Atmospheric Research xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmos

Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty

R. Hostache [a,*], P. Matgen [a], A. Montanari [b], M. Montanari [a], L. Hoffmann [a], L. Pfister [a]

[a] CRP-Gabriel Lippmann, EVA Department, Belvaux, Grand-Duchy of Luxembourg
[b] Faculty of Engineering, University of Bologna, Bologna, Italy

ARTICLE INFO

Article history:
Received 4 June 2009
Received in revised form 6 August 2010
Accepted 20 September 2010
Available online xxxx

Keywords:
Rainfall-runoff model
Forecasting chain
Linear model
Uncertainty
Bivariate meta-gaussian density

ABSTRACT

The pressure on the scientific community to provide medium term flood forecasts with associated meaningful predictive uncertainty estimations has increased in recent years. A technique for assessing this uncertainty in hydro-meteorological forecasting systems is presented. In those, the uncertainties generally propagate from an atmospheric model through a rainfall-runoff model. Consequently, it appears to be difficult to isolate the errors that stem from the individual model components. In this study, the integrated flood forecasting system uses the 7-day rainfall and temperature forecast of the American atmospheric GFS model (deterministic run) as forcing data in a conceptual hydrologic model (deterministic run) coupled with a linear error model in order to predict river discharge. The linear error model is added to the hydrologic model run, in order to take advantage of the correlation in time between forecasting errors, thereby reducing errors that arise from hydrologic simulations. To assess the predictive uncertainty (total uncertainty) of the coupled models, the method makes use of a bivariate meta-gaussian probability density function. The latter allows estimating the probability distribution of the integrated model errors conditioned by the predicted river discharge values. The proposed methodology is applied to the case study of the Alzette river located in the Grand Duchy of Luxembourg. Confidence limits are computed for various lead times of prediction and compared with observations of river discharge.

1. Introduction

Medium range rainfall forecasts are increasingly used in operational flood forecasting applications as they provide an inviting option for extending lead-times of prediction. Nonetheless there is significant uncertainty associated with hydro-meteorological simulations. As a matter of fact, techniques for assessing hydro-meteorological model uncertainty have received a great deal of attention by researchers in recent years (Montanari, 2007). In any flood forecasting system, the predictive uncertainty originates from several causes interacting between each other, namely input uncertainty, model structure uncertainty and parameter uncertainty (Liu and Gupta, 2007). In flood forecasting applications it is well known that the input uncertainty, in particular precipitation uncertainty, affects the total simulation uncertainty very significantly (e.g., Andreassian et al., 2001; Koussis et al., 2003). In order to quantify the uncertainty that is associated with the forcing data, one alternative is to provide the atmospheric model output as an ensemble and not single best estimates. Nowadays most weather services that are running atmospheric models provide rainfall predictions in terms of probabilistic precipitation forecasts. For instance, a probability distribution for the future rainfall can be inferred from ensemble simulation. Individual ensemble members

* Corresponding author. Tel.: +352 470261 417; fax: +352 470264.
E-mail address: hostache@lippmann.lu (R. Hostache).

0169-8095/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.atmosres.2010.09.014

are generated by perturbing the initial conditions of the numerical weather prediction model and by adding a random model error due to parameterization (Buizza et al., 1999).

In operational forecasting systems, the input uncertainty adds up to the parameter and model structure uncertainties that are inherent in hydrologic modelling. The rainfall-runoff models are limited by their representation of flow dynamics and this is as much a problem of knowing the local characteristics (i.e. parameter uncertainty) as it is of representing the dynamics (i.e. model uncertainty) themselves (Pappenberger et al., 2005). A means to assess parameter and model structural uncertainty is the Monte-Carlo approach (e.g., Beven and Binley, 1992), that is carried out by identifying a set of behavioural model structures and parameter sets with respect to observed discharge data. The individual models can be conditioned on observed discharge time series and predictions are made with all the behavioural models. A probability distribution of all the likelihood-weighted model outputs enables a quantitative assessment of the uncertainty that arises from parameter uncertainty.

Hence, each one of the two components of a flood forecasting system is subject to considerable uncertainty and in operational forecasting systems the challenge consists in cascading the uncertainty through the entire modelling system. One straightforward strategy would be to force the sample of behavioural hydrologic models with the rainfall forecasts of all individual ensemble members of the atmospheric model, thereby generating many combinations of rainfall inputs and runoff predictions (Pappenberger et al., 2005). Again, the overall uncertainty is assessed via the predictive percentiles that are obtained from all model realisations. A critical issue related to this approach is to estimate the likelihood of each realisation. An alternative option is the Bayesian Forecasting System introduced by Krzysztofowicz (2002) that allows producing probabilistic river stage forecasts based on probabilistic precipitation forecasts.

This paper aims at estimating the forecast uncertainty by using a third option, that is based on the analysis of the statistical properties of deterministic hydro-meteorological forecast series. The latter are computed with respect to historic time series of observed discharge. The advantage of this approach, with respect to Monte-Carlo techniques, is that a method based on model error statistics leads to the estimation of the total uncertainty in an aggregated forecasting system (e.g., coupled atmospheric–hydrologic models), thereby rendering the assessment of uncertainty originating from the individual contributions unnecessary. The main limitation of such methods is the potential non-stationarity of the statistical properties of forecast series. Indeed, when focusing on extreme flood events, the extrapolation of statistics can be a significant source of error. Moreover, it is difficult to infer statistical properties from the prediction error, since its statistics often appear to be non-stationary. In particular, the prediction error itself is often affected by heteroscedasticity and also by serial correlation which increases the uncertainty of the estimators (Toth et al., 2000).

In this framework, an approach for constructing confidence intervals of hydrologic model simulations was recently proposed by (Montanari and Brath, 2004). In their study, the estimation of probability distributions of runoff simulation errors, conditioned by the value of flow, is performed using a meta-gaussian model.

More specifically, this paper adopts the meta-gaussian approach that was initially developed by Kelly and Krzysztofowicz (1997) to estimate the uncertainty of real-time river flow forecasts, by adapting the application framework that was proposed by (Montanari and Brath, 2004) and (Montanari and Grossi, 2008). Some modifications are actually introduced in the above-mentioned framework. On one hand, a linear error model is used to update the river flow prediction provided by a rainfall-runoff model. On the other hand, the setup of the forecasting chain scheme aims at being widely applicable, providing continuous forecast instead of event forecasting, and providing medium range forecast (prediction lead time until 120 h). In this framework, this study makes use of GFS (Kalnay et al., 1971; Kanamitsu et al., 1991) meteorological forecast, globally available around the world. The approach will be tested by the means of a case study that focuses on a flood forecasting system that was set-up on the Alzette river in Luxembourg.

2. Method

This section aims at explaining the setup of the forecasting system, namely i) the calibration of the hydrologic model, that uses as main input deterministic meteorological rainfall forecasts, ii) the estimation of the linear error model coefficients, and iii) the estimation of the total predictive uncertainty of the forecasting chain defined by a hydrologic model coupled with a linear error model. The objective of this methodology is to assess the total predictive uncertainty of the runoff forecasts using the meta-gaussian bivariate density, without trying to separate the individual sources of uncertainty.

