



HAL
open science

Statistical processing of forecasts for hydrological ensemble prediction: a comparative study of different bias correction strategies

Ioanna Zalachori, Maria-Helena Ramos, Rémy Garçon, Thibault Mathevet,
Joël Gailhard

► To cite this version:

Ioanna Zalachori, Maria-Helena Ramos, Rémy Garçon, Thibault Mathevet, Joël Gailhard. Statistical processing of forecasts for hydrological ensemble prediction: a comparative study of different bias correction strategies. *Advances in Science & Research*, 2012, 8, p. 135 - p. 141. 10.5194/asr-8-135-2012 . hal-00763615

HAL Id: hal-00763615

<https://hal.science/hal-00763615>

Submitted on 11 Dec 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Statistical processing of forecasts for hydrological ensemble prediction: a comparative study of different bias correction strategies

I. Zalachori¹, M.-H. Ramos¹, R. Garçon², T. Mathevet², and J. Gailhard²

¹IRSTEA, Hydrology Group, UR HBAN, 1 rue Pierre-Gilles de Gennes CS 10030, 92761 Antony, France

²EDF-DTG, Electricité de France, Direction Technique de Grenoble, France

Correspondence to: I. Zalachori (ioanna.zalachori@irstea.fr)

Received: 15 January 2012 – Revised: 16 May 2012 – Accepted: 7 July 2012 – Published: 25 July 2012

Abstract. The aim of this paper is to investigate the use of statistical correction techniques in hydrological ensemble prediction. Ensemble weather forecasts (precipitation and temperature) are used as forcing variables to a hydrologic forecasting model for the production of ensemble streamflow forecasts. The impact of different bias correction strategies on the quality of the forecasts is examined. The performance of the system is evaluated when statistical processing is applied: to precipitation and temperature forecasts only (*pre-processing* from the hydrological model point of view), to flow forecasts (*post-processing*) and to both. The pre-processing technique combines precipitation ensemble predictions with an analog forecasting approach, while the post-processing is based on past errors of the hydrological model when simulating streamflows. Forecasts from 11 catchments in France are evaluated. Results illustrate the importance of taking into account hydrological uncertainties to improve the quality of operational streamflow forecasts.

1 Introduction

Probabilistic information is of special importance for users vulnerable to climatic and hydrological hazards at different scales (agriculture and irrigation, navigation, public safety, energy companies, etc.). In hydrology, a common approach to produce probabilistic information is the use of ensemble-based streamflow forecasting systems (see review by Cloke and Pappenberger, 2009). The key advantage of these systems is that they can provide future scenarios of streamflow evolution in time, with information on the uncertainty of the predictions, which can be potentially more useful at longer forecast lead times, notably in terms of increasing preparedness for severe flood events and reducing losses (Bartholmes et al., 2009; Boucher et al., 2012; Verkade and Werner, 2011). However, model output predictions may lack precision and reliability due to several reasons like imperfect numerical representation of physical process or insufficiency account of all sources of uncertainty involved in the system being modelled (e.g. Thirel et al., 2008; Jaun and Ahrens, 2009; Randrianasolo et al., 2010; Velazquez et al., 2011). To im-

prove the quality of probabilistic forecasts and provide reliable estimates of uncertainty, statistical processing of forecasts is recommended (Schaaake et al., 2010). The aim is to remove forecast biases and to improve ensemble dispersion. Several techniques have been proposed in meteorology and hydrology, mainly based on empirical dressing techniques, Bayesian methods or regression analysis (e.g. Krzysztofowicz, 1999; Raftery et al., 2005; Fortin et al., 2006; Hashino et al., 2007; Olsson and Lindström, 2008; Brown and Seo, 2010; Zhao et al., 2011).

