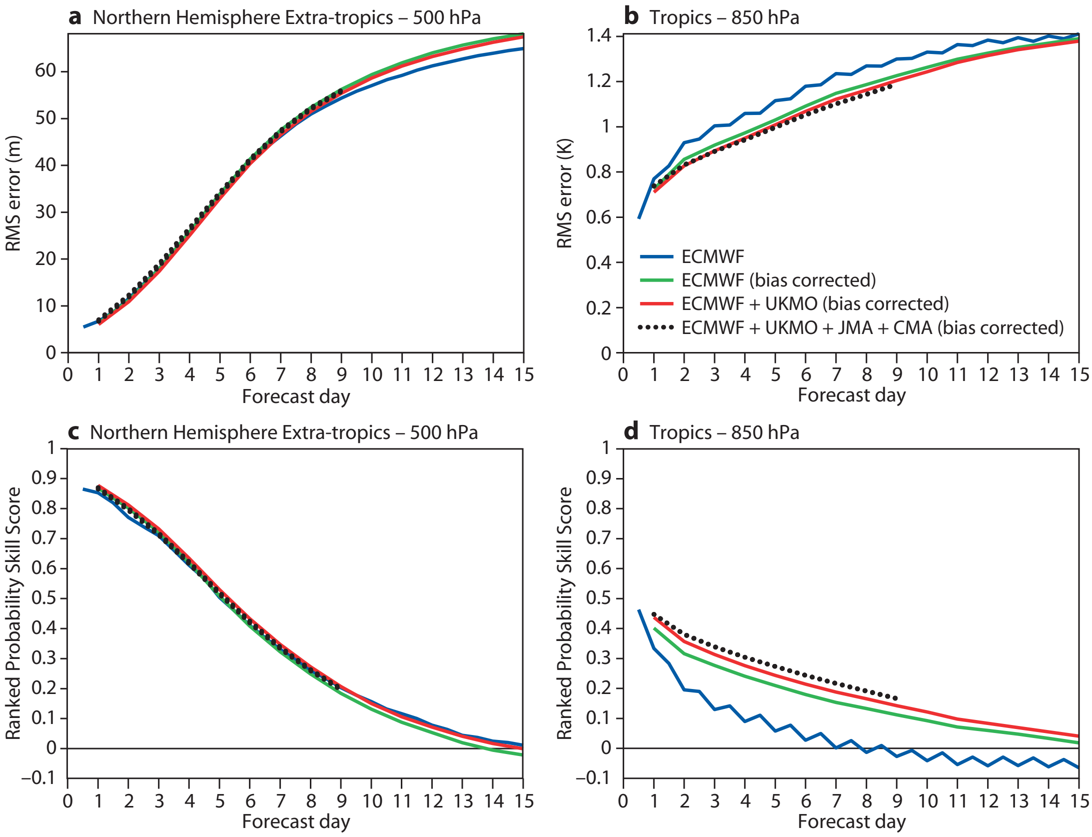


Figure 1 TIGGE combination results for June, July and August 2007 (86 cases): average scores of ECMWF (blue line), bias-corrected ECMWF (green line), bias-corrected combined ECMWF and UKMO (red line) and bias-corrected combined ECMWF, UKMO, JMA and CMA (dotted line), with biases estimated using a 30-day training period. (a) Root-mean-square error of the ensemble mean forecast of 500 hPa geopotential height over the northern hemisphere extra-tropics. (b) Root-mean-square error of the ensemble mean forecast of 850 hPa temperature over the tropics. (c) Rank probability skill score for 500 hPa geopotential height over the northern hemisphere extra-tropics. (d) Rank probability skill score for the 850 hPa temperature over the tropics (from *Park et al.,* 2008).

