Application of Ensembles in Flood Forecasting

Jutta Thielen[1], Jens Bartholmes[1] and Florian Pappenberger[2]

[1]*European Commission DG Joint Research Centre, Institute for Environment and Sustainability, TP 261, via E Fermi, 21027 Ispra, contact: jutta.thielen@jrc.it*

[2] *European Centre for Medium-Range Weather Forecasts, Shinfield Park Reading, RG2 9AX UK*

1. Introduction

A major research challenge of the 21st century is to provide early warning for floods with potentially disastrous consequences. In 2007 floods killed 8349 people, affected 164 million and caused damage in the excess of 21 billion US$ (EM Dat, 2007). Early flood warning several days in advance could provide civil protection authorities and the public with vital preparation time and could reduce the socio-economic impacts of flooding. Unfortunately, precipitation forecasts, in most cases the driving factor for floods, are highly uncertain. In a study based on forecasts from the European Centre for Medium-Range Weather Forecasts, Buizza et al. (1999) showed that although generally the skill in weather forecasting has increased to 5-6 days, e.g. for temperature, it is only of the order of 2-3 days for precipitation. For extreme rainfalls which are of special interest to flood forecasters, the forecasting time for skilled forecasts decreases even further.

2. Increasing flood warning lead time with weather ensembles

The lead time for skilful precipitation forecasting can be, however, extended by exploring ensemble prediction systems (EPS) (e.g.Tracton and Kalnay, 1993, Molteni et al., 1996). Although produced by some meteorological services as early as the 80ies (Molteni et al., 1996), it is only since recently that EPS are being explored for flood forecasting purposes. In Europe, the EFFS project (2000-2003) was one of the first large European research projects to look into the potential of using medium-range weather forecasts, including EPS, for flood forecasting in large trans-national river basins with the aim of extending the early warning time (de Roo et al, 2003, Gouweleeuw et al., 2004 & 2005). In 2004, an international initiative called HEPEX was launched (see http://hydis8.eng.uci.edu/hepex). HEPEX stands for Hydrological Ensemble Prediction Experiment and represents an international effort that brings meteorological and hydrological communities together to develop advanced probabilistic hydrological forecasting techniques that use weather and climate ensemble forecasts (Hamill et al., 2005, Schaake et al., 2006a,b Franz et al., 2005). Recent research results within the framework of HEPEX are encouraging and demonstrate the potential benefit of probabilistic weather forecasts over deterministic ones for flood forecasting in large river basins (Thielen et al., 2007). Roulin (2007) demonstrated that EPS based flood forecasting can also be valuable for small river basins and advances in limited area EPS modelling may provide even better quantitative precipitation estimates also for small basins, (Marsigli et al 2001, Marsigli et al 2005, Tibaldi et al., 2006)

The main benefit of medium-range probabilistic flood forecasts for hydrological services is the increased lead time for warnings of a flood event. In case an early alert is confirmed as the forecasted event approaches the forecast date, flood forecasters will be better prepared to initiate any necessary emergency procedure and there will be a gain in time when analysing the short-term – and more precise – forecasts. Earlier warning can therefore help in reducing the level of stress in the forecasting centres. Research has shown that the negative effects of stress on decision making under time pressure and fatigue due to overwork in the operational centres during a flood event should not be underestimated (Kowaski-Trakofler et al., 2003, Paton and Flin, 1999). In case subsequent forecasts do not confirm the previous alert, forecasters can return to business-as-usual. Adverse effects from earlier warning are therefore minimal.

Although it has been demonstrated that the incorporation of ensemble weather forecasts into a flood warning system can significantly increase forecast lead time (e.g. Ahrens et al., 2007, Gouweleeuw et al., 2004, Bartholmes et al., 2007, Krzysztofowicz 2002) many hydrological services do not include them as they introduce a further degree of uncertainty into their forecasts and thus in their decision making process. However, forecasts from Ensemble Prediction Systems (EPS) embrace some of these uncertainties in the production of multiple weather forecasts for the same period, and, used with a hydrological model, have the potential to provide valuable early flood warning (e.g. Roulin, 2007). Recently there has been a move to integrate EPS into operational flood forecasting systems around the world such as in Bangladesh (Hopson and Webster, 2008), the flood forecasts of the Swedish Hydro-Meteorological Service (*Olsson, J. and Lindström, 2008)*, the European Flood Alert System (Thielen et al., 2008) and many more.

