RACSAM
Rev. R. Acad. Cien. Serie A. Mat.
VOL. **96** (3), 2002, pp. 461–479
Matemática Aplicada–Estadística e Investigación Operativa / Applied Mathematics–Statistics and Operations Research
Artículo panorámico / Survey

Bayesian methods in Hydrology: A review

D. Ríos Insua, R. Montes Díez and J. Palomo

Abstract Hydrology and water resources management are inherently affected by uncertainty in many of their involved processes, including inflows, rainfall, water demand, evaporation, etc. Statistics plays, therefore, an essential role in their study. We review here some recent advances within Bayesian statistics and decision analysis which will have a profound impact in these fields.

Métodos bayesianos en Hidrología. Una revisión.

Resumen. La Hidrología y la gestión de recursos hidrológicos están afectados de forma inherente por la incertidumbre presente en muchos de los procesos asociados, como los de afluencias, lluvia, demanda de agua, evaporación, etc. La estadística desempeña, por tanto, un papel esencial en su estudio. Revisamos aquí algunos avances recientes dentro de la estadística bayesiana y el análisis de decisiones que tendrán un impacto profundo en estos campos.

1. Introduction

Water resources are becoming more and more important. Indeed, its scarcity is a major problem in dry areas, whereas its overabundance may occasionally create flood problems in other regions. Water quality is also getting very relevant with the ever increasing interest in the environment. This is clear by merely looking at the priority given to water issues in national and EU research programs. We could also say that water problems are at the core of many political disputes, as in, e.g., those related with transfers from the Ebro river to South-Eastern Spain.

A peculiarity of water resource problems is that uncertainty is inherent in many of the processes involved, including rainfall and runoff, evaporation, infiltration, water demand for various purposes,... This has been widely acknowledged in the hydrological literature, see e.g. references like Bras and Rodriguez Iturbe [1993], Helsel and Hirsch [1992], Clarke [1994] or Parent *et al* [1997], which describe how statistical methods have been used, among others, to deal with the simulation of hydrologic processes, hydrologic forecasting, the design of data collection networks, or reservoir system management.

Bayesian methods, see e.g. French and Rios Insua [2000], are facing rapid expansion not only from a theoretical point of view, but also affecting many application areas, including hydrology, see e.g. Parent *et al* [1997]. In this paper, stemming from previous work in Berger and Rios Insua [1998], we expose recent developments in Bayesian analysis which have impacted or may impact various areas in hydrology. We shall concentrate on water quantity issues.

Presentado por Jesús Ildefonso Díaz.
Recibido: 7 de Julio de 2002. *Aceptado*: 9 de Octubre de 2002.
Palabras clave / Keywords: Water resources, Bayesian Analysis, Dynamic Models, Spatial Models, Simulation, Water management.
Mathematics Subject Classifications: 62–01, 62C10, 90–01, 90C39

After a brief overview on recent general developments, mainly Bayesian computations and model selection, we then emphasize time issues, spatial issues, space-time issues and decision making. We conclude with a brief discussion.

2. Bayesian analysis

The Bayesian framework provides a unified and coherent approach to solve inference and decision making problems under uncertainty. As such, we would expect a great impact of Bayesian methods in hydrology. To wit, hydrologists gather and analyze data to improve water resource management. Data is usually affected by uncertainty, therefore begging for statistical analysis. In water resources problems, the end use of such analysis is often to make a decision, hence requiring to formally account for uncertainty in the decision making process. Bayesian analysis seems the appropriate approach to inference in hydrology, as it provides a coherent framework that integrates multiple sources of uncertain information within decision making.

Many other advantages would seem to support the use of Bayesian methods in hydrology. For example, it is frequently the case that in this application area there is prior information available, possibly from previous related studies, from which our analysis might benefit, if appropriately modelled. Spatial and time issues are extremely important in water resources and recent Bayesian models provide useful tools to handle them. However powerful we might think the Bayesian approach is, we should admit that most statistical analyses in this field are still performed within the classical framework, though the situation seems to be changing gradually. See some references in Parent *et al* [1997] and Bras and Rodriguez Iturbe [1993].