In hydrologic forecasting systems, statistical correction techniques can be applied to the forecast input of the hydrological model (meteorological variables like precipitation and temperature), to the forecast output of the hydrological model (streamflows) or to both. As shown in Fig. 1, from the hydrological model point of view, a forecasting system can comprise *pre-processing* approaches (statistical correction applied previously to the hydrologic modelling) and *post-processing* approaches (statistical correction applied to flow predictions). In all cases, calibration against

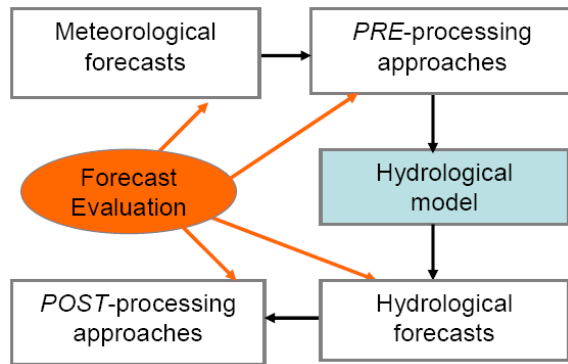


Figure 1. Schematic of a hydro-meteorological forecast chain.

observations and extensive testing over different hydrologic conditions are usually required to develop an operationally robust system. In order to optimize the implementation of post-processing techniques in real-time operational forecasting systems, a better understanding of the propagation of uncertainty from weather forecasts through hydrologic models and on the impact of non-linear hydrological transformations and hydrological updating on the ensemble streamflow forecasts is needed.

The aim of this paper is to investigate the use of statistical correction techniques in hydrological ensemble forecasting. We focus on the evaluation of different correction strategies (*pre-processing*, *post-processing* or both) and on their impact on the quality of operational streamflow forecasts. The context of the study and the modelling framework, including data and model used, are presented in Sect. 2; methodology and verification measures are described in Sect. 3; Sect. 4 presents the results, and in Sect. 5, conclusions are drawn.

2 Study context and modelling framework

The study is based on a modelling framework set up at the French electricity company (EDF) for the forecast of streamflows in France. EDF has produced hydrological forecasts for the past 60 yr (Lugiez and Guillot, 1960). Their operational interest includes flood forecasting (for human safety and dam security), short-term forecasting of water inflows to reservoirs, long-term prediction and reservoir management. For the last decade, EDF has invested into ensemble-based hydrological forecasting, comprising the acquisition of real-time meteorological forecast data and the set up of an appropriate hydrological modelling framework. Moreover, special attention has been paid to the performance of experiments on the use and communication of uncertainty in decision-making (Ramos et al., 2010). EDF forecasting chain follows the schematic description in Fig. 1. In this study, meteorological forecasts come from the 50 perturbed members of the ensemble prediction system produced by the European centre of medium-range weather forecasts (ECMWF-EPS). Opera-

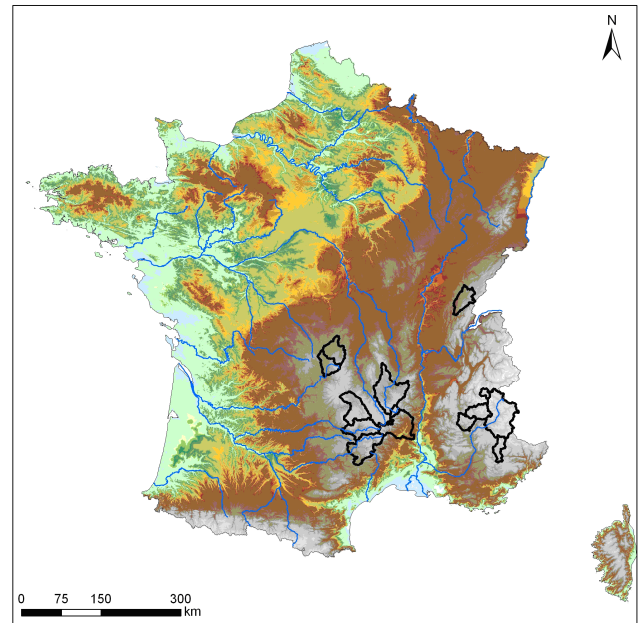


Figure 2. Location of the studied catchments in France.

tionally, EDF also uses the ECMWF-EPS control forecast, as well as deterministic forecasts produced by Météo-France. Forcing data (temperature and precipitation) are spatially aggregated at the catchment scale and evaluated by forecasters before being used as input to the hydrological model. The hydrological model is the MORDOR model. It is a lumped soil-moisture-accounting type rainfall-runoff model developed at EDF (Garçon, 1999). MORDOR has four reservoirs representing the physical processes in a river basin and a snow module that accounts for snow storage and melting in the catchment. The model version used in this study has 11 free parameters that were calibrated against observed data.