Uncertainty in the hydrological forecasting arises not only from the meteorological ensemble input but also for example from the hydrological model parameterisations, crude representation of the physical processes, or observational errors. These errors cascade through the whole system (Pappenberger, 2006) and as a result the flood forecasts can be associated with a wide spread of uncertainty that make it difficult for the endusers to base a decision on such forecasts. One way to quantify the potential value of a forecasting system is to use the concept of potential economic value, estimated using simple cost/loss models (*Richardson* 2000, *Buizza* 2001). So-called cost/loss functions that relate flood depth to economical costs can be developed to provide guidance to decision makers when to act. These cost/loss function can, however, not be universal and depend from case to case and location to location.

3. Skill score assessment of hydrological ensemble forecasts: an analysis from the European Flood Alert System

Literature on skill scores dates back more than 120 years (Peirce, 1884; Gilbert, 1884), but not all skill scores are equally suited and there is no single skill score that can convey all necessary information, thus normally sets of skill scores are used to cover a wider spectrum of properties (Baldwin, 2004). An extensive review can be found in the WMO publication of Stanski *et al.* (1989) and as well as in the works of Murphy (1996, 1997).

Bartholmes has analysed two years of results (2005-2006) from the European Flood Alert System (EFAS) and applied several classical meteorological skill scores. EFAS aims at increasing preparedness for floods in trans-national European river basins by providing local water authorities with medium-range and probabilistic flood forecasting information 3 to 10 days in advance (Thielen et al., 2008). The EFAS research project started in 2003 with the development of a prototype at the European Commission Joint Research Centre (JRC), in close collaboration with the national hydrological and meteorological services. The prototype covers the whole of Europe on a 5 km grid. Flood warning lead-times of 3-10 days are achieved through the incorporation of medium-range weather forecasts from the German Weather Service (DWD) and the European Centre for Medium-Range Weather Forecasts (ECMWF), comprising a full set of 51 probabilistic forecasts from the Ensemble Prediction System (EPS) provided by ECMWF. The ensemble of different hydrographs is analysed and combined to produce early flood warning information, which is disseminated to the hydrological services that have agreed to participate in the development of the system (Ramos et al., 2007)..

EFAS is running pre-operationally since 2005. It uses the hydrological model LISFLOOD (van der Knijff and de Roo, 2007), which is a semi-conceptual rainfall-runoff model applied on a 5km grid across Europe. Each EPS forecast member is propagated through LISFLOOD and the resulting distribution of river discharges is compared against flood warning thresholds provided by four flood warning thresholds (low medium, high and severe). For each pixel the proxy discharge calculated from observed meteorological data has been used to compare against the discharge forecasts based on weather forecasts

Individual case study analyses (e.g. Kalas et al., 2008, Pappenberger et al., 2008) give insight into the different forcings behind individual flood events and to establish decision making rules (Ramos et al., 2007). Case studies are, however, always biased towards events that have taken place and do not allow the determination of the reliability of the probabilistic forecasts. Therefore, ideally, objective skill score analysis on a statistically significant number of events needs to complement the case study analysis. This is not a trivial task when dealing with severe and rare events. For flood forecasting the European Flood Alert System provides an ideal theoretical laboratory since it covers the whole of Europe and the likelihood is high that within a two year period a sufficient number of flood events can be analysed.

For the construction of contingency tables the forecasted discharges were transformed into dichotomous events regarding their exceedance of thresholds (Atger, 2001). In the case of the deterministic forecasts the thresholds were only the EFAS alarm levels. In the case of EPS a second threshold regarding the number of EPS members in a forecast that exceeded a certain EFAS alarm level, was applied (resulting in separate contingency table sets for each EPS threshold). If less EPS members than this threshold forecasted the event it was classified "NOT forecasted".

