Uncertainty assessment via Bayesian revision of ensemble streamflow predictions in the operational river Rhine forecasting system

P. Reggiani *

Section of Hydraulic Structures and Probabilistic Design

Delft University of Technology, P.O. Box 5048, 2600GA Delft, The Netherlands

M. Renner †, A.H.Weerts ‡and P.A.H.J.M. van Gelder §

*also at Deltares-Delft Hydraulics, P.O. Box 177, 2600MH Delft, The Netherlands

†Institute of Hydrology and Meteorology, Dresden University of Technology, Dresden, Germany

‡Deltares-Delft Hydraulics, P.O. Box 177 2600MH Delft, The Netherlands

§Section of Hydr. Struct. and Prob. Design, Delft University of Technology, Delft, The Netherlands

Abstract

Ensemble stream flow forecasts obtained by using hydrological models with ensemble weather products, are becoming more frequent in operational flow forecasting. The uncertainty of the ensemble forecast needs to be assessed, for these products to become useful in forecasting operations. A comprehensive framework for Bayesian revision has been recently developed and applied to operational flood forecasting with deterministic weather forecasts. The Bayesian revision yields a posterior density, conditional on all information available to the forecaster at the onset of a forecast run. This conditional density objectively quantifies the uncertainty.

Here the Bayesian approach is generalized for use with ensemble weather predictions. An end-to-end application of a Bayesian post-processor for ensemble stream flow forecasts in the river Rhine forecasting system is presented. A verification of the post-processor shows good performance when compared in terms of the Ranked Probability Skill Score (RPSS) to non-Bayesian uncertainty assessment, such as ranking threshold exceedance probabilities for members of a stream flow ensemble prediction. In this context it is also addressed how the proposed Bayesian processor can serve in supporting rational decision-making for flood warning under conditions of uncertainty.

Keywords: Ensemble stream flow prediction, uncertainty, decision support, operational flood forecasting, Bayesian revision, river Rhine.

1 Introduction

Operational river flow forecasting is gaining significant importance as non-structural measure for flood protection. Extreme hydro-meteorological events like those registered in the United Kingdom (2005, 2007), Germany (2002), and various Central and Eastern European countries (2002, 2006, 2007) during recent years put further emphasis on the need for enhancing flood forecasting systems and methods.

The fourth IPCC climate assessment report (IPCC, [2007]), concludes that the probability of more frequent meteorological extremes is on the rise throughout the 21st century. Extreme discharges may become more frequent, with a need for countries to invest in structural and non-structural interventions, amongst which real-time flood forecasting systems. The World Meteorological Organization (WMO, [2006]) acknowledges that in several countries flood forecasting remains the only affordable and effective measure that can be realistically implemented to protect life and property in the face of extreme meteorological events. According to UN ISDR [2004], up to 35 percent of flood damage can be mitigated by issuing timely flood warnings.

Real-time flood forecasting systems, as the one referred to in this paper, make use of inter-linked hydrological and hydrodynamic models, that are embedded in a data-management environment. The model chains are run in two principal operational modes: in *i) historical* and *ii) forecast* mode. In the first mode the models are forced by meteorological observations over a limited time period prior to the onset of the forecast. Data assimilation is applied, whereby internal model states are adjusted to optimize model performance. In the second mode the models are forced by quantitative precipitation and temperature forecasts, whereby the internal model

states at the end of the historic run are taken as initial conditions for the forecast run. Depending on whether a deterministic or probabilistic ensemble weather forecasts is used, a single deterministic or an ensemble stream flow or water level forecast is produced. The forecasting lead-time depends on the particular weather forecasting product and commonly extends over a period between 48 and 240 hours ahead.

A deterministic or ensemble streamflow forecast has however limited predictive value for decision-makers (Fortin et al. [2006]), as there remains inherent uncertainty associated with meteorological forcing and the response of the models. The sources of this uncertainty have been widely addressed by Krzysztofowicz [1999]. In the ongoing international Hydrological Ensemble Prediction EXperiment initiative (HEPEX, Schaake et al. [2006]), there is generally agreement between scientists and end-users on the need for systematically processing streamflow forecasts with the aim to objectively quantify the inherent uncertainty. Only in this fashion these forecasting products will find broader acceptance among stakeholders and decision-makers.

In a recent paper Reggiani and Weerts [2008b] applied a Bayesian post-processor (Krzysztofowicz [1999], Krzysztofowicz and Kelly [2000]) to deterministic forecasts in the operational river Rhine forecasting system. The Bayesian processor delivers a posterior probability density function of the expected flow rates or water levels, conditional on all information available at the onset of the forecast. The conditional probability density function has been defined as *predictive uncertainty* (Krzysztofowicz [2001a]). The posterior probability density function results from the revision of an assumed conditional prior density function, derived from historic series, by applying Bayesian updating.

The association of a predictive uncertainty (expressed in terms of a conditional probability density function) with a forecasted flow rate or water level objectively quantifies the uncertainty of the forecast. This uncertainty estimate needs to be explicitly accounted for in the decisional process. One possible solution is to integrate the predictive uncertainty with an appropriate cost function, as suggested by Raiffa and Schlaifer [1961], De Groot [1970] or Todini [2007]. The combination of the two functions provides a concise estimate of expected damage for a forecasted flow rate or water level. An objective assessment of the predictive uncertainty is thus a prerequisite for sound decision support under conditions of uncertainty.

This paper presents an application, where the Bayesian processor is extended from a deterministic to an ensemble streamflow forecast, as suggested by Krzysztofowicz [2001b] and Schaake et al. [2001]. The processor constitutes an Bayesian Ensemble Uncertainty Processor (BEUP). The operational forecasting system of the river Rhine has been chosen as test-bed, whereby an ensemble weather forecasts of 50 members plus 1 control forecast are used to force the hydrological models, yielding as many stream flow forecast at the basin closing section. By employing a conditional prior distribution function, corresponding posteriors are derived for the ensemble member. The ensemble of posterior distributions is then lumped into a single posterior meta-distribution, which is representative for the ensemble forecast. By design the ensemble processor is structured in such a way, that the conditional distributions are modeled off-line, while the processor is executed on-line. This structure makes the approach appealing for operational use. An execution of the processor reduces to the evaluation of parametric functions, for which parameters

have been estimated beforehand. It is also particularly attractive for applications, in which large amounts of information need to be elaborated in real-time, as it is the case with ensemble forecasts. The principles and assumptions underlying the processor will be revisited, the system setup described and the results, including strong points and weaknesses, will be discussed.

The exposure is structured into seven sections: Section 2 describes the principles underlying the Bayesian processor, Section 3 gives a brief overview of the theory, Section 4 describes the forecasting system and the precursory data elaboration steps, Section 5 discusses the numerical experiments and Section 6 discusses the results. Section 7 is dedicated to summary and conclusions.

2 Principles

2.1 Variates

Adopting the notation by Krzysztofowicz [1999] suitable random variates are introduced, that describe the forecasting process. The set

$$\mathbf{H}_n = [H_1, ..., H_n]' \tag{1}$$

called the predictand, are discharge (or water level) observations at forecasting times $1, ..., n$ at one single location, commonly the basin closing section. These quantities lie in the future with respect to those observed at the same location and at an arbitrary number of historical times $1, ..., k$ up to the time $t_0 = 0$ at the onset of

the forecast:

$$\mathbf{H}_{0-k} = [H_0, ..., H_{0-k}]' \tag{2}$$

Finally,

$$\mathbf{S}_{n,j} = [(S_{1,j}, ..., S_{n,j})]' \tag{3}$$

is an ensemble of sets of modeled discharges (or water levels) at the same times of the observations $\mathbf{H}_n$, resulting from an ensemble weather forecast counting $j = 1, ..., m$ members. The realizations of the variates $\mathbf{H}_n$, $\mathbf{H}_{0-k}$ and $\mathbf{S}_{n,j}$ are denoted with the lowercase letters $\mathbf{h}_n$, $\mathbf{h}_{0-k}$ and $\mathbf{s}_{n,j}$, respectively.

2.2 Predictive uncertainty

The predictive uncertainty (Krzysztofowicz [2001a]) can be defined as a measure of the degree of certitude of the occurrence of an event, conditional on all information available in the forecasting process. In operational river flow forecasting, an "event" consists in the exceedance of a critical stream flow at basin control section. The total predictive uncertainty on the forecasted flow (expressed in terms of discharge or water level) for lead time n can be formulated in terms of a conditional probability density function:

$$\phi_j(\mathbf{h}_n|\mathbf{s}_{n,j}, \mathbf{h}_0, \mathbf{h}_{0-1}, ..., \mathbf{h}_{0-k}, \mathcal{M}, ...) \tag{4}$$

This distribution, the predictive uncertainty, represents a family of probability density functions for discharge (or water level) $\mathbf{h}_n$, conditional on the modeled discharge

(or water level) $\mathbf{s}_{n,j}$ resulting from the ensemble weather prediction member j, on the discharge (or water level) observations at times $0-k$ before t_0 and on the particular forecasting model $\mathcal{M}$ used. The convention is adopted to index ϕ_j with the ensemble member subscript, emphasizing that a separate family of probability density functions is derived for each ensemble member used to force the system. The probability distribution can in principle be conditioned on additional information such as the internal state vector of the models and/or the model parameter vector. However, such stochastic dependencies will not be considered explicitly in this application. The absence of any explicit dependence on the meteorological input (e.g. precipitation depth) in Eq. (4) is noted. The reason for this choice is the fact that the modeled flow $\mathbf{s}_{n,j}$ at the forecasting location is a result of uncertain meteorological input, that has been processed through the model chain by forcing the system with the individual members of the ensemble weather forecast. In essence the use of the resulting streamflow forecast $\mathbf{s}_{n,j}$ is equivalent to assuming an auxiliary randomization of $\mathbf{s}_n$ (Krzysztofowicz [2001b]). If deemed necessary this latter assumption can be relaxed and the meteorological forcing made explicit in the formulation, as specified in the Bayesian Forecasting System (BFS) (Krzysztofowicz [1999]). In this case a separate Input Uncertainty Processor would need to be introduced. An Input Uncertainty Processor formulated as Bayesian Processor of Output (BPO) (Krzysztofowicz [2004]) for numerical weather forecasts has been presented by Reggiani and Weerts [2008a].

2.3 Bayesian formulation

A Bayesian Ensemble Uncertainty Processor (BEUP) for water level forecasts consists of an estimator of the conditional probability density function ϕ_j in Eq. (4). The estimator assesses the predictive uncertainty by means of Bayesian revision of prior information. Through Bayesian inference the processor yields a posterior conditional density, whereby model performance against observations, given uncertain meteorological forcing, is brought into the formulation via the likelihood function. The choice of a suitable prior probability distribution function is essential for the quality of the revision. The prior distribution should contain as much information on the behavior of the system as possible.

The basic principles of Bayesian updating in this context is outlined next, whereby the explanation of details is left for subsequent sections. First, a conditional prior distribution function $g(\mathbf{h}_n|\mathbf{h}_0, \mathbf{h}_{0-k})$ is derived from a linear regression model, based on a long climatic series of flow observations at the basin closing section Lobith. Hereby the correlation between flow observations at time t_0 and at the predecessor time t_{0-k} is exploited, with k a time lag expressed in number of days.