This study focuses on 11 catchments located in France, with areas ranging from 220 to 3600 km² (Fig. 2). Meteorological input fields are available at a horizontal resolution of 0.5 × 0.5 degree in latitude/longitude, ca. 50 km over France. The number of grid points falling within each catchment (at different percentages of coverage) varies from 2 to 42, with a median value of 15 grid points inside a catchment. Although some catchments are much smaller than the available meteorological grid scale, they were kept in the study and the sensitivity of the results to the catchment size was evaluated. At each catchment, forecasts are run at the daily time step and for a maximum forecast horizon of 7 days. Verification is performed over a 48-month forecast evaluation period (2005–2008). Forecasts are evaluated against observed daily areal precipitation and daily discharge data available at the outlet of the catchments.

3 Methodology and performance measures

3.1 Statistical correction strategies

Four main scenarios corresponding to different strategies for the processing of forecasts within the hydrologic modelling framework are tested:

- Scenario 1 – *Raw forecasts*: no pre- or post-processing of forecasts is performed and raw model outputs are evaluated;
- Scenario 2 – *Pre-processing of meteorological ensemble forecasts*: meteorological forecasts of temperature and precipitation are corrected using a technique developed at EDF, which is based on the search for analog situations previously archived in a database. At a given forecast day, for each perturbed member of the ECMWF-EPS system, the 50 most analog situations (according to the forecast fields of geopotential height at 700 and 1000 hPa) are retrieved from the database (for details on the use of analog techniques, see Zorita and von Storch, 1999 or Obled et al., 2002). A total of 2500 scenarios (50 ECMWF-EPS \times 50 analogs) is obtained. Each scenario consists of a pair of an analog situation (precipitation and temperature) and its corresponding ECMWF-EPS forecast. A corrected forecast scenario is calculated for each lead time t as indicated in Eq. (1):

$$V'_t = k \cdot V_t^{\text{ECMWF}} (V_t^{\text{ANALOG}} / \bar{V}_t^{\text{ANALOG}})^\alpha \quad (1)$$

V'_t : corrected forecast scenario for lead time t ;

V_t^{ECMWF} : ECMWF-EPS raw forecast for lead time t ;

V_t^{ANALOG} : analog forecast for lead time t ;

$\bar{V}_t^{\text{ANALOG}}$: mean value of the 50 selected analog forecasts for lead time t ;

α, k : parameters to be calibrated.

This combination of ECMWF-EPS and analog-based forecasts aims at perform an ensemble dressing of ECMWF forecasts (to improve reliability and to remove biases). The values of the parameters α and k can vary according to the studied catchment and lead time. In this study, they were considered constant and equal to 0.3 and 1.2, respectively. The 2500 scenarios obtained were sorted in ascending order and 50 scenarios (equidistant quantiles) were selected to be used as input to the hydrological model.

- Scenario 3 – *Post-processing of hydrological ensemble forecasts only*: the statistical correction technique used here takes into account only the errors from the hydrological model. It is thus independent of the meteorological forecasts (raw or pre-processed) used during the forecast evaluation period. Based on the past performance of the model, empirical errors are evaluated by taking the logarithm of the ratio between the observed and the simulated streamflows. Simulations are

obtained using observed precipitation and temperature available in the period 1970–2000 as input to the hydrological model. Subsamples of the errors are defined according to 20 classes of streamflow values (corresponding to a discretization of the cumulative distribution function at steps of 8 % between the 10 % and 90 % quantiles and 2 % for the tails of the distribution) and for each lead time (Mathevet, 2010). During forecasting, to “dress” each hydrological ensemble member, error values are drawn from the subsamples, according to the category to which the forecast discharge belongs, and added to the raw forecast value.

- Scenario 4 – *Pre-processing of meteorological forecasts and post-processing of hydrological ensemble forecasts*: the last scenario tested is the combination of the two correction approaches described above: the pre-processing of meteorological forecasts (scenario 2) and the post-processing of hydrological forecasts (scenario 3).

