ELSEVIER

Journal of Hydrology 268 (2002) 16–40

Journal of Hydrology

www.elsevier.com/locate/jhydrol

Bayesian system for probabilistic river stage forecasting

Roman Krzysztofowicz*

Department of Systems Engineering and Department of Statistics, University of Virginia, P.O. Box 400747, Charlottesville, VA 22904-4747, USA

Received 3 July 2001; revised 3 April 2002; accepted 3 May 2002

Abstract

The purpose of this analytic-numerical Bayesian forecasting system (BFS) is to produce a short-term probabilistic river stage forecast based on a probabilistic quantitative precipitation forecast as an input and a deterministic hydrologic model (of any complexity) as a means of simulating the response of a headwater basin to precipitation. The BFS has three structural components: the precipitation uncertainty processor, the hydrologic uncertainty processor, and the integrator. A series of articles described the Bayesian forecasting theory and detailed each component of this particular BFS. This article presents a synthesis: the total system, operational expressions, estimation procedures, numerical algorithms, a complete example, and all design requirements, modeling assumptions, and operational attributes. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Bayesian analysis; Stochastic processes; Statistical analysis; Probability; Rivers; Floods

1. Introduction

The Bayesian forecasting system (BFS) is a general theoretical framework for probabilistic forecasting via any deterministic hydrologic model (Krzysztofowicz, 1999a). Within this framework a variety of probabilistic forecasting systems suited to different purposes can be developed. One such system was developed for short-term forecasting in headwater basins. A series of articles described the formulation, modeling, and testing of each component of this particular BFS. This article presents a synthesis: the total system, operational expressions, estimation procedures, numerical algorithms, a complete example, and all design requirements, modeling assumptions, and operational attributes. It could serve as a blueprint for operational implementation.

The purpose of this particular BFS is to produce a probabilistic river stage forecast (PRSF) based on a probabilistic quantitative precipitation forecast (PQPF) as an input and a deterministic hydrologic model as a means of simulating the response of a river basin to precipitation. It could be adapted to produce a probabilistic river discharge forecast or a probabilistic runoff volume forecast. It is designed to meet six requirements.

1. The predictand is a time series of river stages at the outlet of a headwater basin.
2. The dominant source of forecast uncertainty is the basin average precipitation amount during the coming period.
3. The PQPF is produced either algorithmically by a model or judgmentally by a meteorologist.
4. The maximum lead time of the PRSF equals, approximately, the lead time of the PQPF plus the concentration time of the basin.
5. The system works in conjunction with any deterministic hydrologic model of a river basin.

* Tel.: +1-434-924-5393; fax: +1-434-982-2972.
E-mail address: rk@virginia.edu (R. Krzysztofowicz).

PII: S0022-1694(02)00106-3

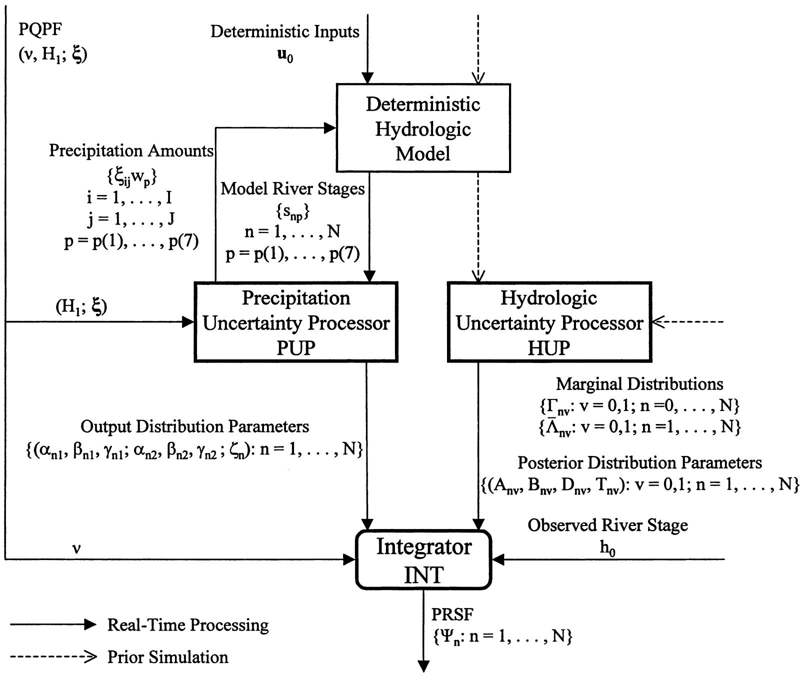

Fig. 1. Structure of the Bayesian forecasting system (BFS) which produces a probabilistic river stage forecast (PRSF) based on a probabilistic quantitative precipitation forecast (PQPF) processed through a deterministic hydrologic model. Symbols denote inputs and outputs during execution of the BFS in real time.

6. The system is computationally efficient and its execution in real time is simple and fast.

In essence, the BFS can be attached to any deterministic hydrologic model for operational forecasting. Its function is to quantify the total uncertainty about the predictand. This uncertainty is defined from the viewpoint of a user of forecasts—a rational decision maker who wants to make optimal decisions that take forecast uncertainty explicitly into account. To achieve the computational efficiency and operational simplicity, the BFS is built of analytic expressions which require minimal numerical calculations. In the classification of all conceivable versions of the BFS according to the methods of computation, this is an analytic-numerical BFS (Krzysztofowicz, 2001a).

Section 2 introduces the decomposition of the total uncertainty, the structure of the system, the time scale, and the example. Section 3 specifies the required format of the PQPF. The next three sections present the system components: Section 4 details the precipitation uncertainty processor, Section 5 details the hydrologic uncertainty processor, and Section 6 details the integrator. Section 7 highlights properties of the PRSF and presents algorithms for updating the PRSF. Section 8 summarizes the assumptions and the attributes of the BFS.

2. Preliminaries

2.1. Decomposition of uncertainty

The sources of uncertainty associated with a river stage forecast can be categorized as operational, precipitation, and hydrologic. *Operational uncertainty* is caused by erroneous or missing data, human processing errors, unpredictable interventions (e.g. changes in reservoir releases not communicated by a dam operator to the forecaster), unpredictable obstacles within a river channel (e.g. ice jams), and the like. These sources of uncertainty are exterior to the forecasting theory. Therefore, the term 'total uncertainty' used henceforth will not encompass operational uncertainty.

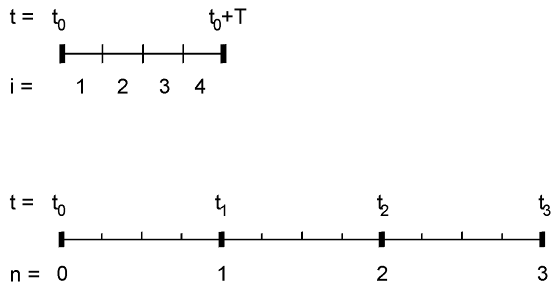

Fig. 2. Example of time scales for the PQPF (upper scale) and the PRSF (lower scale).

In the BFS, the total uncertainty is decomposed into precipitation uncertainty and hydrologic uncertainty. The sources of these uncertainties are as follows. *Precipitation uncertainty* is associated with the total basin average precipitation amount during the period covered by the PQPF. This uncertainty is quantified in terms of a probability distribution specified by the PQPF. *Hydrologic uncertainty* is the aggregate of all uncertainties arising from sources other than the total basin average precipitation amount. These sources include: (i) the deterministic forecast of the temporal disaggregation of a total precipitation amount into subperiods (this forecast is part of the PQPF), (ii) the deterministic forecast of the spatial disaggregation of a total precipitation amount into subbasins (this forecast is part of the PQPF when the basin consists of multiple subbasins), (iii) the deterministic estimates of all other inputs into the hydrologic model (in general, the measurement, estimation, and prediction uncertainties), (iv) the imperfections of the hydrologic model (in general, the model and parameter uncertainties), and (v) the precipitation beyond the period covered by the PQPF.

2.2. System structure

The decomposition of the total uncertainty is justified by principles of Bayesian predictive inference. These principles prescribe the structure of the BFS depicted in Fig. 1. A PQPF, which quantifies precipitation uncertainty, is the primary input. Two processors are attached to the hydrologic model. One processor maps precipitation uncertainty into output uncertainty under the hypothesis that there is no hydrologic uncertainty. Another processor quantifies hydrologic uncertainty under the hypothesis that there is no precipitation uncertainty. Then the two uncertainties are optimally integrated to produce a PRSF. Hence the BFS has three structural components: the precipitation uncertainty processor (PUP), the hydrologic uncertainty processor (HUP), and the integrator (INT).

Fig. 1 also lists symbolically all inputs and outputs within the BFS that must be defined for operational forecasting. The symbols are explained along with the system components in subsequent sections.

2.3. Time scale

As is common in operational forecasting, PRSFs are assumed to be prepared on schedule, once or more times per day. On each forecasting occasion, the time scale t is reset; t_0 denotes the forecast time, which coincides with the last observation time before forecast preparation; and t_n $(n = 1, \ldots, N)$ denotes the time at which the river stage is forecasted and then observed (Fig. 2). PRSFs are prepared for times $t_1, \ldots, t_N$, coinciding with the abscissae of a discrete-time hydrograph calculated by the hydrologic model (though these calculations may be performed on finer time steps).

The lead time of the PRSF prepared at time t_0 for time t_n equals $t_n - t_0$. When $\Delta = t_n - t_{n-1}$ is constant for $n = 1, \ldots, N$, the lead time equals $n\Delta$. For simplicity, index n itself will be referred to as lead time.

On each forecasting occasion, a PQPF must be prepared for period $[t_0, t_0 + T]$, which begins at the forecast time t_0 and ends at time $t_0 + T \in \{t_1, \ldots, t_N\}$. The period is divided into I subperiods indexed by i $(i = 1, \ldots, I)$ and coinciding with the time steps in the hydrologic model (Fig. 2).

The river basin above the forecast point may be represented by a lumped hydrologic model or a semi-distributed (or a distributed) hydrologic model, in which case the basin is partitioned into J subbasins (or subareas) indexed by j $(j = 1, \ldots, J)$.

2.4. Example

The BFS has been tested operationally by the US National Weather Service (NWS). The example reported throughout the article is for the forecast point Eldred, Pennsylvania, located in the headwater of the Allegheny River and closing a drainage area of 550 square miles (1430 km^2). All simulations of river stages were performed at the NWS using the operational forecast system whose description can be found in Hudlow (1988) and Fread et al. (1995). For Eldred, the NWS uses a lumped hydrologic model with a single precipitation input ($J = 1$) in the form of a time series of 6-h basin average precipitation amounts. The model outputs a time series of river stages at 6-h steps. (The time to peak of the unit hydrograph is 30 h.)

In the pilot testing, forecasts are produced daily. The PQPF is prepared for a 24-h period beginning at 1200 UTC (Universal Time Coordinated), divided into four 6-h subperiods ($I = 4$). The PRSFs are prepared based on the input data available at 1200 UTC on the forecast day. The example uses real-time input data from the NWS archives and has two versions. In the complete version, illustrating all system components, the PRSFs are produced for 3 days in 24-h steps ($N = 3$). In the abbreviated version, illustrating only the final output, the PRSFs are produced for 3 days in 6-h steps ($N = 12$).

3. Precipitation input

3.1. Input requirements

A system that produces the PQPF is not part of the BFS. It must be developed separately and may employ any forecasting method. However, it must meet two requirements:

1. The PQPF must be in a specified format.
2. The PQPF system must be well calibrated.

The notion of calibration (from the Bayesian point of view) and the criteria for calibration are described elsewhere (Krzysztofowicz and Sigrest, 1999). The required formats of the predictand and the PQPF are detailed next.

3.2. Precipitation predictand

Let W denote the basin average precipitation amount to be accumulated during the period; let w denote a realization of W. (Hereinafter W is referred to as the total precipitation amount.) Let V denote an indicator of precipitation occurrence, with $V = 0 \Leftrightarrow W = 0$ and $V = 1 \Leftrightarrow W > 0$. Hence, the precipitation event is denoted $V = v$, where $v \in \{0, 1\}$.

Furthermore, let W_{ij} denote the average precipitation amount to be accumulated during subperiod i and over subbasin j. Conditional on the hypothesis that precipitation occurs, $V = 1$, define a factor $\Theta_{ij} = W_{ij}/W$. The matrix of factors $\mathbf{\Theta} = \{\Theta_{ij} : i = 1, \ldots, I;\ j = 1, \ldots, J\}$ defines the spatiotemporal disaggregation of the total precipitation amount $W > 0$. The predictand consists of W and $\mathbf{\Theta}$.

For a lumped hydrologic model, matrix $\mathbf{\Theta}$ reduces to vector $\mathbf{\Theta} = (\Theta_1, \ldots, \Theta_I)$, where $\Theta_i = W_i/W$ is a fraction conditional on the hypothesis $V = 1$, and where W_i is the basin average precipitation amount to be accumulated during subperiod i ($i = 1, \ldots, I$). The vector of fractions $\mathbf{\Theta}$ defines the temporal disaggregation of the total precipitation amount $W > 0$.