The Bayesian framework for inference and decision making problems may be easily described. Indeed, we feel that, at a conceptual level, one of its strengths is, precisely, the ease with which basic concepts are put into place. Rather frequently one of the typical goals in hydrology will be to learn about one (or more) parameters (e.g., mean monthly water inflows, daily rainfall,...) which describe a hydrological phenomenon of interest. Let us designate with θ such parameters. To learn about them, we shall collect data (e.g., inflows at various points,...) and form the likelihood $p(x|\theta)$, which describes how data depends on the parameters. In the Bayesian approach, we also take into account other sources of information about the parameters, based on expert information, past experience, previous studies, the physics of various processes,... which we shall model through the prior distribution $p(\theta)$.

Bayes' formula provides the way to combine both sources of information in the posterior distribution

$$p(\theta|x) = \frac{p(\theta)p(x|\theta)}{\int p(\theta)p(x|\theta)d\theta} \propto p(\theta)p(x|\theta),$$

which summarizes all information available about the parameters and may be used to solve all standard inference problems. For example, for point estimation, we could summarize the posterior through, e.g., its mean, that is

$$E(\theta|x) = \int \theta p(\theta|x)d\theta.$$

For prediction, i.e. forecasting further values y of the phenomenon, we would use the predictive distribution

$$p(y|x) = \int p(y|\theta)p(\theta|x)d\theta.$$

The ultimate aim of statistical research in hydrology will be, most frequently, decision support, e.g. to choose among several release policies. For each action a and each future result y, we would associate a consequence $c(a, y)$. For example, if the water policy is a, we would have incurred in certain costs, created some jobs and recovered a certain piece of land for recreational purposes. Such consequence will be evaluated through its utility $u(a, y)$ and we would choose the alternative a maximizing the expected utility

$$\int u(a, y)p(y|x)dy.$$

In some sense, the recent prevalence of Bayesian methods is due to the development and generalized use of Markov chain Monte Carlo simulation techniques (MCMC). Essentially, we choose a Markov chain on the parameter space, Θ, whose stationary distribution is the posterior. Then, starting at an initial point $\theta^{(0)} \in \Theta$, we generate a sequence of points $\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(m)}$ from the chain. Then, for large m, $\theta^{(m)}$ is (approximately) distributed as $p(\theta|x)$ and we may approximate posterior expectations of the type $\int g(\theta)p(\theta|x)d\theta$ through sample means

$$\frac{1}{m}\sum_{i=1}^{m} g(\theta^{(i)}).$$

By now, there are several standard approaches to build such Markov chains including Gibbs sampling, Metropolis and hybrid algorithms, which we illustrate later on.

Another area which has been fairly fruitful and might be very relevant in hydrology is model selection, see Bertolino and Racugno [1967] for recent references. A typical issue might be whether inflows to a reservoir follow a lognormal or a gamma distribution. In a general model selection problem, the data, x, is assumed to have arisen from one of several possible models $M_1, \ldots, M_m$. Under M_i, the likelihood is $p_i(x|\theta_i)$, where θ_i is an unknown vector of parameters of p_i. The Bayesian approach to model selection begins by assigning prior probabilities, $P(M_i)$, to each model. It is also necessary to choose prior distributions $p_i(\theta_i)$ for the unknown parameters of each model. The analysis then proceeds by computing the posterior probabilities of each model,

$$P(M_i|x) = \frac{P(M_i)m_i(x)}{\sum_{j=1}^{m} P(M_j)m_j(x)},$$

where $m_j(x) = \int p_j(x|\theta_j)p_j(\theta_j)d\theta_j$. Typically, one selects the model (or models) with largest posterior probability.

Our purpose here is to illustrate how the above framework may be used in various hydrology problems.

3. Time issues

Many problems in hydrology require modelling and forecasting time series. In this section, we briefly review classes of models which we have found useful in dealing with hydrological time series.