The effects of various uncertainty sources enter the revision process via a conditional density $f(\cdot|\cdot)$ of the model output $\mathbf{S}_j$ for the j-th ensemble member and conditional on the observations $\mathbf{H}$. The conditional density $f(\mathbf{s}_{n,j}| \cdot, \mathbf{h}_0, \mathbf{h}_{0-k,}), j = \{1, ..., m\}$ is a likelihood on the predictand $\mathbf{h}_n$, conditional on flow forecast $\mathbf{s}_{n,j}$, and streamflow observations. The residuals vector $\xi_j = (\mathbf{s}_{n,j} - \mathbf{h}_n)$ represents the model error attributable to the presence of uncertain meteorological input (indexed by ensemble-members), model conceptualization and sub-optimal initial conditions and internal

model states. The formulation however will be restricted by excluding model and initial condition errors from the formulation.

Secondly, the likelihood function represents a full stochastic characterization of the forecasting error of the flow modelling chain. The likelihood function imports a probabilistic description of the predictive skill of the models into the Bayesian revision process. The total predictive uncertainty is expressed by the revised posterior density:

$$\phi_j(\mathbf{h}_n|\mathbf{s}_{n,j},\mathbf{h}_0,\mathbf{h}_{0-k}) = \frac{f(\mathbf{s}_{n,j}|\mathbf{h}_n,\mathbf{h}_0,\mathbf{h}_{0-k})\; g(\mathbf{h}_n|\mathbf{h}_0,\mathbf{h}_{0-k})}{k_j(\mathbf{s}_{n,j}|\mathbf{h}_0,\mathbf{h}_{0-k})} \tag{5}$$

where

$$k_j(\mathbf{s}_{n,j}|\mathbf{h}_0,\mathbf{h}_{0-k}) = \int_{-\infty}^{+\infty} f(\mathbf{s}_{n,j}|\mathbf{h}_0,\mathbf{h}_{0-k})\; g(\mathbf{h}_n|\mathbf{h}_0,\mathbf{h}_{0-k})\, d\mathbf{h}_n \tag{6}$$

is the expected density on model output, conditional on all information available at the onset of a forecast.

3 The uncertainty processor

The processor in Eq. (5) is specified by deriving parametric expressions for the family of the prior density $g(\mathbf{h}_n|\mathbf{h}_0,\mathbf{h}_{0-k})$ and the family of likelihood functions $f(\mathbf{s}_{n,j}|\mathbf{h}_n,\mathbf{h}_0,\mathbf{h}_{0-k})$. The evaluation of the prior density and the likelihood is carried out following the steps laid out in Reggiani and Weerts [2008b]. The results of the regressions and algebraic manipulations are summarized in the subsequent paragraphs.

3.1 Normal Quantile Transform

As will be shown below, the process variables are realizations of stochastic processes and are best described by probability distributions such as the Weibull or Gamma models. Non-linear relations between sequences of random variable realizations, as for instance used in autoregressive models for river stages forecasts would require the use of genealized linear models with multiple lag terms for adequate fits. A solution to this problem is to transform the process variables into Gaussian variables through the application of the Normal Quantile Transform (NQT), which was first proposed by Van der Waerden [1952, 1953a, 1953b] and has been successively applied in uncertainty analysis of streamflow forecasts. Examples can be found in Kelly and Krzysztofowicz [1997], Krzysztofowicz and Kelly [2000], Montanari and Brath [2004] and Zhiyu et al. [2005].

In practice Gaussian variates are obtained by matching an empirical probability distribution $\Gamma(\cdot)$ with a Gaussian distribution $Q(\cdot)$ and then performing an inversion $Q^{-1}(\cdot)$. Here the convention is used that variables in the original space are indicated with Latin letters, while their transforms in the normal space are indicated with Greek characters. The NQT assures in principle that the marginal distributions of the transformed variables are standard normal. It is however no guarantee that the joint distribution of the variates will be multi-normal, unless a linear dependence between variables in the transformed space can be asserted.

In the Gaussian space multiple variates relate linearly which allows to fit autoregressive models. As will be shown below, a series of assumptions must be verified. A successive back-transformation of the linearly dependent variables into the original

space yields parametric expressions for families of conditional probability densities in Eq. (5). The necessary steps in performing NQT mapping and back transformation into the original space, will be followed up next.

3.2 Prior density

The conditional prior density $g(\mathbf{h}_n|\mathbf{h}_0, \mathbf{h}_{0-k})$ is modeled by pursuing a series of steps: *i)* first the NQT is applied to transform the respective variables h into Gaussian variates η, *ii)* secondly an autoregressive model is applied between the series of Gaussian variables, and *iii)* thirdly a back transformation into the original space is performed. The autoregressive model relates the normalized water level η_n at forecasting day n to the respective normalized levels observed at times $t = 0$ and $t = 0 - k$:

$$\eta_n = \alpha_n \, \eta_0 + \beta_n \, \eta_{0-k} + \gamma_n + \Xi_n \tag{7}$$

The parameters α_n, β_n and γ_n are regression constants, while Ξ_n is the residual, stochastically independent from (η_0, η_{0-k}) and normally distributed with zero mean and variance τ_n^2. For an autoregressive model to be valid, a series of assumptions must be verified: a) independency of the residuals Ξ_n of (η_0, η_{0-k}) b) homoscedasticity of Ξ_n and c) normality of the residuals with zero mean and a fixed variance. Condition *a)* can be verified by analyzing the autocorrelation of the residuals against the lag. Condition b) needs to be verified by analyzing the graph of the residuals against the dependent variable and indicates suitability of the linear model. Condition c) is verified by testing the probability distribution of the residuals via e.g.

quantile-quantile plots.

As lead-time increases however, the regression model starts to break down, as water levels η_n become uncorrelated from levels η_0 and η_{0-k}. It is therefore chosen to abandon the autoregressive model for $n > 3$ and consider an unconditioned prior that is only based on climatic series and therefore unrestricted. Table 1 reports the values for the regression constants α_n, β_n, γ_n,and τ_n for lead times $n = 1, 2, 3, 4, 5$ days, discharge observations h_{0-k} for $k = 1$ day, and selected months. For $n > 3$ days $\alpha_n = \beta_n = \gamma_n = 0$ and $\tau_n \neq 0$. The variance τ_n increases with lead time, as one would expect and remains constant for $n \geq 4$. The fact that α_n and β_n are > 0 indicates that the predicted variable η_n is indeed stochastically dependent from the predictors η_0 and η_{0-k}.

The parametric expression of the conditional prior density in the normal space (Krzysztofowicz and Kelly, [2000]) equates the following expression:

$$g_{Qn}(\eta_n|\eta_0, \eta_{0-k}) = \frac{1}{\tau_n}\, q\left(\frac{\eta_n - \alpha_n\, \eta_0 - \beta_n\, \eta_{0-k} - \gamma_n}{\tau_n}\right) \tag{8}$$

where q is the normal density operator and $\tau_n^2 = VAR\,(\eta_n|\eta_0, \eta_{0-k})$.

3.3 Likelihood

The likelihood is modeled in full analogy to the prior density, assuming again a linear relationship between the transformed normal variables. The transform $\varsigma_{n,j}$ of the discharge forecasted at Lobith for the j-th weather forecast ensemble member is related via an autoregressive model to the normal transform η_n of the level observed

at time t_n, and the transforms η_0 and η_{0-k} of observations at Lobith at times t_0 and t_{0-k} respectively:

$$\varsigma_{n,j} = a_n\ \eta_n + b_n\ \eta_0 + c_n\ \eta_{0-k} + d_n + \Theta_n \tag{9}$$

where the parameters a_n, b_n, c_n and d_n are regression constants, while the residual Θ_n is stochastically independent from $(\eta_n, \eta_0, \eta_{0-k})$ and normally distributed with zero mean and variance σ_n^2. Table 2 summarizes the regression coefficients and standard deviations for lead times of n between 1 and 5 days and discharge observations at t_{0-1}. Some verification results for the autoregressive model will be reported in Section 6. The parametric expression of the likelihood in the normal space becomes:

$$f_{Qn}(\varsigma_{n,j}|\eta_n, \eta_0, \eta_{0-k}) = \frac{1}{\sigma_n}\ q\left(\frac{\varsigma_{n,j} - a_n\ \eta_n - b_n\ \eta_0 - c_n\ \eta_{0-k} - d_n}{\sigma_n}\right) \tag{10}$$

where q is the normal density operator and $\sigma_n^2 = VAR\ (\varsigma_{n,j}|\eta_n, \eta_0, \eta_{0-k})$. For lead times $n > 3$ days it is assumed that the likelihood is no longer conditioned on η_0 and η_{0-k}, but only on η_n , thus $b_n = c_n = 0$. The reason for this choice will become clear in Section 4.2, where the autoregressive models will be discussed.

3.4 Posterior density

The posterior density in the normal space is obtained by combining the prior (8) and the likelihood (10), both normal linear, with the transformed conditional expected density $k_{Qn}(\varsigma_n|\eta_0, \eta_n, \eta_{0-k})$, as stated by Bayes theorem in Eq. (5). The underlying

manipulations are omitted and can be found in Krzysztofowicz and Kelly [2000]). The parameterized density in the normal space results in the expression:

$$\phi_{Qn,j}(\eta_n|\varsigma_{n,j},\eta_0,\eta_{0-k}) = \frac{1}{T_n} q\left(\frac{\eta_n - A_n\,\varsigma_{n,j} - B_n\,\eta_0 - C_n\,\eta_{0-k} - D_n}{T_n}\right) \qquad (11)$$

The expressions A_n, B_n, C_n, D_n and T_n evaluated for lead times $n = 1, 2, 3, 4$ and 5 days and discharge observations at t_{0-1} are summarized in Table 3.

3.5 Transformation into the original space

In the original variable space the prior density of levels at forecasting time t_n, conditional on those observed at forecast start time t_0 and t_{0-1} at Lobith, takes on the form:

$$g_n(h_n|h_0,h_{0-1}) = \frac{\gamma(h_n)}{\tau_n\; q(Q^{-1}\Gamma(h_n))}\; q\left(\frac{Q^{-1}\Gamma(h_n) - \alpha_n\; Q^{-1}\Gamma(h_0) - \beta_n\; Q^{-1}\Gamma(h_{0-k}) - \gamma_n}{\tau_n}\right) \qquad (12)$$

where $\gamma(h_n)$ is the marginal density of h_n at forecast time t_n, Lobith. The posterior density in the original space of discharges to be expected at time t_n at Lobith, conditional on discharge $s_{n,j}$ forecasted for day n and discharges h_0 and h_{0-k} observed at days t_0 and t_{0-k} at Lobith, is given by the following parametric expression:

$$\phi_{n,j}(h_n|s_{n,j},h_0,h_{0-1}) = \frac{\gamma(h_n)}{T_n\ q(Q^{-1}\Gamma(h_n)} \cdot \tag{13}$$
$$\cdot\ q\left(\frac{Q^{-1}\Gamma(h_n) - A_n\ Q^{-1}\Delta(s_{n,j}) - B_n\ Q^{-1}\Gamma(h_0) - C_n\ Q^{-1}\Gamma(h_{0-k}) - D_n}{T_n}\right)$$

The posterior cumulative distribution of h_n finally reads as follows:

$$\Phi_{n,j}(h_n|s_{n,j},h_0,h_{0-1}) = \tag{14}$$
$$Q\left(\frac{Q^{-1}\Gamma(h_n) - A_n\ Q^{-1}\Delta(s_{n,j}) - B_n\ Q^{-1}\Gamma(h_0) - C_n\ Q^{-1}\Gamma(h_{0-k}) - D_n}{T_n}\right)$$

with Q the Gaussian distribution operator. Next we introduce an average density or *"meta-density"*, which is obtained by averaging (13) over the ensemble $s_{n,j}$:

$$\bar{\phi}_n(h_n|\bar{s}_n,h_0,h_{0-1}) = \frac{1}{51}\sum_{j=1}^{51}\phi_{n,j}(h_n|s_{n,j},h_0,h_{0-1}) \tag{15}$$

with $\bar{s}_n = E[s_{n,j}]$. Eq. (15) is conditional on all information available at the onset of a forecast and remains implicitly a function of the stream flow forecast ensemble via the expected value $\bar{s}_n$. The corresponding conditional cumulative probability distribution is stated as:

$$\bar{\Phi}_n(h_n|\bar{s}_n,h_0,h_{0-1},EF) \tag{16}$$

3.6 Direct estimation of the predictive uncertianty

We anticipate that the predictive uncertainty (4) could in principle be estimated directly, without any Bayesian revision of prior information. In analogy to (9) a multi-linear regression model between NQT transformed random variables is applied

in the normal space:

$$\eta_n = \hat{a}_n \, \varsigma_{n,j} + \hat{b}_n \, \eta_0 + \hat{c}_n \, \eta_{0-k} + \hat{d}_n + \hat{\Theta}_n \tag{17}$$