3.2 Forecast evaluation methods

Forecasts were evaluated against observations for a 48-month period from 2005 to 2008. Various measures are available in the literature to evaluate probabilistic forecasts and some have already been applied for the evaluation of hydrological forecasts (Wilks, 2011; Casati et al., 2008; Laio and Tamea, 2007). The scores used here are briefly presented below (see the references for details):

- Normalized RMSE: the root-mean-square error of the ensemble mean is normalized by the mean value of the observations during the forecast evaluation period to allow comparison among catchments of different sizes. Although the RMSE is not a score adapted to ensemble or probabilistic forecasts, we included it here as it is a score commonly used in hydrology.
- Brier Score (BS): one of the most common accuracy measure for forecast verification, the BS is essentially the mean squared error between the predicted probabilities for a set of events and their outcomes ($= 1$, if the event occurs and $= 0$ if it does not occur). The score takes values between 0 and 1; the lower the score, the higher the accuracy. In this study, we focus on the evaluation of severe events given by predictions exceeding the 80 % quantile of the empirical distribution of observed values.
- Rank Probability Score (RPS): the RPS is an extension of the BS to the many-event situation, computed, however, with respect to the cumulative probabilities in the forecast and observations vectors. The score takes values between 0 and 1; a “perfect forecast” receives RPS = 0. Here, we used 10 forecast categories to define

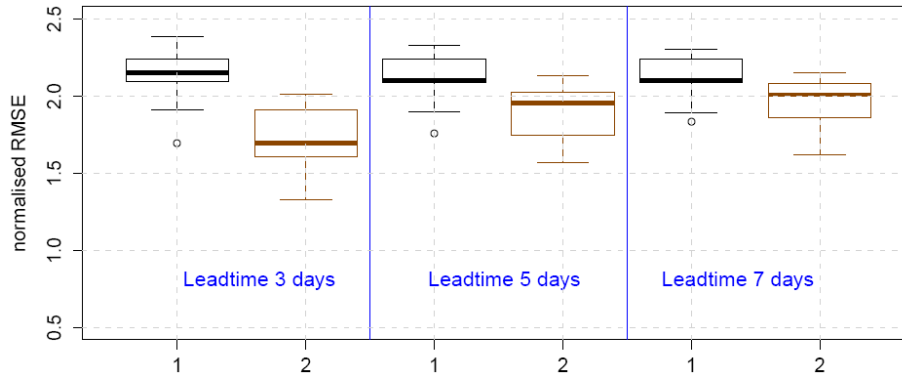


Figure 3. Boxplots of mean normalised RMSE of daily areal precipitation forecasts for lead times of 3 days (left), 5 days (centre) and 7 days (right) using: (1) raw forecasts, (2) pre-processed meteorological forecasts. Forecast evaluation over 48 months (2005–2008) and 11 catchments.

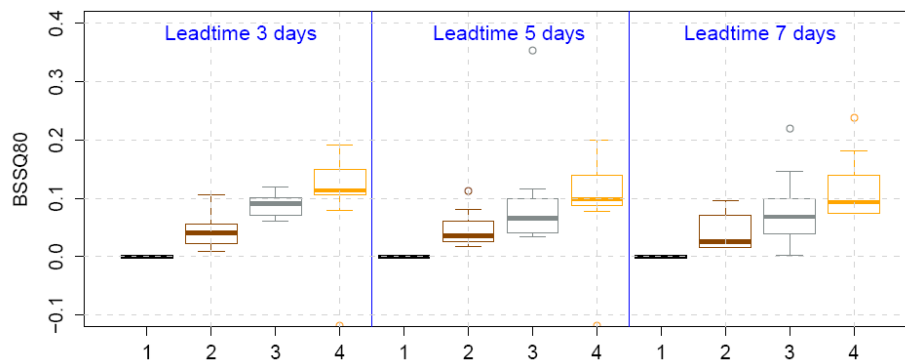


Figure 4. Boxplots of BSS (with reference to raw forecasts) for forecast streamflows exceeding the 80% percentile of observed flows for lead times of 3 days (left), 5 days (centre) and 7 days (right) using: (1) raw forecasts, (2) pre-processed meteorological forecasts, (3) post-processed hydrological forecasts, (4) pre-processed meteorological forecasts and post-processed hydrological forecasts. Forecast evaluation over 48 months (2005–2008) and 11 catchments.

the events, determined by the values of the 10% to the 90% quantiles of the empirical distribution of the observed values during the forecast evaluation period.