2.1. The rainfall-runoff model

The FLEX hydrologic model (Matgen et al., 2005) employed in this study simulates hourly discharge using as input rainfall (R_{tot}) and potential evapotranspiration (ETP) data, derived from the surface temperature according to the Hamon formula (Hamon, 1963). This is a 9-parameter lumped conceptual model (Fig. 1) that can be considered as a modified version of the HBV-96 model (Lindström et al., 1997). The soil reservoir module is characterized by the following parameters, namely the maximal storage capacity $S_{\max}$ [mm], a parameter of non-linearity b [−] controlling the infiltration capacity, a parameter λ [−] giving the fraction of $S_{\max}$ below which the actual evapo-transpiration (ETR) is constrained by $S(t)$ and the percolation rate p [mm. h^{-1}]. The modified soil reservoir module is drained by evaporation and deep percolation fluxes. The rainfall is divided into two fractions. The first one fills the soil reservoir module, while the second part consists of the net rainfall that fills the two routing reservoirs (linear baseflow reservoir and non-linear fast runoff reservoir).

To calibrate this hydrologic model, the procedure proposed by Fenicia et al. (2007), has been applied using

Please cite this article as: Hostache, R., et al., Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty, Atmos. Res. (2010), doi:10.1016/j.atmosres.2010.09.014

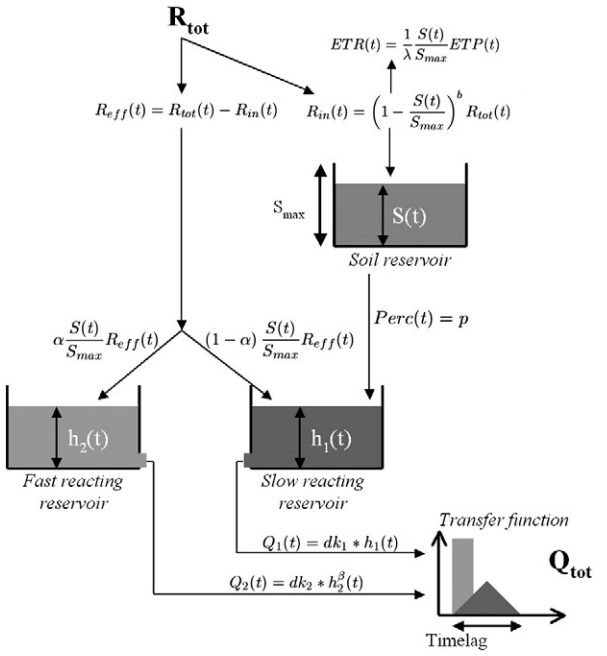

Fig. 1. Schematic representation of the lumped conceptual rainfall-runoff model.

the discharge derived from *in situ* water stage measurements and a rating curve. In their study, Fenicia et al. (2007) showed that using two (instead of one) sets of optimal parameters, one for low flow and one for high flow conditions, improves the model capability of reproducing observed discharge. Hence, two optimal hydrologic parameter sets are sought, one for the low-flow conditions and one for the high-flow conditions. To combine high- and low-flow hydrologic models, two additional parameters δ and γ that govern the transition between the two aforementioned models (one model being defined by one set of optimal parameters) are introduced in the calibration process (see Fig. 2) (Fenicia et al., 2007).

Considering $Q_{HF}(t)$ and $Q_{LF}(t)$ the discharges simulated by the high- and low-flow models respectively, at time t, the membership function shown in Fig. 2 is used as follows (1):

$$Q_s(t) = \begin{cases} Q_{HF}(t) & \text{if } Q_s(t-1) \geq \delta . Q_{o,max} \\ Q_{LF}(t) & \text{if } Q_s(t-1) \leq \gamma . Q_{o,max} \\ \begin{bmatrix} c . Q_{HF}(t) + \\ (1-c) . Q_{LF}(t) \end{bmatrix} & \text{otherwise} \end{cases} \quad (1)$$

$$\text{with } c = \left(\frac{Q_s(t-1)}{Q_{o,max}} - \gamma \right) / (\delta - \gamma)$$

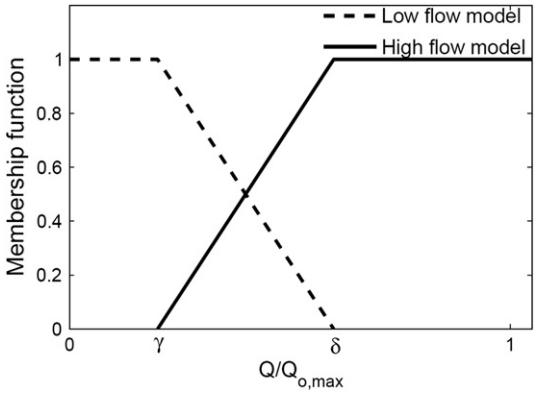

Fig. 2. Membership function used to combine high- and low-flow models as proposed by (Fenicia et al., 2007).

In Eq. (1), $Q_s(t)$ is the river flow simulated at time t, $Q_{o,\max}$ is the maximum observed river flow during the calibration period and δ and γ are calibrated parameters (see below).

The calibration of high- and low-flow hydrologic parameters and the combination function parameters is carried out in two steps using Monte-Carlo simulations. In a first step, hydrologic parameter sets are randomly generated within intervals of physically plausible values. Next, for each generated parameter set, one hydrologic model run is performed using as input temperature and rainfall measurements and the results of these simulations are compared with discharge observations. For this comparison, two performance criteria N_{LF} (2), N_{HF} (3) have been used (Fenicia et al., 2007),

$$N_{LF} = \sqrt{\frac{\sum_{t=1}^{n} \left((Q_s(t) - Q_o(t))^2 . \left(\frac{Q_{o,max} - Q_o(t)}{Q_{o,max}} \right)^2 \right)}{n}} \quad (2)$$

Please cite this article as: Hostache, R., et al., Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty, Atmos. Res. (2010), doi:10.1016/j.atmosres.2010.09.014

$$N_{HF} = \sqrt{\frac{\sum_{t=1}^{n}\left((Q_s(t)-Q_o(t))^2 \cdot \left(\frac{Q_o(t)}{Q_{o,\max}}\right)^2\right)}{n}} \quad (3)$$

In Eqs. (2) and (3), $Q_o(t)$ is the observed river flow at time t.

The two hydrologic parameter sets minimizing respectively N_{LF} and N_{HF} are defined as optimal respectively for low-flow and high-flow conditions. In a second calibration step, combination parameter sets (δ,γ) are randomly generated within intervals of physically plausible values. Next, for each generated parameter set (δ,γ), the "combined" simulated discharge is calculated and compared with discharge observations using the Nash–Sutcliffe efficiency criterion E_{NS} (see Eq. 4) (Nash and Sutcliffe, 1970),

$$E_{NS} = 1 - \frac{\sum_{t=1}^{n}(Q_s(t)-Q_o(t))^2}{\sum_{t=1}^{n}\left(Q_s(t)-\overline{Q_o}\right)^2} \quad (4)$$

In Eq. (4), $\overline{Q_o}$ is the average observed discharge during the calibration time window consisting of n time steps. The (δ,γ) parameter set maximizing the Nash–Sutcliffe efficiency criterion is defined as optimal.