Figure 1 shows the analysis of the absolute numbers of hits, false alarms and misses for the 2 year analysis over all pixels at a lead time of 4 days. If decision makers acted when 10-12 EPS exceed the high alert threshold there would be about as many hits as false alarms and misses. Waiting to act for more EPS would result in more hits than false alarms, however, the number of missed events would be steadily increasing.

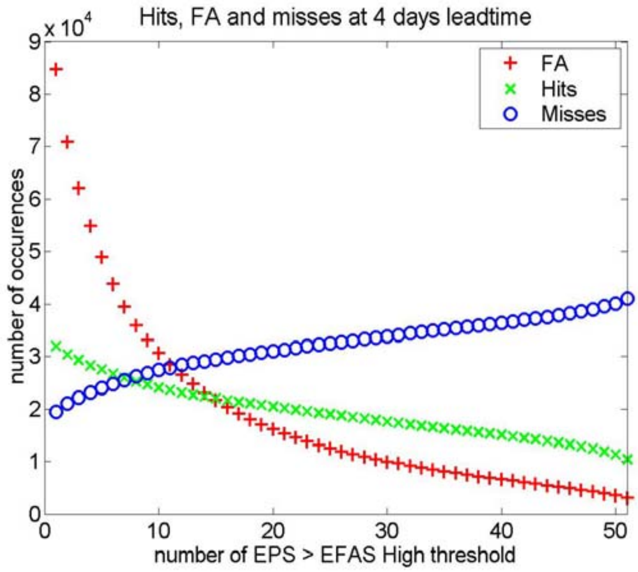

Figure 1: : Absolute numbers reporting the three contingency table fields "hits"(h [x]), "false alerts" (f [+]) and "misses"(m [o]) for a lead time of 4 days.

When dealing with hydrological forecasts and lead times up to 10 days the uncertainty can be quite high. At the expense of decreasing leadtime by one day, the rate of false alerts has been significantly reduced by introducing the criterion of persistence into EFAS forecasts. In this analysis only if the threshold exceedance in a river stretch is forecasted continuously on 2 consecutive dates it is considered as persistent.

Figure 2 illustrates the effect of persistence on the relative frequency of the Brier Skill Score of discharge threshold exceedance.

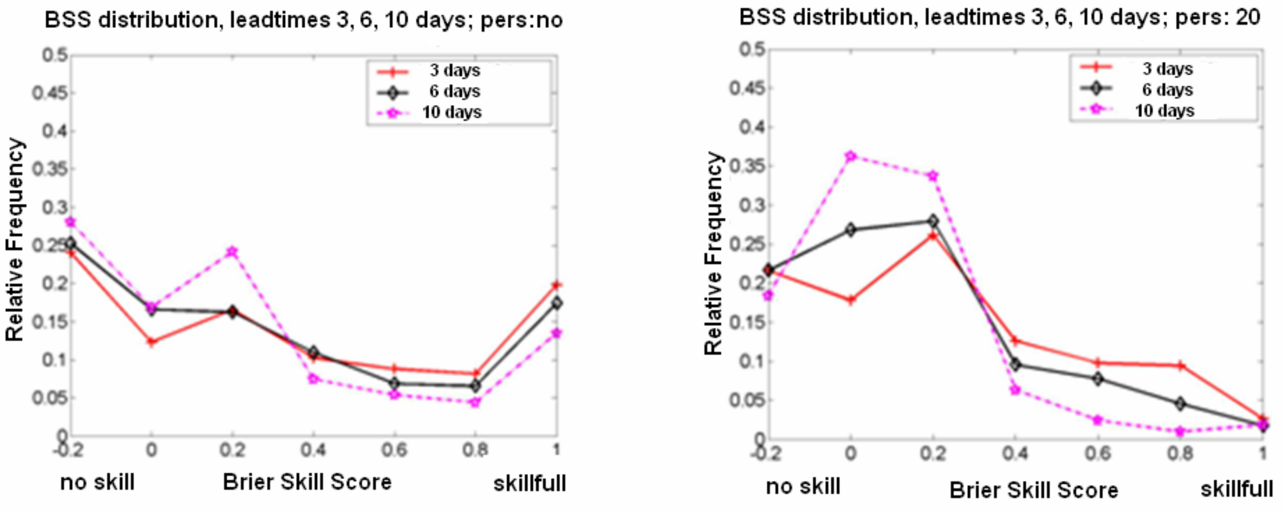

Figure 2: Relative frequency distribution of Brier skill score (BSS) values (0 no skill, 1=perfect forecast) for lead times 3 days (red), 6 days (black) and 10 days (magenta) without persistence (left) and with a persistence of 20 EPS between two forecasts (right).