- Skill Scores (BSS or RPSS): the BS and the RPS are compared to a reference, which in this study corresponds to the raw forecasts (scenario 1). A skill score of 0 indicates a forecast with skill similar to the reference, while a forecast which is less (more) skilful than the reference will result in negative (positive) skill score values.
- Probability Integral Transform (PIT) histogram: the PIT histogram is a continuous analog of the rank histogram (Gneiting et al., 2005; Wilks, 2011), frequently used to verify the consistency of the forecasts, i.e. if the ensemble members of a forecast and the corresponding observations are samples from the same population (Wilks, 2011). If the ensemble consistency condition is satisfied, the relative frequencies given by the ensembles should estimate the actual (observed) prob-

ability. In this case, the PIT histogram shows as a uniform histogram, giving an indication of reliable forecasts. Under-dispersed forecasts will give U-shaped PIT histograms, while over-dispersed forecasts show relative frequencies concentrated in the middle ranks (arch-shaped). Asymmetrical histograms are an indication of over- or under-forecasting bias.

4 Results

The main results obtained from the comparative study of different bias correction strategies applied to meteorological and hydrological forecasts are illustrated in Figs. 3 to 6. For each studied catchment, mean values of the scores RMSE, BSS and RPSS are obtained by averaging the daily scores over the entire evaluation period (2005–2008). In Figs. 3 to 5, results from all catchments are displayed using the boxplot representation: the 75th and the 25th percentiles are represented at the top and the bottom of the box, respectively, while the top and the bottom of the tail indicate the 95th and

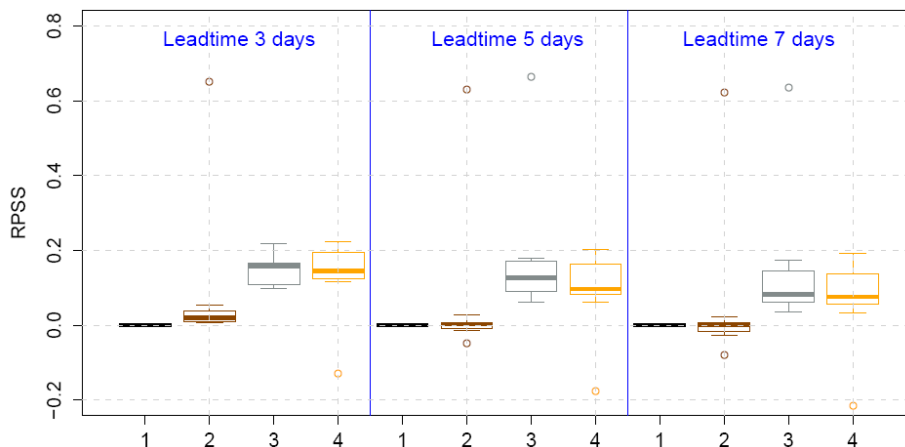


Figure 5. Boxplots of RPSS (with reference to raw forecasts) for forecast streamflows and lead times of 3 days (left), 5 days (centre) and 7 days (right) using: (1) raw forecasts, (2) pre-processed meteorological forecasts, (3) post-processed hydrological forecasts, and (4) pre-processed meteorological forecasts and post-processed hydrological forecasts. Forecast evaluation over 48 months (2005–2008) and 11 catchments.