The parameters $\hat{a}_n$, $\hat{b}_n$, $\hat{c}_n$ and $\hat{d}_n$ are regression constants, the residual $\hat{\Theta}_n$ is stochastically independent and normally distributed with zero mean and variance $\hat{\sigma}_n^2$. A family of conditional probability distributions in the normal space is derived in analogy to (11). A transformation back into the original space yields an expression equivalent to the posterior (13), in which the coefficients A_n, B_n, C_n, D_n and T_n are replaced by $\hat{a}_n$, $\hat{b}_n$, $\hat{c}_n$, $\hat{d}_n$and $\hat{\sigma}_n$. Unlike (13) which is based on Bayesian revision, the conditional distribution obtained from (17) does not implicitly separate information into prior and a stochastic specification of the model error, as performed in applying Bayes theorem. Instead it combines current observations with respective model forecasts, without referring to the observed historical behavior of the system. Under optimal conditions, i.e with sufficiently long data series for appropriate modeling, the directly derived predictive uncertainty is supposed to have an inferior information content than the Bayesian posterior. However, meaningful performance comparisons between (13) and (17) in terms of skill can only be performed in presence of sufficiently long series of forecasts and will not be addressed here.

4 Application

4.1 The test bed system

The ensemble hydrological uncertainty processor is applied to the river Rhine operational flood forecasting system with a basin surface area (upstream of Lobith) of

160,000 km^2. The hydrologic response is simulated with the HBV model (Bergström, [1995]). The hydrological model output is updated for a model run over a historical time window ahead of a forecast run with streamflow observation by means of an error correction method described by Broersen and Weerts [2005]. The hydrological model calculates the runoff from the tributaries towards the main river Rhine channel. The flood wave propagation along the main channel is performed with a simple routing scheme present in the HBV model. **Figure 1** depicts the Rhine basin with its principal tributaries, selected observing stations and the travel time isochrones. The primary focus is on the basin control section Lobith at the Dutch-German border. For Lobith an extended historical series of daily water level and discharge observations is available for the period 01/01/1901 until 01/10/2007.

INSERT FIGURE 1 HERE

The hydrological model is run once daily with hourly timestep in historical and in forecast mode. In historical mode the system is forced by real-time precipitation and temperature observed over a historical period of 196 hours prior to the start time of the fluvial forecast t_0.

In forecast mode the system is forced by probabilistic weather forecasts from the $50 + 1$ member ECMWF Ensemble Prediction System (EPS, Molteni et al. [1996]). The spatial resolution of the model grid is of 80x120 km, the lead-time of the ensemble forecast 240 hours. The weather model is run once daily, with base time T_0 at 12:00 UTC. The period with a continuous availability of ensembles covers three years, starting on 01/06/2004 until 01/06/2007.

4.2 Data elaboration

The available discharge data (observed and forecasted) are grouped into months to account for non-stationarity of modeling errors in the river flow process. By separation into months time series are obtained which are expect to be approximately stationary for the selected period. In summer the regime is characterized predominantly by low flows. High flow rates occur mainly during winter, especially in December and January. For the sake of brevity the analysis is restricted to four selected months of the year. Empirical cumulative distributions $P(h > h^*)$ for the historical and $P(s > s^*)$ for the forecasted discharges are derived from respective time series.

Figure 2 shows cumulative probability distribution functions on flows at Lobith for the 100+ year observing period 01/01/1901-01/06/2007 and the 01/06/2004–01/06/2007 forecasting period for January and July. We note that in the latter distributions the data points are significantly less (one forecast per day) because of the limited period for which forecasts are available. As a result a whole range of events are not represented, leading to differences between the distributions for the historical and the forecasting period. The empirical distributions $P(h > h^*)$ and $P(s > s^*)$ are indicated respectively with circles and crosses. The continuous and dashed curves indicate distributions modelled as 2-parameter Gamma fits $\Gamma(h)$. Here the Gamma distribution has been used as zero-hypothesis in applying the Kolmogorov-Smirnov statistical test (Benjamin and Cornell, [1970]) with a significance level of 10%. As alternative hypotheses it has been assumed the empirical distributions to be either Weibull or Log-Weibull, which had to be rejected in favor

of the Gamma model. For some months also the Gamma model had in principle to be rejected, asking for more complex models with piecewise fits to be applied. Optimal modeling of the empirical data remains however beyond the scope of this paper, which focusses on the functional principles of the processor. Poor distribution modeling is likely to have an adverse effect on the performance of the processor, thus special attention needs to be devoted to this aspect. Especially the tails of the distributions need to be captured adequately, as they represent the extreme events, which the uncertainty processor should be able to predict.

INSERT FIGURE 2 HERE

The variates h_n, h_{0-k}, and $s_{n,j}$, whose NQT transforms are denoted with the Greek letters η_n, η_{0-k}, $\varsigma_{n,j}$ and ς_{0-k}, are shown next. The left column in **Figure 3** shows the normalized discharge observations η_n against the normalized discharge η_0 and η_{0-k} for the 1901-2007 historic series, selected values of n and k at Lobith, January. The continuous line is the intersection of the multiple regression Eq. (7) with the $\eta_n - \eta_0$ and $\eta_n - \eta_{0-k}$ plane. The regression represents a plane in the three-dimensional Gaussian space $\{\eta_n, \eta_0, \eta_{0-k}\}$. The dashed-dotted lines identify the 90% confidence interval intersects. The coefficient of determination R^2 is reported in the legend. It is evident that the variables are linearly related. The error structure for $n \leq 3$ is reasonably homoscedastic, indicating stationarity of the stochastic process. As n increases, the error structure become increasingly heteroscedastic (see analysis

below).

INSERT FIGURE 3 HERE

The right-hand side column in **Figure 3** depicts the data for the prior in the original space, h_n vs. h_0 and h_n vs. h_{0-k}. The transform of the linear regression (continuous line) into the original space is non linear, as are the respective transforms of the 90% confidence intervals. The NQT and its inverse do not preserve the Pearson moment product correlation because of the non-linear transformation, but preserve the Spearman rank correlations, which are reported on the respective figures. For $n < 3$ the correlation between the expected flow h_n (η_0) at Lobith and observations at t_0 and t_{0-k} is linear. For $n > 3$ the correlations diminish progressively, suggesting that unconditioned probability distributions based only on climatic series for the respective month to be used.

INSERT FIGURE 4 HERE

Figure 4 depicts the data for the autoregressive model in the likelihood Eq. (9). The left column consists of plots of the normalized discharge $\varsigma_{n,j}$ forecasted at Lobith via the j-th ensemble member, against the normalized observed discharge η_n, $\varsigma_{n,j}$ vs. η_0 and $\varsigma_{n,j}$ vs. η_{0-k} for selected n and k. The dashed-dotted lines indicate the 90% confidence envelope intersects. The regression parameters are listed in Table 2. The right column shows the forecasted versus the observed discharge in the original space. The continuous curve is the transform of the linear regression into

the original space, and projected onto the $s_{n,j} - h_n$ plane, while the dashed-dotted lines indicated the corresponding transformed 90% confidence envelope. The relation between the data points is reasonably linear.

Residual analysis were performed on the likelihood to test the validity of the autoregressive models. The homoscedasticity of the error structure with residuals Θ_n has been verified as well as the normality of the residuals with zero mean. Tests on the independence of the error were also performed by deriving the autocorrelations of the residuals. It is however noted that as $n > 3$ the error structure becomes increasingly heteroscedastic, implying that the assumption of stationarity is no longer satisfied. Improvements in terms of homoscedasticity could be achieved by censoring the original data sample and eliminating outliers. These outliers are generally attributable to extreme flow events, which distort the linear behavior and have non-stationary characteristics.

The coefficients A_n, B_n, C_n and D_n for the evaluation of the posterior density in the normal space (11) and the original space (13) are evaluated for the months in the year. The algebraic expressions and numerical values for the coefficients and selected months can be found in Table 3. The prior and the posterior densities g_n and $\phi_{n,j}$, and the posterior distribution $\Phi_{n,j}$ are evaluated by inserting the modeled Gaussian probability density function $\gamma(h_n)$ and the probability distribution functions $\Gamma(h_n)$, $\Gamma(h_0)$, $\Gamma(h_{0-k})$ and $\Delta(s_{n,j})$ into expressions (12), (13) and (14), respectively.

5 Experiments

5.1 Application of the processor

In the present application operational forecasting is emulated by starting a daily forecast in hind-cast mode, with a meteorological forecast base-time T_0 at 12:00 UTC. The ensemble streamflow forecast forecast is then executed at 6:00 UTC the following day. The gridded precipitation and temperature weather model output is transformed into sub-basin averaged time series. The HBV model is run in hindcast mode over the 01/06/2004 - 01/06/2007 period. Because the configuration of the EPS system changes frequently due to fast model evolution, only short periods of homogeneous and continuous data are available. For this reason the BEUP is calibrated over the three year period, for which uninterrupted EPS forecasts exist. A verification is subsequently performed for the period 01/06/2007 and 01/10/2007 and will be discussed hereunder.

INSERT FIGURE 5 HERE

Figure 5 shows the predictive uncertainty $\phi_j(\mathbf{h}_n|\mathbf{s}_{n,j}, \mathbf{h}_0, \mathbf{h}_{0-1})$ for EPS members $j = \{1...51\}$ for $n = 5$. The fluvial forecast base time t_0 is 06:00 UTC on the 09/08/2007. The solid green line represents the prior density $g_n(\mathbf{h}_n|\mathbf{h}_0, \mathbf{h}_{0-1})$. The inset boxes indicate discharges. The vertical blue lines in the lower box are the forecasted discharges for the 51 ensemble members, the red vertical line in the upper box the discharge that has actually been observed on the particular day t_{0+n}. The continuous black line is the discharge observed at Lobith at t_0, while the dashed

black line is the discharge observation at t_{0-1}, 24 hours before the forecasting base time. The "meta-density" in Eq. (15) is indicated with a dashed-dotted line. The effect of the Bayesian revision of the prior density is evident. The processor, which has been trained on the basis of additional information from a forecasting model, relies on the historical experience and delivers revised posterior densities. With increasing n the posterior densities are translated, whereby the peak of the distribution (corresponding to maximum probability of occurrence) moves closer towards the observed discharge.

INSERT FIGURE 6 HERE

Figure 6 shows a sample streamflow forecast at Lobith on the 27/02/2007. The figure shows the $50 + 1$ forecasts corresponding to the ensemble members. The continuous and dashed lines are the observed and the forecasted discharges respectively. The shaded areas indicate the uncertainty bands. The light grey area is the prior uncertainty. For $n \leq 3$ the prior uncertainty increases, as evident from the variance τ_n reported in Table 1. For $n > 3$ an unconditioned prior has been chosen, yielding a constant uncertainty. The dark grey area represents the predictive uncertainty as "meta-posteriori" resulting from Bayesian updating. The correcting effect of the Bayesian revision on the posterior probabilities is evident.