3.3. Precipitation forecast

The PQPF for a river basin consists of two parts. The first part is a probabilistic forecast of the total precipitation amount W. This forecast must specify (i) the probability of precipitation occurrence during the period and over the basin,

$$\nu = P(V = 1), \tag{1}$$

such that $0 \le \nu \le 1$, and (ii) the distribution of the total precipitation amount, conditional on the hypothesis that precipitation occurs,

$$H_1(w) = P(W \le w | V = 1), \tag{2}$$

such that $H_1(w) > 0$ if $w > 0$ and $H_1(w) = 0$ if $w = 0$.

The second part of the PQPF is a deterministic forecast of the spatiotemporal disaggregation $\mathbf{\Theta}$. This forecast must specify a matrix of expected factors, conditional on the hypothesis that precipitation occurs:

$$\boldsymbol{\xi} = \{\xi_{ij} : i = 1, \ldots, I;\ j = 1, \ldots, J\}, \tag{3a}$$

$$\xi_{ij} = E[\Theta_{ij} | V = 1], \tag{3b}$$

Table 1
Real-time input to the BFS at the forecast time: probabilistic quantitative precipitation forecast (PQPF) for the basin and river stage observed at the forecast point

Input	Symbol	Value
Precipitation forecast		
Probability of precipitation occurrence	ν	0.85
Conditional distribution H_1 of amount:[a]		
scale parameter	α	1.807
shape parameter	β	1.378
Expected fraction for:		
subperiod 1	ξ_1	0.00
subperiod 2	ξ_2	0.10
subperiod 3	ξ_3	0.40
subperiod 4	ξ_4	0.50
Stage observation		
Observed river stage[b]	h_0	7.90

[a] Weibull distribution, precipitation amount in inches.
[b] Stage in feet.

with the conditions $0 \leq \xi_{ij}$ and

$$\sum_{j=1}^{J} \frac{r_j}{r} \sum_{i=1}^{I} \xi_{ij} = 1, \tag{3c}$$

where r is the area of the basin and r_j is the area of subbasin j.

For a lumped hydrologic model, matrix $\boldsymbol{\xi}$ reduces to vector of expected fractions $\boldsymbol{\xi} = (\xi_1, \dots, \xi_I)$, such that $\xi_i = E[\Theta_i | V = 1]$, with $0 \leq \xi_i \leq 1$ for $i = 1, \dots, I$, and $\xi_1 + \cdots + \xi_I = 1$.

In summary, the PQPF for a river basin must be in the format $(\nu, H_1; \boldsymbol{\xi})$. On those forecasting occasions on which $\nu = 0$, elements H_1 and $\boldsymbol{\xi}$ are undefined. Otherwise, the conditional distribution H_1 may be of any form; however, there is an operational advantage (to be revealed later) when H_1 belongs to a parametric family of distributions which have closed-form expressions for H_1 and its inverse H_1^{-1}. The choice of such family may be guided by the goodness-of-fit between the parametric distribution and the empirical distribution constructed for the given basin and season from a climatic precipitation record (Krzysztofowicz and Sigrest, 1997). In the Eastern United States, the Weibull family with two parameters (scale parameter α and shape parameter β) was found suitable for all seasons. Consequently, the PQPF for a river basin may be in the parametric format $(\nu, \alpha, \beta; \boldsymbol{\xi})$.

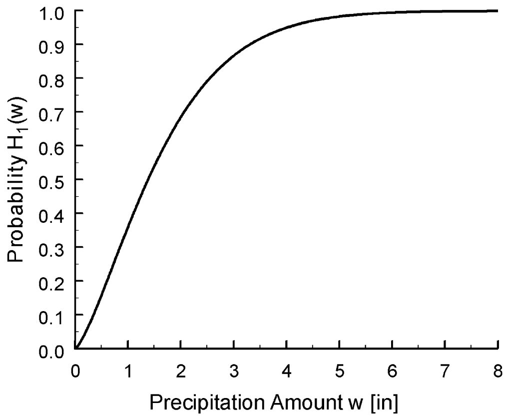

Fig. 3. Distribution H_1 of the basin average precipitation amount W, conditional on the hypothesis that precipitation occurs, $V = 1$, specified by the PQPF for the 24-h period beginning at 1200 UTC on the forecast day. Distribution H_1 is Weibull.

3.4. Equivalence and rescaling

The rationale for forecasting the spatiotemporal disaggregation deterministically is fourfold. First, a probabilistic forecast of the matrix of factors $\boldsymbol{\Theta}$ would be of such a complexity that an operational meteorologist could not prepare or adjust the PQPF judgmentally (Krzysztofowicz et al., 1993), which was one of the requirements for this particular PQPF system. Second, results of an empirical investigation support the *deterministic equivalence principle* (Kelly and Krzysztofowicz, 2000): for a class of basins and under certain conditions a probabilistic forecast of $\boldsymbol{\Theta}$ is unwarranted because a deterministic forecast of $\boldsymbol{\Theta}$ provides equivalent information for river stage forecasting. Third, even if the deterministic equivalence principle does not hold, the BFS accounts for the uncertainty arising from a deterministic forecast of $\boldsymbol{\Theta}$ through a suitably designed HUP (Krzysztofowicz and Herr, 2001). Fourth, the deterministic forecast of $\boldsymbol{\Theta}$ makes it possible to formulate the analytic-numerical PUP which offers significant operational advantages (to be revealed later).

Example. The PQPF for a particular forecast day is reported in Table 1; the conditional distribution H_1 of W is plotted in Fig. 3. This PQPF came from a prototype system of the NWS. In the current implementation, the forecast is for a 24-h period, divided into four 6-h subperiods. A source forecast, which is prepared daily by an operational meteorologist, consists of six gridded fields covering a service area. From the source forecast, a PQPF for any basin within the service area can be computed. The computation begins with the rescaling of the source forecast. At each grid point, the source forecast specifies the PQPF for a nominal area, which is 1930 square miles (5000 km^2) in the current implementation. The PQPF for a nominal area is rescaled to the PQPF for an area equal to the area of the given basin. The operational procedure for rescaling a PQPF is documented by Krzysztofowicz (1999b). Other elements of the NWS prototype PQPF system are partially documented by Antolik (2000), Krzysztofowicz and Pomroy (1997), Krzysztofowicz and Sigrest (1997), Krzysztofowicz et al. (1993), Mills and Krzysztofowicz (1998), and Seo et al. (2000).

4. Precipitation uncertainty processor

The purpose of the PUP is to map the uncertainty associated with the total precipitation amount through the hydrologic model into the uncertainty associated with the model river stage. The mapping takes place at the forecast time after the PQPF has been prepared and all deterministic inputs to the hydrologic model have their values set for the particular forecasting occasion. The PUP presented herein was developed by Kelly and Krzysztofowicz (2000).

4.1. Conditional output distribution

To formulate the PUP, three variables need be defined as follows:

$\mathbf{u}_0$—subvector of deterministic inputs to the hydrologic model at the forecast time; these inputs encompass all internal states (initial conditions) and all exogenous variables (except future precipitation) whose values vary from one forecast time to the next; they exclude parameters of the hydrologic model (whose values remain fixed for a given river basin).

$\mathbf{u}$—vector of all deterministic inputs, which encompasses the subvector $\mathbf{u}_0$ and the matrix of expected disaggregation factors $\boldsymbol{\xi}$; that is, $\mathbf{u} = (\mathbf{u}_0, \boldsymbol{\xi})$.

s_n—model river stage (an estimate of the actual river stage to be observed at time t_n) output from the hydrologic model at the forecast time based on (i) the deterministic input vector $\mathbf{u}$ and (ii) the perfect forecast of the total precipitation amount W; because no perfect forecast of W is available, the model river stage is uncertain and thus is treated as a random variable, denoted S_n.

For every lead time n ($n = 1, \ldots, N$), define the conditional output distribution. This is the distribution of the model river stage, conditional on the hypothesis that precipitation occurs,

$$\Pi_{n1}(s_n) = P(S_n \le s_n | V = 1), \tag{4}$$

such that $\Pi_{n1}(s_n) > 0$ if $s_n > s_{n0}$ and $\Pi_{n1}(s_n) = 0$ if $s_n = s_{n0}$, where s_{n0} is the model river stage resulting from zero total precipitation amount. The task of the PUP is to map the conditional distribution H_1 of total precipitation amount W into the conditional distribution Π_{n1} of model river stage S_n, given the deterministic input vector $\mathbf{u}$.

4.2. Numerical mapping

A numerical procedure for obtaining the conditional distribution Π_{n1} at the forecast time consists of five steps.

1. Define a set of probabilities $\{p(1), \ldots, p(7)\}$ and assign to them values 0, 0.25, 0.50, 0.75, 0.90, 0.95, 0.995; these values were selected experimentally.
2. Calculate a set $\{w_p : p = p(1), \ldots, p(7)\}$, where w_p is the conditional quantile of total precipitation amount W corresponding to probability p under distribution H_1 :
$$w_p = H_1^{-1}(p). \tag{5}$$
3. For each p ($p = p(1), \ldots, p(7)$), use the matrix of expected disaggregation factors $\boldsymbol{\xi}$ to calculate a matrix of precipitation amounts:
$$\{\xi_{ij} w_p : \ i = 1, \ldots, I; \ j = 1, \ldots, J\},$$
where $\xi_{ij} w_p$ is the precipitation amount for subperiod i and subbasin j.
4. For each p ($p = p(1), \ldots, p(7)$), run the hydrologic model. A run takes as input $\mathbf{u}_0$ and $\{\xi_{ij} w_p : \ i = 1, \ldots, I; \ j = 1, \ldots, J\}$ and outputs a time series of model river stages
$$\{s_{np} : \ n = 1, \ldots, N\},$$

Table 2
Conditional quantiles $\{w_p : p = p(1), \ldots, p(7)\}$ of total precipitation amount W determined from distribution H_1, and conditional quantiles $\{s_{np} : p = p(1), \ldots, p(7)\}$ of model river stage S_n ($n = 1, 2, 3$) output from the hydrologic model at the forecast time

Symbol	Probability (p)						
	0	0.25	0.50	0.75	0.90	0.95	0.995
w_p	0.00	0.73	1.39	2.29	3.31	4.01	6.06
s_{1p}	5.99	6.80	7.74	9.17	10.37	12.01	14.44
s_{2p}	5.68	10.54	14.34	18.34	20.80	22.44	25.27
s_{3p}	5.40	8.85	12.19	15.75	18.13	20.04	22.76

Precipitation amount in inches, stage in feet.

where s_{np} is the conditional quantile of S_n corresponding to probability p under distribution Π_{n1} :

$$p = \Pi_{n1}(s_{np}). \tag{6}$$

5. For each n ($n = 1, \ldots, N$), use the set of points $\{(s_{np}, p) : p = p(1), \ldots, p(7)\}$ to estimate a parametric model for the conditional distribution Π_{n1}.

In essence, the mapping requires only seven runs of the hydrologic model. Given the deterministic input vector $\mathbf{u}$ (which encompasses $\boldsymbol{\xi}$), each quantile of W is mapped into the corresponding quantile of S_n for all n ($n = 1, \ldots, N$). The set of seven probabilities was selected via an extensive simulation experiment. This is the smallest set that ensures an accurate representation of Π_{n1} when a parametric model is employed. This model is detailed in Section 4.3.

Example. Table 2 reports the numerical example. The first line shows the seven quantiles $\{w_p\}$ calculated from distribution H_1, which is specified in Table 1. For the lumped hydrologic model, the seven matrices of precipitation amounts reduce to seven time series $\{\xi_i w_p : i = 1, 2, 3, 4\}$ with the expected fractions $\{\xi_i : i = 1, 2, 3, 4\}$ coming from Table 1. The next three lines report the seven time series $\{s_{np} : n = 1, 2, 3\}$ of model river stages output from the hydrologic model.

4.3. Two-piece Weibull distribution

The three-parameter Weibull distribution of variate S, denoted $Wb(s; \alpha, \beta, \gamma) = P(S \le s)$ and having scale parameter $\alpha > 0$, shape parameter $\beta > 0$, and shift parameter $\gamma \in (-\infty, \infty)$, is specified by

$$Wb(s; \alpha, \beta, \gamma) = 1 - \exp\left(-\left(\frac{s - \gamma}{\alpha}\right)^{\beta}\right), \quad \gamma < s, \tag{7}$$

and $Wb(s; \alpha, \beta, \gamma) = 0$ if $s \le \gamma$. The two-parameter Weibull distribution results when $\gamma = 0$.