2.2. Reduction of hydrologic model errors using a linear model

We propose to use a linear model in order to update the output of the hydrologic model thereby reducing its error. The use of linear models for updating discharge forecasts was proposed, among others, by WMO (1992), Toth et al. (1999) and Brath et al. (2002) in order to reduce the bias of the hydrologic model by taking advantage of the significant autocorrelation that usually affects the forecast errors. Linear models are suited to this purpose regarding their simple architecture, which often provides a satisfactory approximation of the structure of the output model error. In detail, we compute the error as follows:

$$e(t) = Q_o(t) - Q_s(t) \quad (5)$$

In Eq, (5), $Q_s(t)$ now indicates the simulated river flow. Then, the linear model applied herein is given by:

$$\hat{e}(t) = a \cdot e(t-1) + b \quad (6)$$

In (6), the coefficients a and b are estimated by using least squares optimisation of the fit of the forecast errors. The least squares technique was preferred because it puts more emphasis on the optimisation of higher river discharges. The multistep ahead prediction is carried out by using the errors estimated at the previous time steps instead of the real error in Eq. (6).

2.3. The forecasting chain and the total predictive uncertainty assessment

The last part of the method aims at assessing the total predictive uncertainty. The forecasting chain developed in this study is presented in Fig. 3. The hydrologic part of this chain consists of the two calibrated hydrologic models (for low- and high-flow conditions respectively) and the combination function (Fig. 3) using as input rainfall and temperature observed over the last 24 h concatenated with forecasts for the next 120 h. As a first test, we assume that predictions are developed every 1 h. As can be seen in Fig. 3, the hydrologic model is initialized for each simulation by using the state of the model at the same date during the last developed prediction. As we mentioned above, to reduce the model error, the discharge forecast is corrected using the linear error model proposed in Section 2.2.

Assuming that an extensive sample of simulated discharges and corresponding errors has been built, this archive contains series of discharge values simulated by the rainfall-runoff model using rainfall and temperature forecasts over 120 h and corresponding series of errors calculated *a posteriori* between observed and simulated discharge values (series of errors calculated for each lead time of forecast: 1 to 120 h).

The assessment of the total predictive uncertainty is then addressed using a meta-gaussian model (Kelly and Krzysztofowicz, 1997), based on the standard Normal Quantile Transform, also called Normal Score (noted NSC hereafter) that can be used to normalise samples of data. As recently argued by Montanari and Grossi (2008), it is assumed herein that the positive and negative errors stem from two different stochastic processes $E^+(t)$ and $E^-(t)$. As a consequence, the error and discharge archives are separated into two parts, the first one corresponding to positive errors, and the second one to negative errors. To calculate the NSC, the method proposed by Montanari and Brath (2004) is used. It consists of two steps, explained here with reference to the positive model error ($e^+(t)$) at time t. First, the cumulative probability $P(E^+(t)<e^+(t))$ is approximated by the cumulative frequency $F(e^+(t))$ in the sample, calculated using the Weibull plotting position based on the ranks in the ascending-order-sorted sample (Stedinger et al., 1993), (see Eq. (7) below):

$$F\left(e^+(t)\right) = \frac{j(t)}{n+1} \quad (7)$$

In Eq. (7), $j(t)$ is the rank of $e^+(t)$ in the ascending-order-sorted sample and n the size of the sample. Next, for each $F(e^+(t))$, the corresponding standard normal quantile $Ne^+(t)$ is calculated and associated with $e^+(t)$. Then, the NSC of e^+ in $\Re$ (called NSC_{e+} hereafter) is provided by linearly interpolating between the points $(Ne^+(t),e^+(t))$ for $t\in[1;n]$ and linearly extrapolating beyond the maximum and the minimum. Furthermore, it is worth noting that the interpolated–extrapolated $(Ne^+(t),e^+(t))$ for $t\in\Re$ provide at the same time the inverse of the standard Normal Quantile Transform of e^+ ($NSC_{e^+}^{-1}$).

Once the normalised series $Ne^+(t)$ and $NQ(t)$ have been calculated by applying respectively the NSC_e and the NSC_Q transformations, it is assumed that the underlying stochastic processes, $Ne^+(t)$ and $N\theta(t)$, respectively, are stationary and ergodic (i.e. the statistical properties of the stochastic process can be estimated from a single, sufficiently long

Please cite this article as: Hostache, R., et al., Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty, Atmos. Res. (2010), doi:10.1016/j.atmosres.2010.09.014

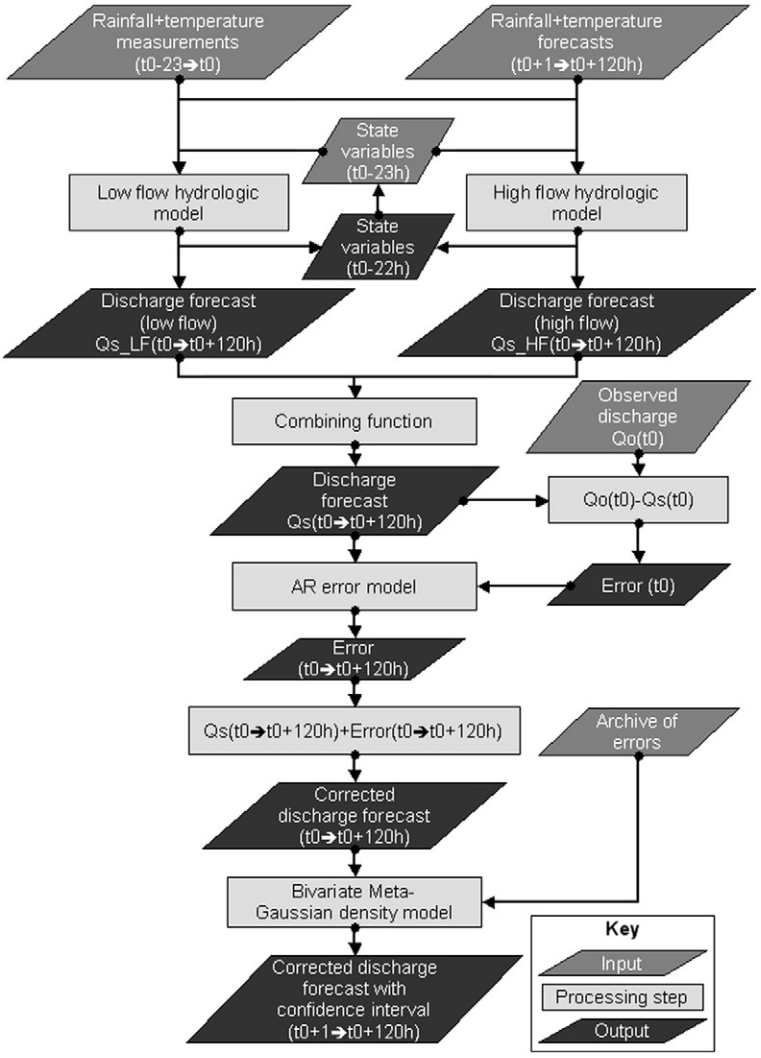

Fig. 3. Schematic representation of the forecasting chain.

realisation), and that their cross dependence is governed by the normal linear equation (Montanari and Brath, 2004):

$$Ne^{+}(t) = \rho_{Ne^{+},NQ}(0)\cdot NQ(t) + \varepsilon^{+}(t) \tag{8}$$

In Eq. (8), $\rho_{Ne^{+},NQ}(0)$ is the zero Pearson's cross correlation coefficient between $Ne^{+}(t)$ and $NQ(t)$, and $\varepsilon^{+}(t)$ is a realisation of a stochastic process $\Upsilon^{+}(t)$ independent of $N\Theta(t)$ and normally distributed, with mean zero and variance $1-\rho^{2}_{Ne^{+},NQ}(0)$ (Montanari and Brath, 2004).