First of all, figure 2 shows that the Brier Skill Score (BSS) calculated from the discharge exceedances is within the range or slightly higher than BSS of precipitation. Considering persistence considerably increases the relative frequency of BSS in the range of 0-0.3. The increase in skill without persistence in the left diagram (0.15-0.2 of BSS=1) as compared to a skill of almost20 EPS persistence (0.025 of BSS=1) is, in fact, an artifact resulting from 1 forecast within the 2 years having a hit and thus artificially increasing skill while these artifacts are largely filtered out by the persistence criterium (two consecutive forecasts need to have predicted a hit).

Analysis of Bartholmes et al. (2008) therefore showed that in the case of rare events such as floods it is very important to look at absolute numbers and simple skill measures like FOH, FOM and FAR that give a direct idea of the ratios between the fields of the contingency table. If less intuitive results of more complex skill scores are taken into consideration one should be aware of their specific, not always intuitive behaviour and their tendency to be strongly influenced by one of the contingency table fields.

4. Summary

Over the recent years Ensemble Prediction Systems, both from global and limited area models, are increasingly applied in hydrology for flood forecasting but also for reservoir management and seasonal studies. International research programs such as the Hydrological Ensemble Prediction Experiment (HEPEX) foster the development of advanced probabilistic hydrologic forecast techniques and of corresponding decision making tools. Increasingly both hydrological and meteorological scientists are involved in the development, testing and operational management of forecasting systems, and end-users.

Since recently the advances in ensemble research start being implemented operationally in flood forecasting. There is a global movement of national water authorities and other related services to run ensemble prediction systems for hydrological applications. However, there are still many science questions that need addressing, for example how to deal with the uncertainty, what are the different sources of uncertainty and how do they cascade through the hydrological system, how to assess reliability of probabilsitic forecasts for rare events, in particular when both the meteorological models and the hydrological structures along the rivers keep changing frequently, which probabilistic skill scores are the most relevant for hydrological studies, how to communicate probabilistic results to different endusers and how and when to act with civil protection measures based on uncertain results.

References

Ahrens, B. and Jaun, S., 1007: On evaluation of ensemble precipitation forecasts with observation-based ensembles. *Advances in Geosciences* **10**, 139-144.

Atger, F., 2001: Verification of intense precipitation forecasts from single models and ensemble prediction systems. *Nonlinear Processes Geophys.,* **8**, 401–417.

Bartholmes J., J. Thielen, M.-H. Ramos and S. Gentilini, 2008: The European Flood Alert System EFAS – Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts, *Hydrol. Earth Syst. Sci. Discuss.,***5**, 289-322.

Buizza R., Hollingsworth, A., Lalaurette F., Ghelli, A., 1999): Probabilistic Predictions of Precipitation Using the ECMWF Ensemble Prediction System. *Weather and Forecasting*. **14**, 168-189.

De Roo, A., Gouweleeuw, B., Thielen, J., Bates, P., Hollingsworth, A. et al., 2003: Development of a European Flood Forecasting System. *International Journal of River Basin Management*, **1**, 49-59.

EM-DAT (2007). *The OFDA/CRED International Disaster Database*. Université Catholique de Louvain, Belgium. www.emdat.be

Franz, K., N. Ajami, J. Schaake, and R. Buizza: Hydrologic Ensemble Prediction Experiment Focuses on Reliable Forecasts, *Eos*, **86**, No. 25, 2005.

Gilbert, G.F., 1884: Finley's tornado predictions. *American Meteorological Journal*, **1**, 166-172.

Gouweleeuw, B., Reggiani, P. & De Roo, A. (Ed.), 2004: A European Flood Forecasting System EFFS. *Full Report. European Report EUR 21208, EC DG JRC & WL Delft Hydraulics*, 304p.