The parametric model for Π_{n1} is built up of two Weibull functions. Called the two-piece Weibull distribution, the model is specified by:

$$\Pi_{n1}(s_n) = Wb(s_n; \alpha_{n1}, \beta_{n1}, \gamma_{n1}), \quad \zeta_n < s_n, \tag{8a}$$

$$\Pi_{n1}(s_n) = Wb(s_n; \alpha_{n2}, \beta_{n2}, \gamma_{n2}), \quad \gamma_{n2} < s_n \le \zeta_n, \tag{8b}$$

$$\Pi_{n1}(s_n) = 0, \quad s_n \le \gamma_{n2}, \tag{8c}$$

where the seven parameters $(\alpha_{n1}, \beta_{n1}, \gamma_{n1}; \alpha_{n2}, \beta_{n2}, \gamma_{n2}; \zeta_n)$ satisfy the following constraints:

$$\max\{\gamma_{n1}, \gamma_{n2}\} < \zeta_n, \tag{9a}$$

$$\left(\frac{\zeta_n - \gamma_{n1}}{\alpha_{n1}}\right)^{\beta_{n1}} = \left(\frac{\zeta_n - \gamma_{n2}}{\alpha_{n2}}\right)^{\beta_{n2}}, \tag{9b}$$

$$\frac{\beta_{n1}}{\zeta_n - \gamma_{n1}} = \frac{\beta_{n2}}{\zeta_n - \gamma_{n2}}. \tag{9c}$$

Constraint (9a) is necessary for the existence of a point ζ_n at which the two Weibull functions meet; constraint (9b) ensures that the distribution Π_{n1} is continuous at ζ_n; constraint (9c) ensures that the corresponding density is continuous at ζ_n. When Eqs. (9b) and (9c) are taken into account, the two-piece Weibull distribution has only five independent parameters.

4.4. Two-piece Weibull density

Whereas the estimation algorithm described below applies to the conditional output distribution Π_{n1}, the integrator of the BFS needs the corresponding conditional output density π_{n1}. This density is readily defined as follows. Let $wb(s; \alpha, \beta, \gamma)$ denote a three-parameter Weibull density of variate S corresponding

Table 3
Parameters of the two-piece Weibull conditional densities π_{n1} of model river stages calculated at the forecast time

Lead time (n)	Parameters[a]						
	α_{n1}	β_{n1}	γ_{n1}	α_{n2}	β_{n2}	γ_{n2}	ζ_n
1	2.935	1.500	5.52	2.758	1.023	5.99	7.00
2	177.590	35.344	− 160.90	13.844	1.210	5.68	11.59
3	38.398	8.291	− 24.19	11.464	1.056	5.40	9.72

[a] Lower bound on S_n is $\gamma_{n2} = s_{n0}$; stage in feet.

to distribution (Eq. (7)) and specified by

$$wb(s;\alpha,\beta,\gamma) = \frac{\beta}{\alpha}\left(\frac{s-\gamma}{\alpha}\right)^{\beta-1} \cdot \exp\left(-\left(\frac{s-\gamma}{\alpha}\right)^{\beta}\right), \quad \gamma < s, \tag{10}$$

and $wb(s;\alpha,\beta,\gamma) = 0$ if $s \le \gamma$. Then a parametric model for π_{n1}, called the two-piece Weibull density, is specified by:

$$\pi_{n1}(s_n) = wb(s_n;\alpha_{n1},\beta_{n1},\gamma_{n1}), \quad \zeta_n < s_n, \tag{11a}$$

$$\pi_{n1}(s_n) = wb(s_n;\alpha_{n2},\beta_{n2},\gamma_{n2}), \quad \gamma_{n2} < s_n \le \zeta_n, \tag{11b}$$

$$\pi_{n1}(s_n) = 0, \quad s_n \le \gamma_{n2}, \tag{11c}$$

where $(\alpha_{n1},\beta_{n1},\gamma_{n1};\alpha_{n2},\beta_{n2},\gamma_{n2};\zeta_n)$ are the parameters satisfying constraint (Eq. (9a)–(9c)).

4.5. Estimation of parameters

The task is to fit model (Eq. (8a)–(8c)) to the set of seven points $\{(s_{np},p): p = 0, 0.25, 0.50, 0.75, 0.90, 0.95, 0.995\}$. The estimation algorithm is partly mathematical and partly heuristic; it is the result of an extensive experimentation. The algorithm is also lengthy. Therefore, it is presented herein only in concept, which boils down to four steps.

1. The class to which Π_{n1} belongs is identified. There are four classes, each defined in terms of the length of the initial steep segment of the distribution.
2. Conditional on the class, three points are selected to calculate parameters $(\alpha_{n1},\beta_{n1},\gamma_{n1})$ of the first Weibull function. The equations are given by Kelly and Krzysztofowicz (2000, Appendix).
3. Conditional on the class, the meeting point ζ_n is determined.
4. Parameters $(\alpha_{n2},\beta_{n2},\gamma_{n2})$ of the second Weibull function are determined as follows. The shift parameter is

$$\gamma_{n2} = s_{n0}; \tag{12a}$$

the shape parameter β_{n2} is calculated from Eq. (9c):

$$\beta_{n2} = \beta_{n1}\left(\frac{\zeta_n-\gamma_{n2}}{\zeta_n-\gamma_{n1}}\right); \tag{12b}$$

the scale parameter α_{n2} is calculated from Eq. (9b):

$$\alpha_{n2} = (\zeta_n-\gamma_{n2})\left(\frac{\alpha_{n1}}{\zeta_n-\gamma_{n1}}\right)^{\beta_{n1}/\beta_{n2}}. \tag{12c}$$

In summary, the parametric model for the conditional output distribution Π_{n1} takes the form of a two-piece Weibull function, whose five independent parameters are estimated based on the output from seven runs of the hydrologic model at the forecast time.

Example. For each lead time n, Table 3 reports the estimates of the seven parameters, while Fig. 4 shows the seven points $\{(s_{np},p)\}$ and the fitted two-piece Weibull conditional output distribution Π_{n1}.

4.6. Update of parameters

The analytic-numerical PUP obviates the need for rerunning the hydrologic model when the forecast of the total precipitation amount is updated between the scheduled forecast times. To wit, suppose the deterministic input subvector $\mathbf{u}_0$ and the matrix of expected disaggregation factors $\boldsymbol{\xi}$ from the last forecast time remain current, while the conditional distribution H_1 of the total precipitation amount W is updated to H_1'. Then the updated conditional output

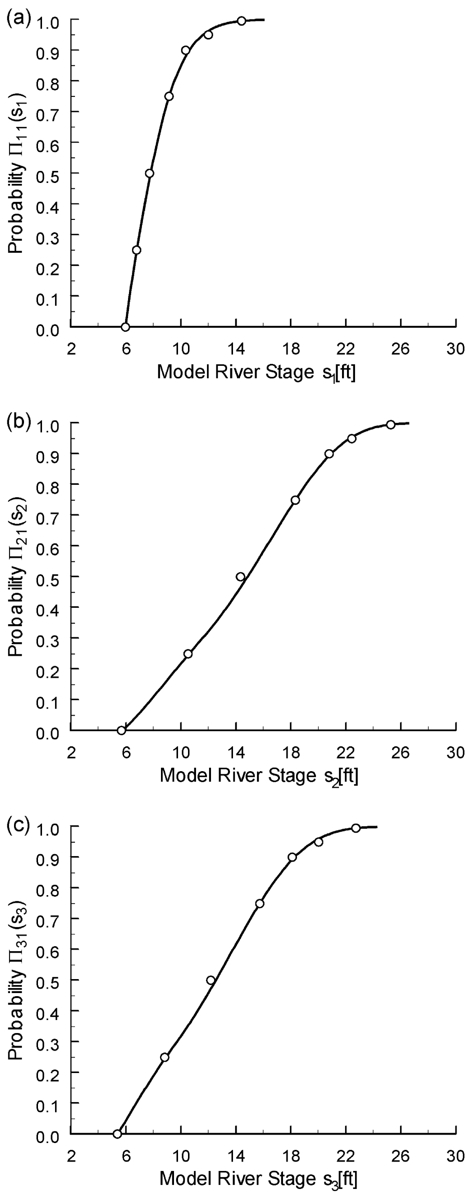

Fig. 4. Output distribution Π_{n1} of model river stage S_n at 1200 UTC on day n, conditional on the hypothesis that precipitation occurs, $V=1$, obtained from the precipitation uncertainty processor (PUP) for 3 days, $n=1,2,3$. A two-piece Weibull model for Π_{n1} is fitted to empirical points.

distribution Π'_{n1} of model river stage S_n is specified by

$$\Pi'_{n1}(s_n) = H'_1(H_1^{-1}(\Pi_{n1}(s_n))), \quad s_n > s_{n0}, \tag{13}$$

and $\Pi'_{n1}(s_n) = 0$ if $s_n = s_{n0}$.

When H_1 is a Weibull distribution with parameters (α, β), H'_1 is a Weibull distribution with parameters (α', β'), and Π_{n1} is a two-piece Weibull distribution with parameters $(\alpha_{n1}, \beta_{n1}, \gamma_{n1}; \alpha_{n2}, \beta_{n2}, \gamma_{n2}; \zeta_n)$, then Π'_{n1} is also a two-piece Weibull distribution with parameters $(\alpha'_{n1}, \beta'_{n1}, \gamma_{n1}; \alpha'_{n2}, \beta'_{n2}, \gamma_{n2}; \zeta_n)$, where

$$\beta'_{n1} = \beta_{n1}\frac{\beta'}{\beta}, \quad \alpha'_{n1} = \alpha_{n1}\left(\frac{\alpha'}{\alpha}\right)^{\beta/\beta_{n1}}, \tag{14a}$$

$$\beta'_{n2} = \beta_{n2}\frac{\beta'}{\beta}, \quad \alpha'_{n2} = \alpha_{n2}\left(\frac{\alpha'}{\alpha}\right)^{\beta/\beta_{n2}}. \tag{14b}$$

Thus the shift parameters $(\gamma_{n1}, \gamma_{n2})$ and the abscissa ζ_n of the meeting point are invariant, whereas the scale parameters $(\alpha_{n1}, \alpha_{n2})$ and the shape parameters (β_{n1}, β_{n2}) are updated via rescaling.

In summary, an operational advantage of this analytic-numerical PUP is that whenever the conditional distribution H_1 of the total precipitation amount is updated, while the deterministic inputs $\mathbf{u}_0$ and $\boldsymbol{\xi}$ remain current, the conditional output distribution Π_{n1} can be updated rapidly by a simple rescaling of its parameters.

5. Hydrologic uncertainty processor

The purpose of the HUP is to quantify hydrologic uncertainty under the hypothesis that there is no precipitation uncertainty. This quantification is based on historical and simulated data samples. It is carried out before any forecast must be prepared, and the results are stored in a form ready for use in real-time forecasting. The HUP presented herein was developed by Krzysztofowicz and Herr (2001).

5.1. Family of posterior distributions

To formulate the HUP, one additional variable need be defined as follows:

h_n—actual river stage at the outlet of the basin at time t_n; when treated as a random variable, it is denoted H_n; at the forecast time, h_0 is the observed river stage, whereas H_n is uncertain for any lead time $n \in \{1, \ldots, N\}$.

The HUP is formulated based on principles of a Bayesian processor and the following three postulates. (i) If there were no hydrologic uncertainty, then one would observe $h_n = s_n$ for $n = 1, \ldots, N$. The presence of hydrologic uncertainty gives rise to a probability distribution of the actual river stage H_n, conditional on a realization of the model river stage $S_n = s_n$ and the observed river stage $H_0 = h_0$. (ii) Hydrologic uncertainty is a nonstationary function of lead time and depends on the precipitation event $V = v$. (iii) The actual river stage process $\{H_n : n = 0, 1, \ldots, N\}$, conditional on the hypothesized precipitation event $V = v$, is a Markov process of order one with nonstationary transition distributions.

For every lead time n $(n = 1, \ldots, N)$ and each hypothesized precipitation event $V = v$ $(v = 0, 1)$, define a family of posterior distributions of actual river stage H_n,

$$\{\Phi_{nv}(\cdot|s_n, h_0) : \text{ all } s_n, h_0\}, \tag{15a}$$

where

$$\Phi_{nv}(h_n|s_n, h_0) = P(H_n \leq h_n|S_n = s_n, H_0 = h_0, V = v). \tag{15b}$$

That is, $\Phi_{nv}(\cdot|s_n, h_0)$ is the distribution of H_n, conditional on the hypotheses that the precipitation event is $V = v$ and the model river stage induced by a perfect forecast of the total precipitation amount is $S_n = s_n$, and given that the river stage observed at the forecast time is $H_0 = h_0$. Each distribution $\Phi_{nv}(\cdot|s_n, h_0)$ has a corresponding density $\phi_{nv}(\cdot|s_n, h_0)$. The family (Eqs. (15a) and (15b)) of the posterior distributions quantifies hydrologic uncertainty for a given hypothesis v and lead time n. The task of the HUP is to supply Eqs. (15a) and (15b).

5.2. Meta-Gaussian posterior distributions

The parametric model for Φ_{nv} takes the form of a family of meta-Gaussian distributions. It is built up of marginal distributions and dependence parameters.