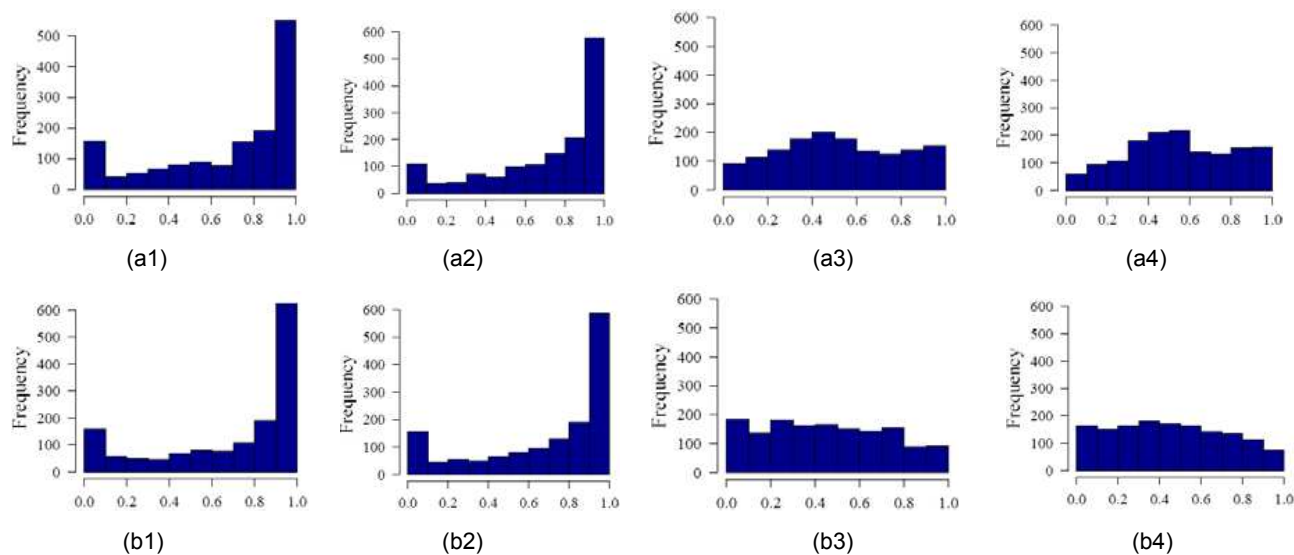


Figure 6. PIT histograms of forecast streamflows for lead time 7 days and the catchments (a) Tarn river at Milau (2138 km²) and (b) Durance river at Serre Ponçon (3580 km²) using: (1) raw forecasts, (2) pre-processed meteorological forecasts, (3) post-processed hydrological forecasts, and (4) pre-processed meteorological forecasts and post-processed hydrological forecasts. Forecast evaluation over 48 months (2005–2008).

the 5th percentile, respectively. PIT histograms are evaluated for each catchment. While Fig. 3 illustrates the results for precipitation (main meteorological input to the hydrological model), the other figures focus on the evaluation of streamflows.

Figure 3 shows the normalized RMSE values obtained from the evaluation of daily areal precipitation forecasts against observed precipitation data. The quality of raw forecasts (scenario 1) is compared to the quality of statistically processed forecasts (according to scenario 2), for leadtimes of 3, 5 and 7 days. The statistical correction applied reduces

significantly the forecast errors and improves forecast precision, especially for short lead times. The results from the other scores (not shown here) indicate the same tendency, confirming the efficiency of the applied statistical correction technique to improve the precision and the reliability of the meteorological input to the hydrological model.

Concerning the impact of statistical correction on the quality of streamflow forecasts, Figs. 4 and 5 show, respectively, the Brier Skill Scores for flow forecasts exceeding the 80 % percentile of observed flows and the Rank Probability Skill Scores for the four scenarios of correction strategies studied.

Both scores show that the use of a pre-processing technique (scenario 2) generally improves the quality of ensemble streamflow forecasts in the studied catchments: BSS and RPSS values are higher than 0, showing an improvement in forecast skill with respect to the raw forecasts (scenario 1). This positive impact of pre-processing meteorological forecasts on the quality of hydrological forecasts is greater at shorter lead times.

Furthermore, results for the scenarios 3 and 4 illustrate the added value of implementing also a post-processing correction approach of the hydrological model outputs. For the prediction of events of high flows (Fig. 4), the implementation of pre- and post-processing techniques together (scenario 4) conducts to the highest score values (better forecast quality). When considering the RPSS values (Fig. 5), differences between scenario 3 and scenario 4 are more significant for shorter lead times. For longer lead times, skill scores achieved for scenario 3, where raw meteorological forecasts are used and only hydrological outputs are post-processed, are basically equivalent to those achieved when using scenario 4, where both pre-processing and post-processing are performed. The analysis of BSS and RPSS as a function of catchment size (not shown) did not indicate a clear sensitivity of the results to the catchment area. Only at longer lead times, hydrological forecasts based only on pre-processed meteorological forecasts showed negative values of skill scores for the largest catchments, i.e. less skilful forecasts comparatively to the reference (raw forecasts). Further analysis, with a larger sample of catchments, would be necessary to better detect any general tendency.