3.1. Hydrological forecasting with Dynamic Linear Models

An essential step in many hydrology problems is the forecasting of time series, as e.g. when forecasting inflows to a reservoir. Numerous recent modelling and computational enhancements have made Dynamic Linear Models (DLMs) readily available for applications, see West and Harrison [1997]. Specific applications to hydrology may be seen in Rios Insua and Salewicz [1995], Rios Insua *et al* [1997], and Muster and Bardossy [1998].

We aim at determining, at time instant t, the next k values of the variable of interest (or a transformation of it), say y_t, from instant $t+1$ to instant $t+k$, given the available information D_t. For that, we use DLMs which, in their simplest formulation, have the following structure, for every instant of time $t = 1, 2, 3, \ldots$, where F_t and G_t are known:

- *Observation equation:*

$$y_t = F_t\theta_t + v_t, \quad v_t \sim N(0, V_t)$$

  where y_t denotes the observed value, which depends linearly on the state variables θ_t, perturbed by a normal noise.

- *System evolution equation:*

$$\theta_t = G_t\theta_{t-1} + w_t, \quad w_t \sim N(0, W_t)$$

  describing the evolution of the state variables, linearly dependent on the variables in the previous time plus a random perturbation.

- *Initial information:*

$$\theta_0|D_0 \sim N(m_0, C_0)$$

  describing the prior beliefs of the forecaster.

The error sequences v_t and w_t are independent, and mutually independent. Moreover, they are independent of θ_0.

DLMs are useful to hydrologists for many reasons. One that is especially important is that they allow for moving away from stationarity assumptions, since process parameters are time varying. Also, they are flexible enough to model the usual behavior of hydrological time series including seasonal patterns and trends, and permit the incorporation of covariates, such as rainfall for inflows, based on the *superposition principle*. As a consequence, we use a model building strategy based on blocks, representing trends, seasonal patterns, dynamic regression (if covariates are available), and, if required, an autoregressive term to improve short term forecasting. Depending on the forecasting horizon different blocks may be used. For example, long-term planning will require monthly forecasts for, say, the next twelve months, and we typically use:

- A term referring to a level or a piecewise linear trend.
- A seasonal effects term, to account for monthly patterns.
- A regression term, based on covariates (like rainfall, snowfall,...).
- An autoregressive term, to improve short term forecasting.

For short-term forecasting, we usually apply a trend term, an autoregressive term and, possibly, a regression term.

DLMs are also computationally fast, enabling real time decision making and the large-scale simulations habitual in hydrology. Finally, they permit the incorporation of all prior information, including that due to interventions, hence incorporating a principle of *management by exception*, fundamental in the Bayesian forecasting philosophy (West and Harrison, 1997): a set of models is routinely used for processing information, making inferences and forecasting, unless exceptional circumstances arise. Examples, would include a sudden rainfall or a big release from a reservoir upstream. In this case, the system is open to external intervention, typically by inclusion of additional subjective information.

Forecasting is performed sequentially based on all available information. Updating procedures and the use of this model for forecasting are described in detail in West and Harrison [1997]. Essentially, inferences about both parameters and forecasts, one or more steps ahead, are based on a normal model, with corresponding parameters computed recursively.

Here we briefly illustrate a DLM that we have used, see Rios Insua *et al* [1997], to forecast inflows to the Kariba-Cahora Bassa system of hydro-power reservoirs at the Zambezi river, the Kariba being upstream.

Example Given the location of both reservoirs, we consider a forecasting model for the inflows to Kariba, and a forecasting model for the inflows to Cahora Bassa. This one will depend on those inflows not coming from Kariba, which we call incremental inflows, and on the releases from Kariba, accounting for interreservoir dynamics.