5.2.1. Marginal distributions

There are no constraints on the marginal distributions: they can be of any form, parametric or nonparametric. They are defined as follows:

Γ_{nv}—marginal prior distribution of actual river stage H_n, conditional on the hypothesis that the precipitation event is $V = v$; it is defined for $v = 0, 1$ and $n = 0, 1, \ldots, N$ such that

$$\Gamma_{nv}(h_n) = P(H_n \leq h_n|V = v). \tag{16}$$

$\bar{\Lambda}_{nv}$—marginal initial distribution of model river stage S_n, conditional on the hypothesis that the precipitation event is $V = v$; it is defined for $v = 0, 1$ and $n = 1, \ldots, N$ such that

$$\bar{\Lambda}_{nv}(s_n) = P(S_n \leq s_n|V = v). \tag{17}$$

If the distributions are parametric, then preferably (but not necessarily) $\bar{\Lambda}_{nv}$ should be of the same type as Γ_{nv} is.

5.2.2. Dependence parameters

The stochastic dependence structure between variates H_n, S_n, and H_0 under the family of posterior distributions Φ_{nv} is characterized parametrically. The dependence parameters are defined in three steps.

1. *Normal quantile transforms* (NQTs). Each original variate is transformed into a normal variate by the composition of the inverse Q^{-1} of the standard normal distribution and the marginal distribution of the original variate:

$$W_n = Q^{-1}(\Gamma_{nv}(H_n)), \quad n = 0, 1, \ldots, N, \tag{18a}$$

$$X_n = Q^{-1}(\bar{\Lambda}_{nv}(S_n)), \quad n = 1, \ldots, N. \tag{18b}$$

2. *Linear regressions*. In the space of the transformed variates, the stochastic dependence structures are characterized parametrically in terms of two linear models. The parameter c_{nv} $(v = 0, 1;\ n = 1, \ldots, N)$ of the prior transition distribution is defined by the following linear regression:

$$E(W_n|W_{n-1} = w_{n-1}, V = v) = c_{nv}w_{n-1}, \tag{19a}$$

$$\mathrm{Var}(W_n|W_{n-1} = w_{n-1}, V = v) = 1 - c_{nv}^2. \tag{19b}$$

The parameters a_{nv}, b_{nv}, d_{nv}, and σ_{nv} $(v = 0, 1;\ n = 1, \ldots, N)$ of the likelihood function are defined by the following linear regression:

$$E(X_n|W_n = w_n, W_0 = w_0, V = v) = a_{nv}w_n + d_{nv}w_0 + b_{nv}, \tag{20a}$$

$$\mathrm{Var}(X_n|W_n = w_n, W_0 = w_0, V = v) = \sigma_{nv}^2. \quad (20b)$$

3. *Dependence parameters*. The parameters A_{nv}, B_{nv}, D_{nv}, and T_{nv} ($v = 0, 1;\ n = 1, \ldots, N$) of the posterior distribution are calculated as follows:

$$A_{nv} = \frac{a_{nv} t_{nv}^2}{a_{nv}^2 t_{nv}^2 + \sigma_{nv}^2}, \quad (21a)$$

$$B_{nv} = \frac{-a_{nv} b_{nv} t_{nv}^2}{a_{nv}^2 t_{nv}^2 + \sigma_{nv}^2}, \quad (21b)$$

$$D_{nv} = \frac{C_{nv}\sigma_{nv}^2 - a_{nv} d_{nv} t_{nv}^2}{a_{nv}^2 t_{nv}^2 + \sigma_{nv}^2}, \quad (21c)$$

$$T_{nv}^2 = \frac{t_{nv}^2 \sigma_{nv}^2}{a_{nv}^2 t_{nv}^2 + \sigma_{nv}^2}, \quad (21d)$$

where

$$C_{nv} = \prod_{i=1}^{n} c_{iv}, \quad t_{nv}^2 = 1 - C_{nv}^2. \quad (22)$$

5.2.3. Posterior distribution and density

Given the marginal distributions (Eqs. (16) and (17)) and the dependence parameters (Eqs. (21a)–(21d)), the posterior distribution (Eq. (15b)) takes the form of a meta-Gaussian distribution:

$$\Phi_{nv}(h_n|s_n, h_o) = Q\left(\frac{Q^{-1}(\Gamma_{nv}(h_n)) - A_{nv}Q^{-1}(\bar{\Lambda}_{nv}(s_n)) - D_{nv}Q^{-1}(\Gamma_{0v}(h_0)) - B_{nv}}{T_{nv}}\right), \quad (23)$$

where Q is the standard normal distribution having inverse Q^{-1}. The corresponding meta-Gaussian posterior density takes the form

$$\phi_{nv}(h_n|s_n, h_0) = \frac{\gamma_{nv}(h_n) q(Q^{-1}(\Phi_{nv}(h_n|s_n, h_0)))}{T_{nv} q(Q^{-1}(\Gamma_{nv}(h_n)))}, \quad (24)$$

where γ_{nv} is the density corresponding to the distribution Γ_{nv}, and q is the standard normal density. Because q has closed form and polynomial approximations to Q and Q^{-1} are available (Abramowitz and Stegun, 1972), expressions (23) and (24) are simple to evaluate.

5.3. Estimation

5.3.1. Estimation framework

The task is to estimate the HUP parameters—the marginal distributions (Eqs. (16) and (17)) and the linear regressions (Eqs. (19a)–(20b)). The estimation is carried out before real-time forecasting begins. The framework for estimation rests on three assumptions.

1. The PRSFs are prepared on schedule, once or more times per day. Each scheduled forecast time defines a different 24-h forecast cycle. For example, there may be two forecast cycles, one with forecast time $t_0 = 0000$ UTC on each day, and another with forecast time $t_0 = 1200$ UTC on each day. Different parameter values may be necessary for each forecast cycle in order to capture the effects of the diurnal cycle of precipitation (Krzysztofowicz and Pomroy, 1997). This may be especially necessary for a small basin and a convective precipitation season.
2. For a given forecast cycle, each day marks the beginning of a separate realization $(v; s_1, \ldots, s_N; h_0, h_1, \ldots, h_N)$ of the precipitation event and the model–actual river stage process.
3. The calendar year is divided into hydrologic seasons such that for a given forecast cycle, the parameters of the HUP remain invariant for all days within a season.

It follows that for each forecast cycle and hydrologic season, the HUP has a different set of parameter values. The estimation procedure, detailed by Krzysztofowicz and Herr (2001), is presented herein only in concept, which involves three main tasks: specification of the candidate parametric models for marginal distributions, estimation of the prior parameters, and estimation of the likelihood parameters.

5.3.2. Specification of parametric models

To ensure fast calculation in real time, each marginal distribution defined in Section 5.2.1 is

Table 4
Marginal prior distributions Γ_{nv} of actual river stages and prior correlation coefficients c_{nv} at forecast time 1200 UTC in November

Precip. indicator (v)	Lead time (n)	Distribution		Parameters			Corr. coeff. (c_{nv})
		Symbol (Γ_{nv})	Type (l_{nv})[a]	Scale (α_{nv})	Shape (β_{nv})	Shift (γ_{nv})	
1	0	Γ_{01}	LW	1.41	2.58	3.45	
	1	Γ_{11}	LW	1.59	3.02	3.45	0.702
	2	Γ_{21}	LW	1.66	3.40	3.45	0.810
	3	Γ_{31}	LW	1.63	3.80	3.45	0.789
0	0	Γ_{00}	LL	3.01	2.93	3.45	
	1	Γ_{10}	LL	2.66	3.02	3.45	0.948
	2	Γ_{20}	LL	2.50	3.23	3.45	0.797
	3	Γ_{30}	LL	2.53	2.98	3.45	0.813

[a] LW: log-Weibull, LL: log-logistic, stage in feet.

represented by a parametric model having closed form expressions for the cumulative distribution function and the probability density function. Three such candidate models are Weibull (WB), log-Weibull (LW), and log-logistic (LL). Each model has three parameters: α–scale, β–shape, and γ–shift. Hence, the basic task is to estimate the parameters of each of the candidate models and next to choose the best model. In effect, each marginal distribution is specified by four parameters: $(\alpha, \beta, \gamma; l)$, where l–distribution type, $l \in \{\text{WB}, \text{LW}, \text{LL}\}$. Usually, the best of these three closed form models offers as good a fit to the marginal distribution of a river stage as does the best of the three popular models: gamma, log-Pearson, and log-normal.

5.3.3. Estimation of prior parameters

Prior parameters are those of the marginal prior distributions (Eq. (16)) and the linear regressions (Eqs. (19a) and (19b)). For a given forecast cycle, the estimation procedure consists of four steps.

1. *Prior sample*. From the historical data record, a joint sample $\{(v; h_0, h_1, \ldots, h_N)\}$ is formed, with one realization per day.
2. *Prior seasons*. Seasons are determined during which the family of the prior transition distributions is assumed invariant. Next, the joint sample is partitioned into seasons.
3. *Marginal distribution parameters*. For each prior season, for every $v \in \{0, 1\}$ and every $n \in \{0, 1, \ldots, N\}$, parameters $(\alpha_{nv}, \beta_{nv}, \gamma_{nv}; l_{nv})$ of the marginal prior distribution (Eq. (16)) are estimated.
4. *Prior dependence parameters*. For each prior season, for every $v \in \{0, 1\}$ and every $n \in \{1, \ldots, N\}$, parameter c_{nv} of the linear regression (Eqs. (19a) and (19b)) is estimated (parameter c_{nv} can be interpreted as the Pearson's product-moment correlation coefficient between W_n and W_{n-1}).

Example. The year was divided into monthly prior seasons. Table 4 lists estimates of all parameters for November. For every n, the two conditional distributions are of different type: Γ_{n0} is log-logistic whereas Γ_{n1} is log-Weibull. For every v, all parameters are nonstationary with n.

5.3.4. Estimation of likelihood parameters

Likelihood parameters are those of the marginal initial distributions (Eq. (17)) and the linear regressions (Eqs. (20a) and (20b)). For a given forecast cycle, the estimation procedure consists of four steps.

1. *Likelihood sample*. The output from a simulation experiment is matched with the historical data record to form a joint sample $\{(v; s_1, \ldots, s_N; h_0, h_1, \ldots, h_N)\}$, with one realization per day. (This sample may be shorter than the prior sample.)
2. *Likelihood seasons*. Seasons are determined during which the family of the likelihood functions is assumed invariant. Next, the joint sample is partitioned into seasons. (Each likelihood season may overlap more than one prior season. When this is the case, the term hydrologic seasons refers to the prior seasons.)

Table 5
Marginal initial distributions $\bar{\Lambda}_{nv}$ of model river stages at forecast time 1200 UTC in November

Precip. indicator (v)	Lead time (n)	Distribution		Parameters		
		Symbol ($\bar{\Lambda}_{nv}$)	Type ($\bar{l}_{nv}$)[a]	Scale ($\bar{\alpha}_{nv}$)	Shape ($\bar{\beta}_{nv}$)	Shift ($\bar{\gamma}_{nv}$)
1	1	$\bar{\Lambda}_{11}$	LW	1.63	3.06	3.32
	2	$\bar{\Lambda}_{21}$	LW	1.96	4.67	2.00
	3	$\bar{\Lambda}_{31}$	LW	1.83	4.63	2.26
0	1	$\bar{\Lambda}_{10}$	LL	2.72	3.41	3.24
	2	$\bar{\Lambda}_{20}$	LL	2.50	3.70	3.14
	3	$\bar{\Lambda}_{30}$	LL	1.79	2.63	3.64

[a] LW: log-Weibull, LL: log-logistic, stage in feet.

3. *Marginal distribution parameters*. For each prior season, for every $v \in \{0, 1\}$ and every $n \in \{1, \ldots, N\}$, parameters $(\bar{\alpha}_{nv}, \bar{\beta}_{nv}, \bar{\gamma}_{nv}; \bar{l}_{nv})$ of the marginal initial distribution (Eq. (17)) are estimated.
4. *Likelihood dependence parameters*. For each likelihood season, for every $v \in \{0, 1\}$ and every $n \in \{1, \ldots, N\}$, parameters $(a_{nv}, b_{nv}, d_{nv}, \sigma_{nv})$ of the linear regression (Eqs. (20a) and (20b)) are estimated.

Example. The year was divided into two likelihood seasons: warm season (June–October) and cool season (November–May). Table 5 lists estimates of the parameters of the marginal initial distributions for November (a prior season). Table 6 lists estimates of the parameters of the linear regressions for the cool season (a likelihood season). Finally, Table 7 lists values of the parameters of the meta-Gaussian posterior distributions for November; these parameter values are calculated via Eqs. (21a)–(22) from the values of c_{nv} for November listed in Table 4 and the values of $(a_{nv}, b_{nv}, d_{nv}, \sigma_{nv})$ for the cool season (which overlaps November) listed in Table 6. For every v, all parameters are nonstationary with n.