Accordingly to Montanari and Grossi (2008), the stochastic framework introduced above provides a means for computing the confidence bands for the predicted river discharge at the α significance level. To this end, one first of all needs to compute the $\alpha/2$ and $1-\alpha/2$ quantiles for the stochastic process $\Upsilon(t)$ obtained by properly mixing the positive and negative stochastic processes $\Upsilon^{+}(t)$ and $\Upsilon^{-}(t)$ defined above, accordingly to the relationships:

$$\begin{aligned} \Upsilon(t)(1-\alpha/2) &= \xi^{+}(\alpha)\cdot\sigma(\varepsilon_t^{+}) \\ &= \xi^{+}(\alpha)\sqrt{1-\rho^{2}_{Ne^{+},NQ}(0)} \\ \Upsilon(t)(\alpha/2) &= \xi^{-}(\alpha)\cdot\sigma(\varepsilon_t^{-}) \\ &= \xi^{-}(\alpha)\sqrt{1-\rho^{2}_{Ne^{-},NQ}(0)} \end{aligned} \tag{9}$$

Let us denote with q^{-1} the inverse of the standard normal quantile and with p^{+} the probability for the forecasting error to be positive, which is assumed to be independent of time and therefore is estimated by computing the percentage of positive errors during an assigned forecasting period. Then,

Please cite this article as: Hostache, R., et al., Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty, Atmos. Res. (2010), doi:10.1016/j.atmosres.2010.09.014

the coefficients $\xi^+(\alpha)$ and $\xi^-(\alpha)$ above can be computed as follows:

$$\xi^+(\alpha) = q^{-1}\left(1-\frac{\alpha}{2p^+}\right)$$
$$\xi^-(\alpha) = q^{-1}\left(\frac{\alpha}{2p^-}\right) \quad (10)$$

Given that p^+ and p^- can be arbitrarily close to 0, in the technical application of Eq. (10), one may obtain values greater than 1 of $\alpha/(2p^+)$ and $\alpha/(2p^-)$, which would lead to an inconsistent quantile computation. From a technical point of view, this means that the probability of getting a positive (or negative) forecast error is small enough so that the width of the corresponding confidence band can be assumed to be negligible at the α significance level (Montanari and Grossi, 2008).

At this stage, the normalised forecast errors can be computed accordingly to the normal linear Eq. (8), depending on the previously computed residual quantiles (9) and the normalised discharge *NQ(t)* forecasted by the hydrological model. Finally, the confidence bands for the forecasted river flow, at the α significance level, can be computed by applying back the NSC and adding the predicted positive and negative errors to the forecasted discharge, accordingly to the relationships:

$$Q^+(t) = Q(t) + \mathrm{NSC}_{e^+}^{-1}\left[\begin{array}{l}\rho_{Ne^+,NQ}(0)\cdot NQ(t) + \\ \xi^+(\alpha)\cdot\sqrt{1-\rho^2_{Ne^+,NQ}(0)}\end{array}\right]$$
$$Q^-(t) = Q(t) + \mathrm{NSC}_{e^-}^{-1}\left[\begin{array}{l}\rho_{Ne^-,NQ}(0)\cdot NQ(t) + \\ \xi^-(\alpha)\cdot\sqrt{1-\rho^2_{Ne^-,NQ}(0)}\end{array}\right] \quad (11)$$

The confidence intervals are calculated independently for each prediction lead time. This means that two bivariate meta-gaussian densities (positive and negative errors) are computed for each lead time of prediction (1–120 h). Consequently, the meta-gaussian density allows for computing a confidence interval for each discharge forecast.

3. Study area and available data

The study area (Fig. 4) that was chosen for applying and testing the method is a sub-catchment of the Alzette basin (Grand-Duchy of Luxembourg). This includes the upstream part of the Alzette catchment until the stream gauge of Pfaffenthal. At this point of the river, the catchment has an area of 356 km^2. The hydrometeorological data used in this study were observed in one hydrometric and two meteorological stations (see Fig. 4) operating since 1996 with a recording time step of 15 min, namely:

- the hydrometric station of Pfaffenthal (collecting river stage observations);
- the meteorological station of Livange (collecting precipitation observations);
- the meteorological station of Mersch (collecting temperature observations).

These stations have been chosen because their data can be downloaded in real-time in operational mode. Mersch is located outside the considered catchment but the measurements therein collected are representative of the study area. Moreover, between October 9, 2006 and February 11, 2009, the half degree medium range rainfall forecasts provided by the Global Forecasting System (GFS) atmospheric model (Kalnay et al., 1971; Kanamitsu et al., 1991) have been recorded as a 120 h rainfall forecast dataset for the geographical coordinates corresponding to the Alzette catchment (latitude 6° east, longitude 49.5° north). These data are provided by the National Oceanic and Atmospheric Administration (NOAA) through a web page every 6 h, as 3 h cumulated rainfall forecasts for the next 7 days. Higher resolution atmospheric models may help to reduce the total prediction uncertainty of discharge forecast. However, we decided to use GFS data in order to propose and test a methodology that is easily transferable. In fact, the GFS rainfall and temperature forecasts are available all over the world.

4. Results and discussion

This section outlines the findings of the study and discusses their implications. The first part of this section explains the set up of the hydro-meteorologic forecasting chain, the second part focuses on the assessment of discharge

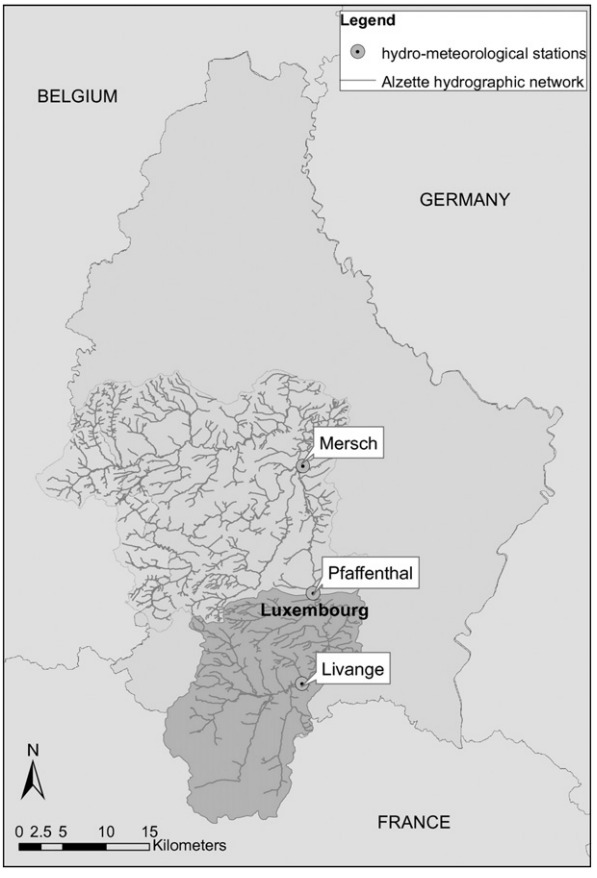

Fig. 4. Study area, Alzette catchment and hydro-meteorological stations.