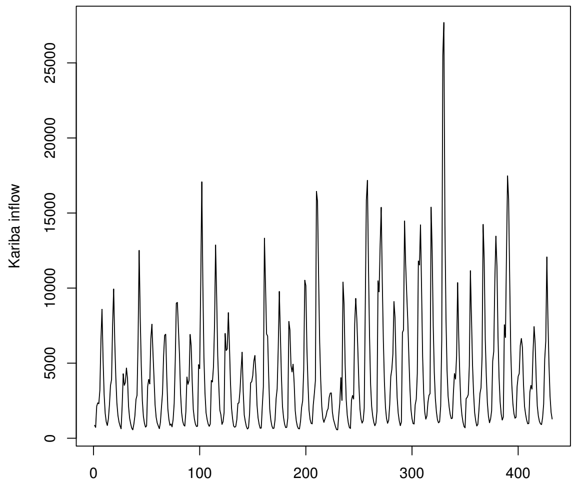

Figure 1. Inflows to Kariba lake

Inflows to Lake Kariba Figure 1 shows monthly inflows to Kariba. After a log transformation, we modeled that time series with a level term, a term representing seasonal (annual) variation and a low coefficient, first order autoregressive term to improve short term forecasts. Consequently, we ended up working with the following DLM:

- *Observation equation.*

$$y_t^k = \theta_t^{1k} + \theta_t^{2k} + \theta_t^{4k} + v_t^k, \quad v_t^k \sim N(0, V^k)$$

where $y_t^k = \log(i_t^k)$ is the logarithm of the inflow to Kariba; θ_t^{1k} designates the level of the series; θ_t^{2k} and θ_t^{3k} refer to the seasonal term; θ_t^{4k} refers to the autoregressive term; and v_t^k designates a Gaussian error term of constant, but unknown, variance V^k.

- *System equation.*

$$\begin{aligned} \theta_t^{1k} &= \theta_{t-1}^{1k} + w_t^{1k} \\ \theta_t^{2k} &= \cos(\pi/6)\theta_{t-1}^{2k} + \sin(\pi/6)\theta_{t-1}^{3k} + w_t^{2k} \\ \theta_t^{3k} &= -\sin(\pi/6)\theta_{t-1}^{2k} + \cos(\pi/6)\theta_{t-1}^{3k} + w_t^{3k} \\ \theta_t^{4k} &= .4\theta_{t-1}^{4k} + w_t^{4k} \end{aligned}$$

with $\mathbf{w}_t^k = (w_t^{1k}, w_t^{2k}, w_t^{3k}, w_t^{4k})$ being an error term such that

$$\mathbf{w}_t^k \sim N\left(\mathbf{0}, \begin{pmatrix} v^k W_t^{*k} & 0 \\ 0 & \sigma_k^2 \end{pmatrix}\right),$$

where σ_k^2 is the autoregressive variance; and W_t^{*k} is the variance matrix (up to v^k) of the first three terms, defined using discounting, with a level discount factor δ_1^k and a seasonal part discount factor δ_2^k.

- *Prior information.*

$$\theta_0^k|\phi^k \sim N(\mathbf{m}_0^k, v^k C^{*k})$$

$$\phi^k \sim Gamma(n_0^k/2, d_0^k/2)$$

with $\theta_0^k = (\theta_0^{1k}, \theta_0^{2k}, \theta_0^{3k}, \theta_0^{4k})$, $\phi^k = \frac{1}{v^k}$ and prior parameters specified judgementally from an expert.

Dynamic regression Kariba–Cahora Bassa Exploratory data analyses suggested regressing dynamically the inflows to Kariba and Cahora Bassa, after a log-transformation. For that, we define the logarithm of the inflow to Cahora Bassa $y_t^c = \log(i_t^c)$ and $z_t = y_t^c - y_t^k$. We ended up working with the following DLM:

- *Observation equation.*

$$z_t = \theta_t^{1c} + \theta_t^{2c} + \theta_t^{4c} + \theta_t^{6c} + v_t^c, \quad v_t^c \sim N(0, V^c)$$

where θ_t^{1c} designates the level of the series; θ_t^{2c} and θ_t^{3c} refer to the first harmonic of the seasonal term; θ_t^{4c} and θ_t^{5c} refer to the second harmonic and θ_t^{6c} refers to the autoregressive term; and v_t^c designates a Gaussian error term of constant, but unknown, variance V^c.