INSERT FIGURES 7a and 7b HERE

5.2 Verification

To corroborate the strength of the Bayesian processor a verification over the test period 01/06/2007 - 01/10/2007 is performed. The verification focuses on the "meta-density" in Eq. (15) acting as proxy for the ensemble. The "meta-density" can be used by a forecaster to assess the uncertainty of the ensemble stream flow forecast for a particular n.

Figures 7a-b show a verification of the processor for the June and September 2007 period, thus extending beyond the period over which the processor has been trained. **Figure 7a** shows the 10%-90% probability interval for lead time $n = 2$, while **Figure 7b** presents a verification for a $n = 5$. Because of the limited period for which flow forecasts $s_{n,j}$ (3 years) are available, the cumulative distributions for the historical and the forecasted series differ significantly (ref. **Figure 2**). This would lead to a poor Bayesian revision in Eq. (13). As an interim solution (until longer forecasted series become available) the modeled distribution $\Delta(s_{n,j})$ in Eq. (13) has been substituted with $\Gamma(h_n)$ for the forecasted series, which significantly improved the performance of the processor for these three months.

The continuous-crossed line indicates the observed discharge, the light-grey shadowed area the probability interval for the prior and the dark-grey area the probability interval for the posterior. In **Figure 7a** the prior for $n = 2$ is conditioned on the observations h_0 and h_{0-1} and follows the observed discharge. In **Figure 7b** the prior is an unconditioned climatic distribution for the respective months and thus changes stepwise between months. The envelope determined by the middle-dark shaded area indicates the uncertainty band for the unprocessed ensemble forecast,

which is derived by simple ranking of threshold exceedance probabilities for members of the stream flow ensemble. It represents the 80% probability interval for the total uncertainty of the model chain, expressed by the total probability $P(s_{n,j} > s^*)$.

Finally a quantitative analysis has been performed by evaluating the ranked probability skill score - RPSS (Wilks [1995]) - for the revised meta-posterior using the conditioned and the unconditioned model. Like the Brier skill score (Brier, 1950), the RPSS is often used as skill criterion of probabilistic forecasts of weather, climate or like in the present case, stream flows. As skill scores, they compare the extent to which a forecast strategy outperforms a reference forecast strategy. The RPSS was compared to that of an unprocessed EPS streamflow forecast. The results are summarized in **Figure 8**. The meta-posterior conditioned on h_0 and on h_{0-1} performs best for $n < 3$, while for $n > 3$ it is outperformed by the unconditioned meta-posterior. A combination of conditioned and unconditioned meta-posterior outperforms the skill for the probability interval derived from the unprocessed EPS forecast.

INSERT FIGURE 8 HERE

6 Discussion

From the previous analysis a series of conclusions can be drawn:

*i)*A first application of the Bayesian processor to an ensemble stream flow forecast is presented. The verification shows that the processor yields a skillful posterior conditional density, with observations located within the 80% probability interval identified by the "meta-density". The "meta-density" is an average density, implicitly dependent on the ensemble streamflow forecast $s_{n,j}$ and conditional on upstream observations at the onset of the forecast and the average forecasted flow $\bar{s}_n$.

ii) The probability interval derived from the unprocessed EPS forecast provides merely an indication of the the total uncertainty of the model chain. The respective uncertainty $P(s_{n,j} > s^*)$ is expressed as total probability density on the forecasted discharges $s_{n,j}$. The Bayesian processor on the other hand delivers a posterior probability density on the expected stream flow h_n, conditional on all information (observations and model forecast) available at the onset of the forecast.

iii) The probability interval tracked by the "meta-density" (dark-grey area) is evidently wider than the one obtained from the unprocessed ensemble stream flow forecasts, indicating that there is a larger uncertainty to be expected from the actual flow h_n conditional on observations, rather than from the unconditioned EPS streamflow forecast $s_{n,j}$ which represents the total uncertainty of the model chain.

iv) The impact of the prior density function on the Bayesian revision is significant in obtaining a valuable posterior density. Here a generic development is proposed, where the prior density is conditioned on two observations at different points in time. The beneficial effect of conditioning the prior on observations diminishes with

increasing lead time n. With $n > 3$ the impact of the conditioning observations on the prior probability is neutralized. In these circumstances the use of an unconditioned climatic prior (ref. Table 1), as shown in **Figure 7b** for $n = 5$ results to be more effective than using the conditional density in Eq. (12).

v) What is valid for the prior density applies equally to the likelihood. The dependency on the earlier observations becomes less influential with increasing n. For $n > 3$ the most significant conditioning variable remains the forecasted discharge $s_{n,j}$ (ref. Table 2), which enters the posterior density via the regression parameter a_n in A_n, B_n, C_n, D_n and T_n (see Eq. (11)).

vi) For $n < 3$ much can be gained by conditioning the likelihood on earlier observations. However, the time lags, at which the observations are taken must be carefully chosen, to ensure that the autoregressive model captures adequately the rising and falling limb of the hydrograph. For example the departures of the probability intervals of the prior and the posterior in **Figures 7** from the observations (e.g. for the peak on the 21/08/2007) are attributable to the low skill of the EPS forecasts, and to a minor degree to poor performance of the autoregressive models (7) and (9).

vii) It is stressed that the present application necessitates additional work. Significant performance gains can be achieved on various fronts. Firstly, the prior needs to be improved by *a)* approximating the empirical distributions Γ with piece-wise functions (low range, mid-range and high flow range), by *b)* selectively removing extreme outliers with the aim to improving the regressions, and *c)* by adequately choosing the time lags. It is important to note that outliers actually represent extreme events, which should be predicted by the forecasting system. Any removal of

outliers must thus be weighted carefully.

Secondly, criteria *a)*, *b)* and *c)* apply equally to the likelihood function. The importance of using sufficiently long forecasted flow time series for performing the regressions between observations and forecasts is also highlighted. The three years daily data record used for the present application can be considered a rather short training period.

7 Summary and Conclusions

An application of the Hydrological uncertainty processor (Krzysztofowicz [1999], Krzysztofowicz [2000], Krzysztofowicz and Kelly [2000]) for an ensemble streamflow forecast by extending the work by Reggiani and Weerts [2008a,b] has been presented. The processor translates the prior probability derived from a long climatic series into an ensemble of posterior densities on water levels or discharges, conditional on model forecasts and multiple observations. The ensemble of revised probability density and distribution functions are subsequently averaged into a posterior "meta-density". The "meta-density" can be used by forecasters as a proxy to identify the probability of occurrence of a predicted water level or discharge. As such the BEUP provides an uncertainty assessment tool, which effectively decouples the responsibility of the forecaster from that of the decision-maker. In a decision support system for evacuation, the expected damage D relative to a forecasted flow rate can be estimated by combining the "meta-density" with a suitable cost function $\chi(h_n)$. The cost function expresses the economic value of issuing/not-issuing

an alert in terms of the forecasted water level at Lobith (Todini [2007]):

$$D_{\bar{s}_n,h_0,h_{0-1}} = \int_0^\infty \chi(h)\ \bar{\phi}(h|\bar{s}_n, h_0, h_{0-1})\ dh \tag{18}$$

The determination of the cost function needs to be explored on a case-by-case basis. The total economical value of a forecast can then expressed in terms of a cost-loss ratio (e.g. Mylne [2002] Palmer, [2002]), by evaluating the cost of mitigating action against the potential losses incurred in absence of taking actions. These aspects need to be investigated in more depth in sequel research. The processor has been validated for selected events extracted from the three-year period for which continuous forecasts are available. To expand the understanding of the processor and the shortcomings of underlying assumptions, additional verification is needed. However, due to the rapid evolution in operational weather models and respective forecasting products, significant challenges have to be met in obtaining long and homogeneous forecasting time series, which are required for the investigations.

8 Acknowledgment

The authors would like to thank the forecasting office of the Dutch Ministry of Transport and Waterways and the German Federal Office of Hydrology for granting permission for the use of the operational river Rhine Forecasting system in this research.

9 Abbreviations

BFS: Bayesian Forecasting System

BEUP: Bayesian Ensemble Uncertainty Processor

BPO: Bayesian Processor of Output

EF: Ensemble Forecast

EPS: Ensemble Prediction System

ECMWF: European Centre for Medium-Range Weather Forecasting

FEWS: Flood Early Warning System

HEPEX: Hydrological Ensemble Prediction Experiment

NQT: Normal Quantile Transform

UTC: Coordinated Universal Time

10 References

Bergström, S., (1995). The HBV Model. In Computer Models of Watershed Hydrology. Singh, V.P., (Ed.), Water Resources Publications,: Highlands Park, CO, pp. 443-476.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability, Monthly Weather Review, 75, 1-3.

Broersen, P.M.T. and A.H. Weerts, (2005). Automatic error correction of rainfall-runoff models in flood forecasting systems. Instrumentation and Measurement Tech-

nology Conference (IMTC), 2005, Proc. of the IEEE, 16-19 May 2005, Vol. 2, 963–968, ISBN: 0-7803-8879-8.

Fortin, V., A.C. Favre and M. Saïd (2006), Probabilisitc forecasting from ensemble prediction systems: Improving upon best-member method by using a different weight and dressing kernel for each member, Q.J.R. Meteorol. Soc. 132, 1349–1369, doi: 10.1256/qj.05.167.

De Groot, M. H. (1970), *Optimal Statistical Decisions*, McGraw-Hill, New York.

Helsel, D.R. and R. M. Hirsch, (2002). Statistical Methods in Water Resources Techniques of Water Resources Investigations, Book 4, chapter A3. U.S. Geological Survey. pp. 522.

IPCC: Fourth Climate Assessment report, IPCC, Geneva 2007.

Kelly, K.S. and R. Krzysztofowicz,(1997). A bi-variate meta-Gaussian density for the use in hydrology, *Stochastic Hydrol. Hydraul*, *11*, 17–31.

Krzysztofowicz, R. (1999). Bayesian theory of probabilistic forecasting via deterministic hydrologic model, *Water Resour. Res.*, *35(9)*, 2739–2751.

Krzysztofowicz, R. and K.S. Kelly (2000). Hydrologic uncertainty processor for probabilistic river stage forecasting, *Water Resour. Res.*, *36(11)*, 3265–3277.

Kelly, K.S. and R. Krzysztofowicz,(2000). Precipitation uncertainty processor for probabilistic river stage forecasting, *Water Resour. Res.*, *36(9)*, 2643–2653.

Krzysztofowicz, R. (2001a), The case for probabilistic forecasting in hydrology, *J. of Hydrol*, *249*, 2–9.

Krzysztofowicz, R. (2001b), Reply, *Water Resour. Res.*, *37(2)*, 441-442.

Krzysztofowicz, R. (2004), Bayesian processor of output: a new technique for probabilistic weather forecasting, Preprints of the 17th Conference on Probablity and Statistics in the Atmospheric Sciences, Seattle, Washington, American Meteorological Society; paper no. 4.2.

Montanari, A., Brath, A., 2004. A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. Water Resour. Res., 40, doi:10.1029/2003WR002540.

Molteni, F., R. Buizza, et al. (1996). "The ECMWF Ensemble Prediction System: Methodology and validation." Quarterly *J. of the Royal Met. Soc.*, 122:73–119.

Mylne, K.R. (2002). Decision-making from probability forecasts based on forecast value, Meteorological Applications, doi:10.1017/S1350482702003043, 9 307-315.

Palmer, T.N. (2002). The economic value of ensemble forecasts as a tool for risk assessment: from days to decades, *Quart. J. Roy. Meteor. Soc.*, 128 (2002), pp. 747-774.