Table 6
Likelihood parameters at forecast time 1200 UTC in cool season

Precip. indicator (v)	Lead time (n)	Parameters			
		a_{nv}	b_{nv}	d_{nv}	σ_{nv}
1	1	0.95	0.00	0.00	0.323
	2	0.84	0.00	0.00	0.545
	3	0.47	0.00	0.37	0.669
0	1	1.00	0.00	0.00	0.065
	2	0.27	0.00	0.74	0.188
	3	0.16	0.00	0.86	0.189

5.4. Parameters for forecasting

To recapitulate, the precipitation-dependent meta-Gaussian HUP requires the estimation of marginal distributions and dependence parameters for each forecast cycle and hydrologic season. When the marginal distributions are parametric and forecasts are for N steps ahead, the HUP for a given forecast cycle and hydrologic season is specified by $24N + 8$ parameters:

(i) parameters of the marginal prior distributions Γ_{nv} of actual river stages,

$$\{(\alpha_{nv}, \beta_{nv}, \gamma_{nv}; l_{nv}) : v = 0, 1; \;\; n = 0, 1, \ldots, N\},$$

(ii) parameters of the marginal initial distributions $\bar{\Lambda}_{nv}$ of model river stages,

$$\{(\bar{\alpha}_{nv}, \bar{\beta}_{nv}, \bar{\gamma}_{nv}, \bar{l}_{nv}) : v = 0, 1; \; n = 1, \ldots, N\},$$

Table 7
Dependence parameters of the meta-Gaussian families of posterior distributions Φ_{nv} of actual river stages at forecast time 1200 UTC in November

Precip. indicator (v)	Lead time (n)	Parameters			
		A_{nv}	B_{nv}	D_{nv}	T_{nv}
1	1	0.857	0.000	0.130	0.307
	2	0.734	0.000	0.218	0.509
	3	0.602	0.000	0.099	0.757
0	1	0.960	0.000	0.038	0.064
	2	1.739	0.000	−0.886	0.477
	3	1.928	0.000	−1.234	0.656

(iii) dependence parameters of the posterior distributions,

$$\{(A_{n\nu}, B_{n\nu}, D_{n\nu}, T_{n\nu}) : \nu = 0, 1;\;\; n = 1, \ldots, N\}.$$

Once acquired, the values of these parameters are stored for future use in real-time forecasting. At the forecast time, the parameter values for the current forecast cycle and hydrologic season are sent to the integrator of the BFS.

6. Integrator

The purpose of the INT is to integrate precipitation uncertainty with hydrologic uncertainty according to the scheme prescribed by the Bayesian theory. The integration takes place at the forecast time and the result is the PRSF. The INT presented herein was developed by Krzysztofowicz (2001b).

6.1. Predictive distribution

For every lead time n $(n = 1, \ldots, N)$ define the conditional predictive distribution. This is the distribution of actual river stage H_n at time t_n, conditional on river stage $H_0 = h_0$ observed at the forecast time t_0 :

$$\Psi_n(h_n) = P(H_n \le h_n | H_0 = h_0). \tag{25}$$

This predictive distribution is a mixture of two distributions:

$$\Psi_n(h_n) = \frac{\gamma_{00}(h_0)(1-\nu)}{\gamma_0(h_0)} \Phi_{n0}(h_n | s_{n0}, h_0) + \frac{\gamma_{01}(h_0)\nu}{\gamma_0(h_0)} \int_{s_{n0}}^{\infty} \Phi_{n1}(h_n | s_n, h_0)\pi_{n1}(s_n)\mathrm{d}s_n, \tag{26}$$

where γ_0 is the marginal density of initial river stage H_0 :

$$\gamma_0(h_0) = \gamma_{00}(h_0)(1-\nu) + \gamma_{01}(h_0)\nu. \tag{27}$$

All the elements of Eq. (26) are already known. The PQPF supplies the probability of precipitation occurrence ν, which is defined by Eq. (1). The PUP supplies Eqs. (11a)–(11c) and parameter values for the conditional output density π_{n1}, including the model river stage s_{n0} resulting from zero total precipitation amount. The HUP supplies Eq. (23) and parameter values for the families of the posterior distributions Φ_{n0} and Φ_{n1}, including the expressions and parameter values for the marginal prior densities γ_{00} and γ_{01}. Finally, the river gauge at the forecast point supplies h_0.

The predictive distribution Ψ_n quantifies the total uncertainty about actual river stage H_n at time t_n, given all information (the PQPF, the deterministic input subvector $\mathbf{u}_0$, and the hydrologic model) utilized at the forecast time t_0. As such, Ψ_n constitutes the PRSF with lead time $t_n - t_0$. The task of the INT is to output Ψ_n for $n = 1, \ldots, N$.

6.2. Algorithm for predictive distribution

In the numerical algorithm, Eq. (23) for $\Phi_{n\nu}$ is used as an external function, but Eqs. (10)–(11c) for π_{n1} are inserted into Eq. (26). After some rearrangement of Eqs. (26) and (27), the following numerical algorithm for calculating the value of the predictive distribution $\Psi_n(h_n)$ for any river stage h_n is obtained.

1. Calculate two constants, the posterior probability of precipitation occurrence

$$\mu = \frac{\gamma_{01}(h_0)\nu}{\gamma_{00}(h_0)(1-\nu) + \gamma_{01}(h_0)\nu}, \tag{28}$$

and the integration limit

$$u_n = \left(\frac{\zeta_n - \gamma_{n1}}{\alpha_{n1}}\right)^{\beta_{n1}}. \tag{29}$$

2. For the given h_n, evaluate numerically two integrals,

$$I_{n1}(h_n) = \int_{u_n}^{\infty} \Phi_{n1}(h_n | \alpha_{n1}u^{1/\beta_{n1}} + \gamma_{n1}, h_0)e^{-u}\,\mathrm{d}u, \tag{30a}$$

$$I_{n2}(h_n) = \int_{0}^{u_n} \Phi_{n1}(h_n | \alpha_{n2}u^{1/\beta_{n2}} + \gamma_{n2}, h_0)e^{-u}\,\mathrm{d}u, \tag{30b}$$

and calculate the sum

$$I_n(h_n) = I_{n1}(h_n) + I_{n2}(h_n). \tag{31}$$

3. For the given h_n, calculate the predictive probability

$$\Psi_n(h_n) = (1-\mu)\Phi_{n0}(h_n|s_{n0},h_0) + \mu I_n(h_n). \quad (32)$$

6.3. Algorithm for predictive density

Let ψ_n denote the predictive density corresponding to the predictive distribution Ψ_n. The numerical algorithm for calculating the value of the predictive density $\psi_n(h_n)$ for any river stage h_n is obtained by modifying the above algorithm as follows. Eqs. (28) and (29) remain intact. The posterior distribution functions Φ_{n1} on the right sides of Eqs. (30a) and (30b) are replaced by the posterior density functions ϕ_{n1}, which are given by Eq. (24); the resultant left sides are denoted $i_{n1}(h_n)$ and $i_{n2}(h_n)$, respectively; and Eq. (31) is replaced by $i_n(h_n) = i_{n1}(h_n) + i_{n2}(h_n)$. Finally, Eq. (32) is replaced by

$$\psi_n(h_n) = (1-\mu)\phi_{n0}(h_n|s_{n0},h_0) + \mu i_n(h_n). \quad (33)$$

6.4. Bounds and discretization

Because each marginal distribution has support bounded below by a shift parameter γ, the following constraints must be satisfied for effective numerical calculations: $h_n > \max\{(\gamma_{n0} \text{ of } \Gamma_{n0}), (\gamma_{n1} \text{ of } \Gamma_{n1})\}$, and $s_{n0} > \max\{(\bar{\gamma}_{n0} \text{ of } \bar{\Lambda}_{n0}), (\bar{\gamma}_{n1} \text{ of } \bar{\Lambda}_{n1})\}$. An algorithm for operational forecasting discretizes h_n so that the numerical representation of Ψ_n consists of more than 100 values covering the probability range (0.01, 0.99) in steps of the river stage smaller than 0.5 foot. Some distributions and densities reported herein cover the probability range (0.0001, 0.9999) in order to better depict the tails.

Example. With ν and h_0 given in Table 1, the parameter values for π_{n1} given in Table 3, and the parameter values for $\Phi_{n\nu}$ (and $\phi_{n\nu}$) given in Tables 4, 5, and 7, the analytic-numerical INT outputs the numerical representations of Ψ_n (and ψ_n). Fig. 5 displays the predictive distributions Ψ_n for $n = 1, 2, 3$; the lead times are 24, 48, and 72 h, respectively. Fig. 6 displays the corresponding predictive densities ψ_n for $n = 1, 2, 3$.

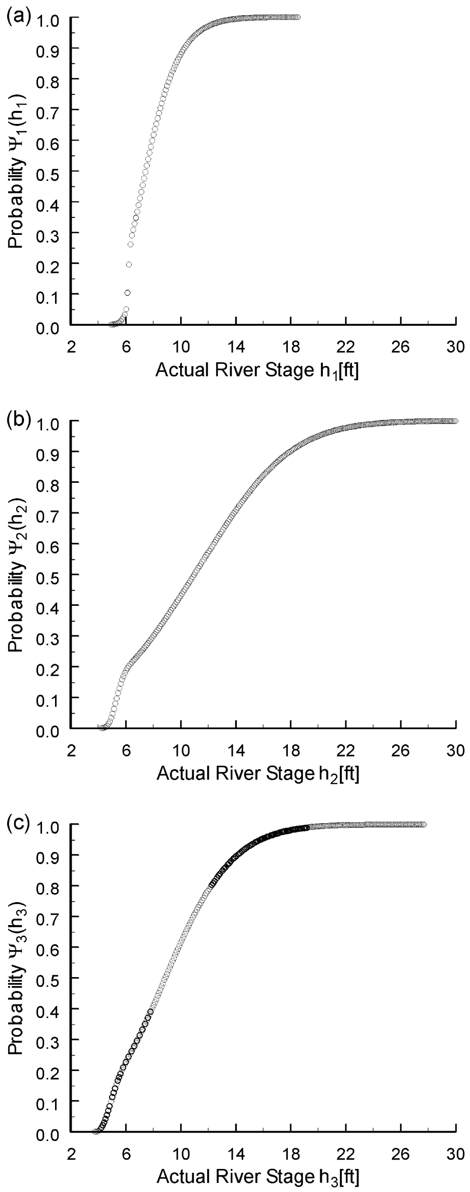

Fig. 5. Predictive distribution Ψ_n of actual river stage H_n at 1200 UTC on day n, calculated numerically from the Bayesian integrator for three days, $n = 1, 2, 3$; the corresponding lead times are 24, 48, and 72 h.

7. Probabilistic river stage forecast

This section highlights three aspects of the BFS: theoretical and empirical properties of the PRSF, algorithms for updating the PRSF based on an updated

PQPF, and displays of the PRSF that convey the evolution of the total uncertainty in time.

7.1. Forecast properties

The theoretical properties of the predictive distribution Ψ_n, an analysis of the behavior of Ψ_n, and an explanation of the shapes of Ψ_n can be found in Krzysztofowicz (2001b). The following synopsis of four properties provides a background for the subsequent sections.

The predictive distribution Ψ_n quantifies the total uncertainty about actual river stage H_n at time t_n. The total uncertainty is the union of precipitation uncertainty and hydrologic uncertainty, each of which is quantified based on information utilized at the forecast time t_0.

The predictive distribution Ψ_n is a mixture of two distributions. As Eq. (26) shows, these are (i) the posterior distribution of actual river stage H_n, conditional on the hypothesis that precipitation does not occur, given model river stage $S_n = s_{n0}$ for time t_n, and given observed river stage $H_0 = h_0$ at the forecast time t_0, and (ii) the expected posterior distribution of actual river stage H_n, conditional on the hypothesis that precipitation occurs, and given observed river stage $H_0 = h_0$ at the forecast time t_0. The expectation, calculated as the integral over all possible realizations of S_n, accounts for uncertainty about the model river stage that results from uncertainty about the total precipitation amount. The weights of the mixture are determined by the posterior probability of precipitation occurrence, given observed river stage $H_0 = h_0$. In other words, it is the uncertainty associated with the intermittence of the precipitation process that determines the basic structure of the predictive distribution of the river stage.

The predictive density ψ_n is bimodal, as illustrated in Fig. 6. This is a theoretical property—the implication of the structure of ψ_n, which is a mixture whenever the occurrence of precipitation is uncertain ($0 < \nu < 1$).

The predictive density ψ_n is asymmetric, as illustrated in Fig. 6. This is an empirical property—the outcome of several interacting factors: the asymmetry of the conditional density of the total precipitation amount, the nonlinearity of the hydrologic model, the asymmetry of the prior density of the river stage, and the asymmetry of the likelihood functions characterizing the hydrologic uncertainty.