- *System equation.*

$$\begin{aligned}
\theta_t^{1c} &= \theta_{t-1}^{1c} + w_t^{1c} \\
\theta_t^{2c} &= \cos(\pi/6)\theta_{t-1}^{2c} + \sin(\pi/6)\theta_{t-1}^{3c} + w_t^{2c} \\
\theta_t^{3c} &= -\sin(\pi/6)\theta_{t-1}^{2c} + \cos(\pi/6)\theta_{t-1}^{3c} + w_t^{3c} \\
\theta_t^{4c} &= \cos(\pi/3)\theta_{t-1}^{4c} + \sin(\pi/3)\theta_{t-1}^{5c} + w_t^{4c} \\
\theta_t^{5c} &= -\sin(\pi/3)\theta_{t-1}^{4c} + \cos(\pi/3)\theta_{t-1}^{5c} + w_t^{5c} \\
\theta_t^{6c} &= .4\theta_{t-1}^{6c} + w_t^{6c}
\end{aligned}$$

with $\mathbf{w}_t^c = (w_t^{1c}, w_t^{2c}, ..., w_t^{6c})$ an error term such that

$$\mathbf{w}_t^c \sim N\left(\mathbf{0}, \begin{pmatrix} v^c W_t^{*c} & 0 \\ 0 & \sigma_c^2 \end{pmatrix}\right),$$

with σ_c^2 the autoregressive variance; and W_t^{*c} the variance matrix (up to term v^c) of the first five terms. This matrix was defined using discounting, with a level discount factor δ_1^c and a seasonal part discount factor δ_2^c.

- *Prior information.*

$$\theta_0^c|\phi^c \sim N(\mathbf{m}_0^c, v^c C^{*c})$$

$$\phi^c \sim Gamma(n_0^c/2, d_0^c/2)$$

with $\theta_0^c = (\theta_0^{1c}, \theta_0^{2c}, \ldots, \theta_0^{6c})$ and $\phi^c = \frac{1}{v^c}$, assessed judgamentally

Inflows to Cahora Bassa The model we propose now for Cahora Bassa depends on the release a_t^k from Kariba and the incremental inflow (inc_t) to Cahora Bassa from the tributaries and basin between Kariba and the inlet to Cahora Bassa. Taking into account that water travelling time, from Kariba to Cahora, is less than a month, we use the following relation:

$$i_t^c = a_t^k + inc_t.$$

Given that releases will depend on the release policy adopted, we shall estimate incremental inflows as follows. First, if there was no reservoir, we would have:

$$i_t^c = i_t^k + inc_t.$$

The second relation describes a dynamic regression between inflows to Cahora and Kariba, and reflects the physical relation that inflow is related with basin size:

$$i_t^c = \beta_t i_t^k.$$

Simple computations suggest modelling incremental inflows by

$$inc_t = (\beta_t - 1) i_t^k,$$

and inflows to Cahora by

$$i_t^c = a_t^k + (\beta_t - 1) i_t^k.$$

Observe that we no longer have a dynamic linear model, since both the regression weights β_t and regressors i_t^k are subject to uncertainty. Besides, the distribution of a_t^k is not standard and is controlled by the reservoir manager. Forecasting i_t^c is easily done by simulation, since we know or can easily sample the distributions involved: that of i_t^k is obtained from the inflow to Kariba model above; a_t^k is obtained from a release model from Kariba; the distribution on β_t comes from the inversion of the transformation and the dynamic regression Kariba-Cahora Bassa above.

3.2. Forecasting nonlinear series

As our example showed, we may not always use DLMs to forecast time series. In some cases, we may appeal to nonlinear models. One class is that based on neural network. Here we describe their use as nonlinear autoregressions. A full description may be seen in Menchero *et al* [2002]. We use this to illustrate model selection and MCMC methods.