Raiffa, H. and Schlaifer, R. (1961). Applied statistical decision theory, The MIT Press, Cambridge, Massachusetts.

Reggiani, P. and A.H. Weerts (2008a). Probabilistic Quantitative Precipitation Forecast for Flood Prediction: an Application, *Journal of Hydrometeorology, in press*, DOI:10.1175/2007JHM858.1.

Reggiani, P. and A.H. Weerts (2008b). A Bayesian approach to decision-making under uncertainty: an application to real-time forecasting in the river Rhine, submitted to *Journal of Hydrology*.

Schaake, J. C., E. Welles, and T. Graziano (2001), Comment on Bayesian theory of probabilistic forecasting via deterministic hydrologic model by Roman Krzysztofowicz, *Water Resour. Res.*, 37(2), 439-440.

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

Todini, E. (2007). Hydrological modeling: past, present and future, Hydrol. Earth Sys. Sci., 11(1), 468–482.

United Nations, International Strategy for Disaster(2004), Guidelines for Reducing Flood Losses, UN ISDR;

Van der Waerden, B.L. (1952). Order tests for two-sample problem and their power 1. *Indagationes Mathematicae*, 14, 453–458.

Van der Waerden, B.L. (1953a). Order tests for two-sample problem and their power 2. *Indagationes Mathematicae*, 15, 303–310.

Van der Waerden, B.L. (1953b). Order tests for two-sample problem and their power 3. *Indagationes Mathematicae*, 15, 311–316.

Wilks, D. S. (1995). Statistical methods in the atmospheric sciences : an introduction. International geophysics series, Vol 59, Academic Press, San Diego.

WMO, Flood Forecasting Initiative, Report On The Technical Conference On Improved Meteorological And Hydrological Forecasting Geneva, Switzerland, 20–23 November 2006.

Zhiyu, L., M.L.V. Martina and E. Todini (2005). Flood forecasting using a fully distributed model: Application of the TOPKAPI model to the Upper Xixian Catchment, Hydrol. Earth Sys. Sci., 9(4) 347–364.

11 Figure Captions

Figure 1: The river Rhine basin with travel time isochrones expressed in number of days.

Figure 2: Empirical and modelled cdfs, January and July, Lobith.

Figure 3: Prior: linear regressions in the Gaussian (left) and the original space (right) for selected lead times, January, Lobith. The dashed line identifies the 90% confidence interval, the continuous line the linear regression.

Figure 4: Likelihood: linear regressions in the Gaussian (left) and the original space (right) for selected lead times, January, Lobith. The dashed line identifies the 90% confidence interval, the continuous line the regression.

Figure 5: Prior and revised posterior conditional probability density functions on discharge for changing lead time, Lobith, 10/02/2007. The boxes show the observed (upper box) and the forecasted (lower box) flow rates.

Figure 6: Ensemble stream-flow forecast, Lobith, 27/02/2007 with 10%-90% confidence intervals

Figure 7a: Verification: 10%-90% probability interval for the conditional prior and posterior densities around observations, lead time of 2 days, June- September 2007, Lobith.

Figure 7b: Verification: 10%-90% probability interval for the conditional prior and posterior densities around observations, lead time 5 days, June-September 2007, Lobith.

Figure 8: Ranked Probability Skill Score for Bayesian processor and EPS probability ranking.

12 Table Captions

Table 1: Coefficients and standard deviation in the prior distribution function for h0-k, k=24 hours, lead times 24, 48 and 72 hours, selected months.

Table 2: Coefficients and standard deviation in the likelihood function for h0-k , k=24 hours, lead times 24, 48 and 72 hours, selected months.

Table 3: Definition of coefficients and numerical values in the posterior distribution function for h0-k, k=24 hours, lead times 24, 48 and 72 hours, selected months.

lead time (days)	coeff.	January	April	July	October
1	α_1	1.7470	1.65670	1.67460	1.3987
	β_1	-0.771	-0.67173	-0.68543	-0.41366
	γ_1	0.0000	0.0000	0.0000	-0.00027
	τ_1	0.0107	0.009852	0.0066317	0.016745
2	α_2	2.1832	2.0286	1.9987	1.6169
	β_2	-1.2477	-1.0681	-1.0267	-0.65337
	γ_2	0.0000	0.0000	0.0000	-0.000457
	τ_2	0.051011	0.037799	0.028388	0.048031
3	α_3	2.3722	2.1655	2.1156	1.7147
	β_3	-1.4874	-1.2332	-1.1633	-0.77533
	γ_3	0.00000	0.000131	0.00000	0.00000
	τ_3	0.12115	0.082193	0.06101	0.090346
4	τ_4	0.99601	0.99481	0.99598	0.99018
5	τ_5	0.9960	0.99481	0.99598	0.99018

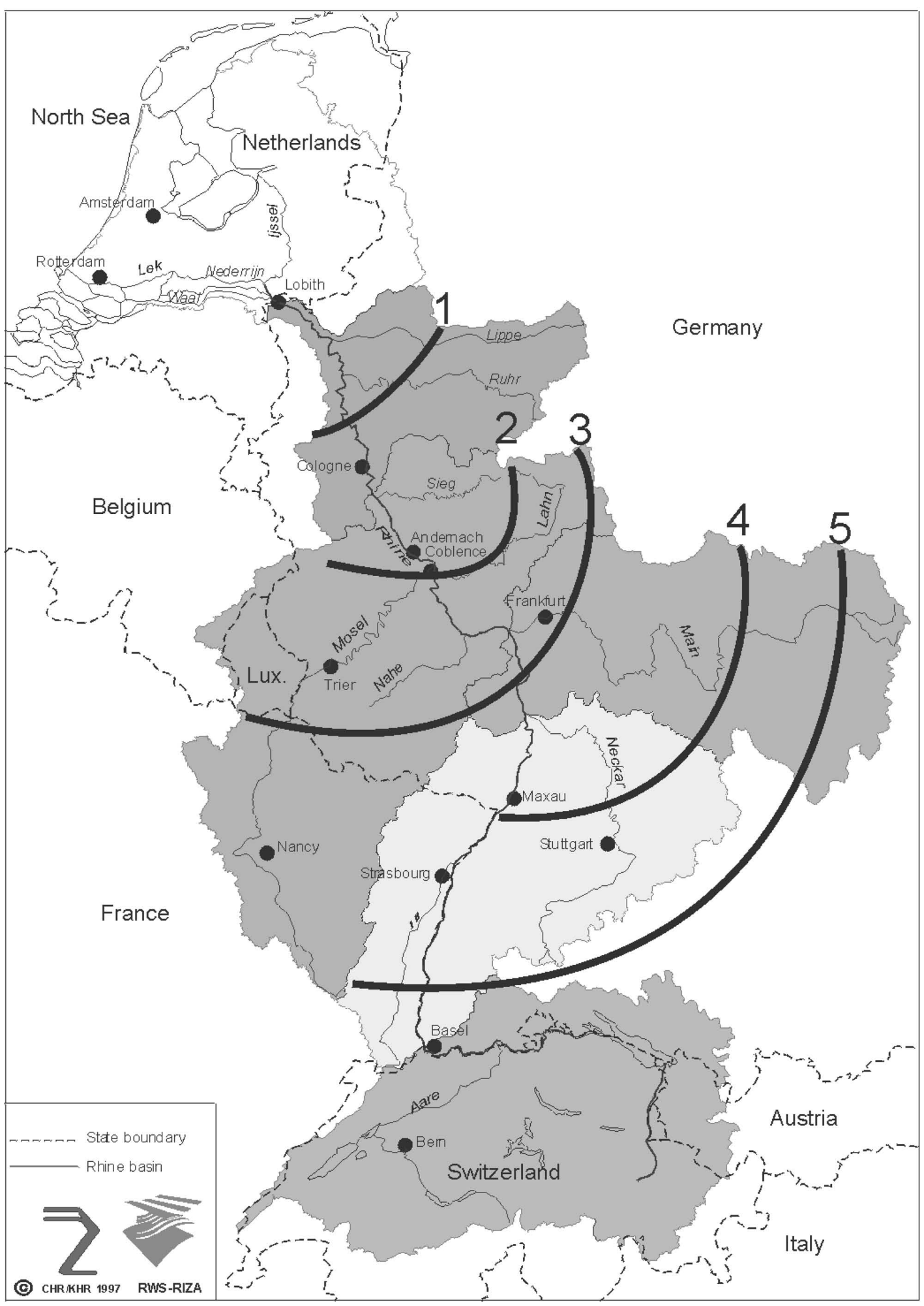
Table 1: Coefficients and standard deviation in the prior distribution for h_n, lag k=1 day, lead times 1,2,3,4 and 5 days, selected months.

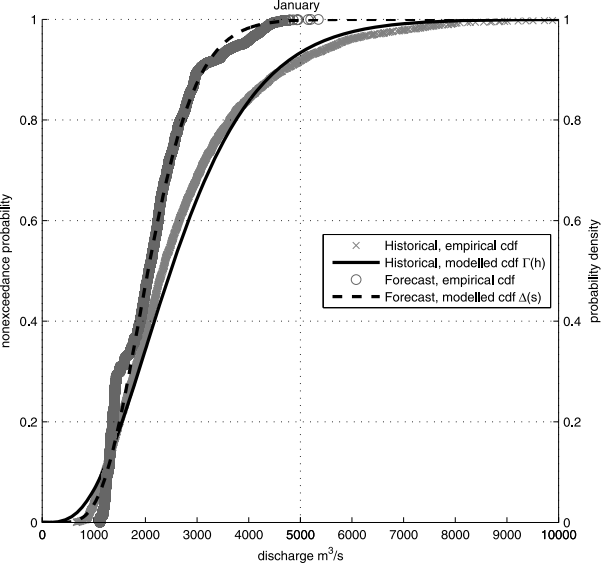
lead time (days)	coeff.	January	April	July	October
1	a_1	0.72893	0.61772	0.18364	0.63067
	b_1	0.38736	0.5725	0.88763	0.046669
	c_1	-0.12435	-0.19624	-0.13938	0.17073
	d_1	0.0000	0.0000	0.0000	0.002007
	σ_1	0.010413	0.007725	0.11756	0.37243
1	a_2	0.79367	0.70006	0.21535	0.71037
	b_2	0.34014	0.50521	0.60079	0.0093302
	c_2	-0.14908	-0.21212	0.075413	0.15483
	d_2	0.0000	0.0000	0.0000	0.04282
	σ_2	0.032016	0.016355	0.2355	0.34471
3	a_3	0.78832	0.76938	0.30164	0.71616
	b_3	0.50149	0.44434	0.40702	-0.019476
	c_3	-0.3166	-0.21827	0.18574	0.18546
	d_3	0.00000	0.0000	0.0000	0.06965
	σ_3	0.046071	0.017201	0.27702	0.32093
4	a_4	0.95738	0.98552	0.73495	0.71527
	d_4	0.00000	0.00000	0.00000	0.001833
	σ_4	0.076648	0.026348	0.42248	0.46749
5	a_5	0.94157	0.97726	0.71942	0.70765
	d_5	0.00000	0.00000	0.00000	0.014182
	σ_5	0.10422	0.041209	0.44323	0.47671