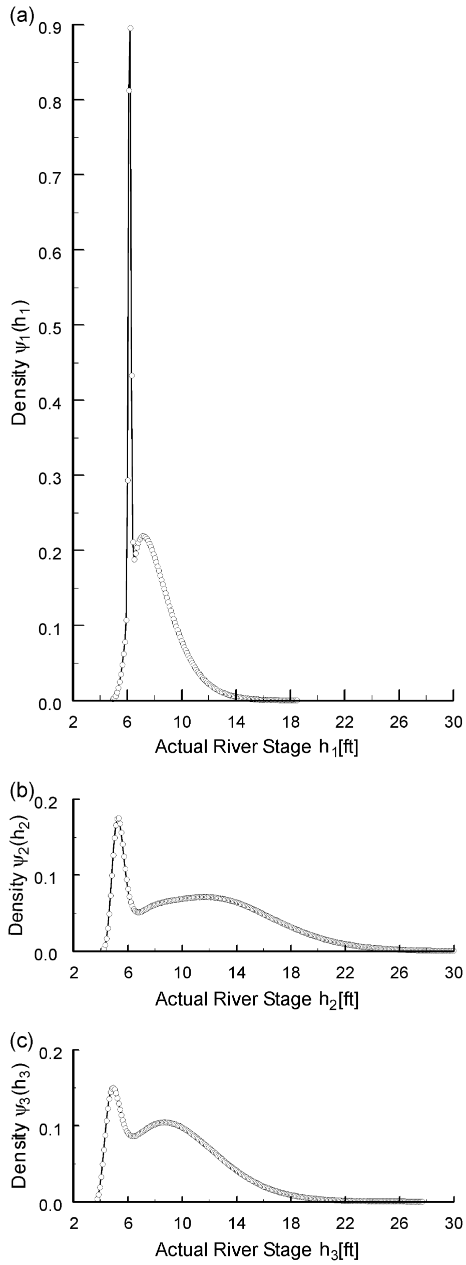

Fig. 6. Predictive density ψ_n of actual river stage H_n at 1200 UTC on day n, calculated numerically from the Bayesian integrator for three days, $n = 1, 2, 3$; the corresponding lead times are 24, 48, and 72 h.

7.2. Forecast updating

An operational advantage of the analytic-numerical BFS is the ease of updating the PRSF whenever the probabilistic part of the PQPF is updated between the scheduled forecast times. Specifically, let $(\nu, H_1; \boldsymbol{\xi})$ denote the PQPF used to produce the PRSF at the last forecast time. Suppose that thereafter the probability of precipitation occurrence ν is updated to ν', and/or the conditional distribution H_1 of the total precipitation amount is updated to H_1', while the matrix of expected disaggregation factors $\boldsymbol{\xi}$ remains current; thus the updated PQPF is $(\nu', H_1'; \boldsymbol{\xi})$. When the deterministic input subvector $\mathbf{u}_0$ from the last forecast time also remains current, an updated PRSF can be produced easily (without the need for rerunning the hydrologic model). We describe two updating algorithms, first using ν' and second using H_1'.

7.2.1. Updated probability of precipitation occurrence

At the forecast time, one should store values $\Phi_{n0}(h_n|s_{n0}, h_0)$ and $I_n(h_n)$ calculated as part of Eq. (32) and values $\phi_{n0}(h_n|s_{n0}, h_0)$ and $i_n(h_n)$ calculated as part of Eq. (33). Then given the updated ν', the updated PRSF can be obtained as follows:

1. Replace ν by ν' in Eq. (28), and calculate the updated μ.
2. Insert into Eq. (32) the updated μ and the stored $\Phi_{n0}(h_n|s_{n0}, h_0)$ and $I_n(h_n)$, and calculate the updated $\Psi_n(h_n)$.
3. Insert into Eq. (33) the updated μ and the stored $\phi_{n0}(h_n|s_{n0}, h_0)$ and $i_n(h_n)$, and calculate the updated $\psi_n(h_n)$.

Example. Fig. 7 shows the updated predictive distributions Ψ_n for $n = 1, 2, 3$ based on six different updates of the probability of precipitation occurrence ν. It is apparent that ν alone exerts a phenomenal influence on the shape of Ψ_n. In general, the shape of Ψ_n varies from cliffy, when $\nu = 0$, to convex–concave when $\nu = 1$. In between, when $0 < \nu < 1$, the shape of Ψ_n includes both a cliffy left tail and a concave right tail. The cliff is located in the vicinity of s_{n0} and has a height proportional to the probability of no precipitation, $1 - \nu$. Fig. 8 shows the updated predictive densities ψ_2 corresponding to the updated predictive distributions Ψ_2 from Fig. 7b.

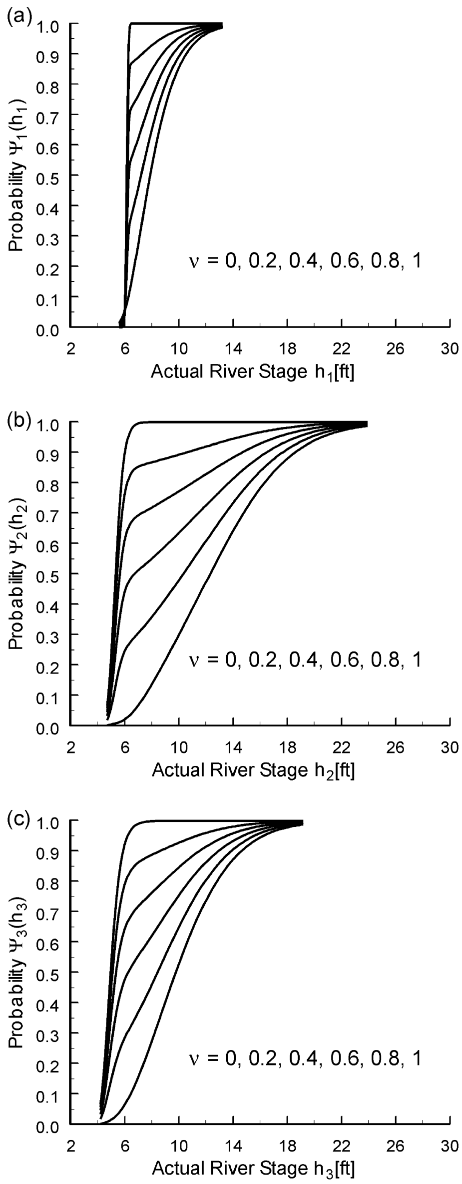

Fig. 7. Updated predictive distribution Ψ_n for day n $(n = 1, 2, 3)$, given the updated probability of precipitation occurrence ν.

7.2.2. Updated conditional distribution of amount

At the forecast time, one should store parameter values $(\alpha_{n1}, \beta_{n1}, \gamma_{n1}; \alpha_{n2}, \beta_{n2}, \gamma_{n2}; \zeta_n)$ of the two-piece Weibull distribution (Eq. (8)), values $\Phi_{n0}(h_n|s_{n0}, h_0)$ calculated as part of Eq. (32), and values

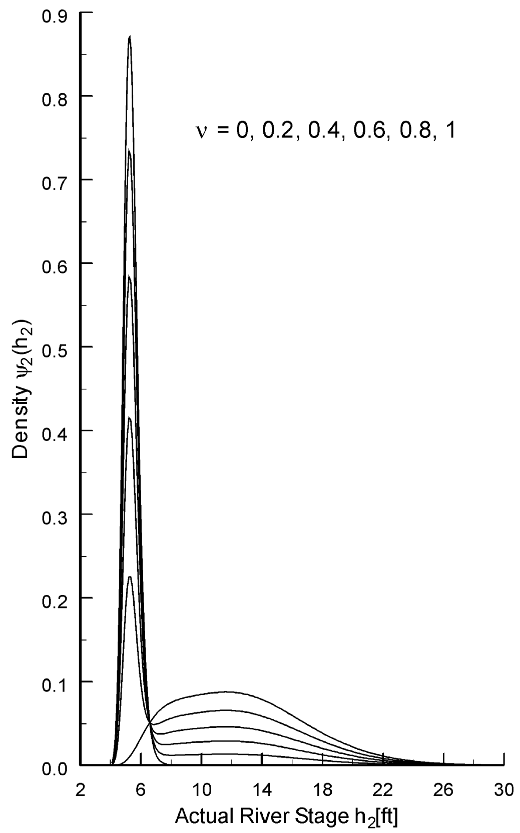

Fig. 8. Updated predictive density ψ_2 for day 2, given the updated probability of precipitation occurrence ν.

$\phi_{n0}(h_n|s_{n0}, h_0)$ calculated as part of Eq. (33). When H_1 and H_1' are arbitrary distributions, an updating algorithm can be devised based on Eq. (13). When H_1 is a Weibull distribution with parameters (α, β), and H_1' is a Weibull distribution with parameters (α', β'), the updated PRSF can be obtained as follows:

1. Via Eqs. (14a) and (14b) calculate the updated parameter values $(\alpha_{n1}', \beta_{n1}', \alpha_{n2}', \beta_{n2}')$ of the two-piece Weibull distribution.
2. Replace $(\alpha_{n1}, \beta_{n1}, \alpha_{n2}, \beta_{n2})$ by $(\alpha_{n1}', \beta_{n1}', \alpha_{n2}', \beta_{n2}')$ in Eqs. (29)–(30b), and repeat all calculations using Eqs. (29)–(31).
3. Insert into Eq. (32) the updated $I_n(h_n)$ and the stored $\Phi_{n0}(h_n|s_{n0}, h_0)$, and calculate the updated $\Psi_n(h_n)$.
4. Insert into Eq. (33) the updated $i_n(h_n)$ and the stored $\phi_{n0}(h_n|s_{n0}, h_0)$, and calculate the updated $\psi_n(h_n)$.

Example. Four different updates of the conditional distribution H_1 of the total precipitation amount are plotted in Fig. 9. Case B is a reference, as the updated distribution H_1' is nearly identical to the distribution H_1 plotted in Fig. 3 and used in all previous examples. Fig. 10 shows the updated predictive distributions Ψ_n for $n = 1, 2, 3$ based on the four updates of H_1. Again, it is apparent that H_1 alone exerts a phenomenal influence on the shape of Ψ_n. In general, the shape of Ψ_n varies from an inverted J, in case A, to an asymmetric double S, in case D. Again, these shapes manifest the basic structure of the total uncertainty about the river stage. The cliffy left tail of Ψ_n reflects mostly hydrologic uncertainty under the hypothesis of precipitation nonoccurrence, whose probability is $1 - \nu = 0.15$, whereas the S-shaped part of Ψ_n reflects uncertainty about the total precipitation amount under the hypothesis of precipitation occurrence, whose probability is $\nu = 0.85$. As the conditional distribution H_1 of the total precipitation amount moves to the right (from case A to case D), the S-shaped part of the predictive distribution Ψ_n of the actual river stage moves to the right as well. In effect, the mixed structure of Ψ_n is accentuated. Fig. 11 shows the updated predictive densities ψ_2 corresponding to the updated predictive distributions Ψ_2 from Fig. 10b. The mixed structure of Ψ_2 is manifested by the bimodality of ψ_2.

7.2.3. Attributes of updating

The algorithms presented above make it obvious that the updating is computationally simple. The examples shown in Figs. 7, 8, 10 and 11 make it clear that the updating is potentially important. It should thus be a part of any operational implementation of the BFS.

From a hydrologic point of view, the updating may be especially useful in forecasting hurricane-induced or convection-induced floods, when the degree of uncertainty about the storm track and basin coverage changes rapidly. Once the elements ν and H_1 of the PQPF are updated, an updated PRSF may be prepared also rapidly.

From an organizational point of view, the updating may be especially useful in decentralized systems

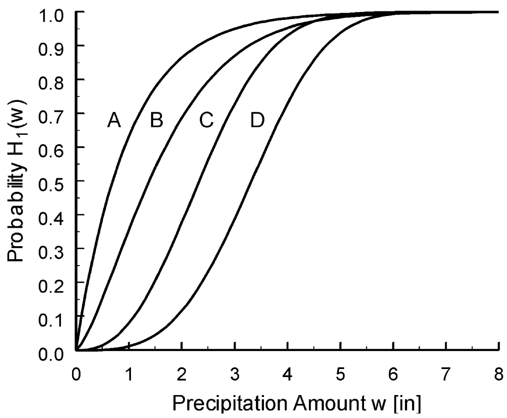

Fig. 9. Conditional distribution H_1 of the basin average precipitation amount W used in the updating. Distribution H_1 is Weibull with parameters (α, β). Case A: (1.0,1.0); case B: (1.8,1.4); case C: (2.7,2.5); case D: (3.7,3.4).

such as that of the NWS (Stallings and Wenzel, 1995). Its hydrologic computing resources are concentrated in 13 River Forecast Centers (RFCs). Each RFC executes the hydrologic models for all basins within its service area and then transmits the results to Weather Forecast Offices (WFOs), each serving a smaller area. The WFO is responsible for disseminating forecasts and issuing flood warnings. During rapidly evolving storms, more timely PRSFs could be produced if the WFO could update the PQPF and the PRSF locally, without having to call upon the RFC to rerun the hydrologic model.