Suppose we have univariate time series data $\{y_1, y_2, \ldots, y_n\}$, like rainfall or inflows. We model the generating stochastic process in an autoregressive fashion,

$$p(y_1, y_2, \ldots, y_n) = p(y_1, y_2, \ldots, y_q) \prod_{t=q+1}^{n} p(y_t | y_{t-1}, y_{t-2}, \ldots, y_{t-q})$$

Specifically, we shall assume that each y_t is modelled by a nonlinear autoregression function of q past values plus a normal error term:

$$\begin{aligned} y_t &= f(y_{t-1}, y_{t-2}, \ldots, y_{t-q}) + v_t, \; t = q+1, \ldots, n \\ v_t &\sim N(0, \sigma^2). \end{aligned}$$

For f we propose a mixed model as a linear combination of a linear autoregression term and a feedforward neural network.

A feedforward neural network model with q input nodes, one hidden layer with M hidden nodes, one output node and activation function φ is a model relating a response variable $\widehat{y}_t$ and q explanatory variables, in our case $x_t = (y_{t-1}, y_{t-2}, \ldots, y_{t-q})$:

$$\widehat{y}_t(x_t) = \sum_{j=1}^{M} \beta_j \varphi(x_t' \gamma_j + \delta_j)$$

with $\beta_j \in \mathcal{R}$, $\gamma_j \in \mathcal{R}^q$. The biases δ_j may be assimilated to the rest of the γ_j vector if we consider an additional input with constant value one, say $x_t = (1, y_{t-1}, \ldots, y_{t-q})$. Without loss of generality, we assume the basis function φ to be $\varphi(z) = \exp(z)/(1+\exp(z))$. In our mixed model, the linear term will account for linear features and the FFNN term for nonlinear ones:

$$\widehat{y}_t(x_t) = x_t'\lambda + \sum_{j=1}^{M} \beta_j \varphi(x_t'\gamma_j),\ t = q+1,\ldots,n \tag{1}$$

The parameters in our model are the linear coefficients $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_q) \in \mathcal{R}^{\,q+1}$, the hidden to output weights $\beta = (\beta_1, \beta_2, \ldots, \beta_M)$, the input to hidden weights $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_M)$, the error variance σ^2 and the number M of hidden nodes.

Note that, as particular cases, we include the simple linear autoregression model (which may be viewed as a DLM)

$$\widehat{y}_t = f(y_{t-1}, y_{t-2}, \ldots, y_{t-q}) = x_t'\lambda,\ t = q+1,\ldots,n$$

and a nonlinear autoregression feedforward neural net model

$$\widehat{y}_t = f(y_{t-1}, y_{t-2}, \ldots, y_{t-q}) = \sum_{j=1}^{M} \beta_j \varphi(x_t'\gamma_j),\ t = q+1,\ldots,n$$

for each value of M.

The prior over model parameters will be:

$$\begin{aligned}\beta_j &\sim N(\mu_\beta, \sigma_\beta^2),\ \ \lambda \sim N(\mu_\lambda, \sigma_\lambda^2 I),\\ \gamma_j &\sim N(\mu_\gamma, \Sigma_\gamma),\ \ \sigma^2 \sim InvGamma(a_\sigma, b_\sigma).\end{aligned}$$

When there is non-negligible uncertainty about prior hyperparameters, we may complete the prior model with a hyperprior over them, by using, for example, standard choices in hierarchical models for $\chi = (\mu_\beta, \sigma_\beta^2, \mu_\lambda, \sigma_\lambda^2, \mu_\gamma, \Sigma_\gamma, a_\sigma, b_\sigma)$. Since the likelihood is invariant to relabelings, we include an order constraint to avoid trivial posterior multimodality, due to permutations of indexes. For example, we could use $\gamma_{1p} \leq \gamma_{2p} \cdots \leq \gamma_{Mq}$.