Please cite this article as: Hostache, R., et al., Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty, Atmos. Res. (2010), doi:10.1016/j.atmosres.2010.09.014

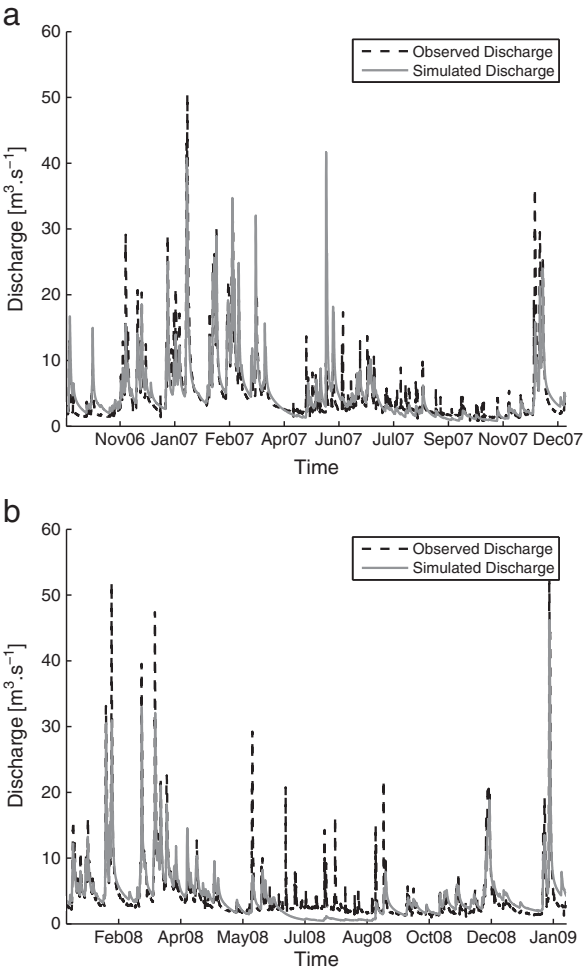

Fig. 5. Observed and simulated hydrographs during (a) the calibration period and (b) the validation period.

forecast uncertainty and the last part discusses the performance, benefits and limitations of the methodology.

4.1. Model calibration

As proposed in Section 2.1, the offline calibration of the FLEX hydrologic model has been done through Monte-Carlo simulations. For this procedure, sets of hydrologic parameters (100,000 sets) were randomly generated within intervals of physically plausible values. Next, for each generated set of parameters, one hydrologic model run (between October 1, 2006 and December 31, 2007) was performed using as inputs the rainfall and the surface temperature measured respectively at the meteorological stations of Livange and Mersch. It is worth noting that the hydrologic model is calibrated offline, by simulating the discharge at Pfaffenthal using rainfall and temperature measurements. The results of these simulations were compared with discharges recorded at the Pfaffenthal gauging station. In this context, for each model, the N_{LF} (Eq. 2) and N_{HF} (Eq. 3) performance measures have been calculated. The parameter sets providing the lowest values for respectively N_{LF} and N_{HF} have been selected as optimal with respect to low and high flows.

Once these two optimal hydrologic parameter sets were defined, additional Monte-Carlo simulations were done to calibrate the combination function of the two hydrologic models. For this, 10,000 sets of combination parameters (δ,γ) were randomly generated, defining 10,000 combination functions. Next, for each of these parameter sets, the combination functions were applied to the simulated discharges and the Nash–Sutcliffe efficiency (Eq. 4) has been calculated with respect to the resulting simulated discharge $Q_s(t)$. The optimization of the combination parameters (δ,γ) has been done by selecting the parameter set providing the highest value of the Nash–Sutcliffe efficiency criterion (E_{NS} (calib.) = 75%). The model was tested in validation mode for the period between January 1, 2008 and February 11, 2009. We found that the Nash–Sutcliffe efficiency remains satisfactory (E_{NS} = 83%). Fig. 5 shows the discharge hydrographs observed and simulated using the optimized parameter sets during the calibration (Fig. 5a) and the validation periods (Fig. 5b).

As proposed in Section 2.2, the coefficients of the linear error model have been calculated. For the whole calibration period (between October 1, 2006 and December 31, 2007) hydrologic model errors have been calculated as the

Please cite this article as: Hostache, R., et al., Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty, Atmos. Res. (2010), doi:10.1016/j.atmosres.2010.09.014

difference between simulation and measurements of discharge: $e(t) = Q_o(t) - Q_s(t)$. A linear regression between errors at time t ($e(t)$) and errors at time $t-1$ ($e(t-1)$) provides the error model coefficients a and b of Eq. (6). By cascading the latter as proposed in Section 2.2, it is possible to estimate the error ($\hat{e}(t)$) at the subsequent time step t. Added to the value of the simulated discharge $Q_s(t)$, $\hat{e}(t)$ allows for correcting part of the past hydrologic model errors.

To evaluate the efficiency of the linear error correction model, depending on the elapsed time since the last discharge observation, the following test has been used: for each time step of the simulation, the linear error model has been applied for predicting the next runoff errors, for increasing elapsed time since the last discharge observation. Next, the Nash–Sutcliffe efficiency has been calculated between the corrected discharges and the observed discharge. The results of this test are shown in Fig. 6. As it is expected, the efficiency of the linear error model decreases for increasing elapsed time since the last discharge observation. In fact, the Nash–Sutcliffe coefficient is 98.7% for a lead time of 1 h, and decreases for increasing horizon of the prediction. For high elapsed time since the last discharge observation, the simulations with and without linear error correction show similar performances (see Fig. 6).

4.2. Forecasting uncertainty assessment

This part aims at assessing the total predictive uncertainty of a hydro-meteorological forecasting chain by making use of the bivariate meta-gaussian density. As proposed in Fig. 3 and Section 2.3, the forecasting chain consists of the two hydrologic models (for low- and high-flow conditions), the combination function and the linear error model. This sequence of models is used in this part in forecasting mode. The simulations of the forecasting chain have been performed for the time window between October 9, 2006 and February 11, 2009. During this time period, one model simulation is performed every 1 h using as input a concatenation of rainfall and temperature recorded during the previous 24 h and of GFS forecasts of rainfall and temperature for the subsequent 120 h. These runs provide a two-dimensional matrix of discharge forecasts $Q_{tt}(t)$, where t is the number of the model run (one operational model run every 1 h) and tt is the prediction lead time (from 1 h to 120 h).

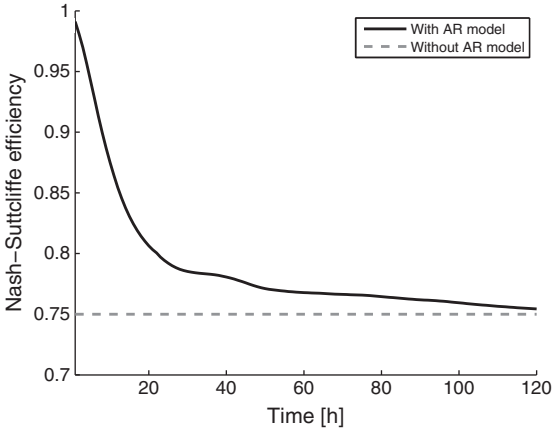

Fig. 6. Progress of the Nash–Sutcliffe efficiency with and without linear-error-model based correction of simulated discharge for increasing elapsed time since the last discharge observation.

To assess the uncertainty of the discharge forecasts for the various lead times and in order to evaluate the performance and the validity of the meta-gaussian model, the data set has been split into two parts. The first one (between October 9, 2006 and December 31, 2007), called hereafter calibration period, is used for building the discharge/error sample that is necessary for the meta-gaussian model. The second one (between January 1, 2008 and February 11, 2009), called hereafter the validation period, is used for evaluating and eventually validating the method. This means that the sample of discharge forecasts and corresponding errors is built during the calibration period. Then, the bivariate meta-gaussian densities are calculated using this sample, with separated errors (positive and negative).