7.3. Forecast time step

One of the purposes of the PRSF is to convey to the decision maker the evolution of the total uncertainty in time. To illustrate this capability, the PRSFs were produced for 3 days in 6-h steps ($N = 12$) using the same PQPF and the same subvector $\mathbf{u}_0$ of deterministic inputs to the hydrologic model as in the previous example. Thus lead times $n = 1, 2, 3$ of the previous example correspond to lead times $n = 4, 8, 12$ of the new example.

Fig. 12 displays the sequence $\{\Psi_n : n = 1, \ldots, 12\}$ of the predictive distributions of the river stages. Fig. 13 displays the sequence $\{\psi_n : n = 1, \ldots, 12\}$ of the predictive densities of the river stages. Fig. 14 displays the isoprobability time series $\{h_{np} : n = 1, \ldots, 12\}$ of quantiles having probability $p = \Psi_n(h_{np})$; there are seven time series corresponding to $p = 0.005, 0.05, 0.25, 0.50, 0.75, 0.95, 0.995$.

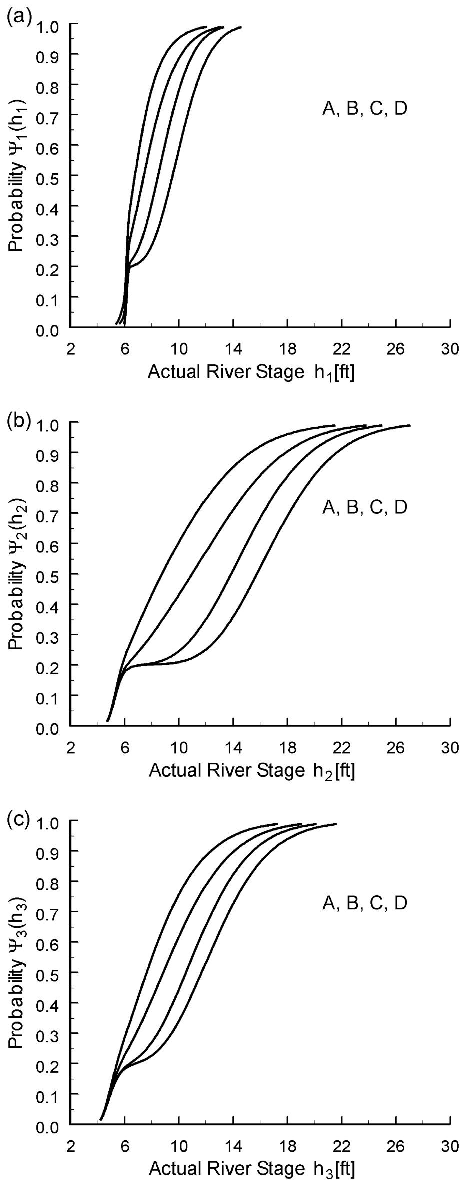

Fig. 10. Updated predictive distribution Ψ_n for day n ($n = 1, 2, 3$), given the updated conditional distribution H_1 of the basin average precipitation amount. Distribution H_1 is Weibull with parameters (α, β). Case A: (1.0,1.0); case B: (1.8,1.4); case C: (2.7,2.5); case D: (3.7,3.4).

The PQPF (Table 1) specifies $\xi_1 = 0$ and $\xi_2 = 0.10$; that is no precipitation is expected in the first 6-h

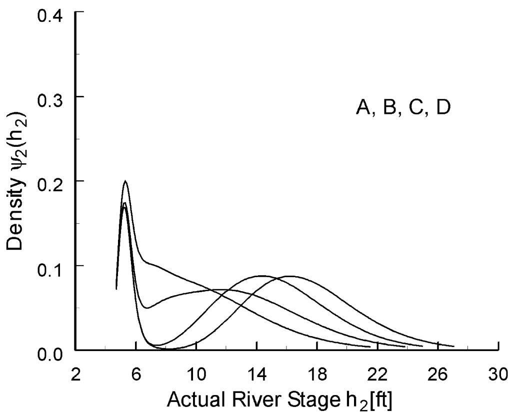

Fig. 11. Updated predictive density ψ_2 for day 2, given the updated conditional distribution H_1 of the basin average precipitation amount. Distribution H_1 is Weibull with parameters (α, β). Case A: (1.0,1.0); case B: (1.8, 1.4); case C: (2.7, 2.5); case D: (3.7, 3.4).

subperiod and only 10% of the 24-h basin average precipitation amount (whatever this amount might be) is expected in the second 6-h subperiod. Consequently, for $n = 1, 2$ the total uncertainty is essentially equal to the hydrologic uncertainty. As n increases from 3 to 8, so does the contribution of the precipitation uncertainty.

The total uncertainty first increases with lead time, from $n = 1$ to about $n = 8$, and then slowly decreases with lead time. Eventually, for very long lead times, each predictive distribution within the 24-h forecast cycle would converge to a prior (climatic) distribution of the river stage for the corresponding time within the forecast cycle. This is one of the unique properties of the BFS (Krzysztofowicz, 2001b).

8. Assumptions and attributes

Like any operational system, this particular analytic-numerical BFS rests on various assumptions (structural and distributional) and offers various attributes (limitations and advantages). The assumptions and attributes reflect the trade-offs made by the system designers. These trade-offs need to be understood by the potential users of the BFS; they may also be instructive to future designers of forecasting systems, which may have to meet different requirements.

8.1. Assumptions

8.1.1. Structural assumptions

The structural assumptions cannot be relaxed within the current formulation of the BFS; therefore, they limit the domain of the applicability.

1. The total precipitation amount (over a basin and during a period) is the dominant source of uncertainty. Based on this assumption, the total precipitation amount is forecasted probabilistically whereas the aggregate of all other uncertainties is quantified statistically. In theory, such a forecasting scheme is suboptimal in comparison with a scheme in which every input uncertainty would be forecasted probabilistically. In application, the suboptimal scheme is much simpler than the optimal scheme and, as tested by Kelly and Krzysztofowicz (2000), may be near optimal. However, the assumption of a single dominant source of uncertainty is not universally valid. Consider, for example, warm rain falling on a deep snow cover. The outcome may be no river rise or a flood (as occurred in the Ohio River basin in January 1997), depending on the rain amount and air temperature in a period following rain. Thus, prior to the event, there may be two dominant sources of uncertainty, precipitation and temperature.

2. The PQPF has a disaggregative structure. The PQPF decomposed into a probabilistic forecast of the total precipitation amount and a deterministic forecast of the spatiotemporal disaggregation can be highly informative but only within a certain space–time domain. It is known that the informativeness of the total precipitation amount forecast increases with the size of the basin area and the length of the time period (Antolik, 2000; Grecu and Krajewski, 2000; Krzysztofowicz et al., 1993), but obviously there must be (as yet unexplored) limits beyond which the informativeness decreases.

3. The HUP has a two-branch, Markovian, nonstationary dependence structure. A possible limitation of this structure is the order of the conditional Markov process, which is assumed to be one. As tested by Krzysztofowicz and Herr (2001), the assumption that, conditional on the precipitation event, the daily river stage process is Markov of order one with nonstationary transition densities appears to be valid. For other time steps and hydrologic regimes, the validity of the assumption

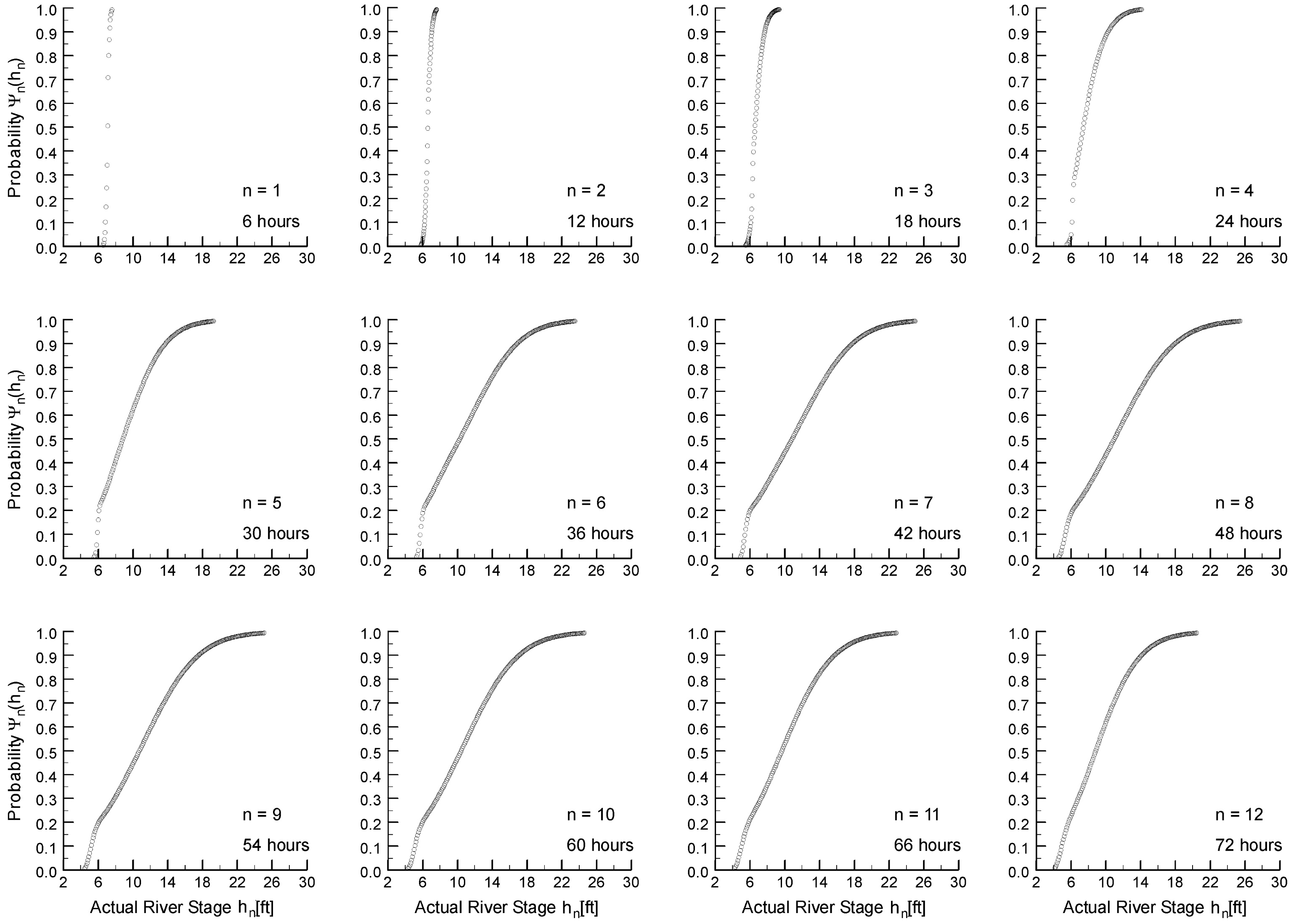

Fig. 12. Sequence of predictive distributions $\{\Psi_n : n = 1, \ldots, 12\}$ of actual river stages at 6-h steps, counting from 1200 UTC on the forecast day.