MCMC posterior inference with fixed model Consider first the mixed model (1). With our prior assumptions, the joint posterior distribution is given by:

$$p(\lambda, \beta, \gamma, \sigma^2, \chi \mid D') \propto p(y_{q+1}, \ldots, y_N \mid y_1, \ldots, y_q, \lambda, \beta, \gamma, \sigma^2) p(\lambda, \beta, \gamma, \sigma^2, \chi) M!$$

where

$$\begin{aligned}p(\lambda, \beta, \gamma, \sigma^2, \chi) &= p\left(\mu_\lambda, \sigma_\lambda^2, \mu_\beta, \sigma_\beta^2, \mu_\gamma, \Sigma_\gamma\right) p(\sigma^2)\\ &\qquad p(\lambda \mid \mu_\lambda, \sigma_\lambda^2 I) p(\beta \mid \mu_\beta, \sigma_\beta^2 I) \prod_{i=1}^{M} p(\gamma_i \mid \mu_\gamma, \Sigma_\gamma)\end{aligned}$$

is the joint prior distribution, $\chi = (\mu_\lambda, \sigma_\lambda^2, \mu_\beta, \sigma_\beta^2, \mu_\gamma, \Sigma_\gamma)$ is the set of hyperparameters and $M!$ takes into account the order constraint on γ .

Exact analytic inference in such model is not possible and we need to appeal to MCMC methods to obtain a sample from the posterior. In this case, we propose a hybrid, partially marginalized sampling scheme, see Menchero *et al* [2002] for a detailed description:

1. Use a Metropolis step to update the input to hidden weights γ_j using the marginal likelihood over (β, λ): $p(D' \mid y_1, \ldots, y_q, \gamma, \sigma^2)$ to partly avoid the random walk nature of Metropolis algorithm.
2. Generate new values for parameters β, λ and σ^2 by drawing from the corresponding full conditional posteriors.
3. Given current values of $\left(\gamma, \beta, \lambda, \sigma^2\right)$, generate a new value for each hyperparameter $\mu_\lambda, \sigma_\lambda^2, \mu_\beta, \sigma_\beta^2, \mu_\gamma$ and Σ_γ drawing from their complete conditional posterior distributions.

Model Selection Once defined how to perform inference with any of the models considered above, we propose a reversible jump MCMC (Green, 1995) in order to perform model selection. We use a pair of indexes (h, k), to index models: $h = \{0, 1\}$, denoting whether or not the linear autoregression term is included and $k = 0, 1, \ldots$, denoting the number of hidden nodes in the FFNN term. The algorithm, described in Menchero *et al* [2002], evolves through various models as described in Figure 2.

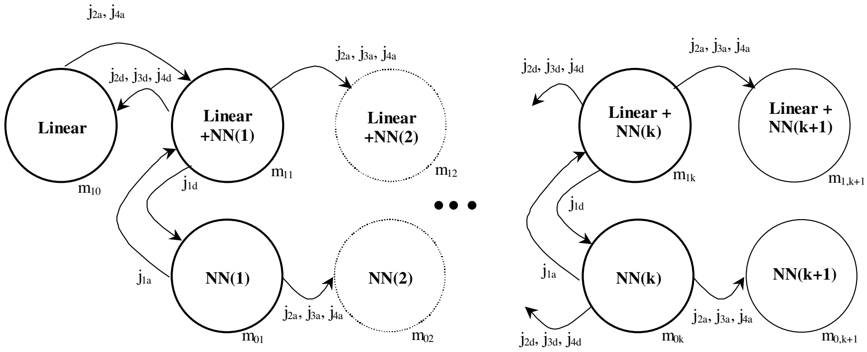

Figure 2. Possible models along with moves available from each one.

Example We consider the time series of monthly inflows to Lake Kariba (see Figure 1). We split the Kariba data set into a training data set (first 350 observations) and a test data set (last 82). Exploratory data analysis suggests seasonal behavior of the time series. We shall therefore model each y_t on the basis of the immediately past value, y_{t-1}, as well as the corresponding previous year value y_{t-12}.

Prior distributions for the unknown parameters and hyperparameters in the model, are chosen as follows

$$\begin{aligned} \mu_\beta &\sim N(0,3), \;\; \sigma^2_\beta \sim InvGamma(9,1) \\ \mu_\lambda &\sim N((0,0,0),3I), \;\; \sigma^2_\lambda \sim InvGamma(9,1) \\ \mu_\gamma &\sim N((0,0,0),3I), \;\; \Sigma_\gamma \sim InvWishart(10,2.5I) \\ \sigma^2 &\sim InvGamma(1,1) \end{aligned}$$

The number of nodes M is given a geometric prior distribution with parameter 2.