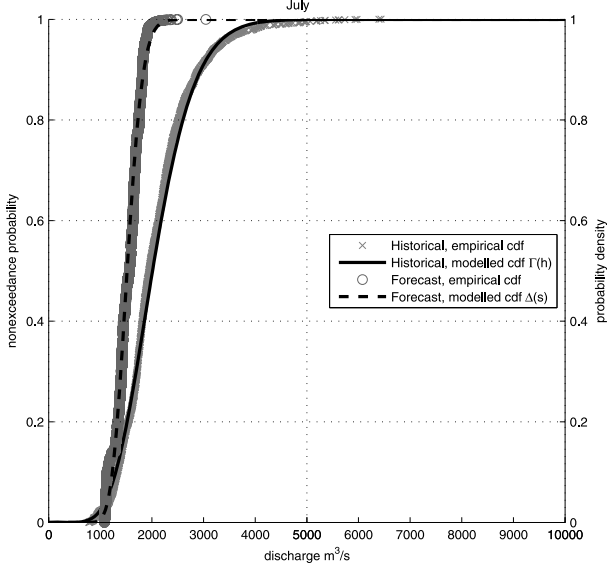
Table 2: Coefficients and standard deviation in the likelihood for h_n, lag k=1 day, lead

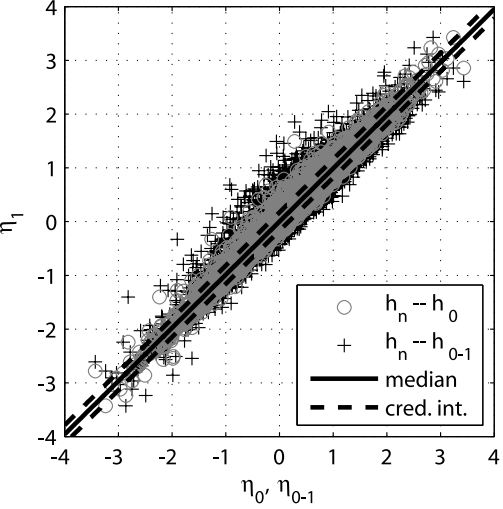
	coefficient	A_n	B_n	C_n	D_n	T_n^2
lead time (days)		$\frac{a_n\tau_n^2}{a_n^2\tau_n^2+\sigma_n^2}$	$\frac{\alpha_n\sigma_n^2-a_nb_n\tau_n^2}{a_n^2\tau_n^2+\sigma_n^2}$	$\frac{\beta_n\sigma_n^2-a_nc_n\tau_n^2}{a_n^2\tau_n^2+\sigma_n^2}$	$\frac{\gamma_n\sigma_n^2-a_nd_n\tau_n^2}{a_n^2\tau_n^2+\sigma_n^2}$	$\frac{\tau_n^2\sigma_n^2}{a_n^2\tau_n^2+\sigma_n^2}$
1	January	0.48505	0.94143	−0.43801	0.0000	0.00693
	April	0.52992	0.81099	−0.34785	0.0000	0.00663
	July	0.01034	1.66220	−0.68269	0.0000	0.00662
	October	0.02786	1.3728	−0.41115	−0.0003	0.01645
2	January	0.63113	0.87493	−0.52863	0.0000	0.02546
	April	0.75865	0.75865	−0.3399	0.0000	0.017724
	July	0.025815	0.025815	−1.0229	0.0000	0.02823
	October	0.09248	0.09248	−0.62477	−0.00438	0.044876
3	January	0.78696	0.5059	−0.31549	0.0000	0.04599
	April	0.96026	0.13893	−0.1125	0.0000	0.02147
	July	0.065126	2.0475	−1.1526	0.0000	0.05981
	October	0.17617	1.5018	−0.71019	−0.0123	0.07895
4	January	0.96361			0.0000	0.077147
	April	0.98775			0.0000	0.026407
	July	0.76213			0.0000	0.4381
	October	0.7271			−0.00133	0.47522
5	January	0.94993			0.0000	0.10515
	April	0.98073			0.0000	0.04135
	July	0.74739			0.0000	0.46046
	October	0.72047			−0.01021	0.48535

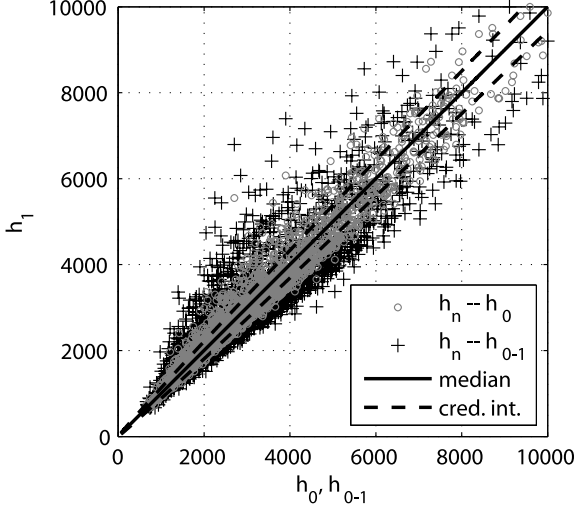
Table 3: Definition of coefficients and numerical values for observations at t_{0-1} days, lead

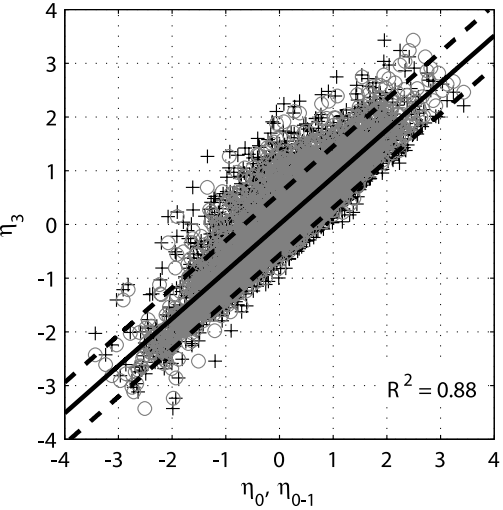

North Sea
Netherlands
Amsterdam
Rotterdam
Lek
Nederrijn
Waal
Ijssel
Lobith
1
Germany
Lippe
Ruhr
2
3
Cologne
Sieg
Lahn
Belgium
Rhine
Andernach
Coblence
4
5
Frankfurt
Main
Mosel
Lux.
Trier
Nahe
Neckar
Maxau
Nancy
Stuttgart
Strasbourg
France
Basel
Aare
Austria
Bern
Switzerland
Italy
State boundary
Rhine basin
© CHR/KHR 1997
RWS-RIZA

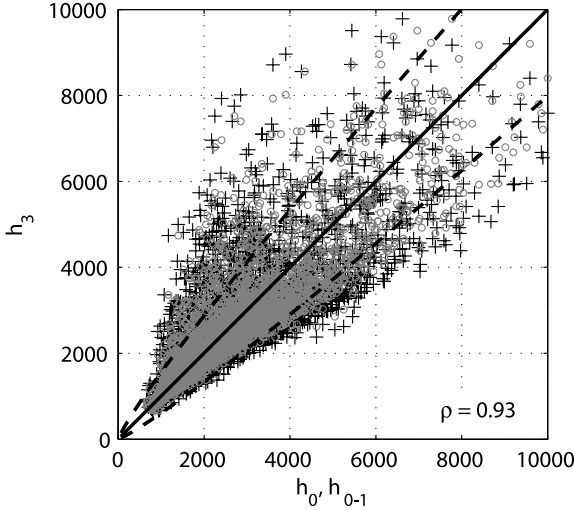

January
nonexceedance probability
probability density
discharge m3/s
Historical, empirical cdf
Historical, modelled cdf Γ(h)
Forecast, empirical cdf
Forecast, modelled cdf Δ(s)
0
0.2
0.4
0.6
0.8
1
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000

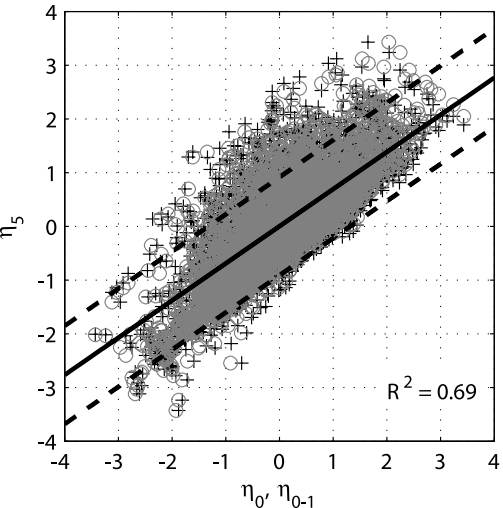

July
nonexceedance probability
probability density
discharge m3/s
Historical, empirical cdf
Historical, modelled cdf Γ(h)
Forecast, empirical cdf
Forecast, modelled cdf Δ(s)
0
0.2
0.4
0.6
0.8
1
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000

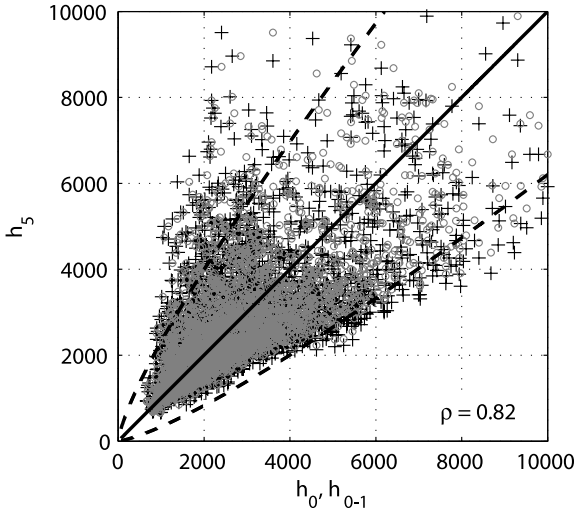

η_1
η_0, $\eta_{0\text{-}1}$
h_n -- h_0
h_n -- $h_{0\text{-}1}$
median
cred. int.

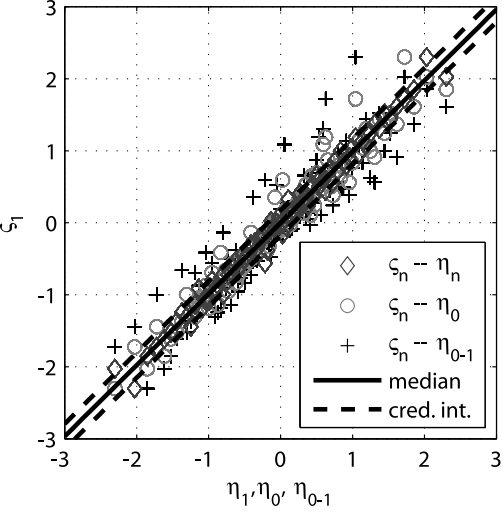

h_1
h_0, $h_{0\text{-}1}$
h_n -- h_0
h_n -- $h_{0\text{-}1}$
median
cred. int.