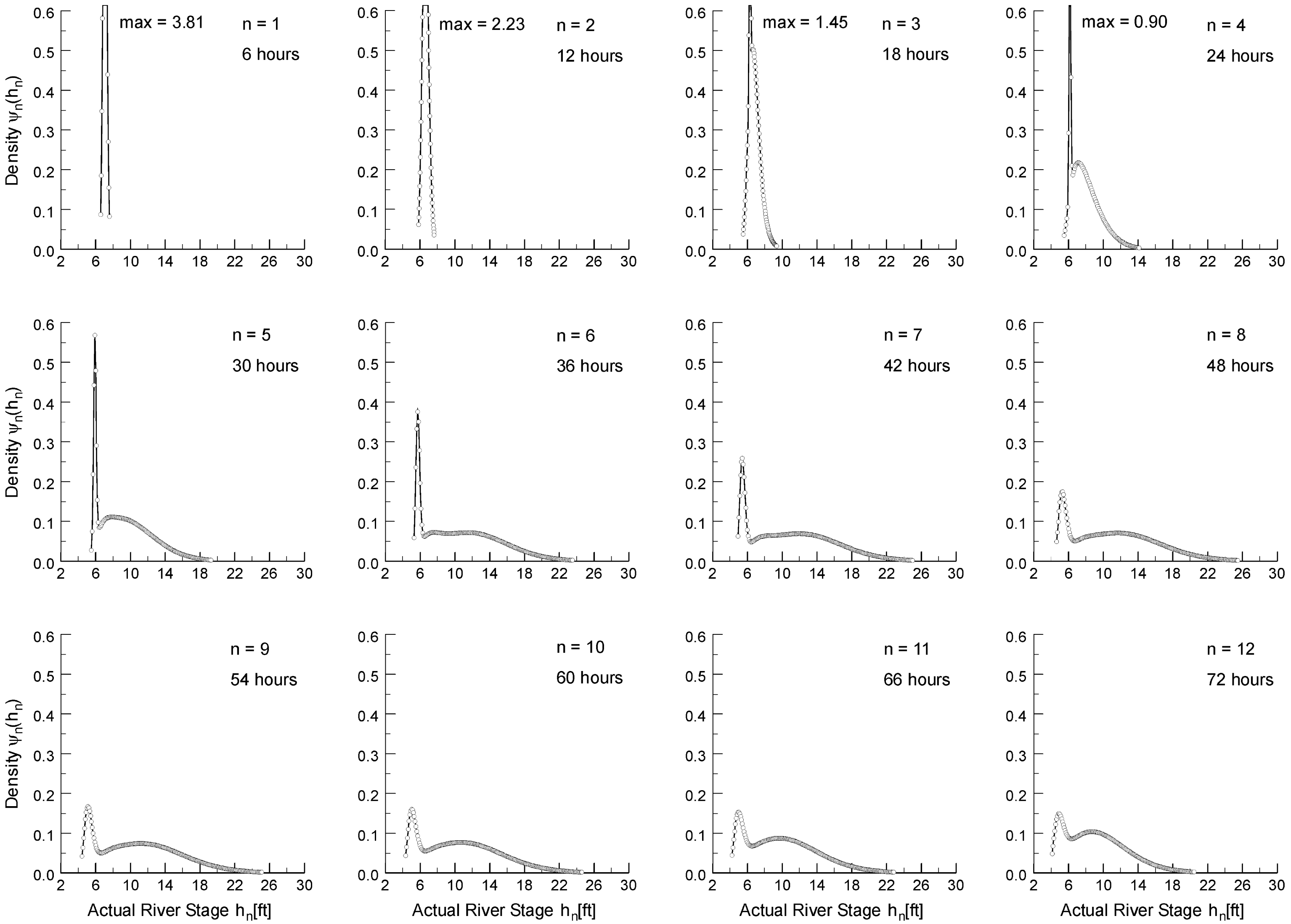

Fig. 13. Sequence of predictive densities $\{\psi_n : n = 1, \ldots, 12\}$ of actual river stages at 6-h steps, counting from 1200 UTC on the forecast day.

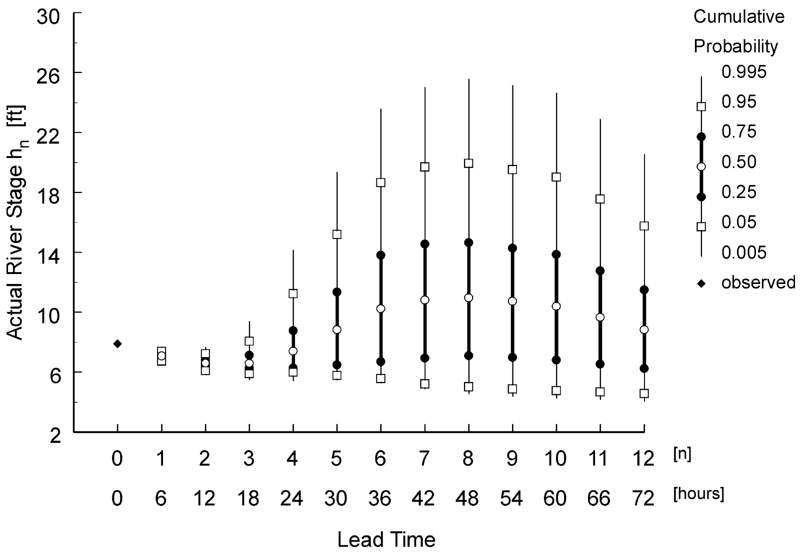

Fig. 14. Isoprobability time series $\{h_{np} : n = 1, \ldots, 12\}$ of quantiles having probability $p = \Psi_n(h_{np})$ on the predictive distributions $\{\Psi_n : n = 1, \ldots, 12\}$ of actual river stages at 6-h steps; the seven time series correspond to $p = 0.005, 0.05, 0.25, 0.50, 0.75, 0.95, 0.995$.

would have to be tested. Should a higher order of the conditional Markov process be necessary, a suitable HUP can be formulated.

4. The PRSF specifies only univariate predictive distributions. Specifically, for each n ($n = 1, \ldots, N$), the PRSF specifies the predictive distribution of actual river stage H_n, which is the n-step transition distribution from observed river stage $H_0 = h_0$ at time t_0 to uncertain river stage H_n at time t_n, given all information utilized by the forecasting system. Such a PRSF may be sufficient to support flood warning decisions, navigation operations, and information needs of the general public. To support reservoir control and other multistage decision processes, one usually requires a forecast that characterizes the stochastic dependence among $H_1, \ldots, H_N$. Specifically, one requires a multivariate predictive distribution or, under the Markovian assumption, a family of one-step transition predictive distributions.

8.1.2. Distributional assumptions

The distributional assumptions can be modified, if necessary, with the implication that new analytic expressions and estimation methods will have to be derived.

1. In the PQPF, the distribution of the total precipitation amount, conditional on the hypothesis that precipitation occurs, is Weibull.
2. In the PUP, the distribution of the model river stage, conditional on the hypothesis that precipitation occurs, is two-piece Weibull.
3. In the HUP, the family of transition densities characterizing the prior uncertainty about the actual river stage process, and the family of likelihood functions characterizing the hydrologic uncertainty, each is meta-Gaussian.

8.2. Attributes

8.2.1. System attributes

The theoretical attributes of the BFS are proven and discussed elsewhere (Krzysztofowicz, 1999a, 2001b). The operational attributes are summarized herein.

1. The BFS can be attached to any deterministic hydrologic model used for operational forecasting without imposing on that model any structural (e.g. linearizing) or distributional (e.g. normalizing) assumptions.
2. The system structure is simple, and the information flow is transparent (Fig. 1).
3. The precipitation uncertainty is quantified in real

time by a PQPF, which must meet two requirements but otherwise may be produced via any forecasting method.

4. The hydrologic uncertainty is quantified before real-time forecasting begins; the quantification requires only historical data, a simulation experiment, estimation of univariate distributions, and estimation of linear regressions.
5. The structural assumptions and the distributional assumptions are all testable directly on data. (For example, there are no assumptions about unobservable quantities such as parameters of the hydrologic model for which no empirical prior distributions exist.) Thus far, all component models have been validated on data from the operational forecast system of the NWS for a 1430 km^2 headwater basin (results reported herein and elsewhere) and for three other headwater basins of sizes 480, 1860, and 2370 km^2 (results not yet reported).
6. The computational effort at the forecast time is small—just seven runs of the hydrologic model and straightforward numerical calculations. Yet the predictive probability distribution of the actual model river stage is (essentially) exact. Thus, the analytic-numerical BFS is computationally more efficient (for this particular forecasting problem) than a Monte Carlo simulation would be if both methods were to produce the predictive probability distributions with the same degree of accuracy. The fast and simple execution makes the BFS well suited to forecasting rainfloods in headwater basins.
7. The PRSF can be rapidly updated based on an updated PQPF between the scheduled forecast times without the need for rerunning the hydrologic model.

8.2.2. Forecast attributes

The Bayesian approach to quantification and integration of uncertainties bestows upon the PRSF two general attributes.

1. The PRSF quantifies the total uncertainty that exists at the forecast time, given all knowledge embodied in the hydrologic model and all information utilized by the BFS.
2. The PRSF possesses a *self-calibration property*: provided the PQPF is well calibrated, the PRSF is also well calibrated. Loosely speaking, it means that in the long run, each exceedance probability of the river stage specified by the PRSF should verify as a conditional relative frequency of that stage being exceeded.

The fact that the total uncertainty comprises both the precipitation uncertainty and the hydrologic uncertainty manifests itself to the user in the following two situations.

3. When the PQPF specifies zero probability of precipitation occurrence during the period and over the basin, so that there is no precipitation uncertainty, the PRSF still indicates uncertainty about the river stage; this is the always-present hydrologic uncertainty.
4. When the PQPF is prepared for period $[t_0, t_0 + T]$ that is shorter than period $(t_0, t_n]$ within which the PRSF is computed, no presumption of zero precipitation beyond time $t_0 + T$ fills the gap. Instead, the PRSF converges to the prior (climatic) distributions of river stages that reflect all possible realizations of the precipitation process beyond time $t_0 + T$. Consequently, the predictive distributions of river stages beyond time $t_0 + T$ always convey larger uncertainty than would be the case if a PQPF specified zero probability of precipitation occurrence during period $(t_0 + T, t_n]$.

9. Closure

The BFS described herein constitutes one of the many possible operational systems that could be developed within the Bayesian theory of probabilistic forecasting via deterministic hydrologic model. The purpose of this particular BFS is to produce a short-term PRSF at the outlet of a headwater basin based on a PQPF. The overall system design offers (i) the theoretically derived structure for integrating all uncertainties, (ii) the empirically validated models for characterizing the component uncertainties, and (iii) the parsimonious analytic-numerical method of computation for real-time forecasting.

The theoretically derived structure and the empirically validated models are essential to advancing the science of probabilistic forecasting. In particular, they have enabled us to show that the predictive probability density function of a river stage can take many different and unusual shapes, which are predominantly asymmetric and bimodal. These shapes are the outcome of several interacting factors: the intermittence of the precipitation process, the conditional nonstationarity of the river stage process, the asymmetry of the densities characterizing the component uncertainties, the nonlinearity and the heteroscedasticity of the stochastic dependence structures, and the nonlinearity of the hydrologic model. While any operational system always involves some degree of approximation, a good system should necessarily capture the bulk of these key factors. The Bayesian theoretic structure helps the modeler to achieve this goal.

Acknowledgments

Thomas E. Adams of the Ohio River Forecast Center, National Weather Service, performed all simulations. Henry D. Herr wrote software implementing the BFS. Kelsey LeBeau and Coire Maranzano performed the numerical calculations and made the drawings. Research leading to this article was funded by the National Oceanic and Atmospheric Administration under award no. NA07WD0097 'Probabilistic Hydrometeorological Forecast System' to the University of Virginia.

References

Abramowitz, M., Stegun, I.A. (Eds.), 1972. Handbook of Mathematical Functions, Dover Publications, Mineola, New York.

Antolik, M.S., 2000. An overview of the National Weather Service's centralized statistical quantitative precipitation forecasts. Journal of Hydrology 239 (1–4), 306–337.

Fread, D.L., Shedd, R.C., Smith, G.F., Farnsworth, R., Hoffeditz, C.N., Wenzel, L.A., Wiele, S.M., Smith, J.A., Day, G.N., 1995. Modernization in the National Weather Service river and flood program. Weather and Forecasting 10, 477–484.

Grecu, M., Krajewski, W.F., 2000. A large-sample investigation of statistical procedures for radar-based short-term quantitative precipitation forecasting. Journal of Hydrology 239 (1–4), 69–84.

Hudlow, M.D., 1988. Technological developments in real-time operational hydrologic forecasting in the United States. Journal of Hydrology 102, 69–92.

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

Krzysztofowicz, R., 1999a. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. Water Resources Research 35 (9), 2739–2750.

Krzysztofowicz, R., 1999b. Point-to-area rescaling of probabilistic quantitative precipitation forecasts. Journal of Applied Meteorology 38 (6), 786–796.

Krzysztofowicz, R., 2001a. Reply. Water Resources Research 37 (2), 441–442.

Krzysztofowicz, R., 2001b. Integrator of uncertainties for probabilistic river stage forecasting: precipitation-dependent model. Journal of Hydrology 249 (1–4), 69–85.

Krzysztofowicz, R., Herr, H.D., 2001. Hydrologic uncertainty processor for probabilistic river stage forecasting: precipitation-dependent model. Journal of Hydrology 249 (1–4), 46–68.

Krzysztofowicz, R., Pomroy, T.A., 1997. Disaggregative invariance of daily precipitation. Journal of Applied Meteorology 36 (6), 721–734.

Krzysztofowicz, R., Sigrest, A.A., 1997. Local climatic guidance for probabilistic quantitative precipitation forecasting. Monthly Weather Review 125 (3), 305–316.

Krzysztofowicz, R., Sigrest, A.A., 1999. Calibration of probabilistic quantitative precipitation forecasts. Weather and Forecasting 14 (3), 427–442.

Krzysztofowicz, R., Drzal, W.J., Drake, T.R., Weyman, J.C., Giordano, L.A., 1993. Probabilistic quantitative precipitation forecasts for river basins. Weather and Forecasting 8 (4), 424–439.

Mills, H.W. II, Krzysztofowicz, R., 1998. Interactive software for probabilistic quantitative precipitation forecasting, 14th International Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, Preprint Volume. American Meteorological Society, pp. J12–J16, Boston, MA, January.

Seo, D.-J., Perica, S., Welles, E., Schaake, J.C., 2000. Simulation of precipitation fields from probabilistic quantitative precipitation forecast. Journal of Hydrology 239 (1–4), 203–229.

Stallings, E.A., Wenzel, L.A., 1995. Organization of the river and flood program in the National Weather Service. Weather and Forecasting 10 (3), 457–464.