The PIT histograms in Fig. 6 illustrate the impact of statistical correction strategies on the reliability of streamflow forecasts for two catchments representative of the studied sample and for forecast lead time of 7 days. For both catchments, scenario 1 with no bias correction strategy (raw forecasts) displays biased under-dispersive streamflow ensemble forecasts. The use of scenario 2 (statistical correction applied only to meteorological forecasts) does not improve the PIT histograms of streamflow forecasts. Since the PIT histograms of statistical corrected precipitations (not shown here) do not display significant under-dispersion problems, the examples shown in Fig. 6 illustrate the impact of the rainfall-runoff transformation on the spread of the ensemble streamflow forecasts. It is possible that the added value of pre-processing meteorological input forcings in these cases has been obscured by the mixed evaluation of high and low (more frequent) streamflow periods. Biases in the modelling of the recession part of hydrographs can be of different nature from biases in the modelling of high flows. It would be interesting to separate the evaluation of hydrological forecasts by considering separately flood and recession periods. This is part of an ongoing study and is beyond the scope of this paper. Furthermore Fig. 6 presents also the impact of applying the post-processing approaches described in scenarios 3 and

4. For these scenarios a significant improvement of forecast reliability is observed for both catchments studied.

5 Conclusions

This paper investigates the use of statistical bias correction techniques in hydrological ensemble forecasting. Our main focus is on evaluating of the impact of different strategies of statistical bias correction on the quality of operational streamflow forecasts. From the hydrological model point of view, forecasters can use *pre-processing* approaches (statistical corrections applied prior to the hydrologic modelling, i.e. on the meteorological forcing), *post-processing* approaches (statistical corrections applied only to the output of the hydrological model, i.e. streamflow predictions) or both. We compared performance measures obtained for 11 catchments in France during a 48-month evaluation period (2005–2008) according to four scenarios of statistical bias correction: raw forecasts, only pre-processed meteorological forecasts, only post-processed hydrological forecasts and with statistical processing applied to both meteorological and hydrological forecasts.

Results show that even though correcting the meteorological uncertainties is of high importance to obtain precise and reliable inputs to the hydrological model, the errors linked to hydrological modelling remain a key-component of the total predictive uncertainty of hydrological ensemble forecasts. Statistical corrections made to precipitation forecasts can lose their effect when propagated through the hydrological model. As a result efforts to also implement a post-hydrological model correction may be necessary. In this paper we showed that even a relatively simple empirical post-processing approach can be useful to achieve reliable hydrological forecasts for operational needs. Future work should include the application of other statistical correction techniques and the use of other hydrological models and performance measures on a larger set of catchments.

Acknowledgements. The authors would like to thank Massimiliano Zappa and the anonymous referee for their constructive comments on this paper.

Edited by: B. Ahrens

Reviewed by: M. Zappa and one anonymous referee



The publication of this article is sponsored by the European Meteorological Society.