Gouweleeuw, B., Thielen, J., Franchello, G., De Roo, A., Buizza, R., 2005: Flood forecasting using medium-range probabilistic weather prediction. *Hydrological and Earth System Sciences*, **9(4)**, p.365-380,

Hamill, T., Clark, M., Schaake, J., Buizza, R, 2005: *2nd HEPEX Workshop Summary Report*, Boulder, Colorado (http://hydis8.eng.uci.edu/hepex/scndwksp/HEPEX05-Summary.pdf).

Krzysztofowicz, R., 2002: Bayesian system for probabilistic river stage forecasting. *J. Hydrol*. **268**, 16-40.

Marsigli, C., Montani, A., Nerozzi, F., Paccagnella, T., Tibaldi, S., Molteni, F., Buizza, R., 2001: A strategy for high-resolution ensemble prediction. II: Limited-area experiments in four Alpine flood events. *Q. J. R. Meteorol Soc*, **127**, 2095-2115.

Marsigli, C., Boccanera, F., Montani, A., Paccagnella, T., 2005: The COSMO-LEPS mesoscale ensemble system: validation of the methodology and verification. *Nonlinear Processes in Geophysics*, **12**, 527–536,

Molteni, F., R. Buizza, T. N. Palmer, & T. Petroliagis, 1996: The ECMWF ensemble prediction system: methodology and validation. *Q. J. R. Meteorol. Soc.*, **122**, 73-119

Murphy, A H., 1996: The Finlay Affair. A signal event in the history of forecast verification. *Weather and Forecasting*, **11**, 3-20.

Murphy, A.H., 1997.Forecast verification. In*: Economic Value of Weather and Forecasts, ed. R.W. Katz and A.H. Murphy*. Cambridge:Cambridge University Press, 19-74.

Pappenberger, F., Matgen, P., Beven, K. J., Henry, J.-B., Pfister, L., and de Fraipont, P., 2006: Influence of uncertain boundary conditions and model structure on flood inundation predictions. *Advances in Water Resources* **29** (10), 1430-1449.

Peirce, C.S., 1884: The numerical measure of success of predictions. *Science*, **4**, 453-454.

Ramos, M.H., Bartholmes J., Thielen-del Pozo, J., 2007: Development of decision support products based on ensemble weather forecasts in the European Flood Alert System. *Atmospheric Science Letters*, **8**, 113-119.

Roulin E., 2007: Skill and relative economic value of medium-range hydrological ensemble predictions; *Hydrol. Earth Syst. Sci*., **11**, 725-737.

Schaake, J., Franz K., Bradley A. and Buizza R., 2006: The Hydrologic Ensemble Prediction Experiment (HEPEX). *Hydrol. Earth Syst. Sci. Discuss.*, **3**, 3321-3332.

Schaake, J. C., T. M. Hamill, R. Buizza and M. Clark, 2007: HEPEX, the Hydrological Ensemble Prediction Experiment. *Bul. Amer. Meteo. Soc*., 88, 1541-1547.

Stanski, H.R., L.J.Wilson and W.R. Burrows, 1989: Survey of common verification methods in meteorology. WMO Technical Report No. 8 Geneva : *WMO, (WMO/TD 358)*

Tibaldi, S., Paccagnella, T., Marsigli, C., Montani, A., Nerozzi, F., 2006: Limited-Area Ensemble Forecasting: the COSMO-LEPS System. In: *Predictability of Weather and Climate* (Palmer, T. and R. Hagedorn, eds). Cambridge University Press, 489-513.

Thielen J., Bartholmes J., and Schaake, J. (Eds.) 2007: 3rd HEPEX workshop, *Book of Abstracts, European Commission, Report, EUR22861 E.* http://hydis8.eng.uci.edu/hepex/thrdwksp/thrdwksp.html

Thielen, J., Bartholmes J., Ramos M.-H., and de Roo A., 2008: The European Flood Alert System – Part 1: Concept and development, *Hydrol. Earth Syst. Sci. Discuss*., **5**, 257-287.

Tracton, M. S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Wea. Forecasting*, **8**, 379-398.