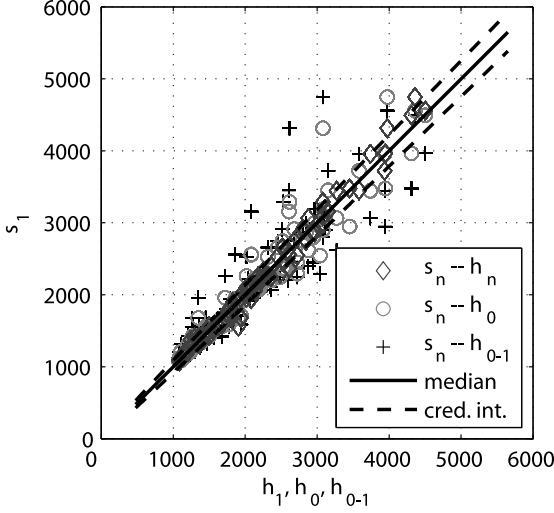

η_3
η_0, $\eta_{0\text{-}1}$
$R^2 = 0.88$

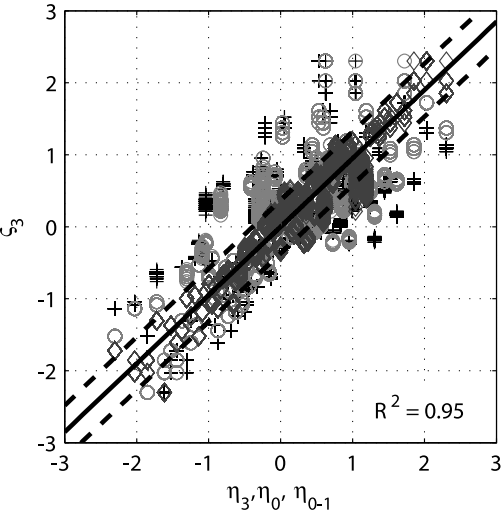

h_3
h_0, $h_{0\text{-}1}$
$\rho = 0.93$

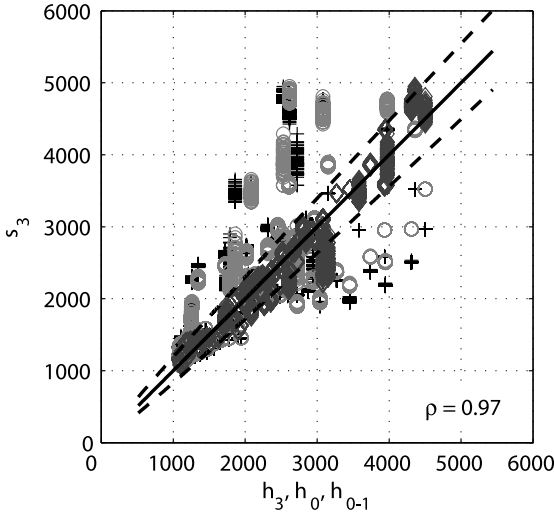

η_5
η_0, $\eta_{0\text{-}1}$
$R^2 = 0.69$

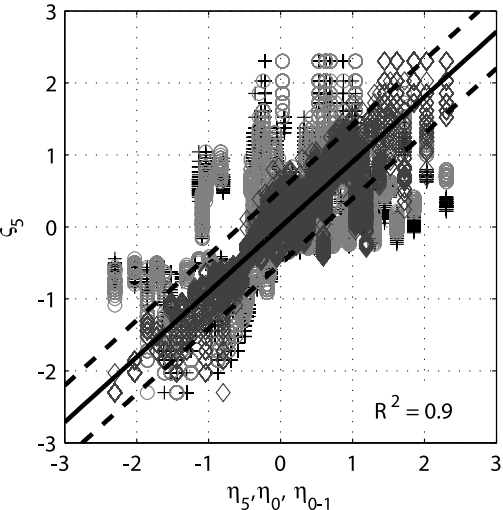

h_5
h_0, $h_{0\text{-}1}$
$\rho = 0.82$

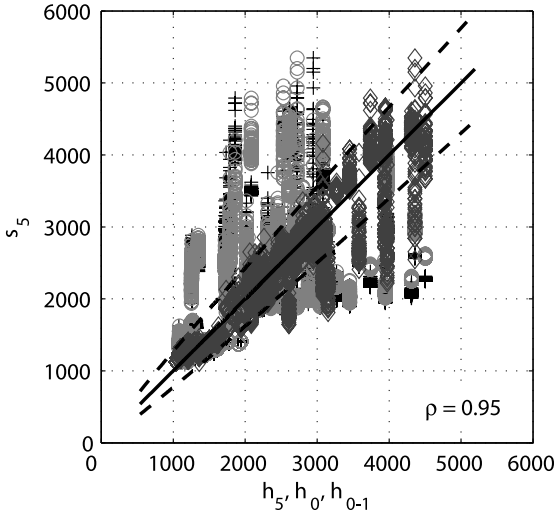

ζ1
η1, η0, η0-1
ζn -- ηn
ζn -- η0
ζn -- η0-1
median
cred. int.

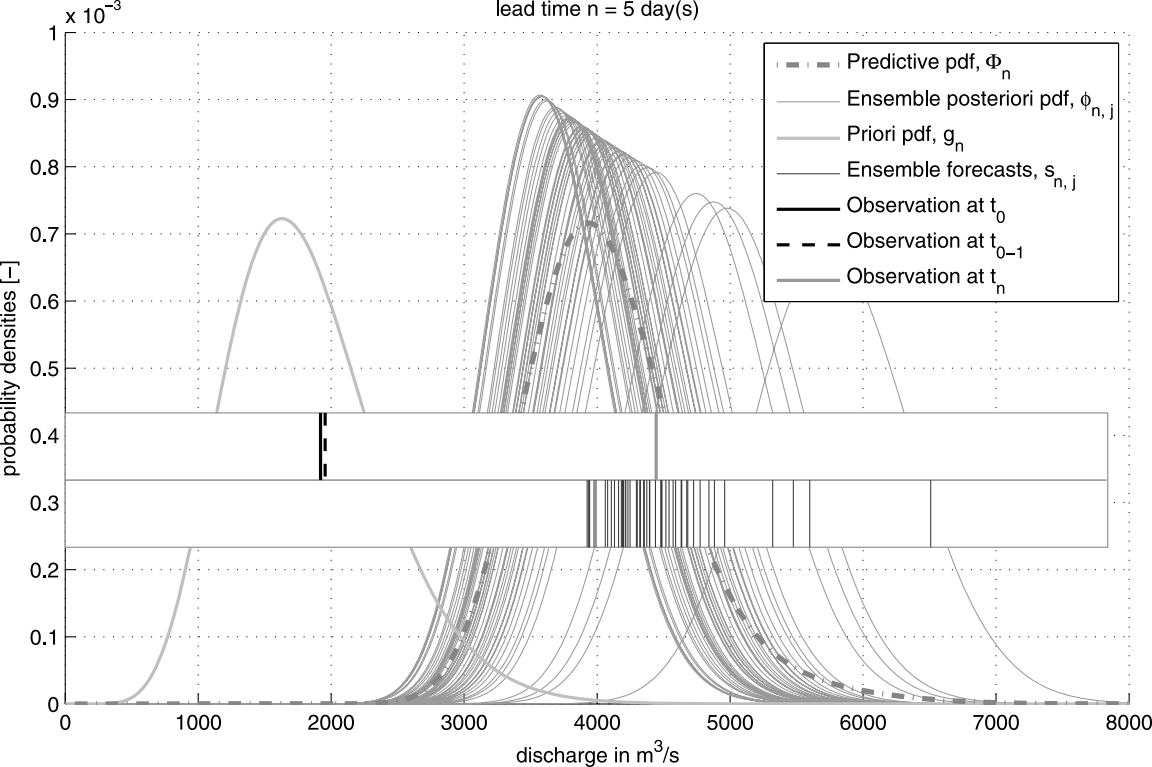

s1
h1, h0, h0-1
sn -- hn
sn -- h0
sn -- h0-1
median
cred. int.