doi:10.21957/j9qbbbftjw

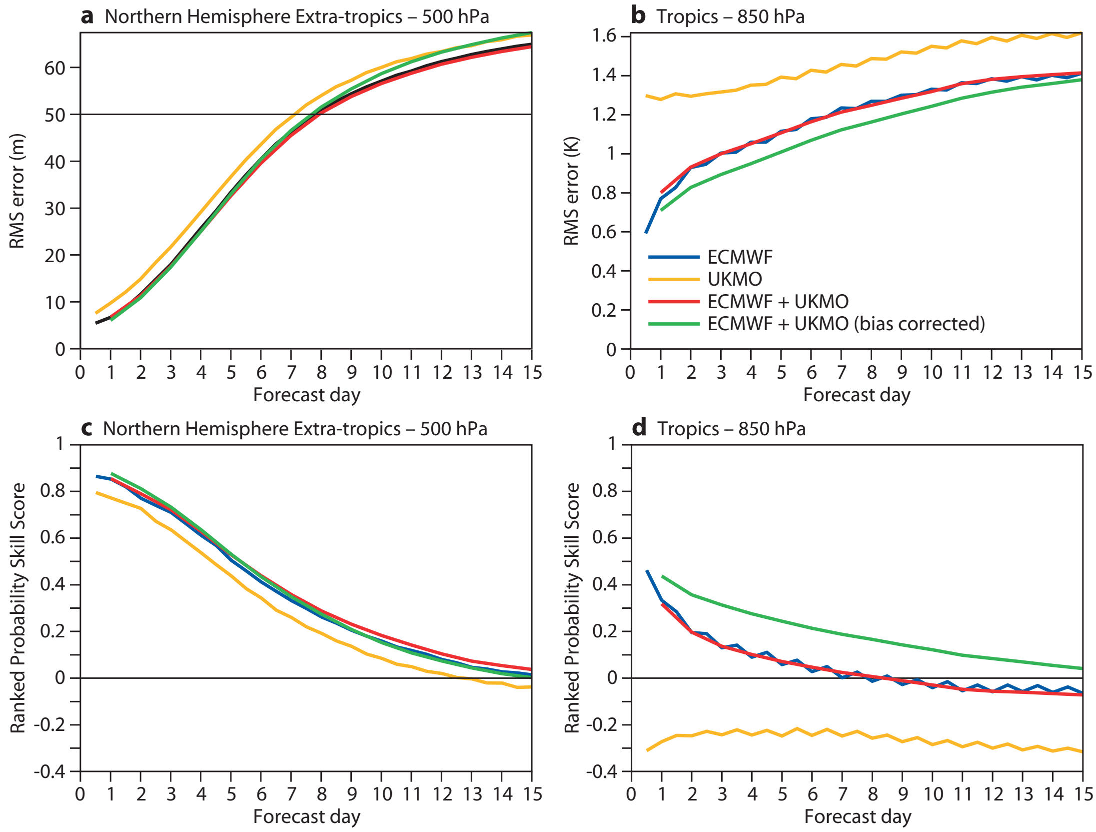


Figure 2 TIGGE combination results for June, July and August 2007 (86 cases): average scores of ECMWF (blue line), UKMO (yellow line), combined ECMWF and UKMO (red line) and bias-corrected combined ECMWF and UKMO (green line), with biases estimated using a 30-day training period. (a) Root-mean-square error of the ensemble mean forecast of the 500 hPa geopotential height over the northern hemisphere extra-tropics. (b) Root-mean-square error of the ensemble mean forecast of the 850 hPa temperature over the tropics. (c) Rank probability skill score for the 500 hPa geopotential height over the northern hemisphere extra-tropics. (d) Rank probability skill score for 850 hPa temperature over the tropics (from *Park et al.,* 2008).

Using TIGGE data in hydrological probabilistic prediction

Pappenberger et al. (2008) illustrated how TIGGE weather forecasts can be used to drive the European Flood Alert System (EFAS), a flood prediction model developed as an initiative of the European Commission. EFAS has been running on a pre-operational basis at the EU's Joint Research Centre (JRC) since 2005, driven by Deutscher Wetterdiens (DWD) and ECMWF rainfall forecasts. It provides local water authorities with probabilistic flood forecasting information up to 10 days in advance based on four warning thresholds (low, medium, high and severe). More precisely, *Pappenberger et al.* (2008) discussed the potential of using grand-ensembles for early flood warnings by evaluating flood forecasts based on TIGGE data for the floods on several tributaries to the Danube in Romania (Siret, Jiu, Olt and Arges) in October 2007.

Figure 3 shows EFAS discharge forecasts driven from the single TIGGE ensembles and for a grand-ensemble which includes all the single ensemble forecasts for a 5-day lead time. It is essential to correctly represent the onset of the rising limb of the flood hydrograph to allow flood preparedness and disaster mitigation. All six single ensemble distributions and the grand-ensemble distribution predict the onset of the rising limb correctly in terms of timing and discharge thresholds. However, some of them fail to include the observed hydrograph at the onset and peak time. None of the forecasts perform very well for the lower end of the recession limb.