It is worth noting that the residuals $\varepsilon^+(t)$ and $\varepsilon^-(t)$ of the regression of $Ne^+(t)$ and $Ne^-(t)$ on $NQ(t)$ turned out to be non-gaussian and heteroscedatic. To solve this problem, as proposed by Montanari and Grossi (2008), the computation of the forecast uncertainty was carried out for river flows higher than 6.5 m^3/s.

Fig. 7 presents an example of the forecasting chain results obtained during the March 2008 flood event, in validation mode. Each subplot corresponds to the prediction obtained at a given date. The light gray area corresponds to the 90% confidence interval of the discharge forecast (light gray line). The observed discharge is the black dotted line. The black and gray bars correspond to the rainfall observation (Livange station) and forecast (GFS), respectively. The vertical black line indicates the forecast time. The simulated discharge obtained using rainfall and temperature observation is plotted as a baseline (black curve). On Fig. 7, one can note that the 90% confidence intervals encompass the observed discharge at the majority of prediction lead time (even for large prediction lead time, see Fig. 7a). This result illustrates the efficiency of the presented method in forecasting mode. The peak discharge is significantly underestimated in Fig. 7b, mainly because of a corresponding underestimation by GFS of the future precipitation amount. When the prediction lead time decreases (see Fig. 7c and d), the predicted precipitation amounts increase so that the peak discharge is better predicted. A delay can be noted between true and predicted time to peak, but nevertheless the 90% confidence intervals encompass adequately the observed discharge in Fig. 7c and d, therefore remarking the reliability of the uncertainty assessment.

4.3. Method performance and reliability assessment

This section aims at evaluating the validity of the statistical hypothesis introduced above, and discusses the benefits and limitations of the proposed methodology. To this end, three different structures of the forecasting chain were tested, namely:

1. FC1: full forecasting chain, i.e. with linear error correction and with separation between positive and negative errors in the meta-gaussian model;

Please cite this article as: Hostache, R., et al., Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty, Atmos. Res. (2010), doi:10.1016/j.atmosres.2010.09.014

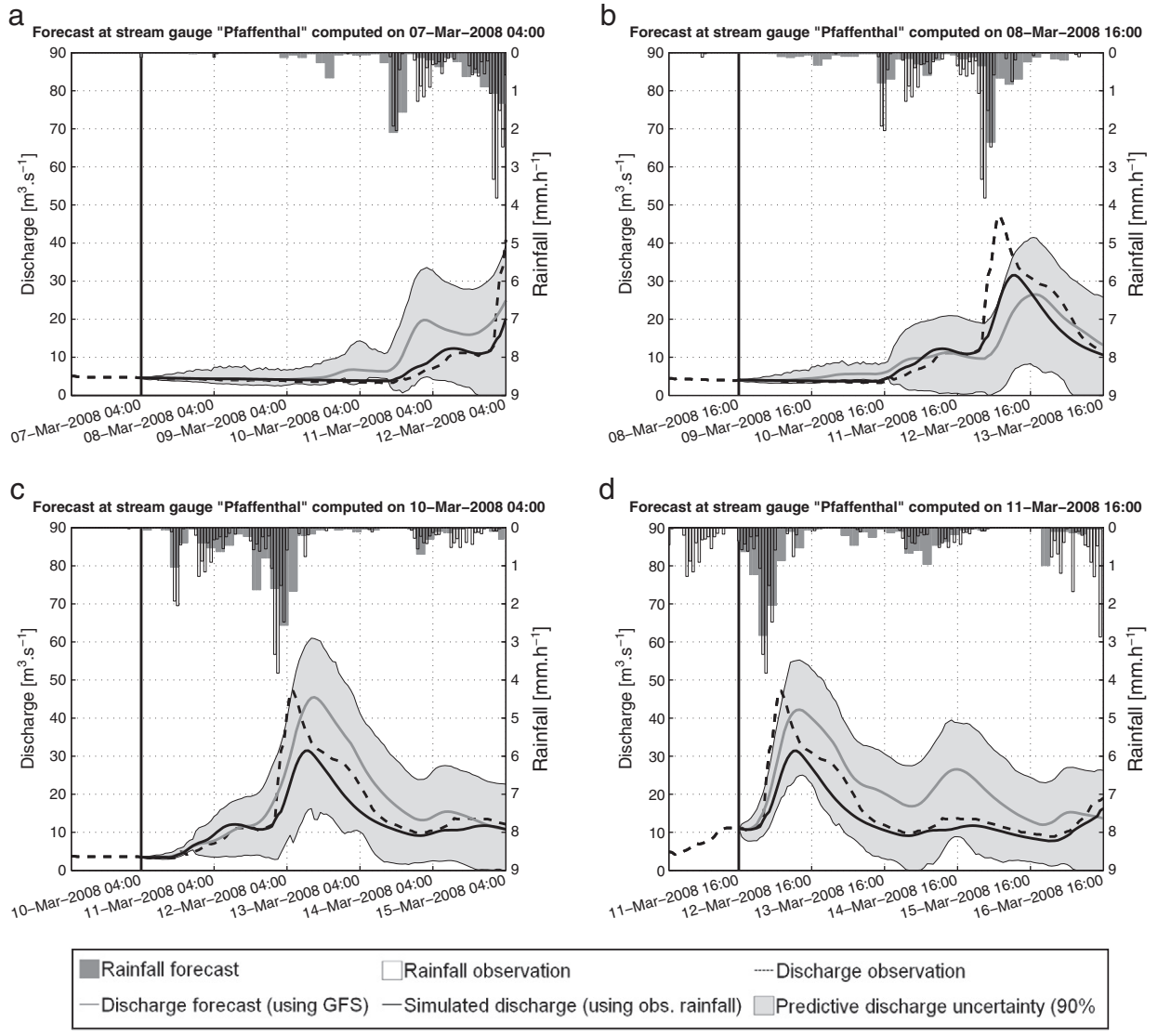

Fig. 7. Sequence of discharge forecasts with associated 90% confidence intervals during the March 2008 flood event.

2. FC2: with linear error correction, but without separation between positive and negative errors in the meta-gaussian model;
3. FC3: without linear error correction, but with separation between positive and negative errors in the meta-gaussian model.

The forecasting chain FC2 aims at assessing the value of separating positive and negative errors for the computation of the meta-gaussian densities. FC3 aims at evaluating the efficiency of the linear error model that is used to update the discharge forecast.

As a first evaluation, we assessed the mean uncertainty of the discharge forecasts. As a matter of fact, the mean relative deviations of the confidence intervals (time averaged ratio between the upper-lower confidence bounds distance and the discharge forecasts $Q_{tt}(t)$), are presented in Fig. 8, for all prediction lead times. As one would expect, the uncertainty of the discharge forecast increases for increasing prediction lead time. More specifically, one can see in Fig. 8a and b that for small lead times the uncertainty is one order of magnitude lower than the discharge forecasts. For larger lead times, for instance from 60 to 120 h, the latter becomes higher than the order of magnitude of the forecast. Without the separation of positive and negative errors in the meta-gaussian model (FC2, Fig. 8b), the mean relative deviation increases very quickly. Without the linear error correction, the FC3 mean relative deviation is significantly larger for the lower lead times (up to 65 h). This shows the benefit of the linear error correction for relatively small prediction lead times.