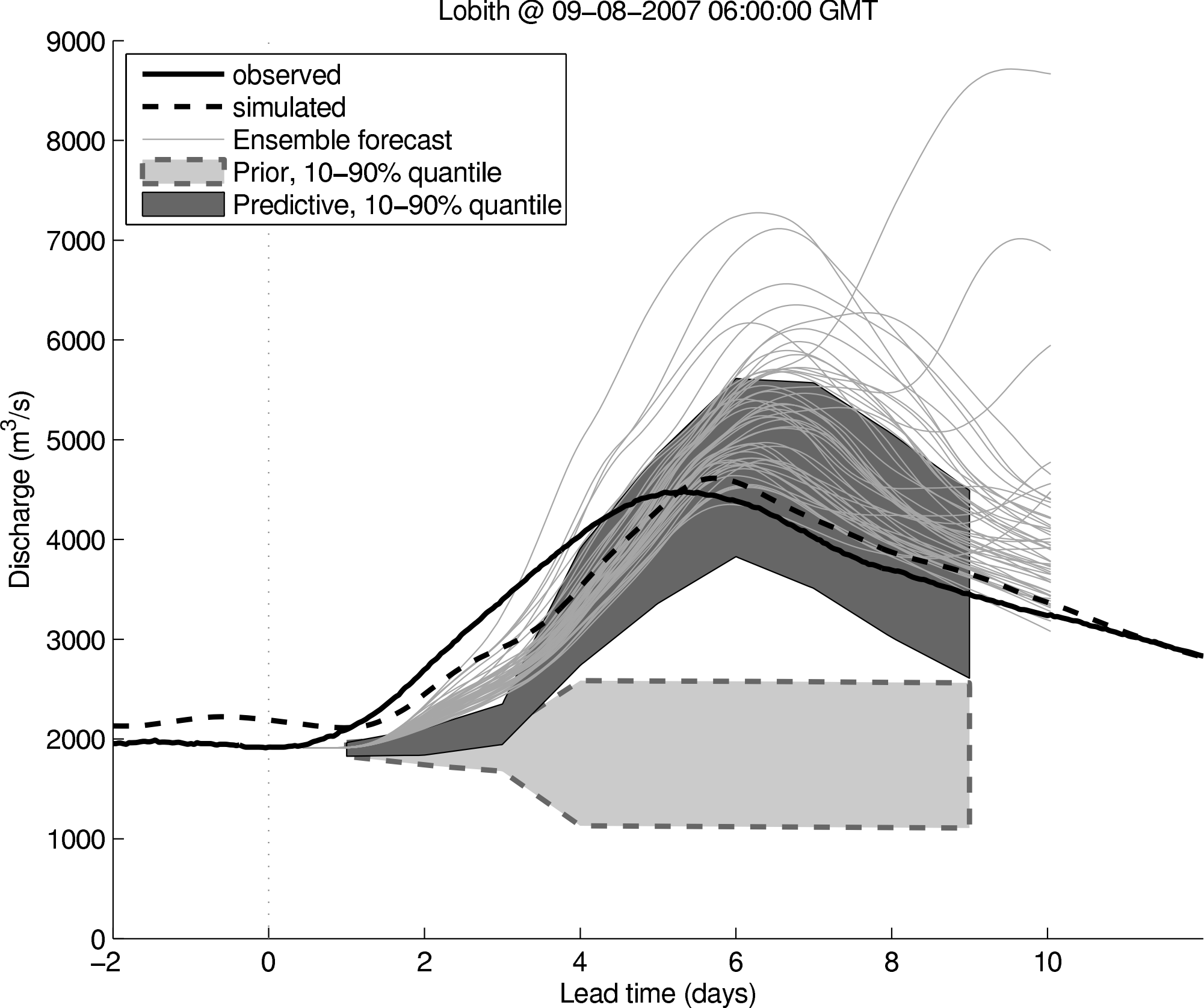

ζ3
η3, η0, η0-1
R2 = 0.95

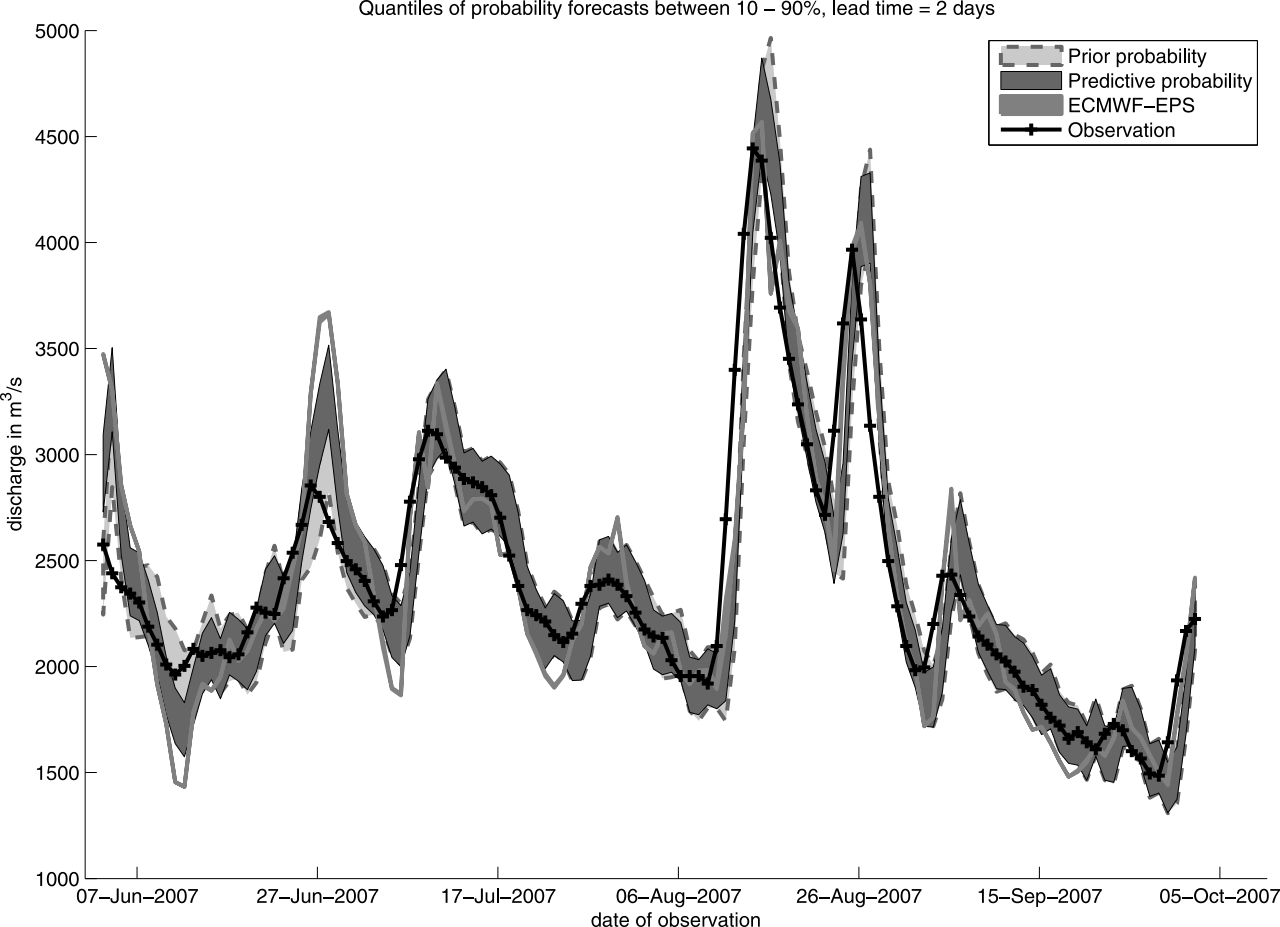

s3
h3, h0, h0-1
ρ = 0.97

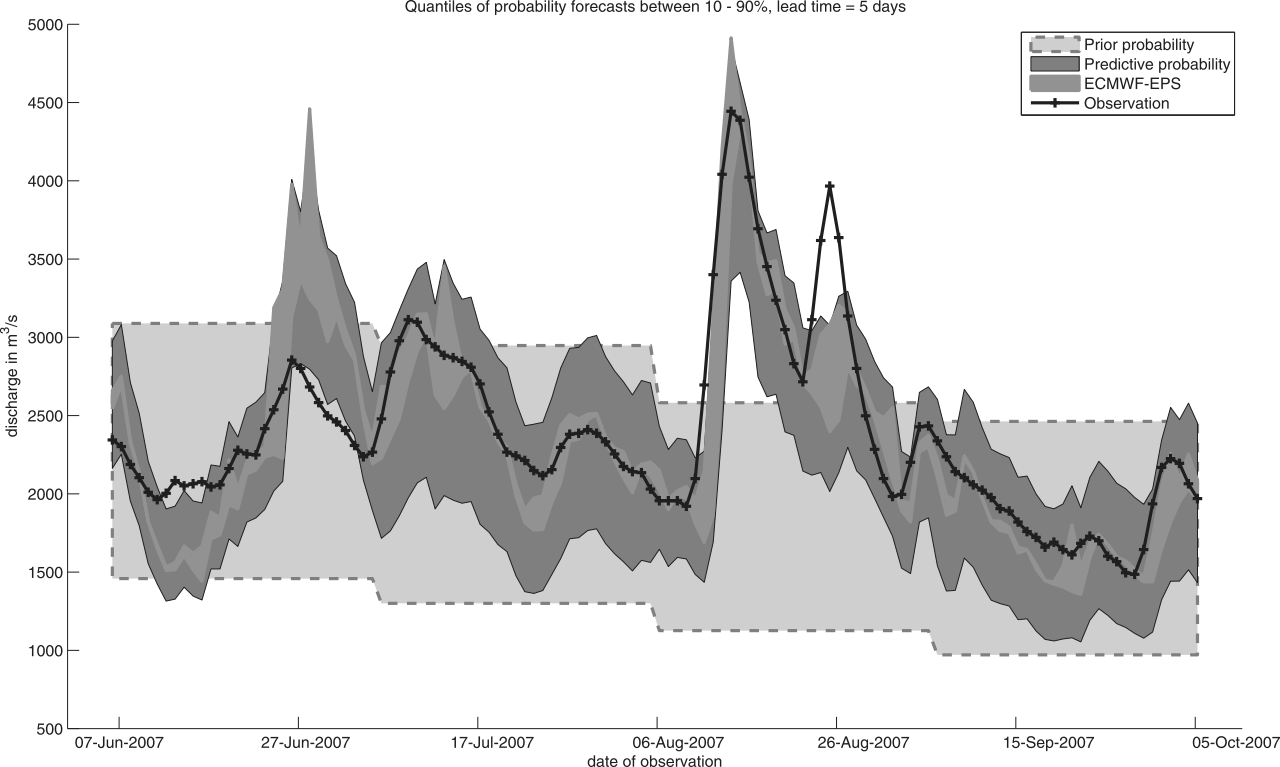

ζ5
η5, η0, η0-1
R2 = 0.9

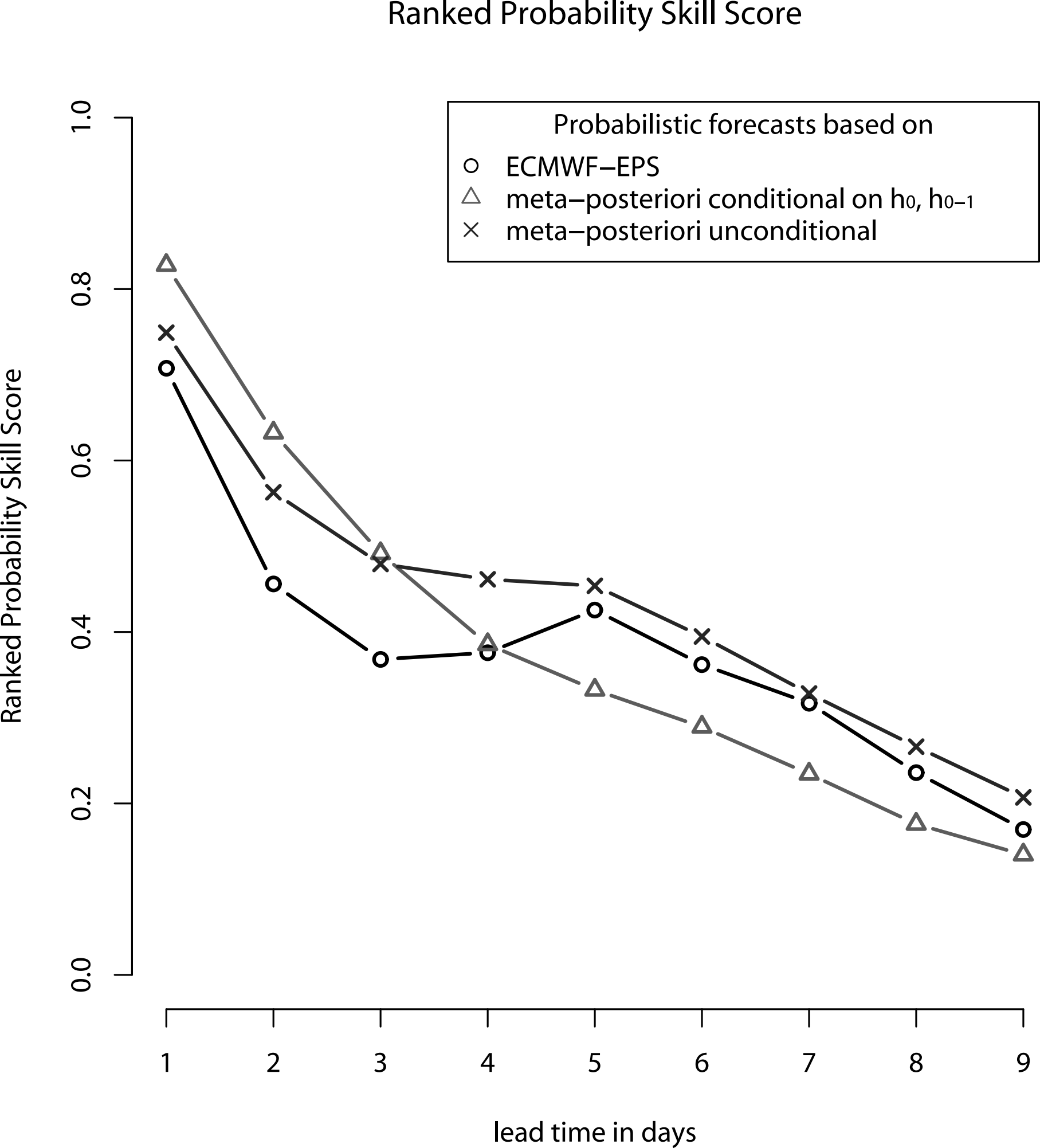

s5
h5, h0, h0-1
ρ = 0.95


lead time n = 5 day(s)
x 10^-3
probability densities [-]
discharge in m^3/s
Predictive pdf, Φ_n
Ensemble posteriori pdf, φ_n, j
Priori pdf, g_n
Ensemble forecasts, s_n, j
Observation at t_0
Observation at t_0−1
Observation at t_n
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
1000
2000
3000
4000
5000
6000
7000
8000


Lobith @ 09–08–2007 06:00:00 GMT
observed
simulated
Ensemble forecast
Prior, 10–90% quantile
Predictive, 10–90% quantile
Discharge (m3/s)
Lead time (days)
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
–2
0
2
4
6
8
10


Quantiles of probability forecasts between 10 – 90%, lead time = 2 days
Prior probability
Predictive probability
ECMWF–EPS
Observation
discharge in m3/s
5000
4500
4000
3500
3000
2500
2000
1500
1000
07–Jun–2007
27–Jun–2007
17–Jul–2007
06–Aug–2007
26–Aug–2007
15–Sep–2007
05–Oct–2007
date of observation


Quantiles of probability forecasts between 10 - 90%, lead time = 5 days
Prior probability
Predictive probability
ECMWF-EPS
Observation
discharge in m3/s
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
07-Jun-2007
27-Jun-2007
17-Jul-2007
06-Aug-2007
26-Aug-2007
15-Sep-2007
05-Oct-2007
date of observation


Ranked Probability Skill Score
Probabilistic forecasts based on
ECMWF–EPS
meta–posteriori conditional on h0, h0–1
meta–posteriori unconditional
Ranked Probability Skill Score
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
8
9
lead time in days