It is found that the ensemble spread widens with lead time and thus more observations are bracketed. The widening distribution also means that a lower percentage of discharge predictions are above the warning thresholds. This has important implications for issuing warnings – at long lead times there will be fewer ensemble members that will be able to trigger such warnings and this should be taken into account (this is recognised in the EFAS system). Overall, using the grand-ensemble consistently gives a good prediction of the flood hydrograph, apart from the falling recession limb.

 doi:10.21957/j9qbbbftjw

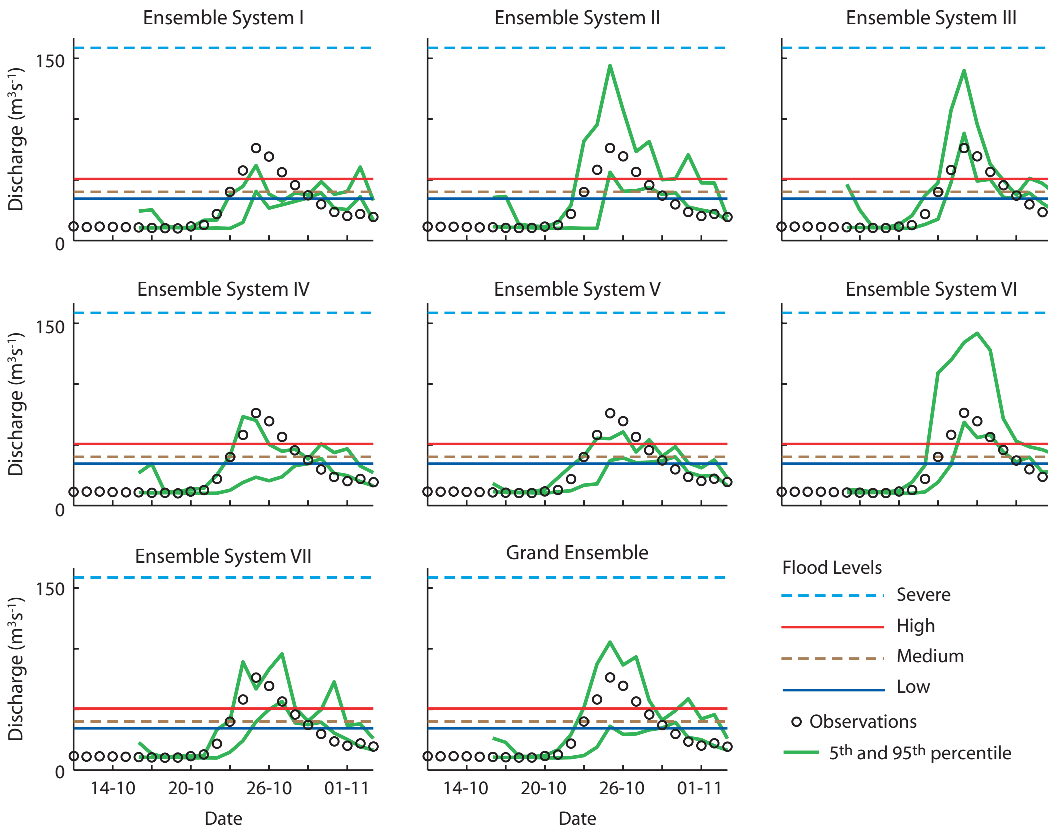


Figure 3 The 5th and 95th percentile of discharge predictions of the different forecasts with a 5 day lead time. The horizontal lines indicate the four warning thresholds (from *Pappenberger et al.,* 2008).

The potential value of TIGGE data for predictability studies

Scientists are encouraged to access the TIGGE database, and try to answer some of the key questions which are still outstanding in ensemble prediction. TIGGE gives users access to many ensemble systems, but it is still not clear whether there is one method which is most efficient to calibrate and combine single ensembles to generate grand-ensemble products, or whether different methods should be used for different variables, areas and users. Also there is still the unanswered question about whether a single system can perform as well as a grand-ensemble system. This brings us to a fundamental question about the simulation of model approximations in ensemble prediction: is using a multi-system ensemble the best approach to simulate the effect of model uncertainty on forecast quality? Alongside these and other key predictability questions, the TIGGE database can also help users become more familiar with ensemble products, test the development of new products and foster a more intense exploitation of a probabilistic approach to weather prediction.

Further Reading

Pappenberger, F., J. Bartholmes, J. Thielen, H.L. Cloke, R. Buizza & **A. de Roo,** 2008: New dimensions in early flood warning across the globe using grand-ensemble weather predictions. *Geophys. Res. Lett.,* doi:10.1029/ 2008GL033837, in press, available online (also published as *ECMWF Tech. Memo No. 558*).

Park, Y.-Y., R. Buizza & **M. Leutbecher,** 2008: TIGGE: preliminary results on comparing and combining ensembles. *Q. J. R. Meteorol. Soc,* submitted (also published as *ECMWF Tech. Memo. No. 548*).