Moreover, to evaluate the uncertainty that can be ascribed to the GFS rainfall forecasts, a perfect foreknowledge scenario has been considered. Accordingly, the whole method has been applied using the rainfall measurements at Livange

Please cite this article as: Hostache, R., et al., Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty, Atmos. Res. (2010), doi:10.1016/j.atmosres.2010.09.014

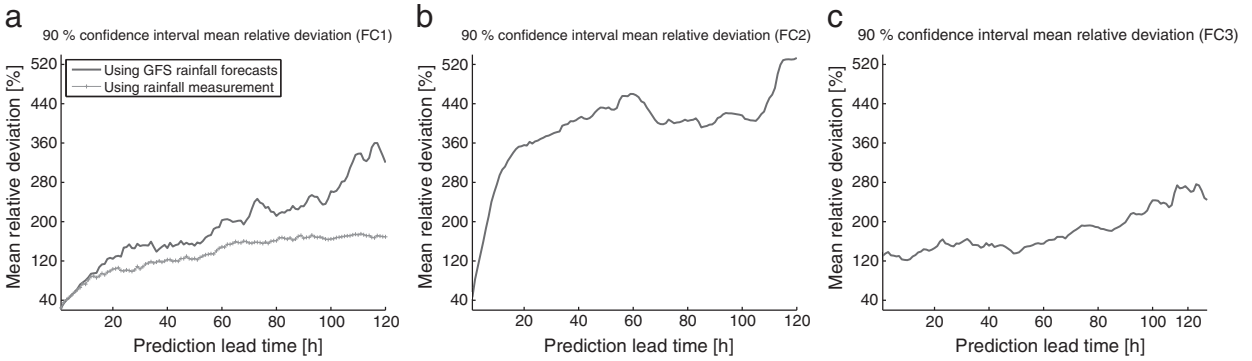

Fig. 8. Mean relative deviation of the 90% confidence interval, for the three considered structures of the forecasting chain.

meteorological station, instead of the GFS forecasts. The mean relative deviations obtained with these calculations are plotted in Fig. 8a (gray line).

Furthermore, we evaluate the 90% confidence interval reliability, for each lead time of prediction, by calculating the percentage of simulation time steps for which the observed discharge is included in the 90% confidence intervals provided by the meta-gaussian models. The related results are presented in Fig. 9 for each structure of the forecasting chain.

For FC1 in Fig. 9a, during the calibration period the meta-gaussian model provides satisfactory performances since the percentage of true runoff values encompassed by the confidence intervals is of the order of 90%. For the validation period, the reliability of the confidence intervals remains acceptable, since the above percentage ranges from 87% to 97%. Without the linear error correction (FC3, Fig. 9c), the reliability of the 90% confidence interval is comparable with the one of FC1. For FC2 (see Fig. 9b) the 90% confidence intervals turns out to be less reliable during the validation period since the above percentage is higher than 95% for every lead time of prediction.

Finally, to statistically evaluate the result of the meta-gaussian model, the same approach has been applied for various confidence interval levels (5% to 99%). This allowed to calculate the percentage of time steps for which the confidence interval at a given significance level intercepts the observed discharge. In Fig. 9 these percentages are plotted against the considered confidence interval, respectively during the calibration and the validation periods. Fig. 10a proves the statistical validity of the method during both the calibration and the validation periods since the confidence intervals intercept the observed discharge with percentages close to their theoretical values. However, Figs. 10b and 10c show coarser results during the validation period.

To verify the hypotheses of gaussianity and homoscedasticity (homogeneity of variance) of the residuals ($\varepsilon^{\pm}(t)$) of the regression of $Ne^{\pm}(t)$ on $NQ(t)$, the Kolmogorov–Smirnov (K-S) (Snedecor and Cochran, 1989) and Bartlett (Chakravarti and Laha, 1967) tests have been performed respectively. To perform the Bartlett test, the samples of the residuals $\varepsilon^{\pm}(t)$ have been split in ten parts, depending on the associated normalised simulated discharge value NQ. This means that ten equally sized intervals of NQ values have been defined and the corresponding values of $\varepsilon^{\pm}(t)$ have been gathered in ten subsamples. Then the Bartlett test is performed over the subsamples in order not to reject the null hypothesis that the variance is unchanging from one subsample to another. For FC1, for positive errors, the gaussianity and homoscedasticity of the residuals are verified since most of the p values for the K-S and Bartlett tests are higher than 5%. Nevertheless, for the negative errors, the gaussianity of residuals is not fully satisfactory when the prediction lead time exceeds 6 h. However, this lack of fit is deemed to be not significant in

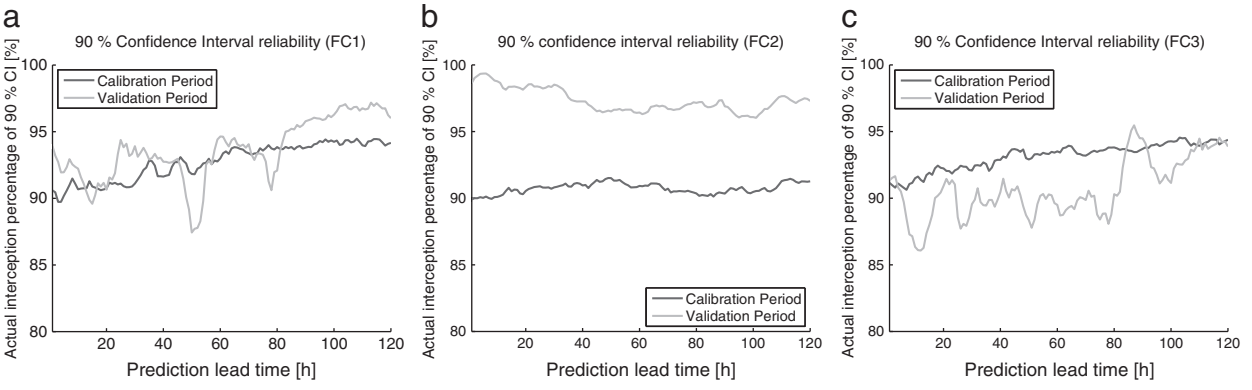

Fig. 9. Reliability of the 90% confidence interval for the three structures of the forecasting chain.