References

- Bartholmes, J. C., Thielen, J., Ramos, M. H., and Gentilini, S.: The European flood alert system EFAS – Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts, *Hydrol. Earth Syst. Sci.*, 13, 141–153, doi:10.5194/hess-13-141-2009, 2009.
- Boucher M.-A., Tremblay, D., Delorme, L., Perreault, L., and Anctil, F.: Hydro-economic assessment of hydrological forecasting systems, *J. Hydrol.*, 416–417, 133–144, 2012.
- Brown, J. D. and Seo, D.-J.: A nonparametric post-processor for bias-correction of hydrometeorological and hydrologic ensemble forecasts, *J. Hydrometeorol.*, 11, 642–665, 2010.
- Casati, B., Wilson, L. J., Stephenson, D. B., Nurmi, P., Ghelli, A., Pocerich, M., Damrath, U., Ebert, E. E., Brown, B. G., and Mason, S.: Forecast verification: current status and future directions, *Meteorol. Appl.*, 15, 3–18, 2008.
- Cloke, H. and Pappenberger, F.: Ensemble Flood Forecasting: A Review, *J. Hydrol.*, 375, 613–626, 2009.
- Fortin, V., Favre, A. C., and Eriem, S.: Probabilistic forecasting from ensemble prediction systems: Improving upon the best-member method by using a different weight and dressing kernel for each member, *Q. J. Roy. Meteorol. Soc.*, 132, 1349–1369, 2006.
- Garçon, R.: Modèle global Pluie-Débit pour la prévision et la prédétermination des crues (Lumped rainfall-runoff model for flood forecasting and design flood estimation), *La Houille Blanche*, 7/8, 88–95, 1999 (in French).
- Gneiting, T., Raftery, A. E., Westveld, A. H., and Goldman, T.: Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation, *Mon. Weather Rev.*, 133, 1098–1118, 2005.
- Hashino, T., Bradley, A. A., and Schwartz, S. S.: Evaluation of bias-correction methods for ensemble streamflow volume forecasts, *Hydrol. Earth Syst. Sci.*, 11, 939–950, doi:10.5194/hess-11-939-2007, 2007.
- Jaun, S. and Ahrens, B.: Evaluation of a probabilistic hydrometeorological forecast system, *Hydrol. Earth Syst. Sci.*, 13, 1031–1043, doi:10.5194/hess-13-1031-2009, 2009.
- Krzysztofowicz, R.: Bayesian theory of probabilistic forecasting via deterministic hydrologic model, *Water Resour. Res.*, 35, 2739–2750, 1999.
- Laio, F. and Tamea, S.: Verification tools for probabilistic forecasts of continuous hydrological variables, *Hydrol. Earth Syst. Sci.*, 11, 1267–1277, doi:10.5194/hess-11-1267-2007, 2007.
- Lugiez, F. and Guillot, P.: Dix années de prévisions d'apports à Electricité de France, *IAHS-AISH P.*, 51, 558–566, 1960.
- Mathevet, T.: Erreur empirique de modèle, Note technique internet EDF-DTG D4165/NT/2010-00395-A, Internal Technical Report, p. 11, 2010.
- Obed, C., Bontron, G., and Garçon, R.: Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach, *Atmos. Res.*, 63, 303–324, 2002.
- Olsson, J. and Lindström, G.: Evaluation and calibration of operational hydrological ensemble forecasts in Sweden, *J. Hydrol.*, 350, 14–24, 2008.
- Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M.: Using Bayesian model averaging to calibrate forecast ensembles, *Mon. Weather Rev.*, 133, 1155–1174, 2005.
- Ramos, M. H., Mathevet, Th., Thielen, J., and Pappenberger, F.: Communicating uncertainty in hydro-meteorological forecasts: mission impossible?, *Meteorol. Appl.*, 17, 223–265, 2010.
- Randrianasolo, A., Ramos, M.-H., Thirel, G., Andréassian, V., and Martin, E.: Comparing the scores of hydrological ensemble forecasts issued by two different hydrological models, *Atmos. Sci. Lett.*, 11, 100–107, 2010.
- Schaake, J., Pailleux, J., Thielen, J., Arritt, R., Hamill, T., Luo, L., Martin, E., McCollor, D., and Pappenberger, F.: Summary of recommendations of the first workshop on Postprocessing and Downscaling Atmospheric Forecasts for Hydrologic Applications held at Météo-France, Toulouse, France, 15–18 June 2009, *Atmos. Sci. Lett.*, 11, 59–63, 2010.
- Thirel, G., Rousset-Regimbeau, F., Martin, E., and Habets, F.: On the impact of short-range meteorological forecasts for ensemble stream flow predictions, *J. Hydrometeorol.*, 9, 1301–1317, 2008.
- Velazquez, J. A., Anctil, F., Ramos, M. H., and Perrin, C.: Can a multi-model approach improve hydrological ensemble forecasting? A study on 29 French catchments using 16 hydrological model structures, *Adv. Geosci.*, 29, 33–42, 2011, <http://www.adv-geosci.net/29/33/2011/>.
- Verkade, J. S. and Werner, M. G. F.: Estimating the benefits of single value and probability forecasting for flood warning, *Hydrol. Earth Syst. Sci.*, 15, 3751–3765, doi:10.5194/hess-15-3751-2011, 2011.
- Wilks, D. S.: *Statistical Methods in the Atmospheric Sciences*, 3rd Edn., Academic Press, Amsterdam, 627 pp., 2011.
- Zhao, L., Duan, Q., Schaake, J., Ye, A., and Xia, J.: A hydrologic post-processor for ensemble streamflow predictions, *Adv. Geosci.*, 29, 51–59, 2011, <http://www.adv-geosci.net/29/51/2011/>.
- Zorita, E. and von Storch, H.: The analog method as a simple statistical downscaling technique: Comparison with more complicated methods, *J. Climate*, 12, 2474–2489, 1999.