Please cite this article as: Hostache, R., et al., Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty, Atmos. Res. (2010), doi:10.1016/j.atmosres.2010.09.014

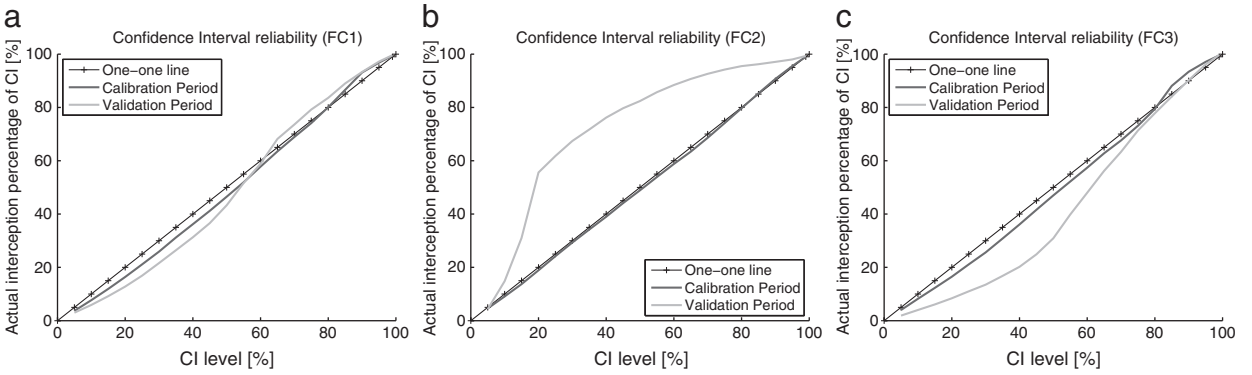

Fig. 10. Statistical reliability of various level confidence interval, for the three variants of the forecasting chain.

terms of reliability of the results. Indeed, as shown in Figs. 9a and 10a, the percentage of observed discharges included in the confidence bands is satisfactory for both the calibration and the validation period. For FC2 and FC3, the gaussianity and the homoscadasticity are rarely satisfied. This explain especially the coarse result obtained in Figs. 10b and 10c. This shows that using both a linear error correction model and a separation of positive and negative error to compute the meta-gaussian models allow to obtain a more reliable method.

5. Conclusion

This paper presents an innovative approach for setting up a flood forecasting system and for assessing the total uncertainty of coupled atmospheric-hydrologic-error correction models.

The proposed flood forecasting system is composed of a rainfall-runoff hydrologic model coupled with a linear error model that reduces output errors of the hydrologic model, knowing the error at the initial computation time. This system uses as input the GFS (Global Forecasting System) forecasts of rainfall and temperature over the subsequent 5 days.

To evaluate the predictive uncertainty of the forecasting system, instead of computing the uncertainties generated by individual model components, the approach that was followed in this study focused on the analysis of the statistical properties of the discharge forecast errors. The bivariate meta-gaussian model that has been used provides a framework that enables the computation of confidence intervals in a normalised space.

During the calibration period of the meta-gaussian model, it was shown that the confidence intervals that are associated with the streamflow simulations encompass the measured discharge with the expected probability, thereby demonstrating the validity of this model. In validation mode, although the skill of the meta-gaussian model of course decreases, especially for the largest prediction lead times, the reliability of the computed confidence intervals remains acceptable.

Moreover, the method presented in this study, making use of GFS forecasts as forcing of the forecasting system, is fully transferable to any other source of weather forecasts, like for instance those that are published by the ECMWF (European Centre for Medium-Range Weather Forecasts). The only need for applying the methodology is a relatively representative archive of observed discharge.

Acknowledgements

The research is in the framework of the MAPRISK project, supported by the National Research Fund (FNR) of Luxembourg through the VIVRE program (FNR/VIVRE/06/36/04).

References

Andreassian, V., Perrin, C., Michel, C., Usart-Sanchez, I., Lavabre, J., 2001. Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. J. Hydrol. 250 (1–4), 206–223.

Beven, K., Binley, A., 1992. The future of distributed models—model calibration and uncertainty prediction. Hydrol. Proc. 6 (3), 279–298.

Brath, A., Montanari, A., Toth, E., 2002. Neural networks and non-parametric methods for improving real-time flood forecasting through conceptual hydrological models. Hydrol. Earth Syst. Sci. 6 (4), 627–640.

Buizza, R., Miller, M., Palmer, T.N., 1999. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. Quart. J. Roy. Meteo. Soc. 125, 2887–2908.

Chakravarti, I., Laha, R., 1967. Handbook of Methods of Applied Statistics, Volume I. John Wiley and Sons.

Fenicia, F., Solomatine, D.P., Savenije, H.H.G., Matgen, P., 2007. Soft combination of local models in a multi-objective framework. Hydrol. Earth Syst. Sci. 11, 1797–1809.

Hamon, W., 1963. Computation of direct runoff amounts from storm rainfall. Int. Assoc. Sci. Hydrol. Publ. 63, 52–62.

Kalnay, E., Kanamitsu, M., Baker, W., 1971. Global numerical weather prediction at the national meteorological center. Bull. Amer. Meteor. Soc. 71 (10), 1410–1428.

Kanamitsu, M., Alpert, J., Campana, K., Caplan, P., Deaven, D., Iredell, M., Katz, B., Pan, H.-L., Sela, J., White, G., 1991. Recent changes implemented into the global forecast system at NMC. Weather Forecast. 6, 425–435.

Kelly, K.S., Krzysztofowicz, R., 1997. A bivariate meta-gaussian density for use in hydrology. Stoch. Hydrol. Hydraul. 11 (1), 17–31.

Koussis, A.D., Lagouvardos, K., Mazi, K., Kotroni, V., Sitzmann, D., Lang, J.-G., Zaiss, H., Buzzi, A., Malguzzi, P., 2003. Flood forecasts for an urban basin with integrated hydro-meteorological model. J. Hydrol. Eng. 8, 1–11.

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

Lindström, G., Johansson, B., Persson, M., Gardelin, M., Bergström, S., 1997. Development and test of the distributed HBV-96 hydrological model. J. Hydrol. 201, 272–288.

Liu, Y., Gupta, H.V., 2007. Uncertainty in hydrologic modeling: toward an integrated data assimilation framework. Water Resour. Res. 43 (W07401), 1–18.

Matgen, P., Henry, J.-B., Hoffmann, L., Pfister, L., 2005. Assimilation of remotely sensed soil saturation levels in conceptual rainfall-runoff models. Predictions in Ungauged Basins: Promises and Progress, Proceedings of symposium S7 held during the Seventh IAHS Scientific Assembly at Foz do Iguazu, Brazil. IAHS, pp. 226–234.

Please cite this article as: Hostache, R., et al., Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty, Atmos. Res. (2010), doi:10.1016/j.atmosres.2010.09.014

Montanari, A., 2007. What do we mean by 'uncertainty'? The need for a consistent wording about uncertainty assessment in hydrology. Hydrol. Proc. 21 (6), 841–845.
Montanari, A., Brath, A., 2004. A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. Water Resour. Res. 40 (W01106), 1–11.
Montanari, A., Grossi, G., 2008. Estimating the uncertainty of hydrological forecasts: a statistical approach. Water Resour. Res. 44 (W00B08), 1–9.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part I: A discussion of principles. J. Hydrol. 10 (3), 282–290.
Pappenberger, F., Beven, K.J., Hunter, N.M., Goudeweleeuw, B.T., Thielen, J., de Roo, A.P.J., 2005. Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). Hydrol. Earth Syst. Sci. 9 (4), 381–393.
Snedecor, G., Cochran, W., 1989. Statistical Methods. Iowa State University Press, Eighth Edition.
Stedinger, J., Vogel, R., Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. In: Maidment, D. (Ed.), Handbook of hydrology. McGraw-Hill, New York, pp. 18.1–18.66.
Toth, E., Brath, A., Montanari, A., 2000. Comparison of short-term rainfall prediction models for real-time flood forecasting. J. Hydrol. 239, 132–147.
Toth, E., Montanari, A., Brath, A., 1999. Real-time flood forecasting via combined use of conceptual and stochastic models. Phys. Chem. Earth, Pt B: Hydrology, Oceans and Atmosphere 24 (7), 793–798.
WMO, 1992. Simulated real-time intercomparison of hydrological models. World Meteorological Organization.

Please cite this article as: Hostache, R., et al., Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty, Atmos. Res. (2010), doi:10.1016/j.atmosres.2010.09.014