Two runs of the algorithm were carried out from different starting points, namely $M = 0$ and $M = 15$. A burn–in of 1000 iterations was used and then 9000 additional iterations were monitored for inference purposes. Figure 3 shows the histogram of the posterior distribution of M, respectively, suggesting $M = 2$ as the most likely number of nodes for the hidden layer of the FFNN term, although 1, 3 and 4 nodes receive nonnegligible mass. There is, clearly, some nonlinearity in the series.

Figure 4 shows the time series data, after log-transformation and the predicted values for the test data set, showing good performance of the Bayesian nonlinear autoregression model developed.

3.3. Modelling long term effects

Long-memory time series models are used to formalize the notion of strongly dependent series of observations. In hydrology, this is known as Hurst effect (Hurst, 1951). While studying the issue of reservoir capacity storage, Hurst empirically observed that if i_t is the inflow to a reservoir at time t and $y_j = \sum_{t=1}^{j} i_t$,

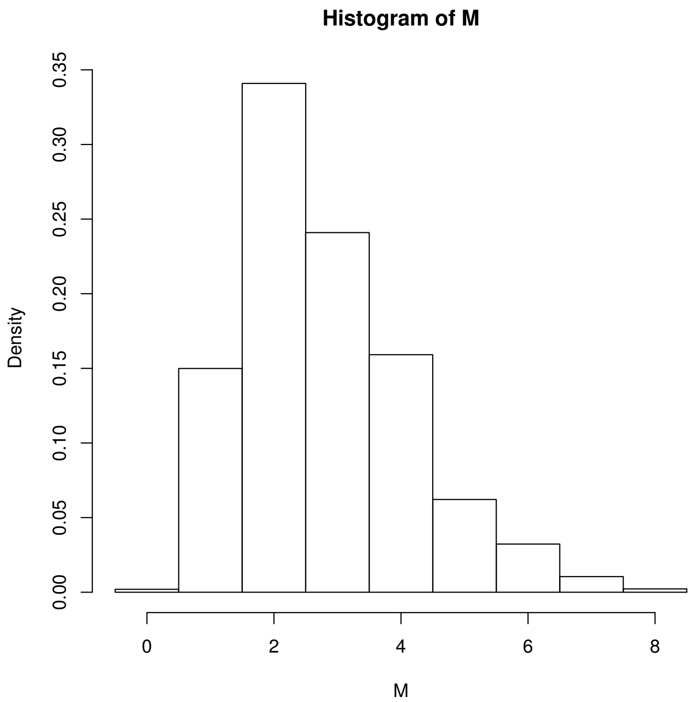

Figure 3. Histogram of the posterior distribution of M, suggesting $M = 2$ nodes for the hidden layer of the FFNN term.

the rescaled adjusted range, or R/S statistic, defined by

$$R/S = \frac{\max_t(y_t - y_0 - \frac{t}{k}(y_k - y_0)) - \min_t(y_t - y_0 - \frac{t}{k}(y_k - y_0))}{\sqrt{\frac{1}{k}\sum(i_t - \bar{\imath})^2}},$$

behaves, for large values of k approximately as a constant times k^H for some $H > 1/2$, whereas for standard short memory models we get $H = 1/2$. The phenomenon is also known as Noah and Joseph effect (Mandelbrot and Wallis 1968).

The use of standard short memory models, such as ARMA, to analyze strongly correlated data generally results in underestimating uncertainty inherent to the results of the analysis. This would result, for example, in illusory nonstationarity. Beran [1994] is a good introduction to the statistics of long-memory processes. Petris [1997] and Petris and West [1998] provide Bayesian models for long-memory time series.

4. Spatial issues

Spatial issues are also extremely relevant in hydrology, as illustrated in e.g. Bras and Rodriguez Iturbe [1993]. For example, an appropriate spatial model may aid in using regional information to enlarge scarce data sets. They are also relevant in rainfall generation models. Kriging is also a spatial technique used to forecast a certain phenomenon based on spatial information. To give a feeling of the ideas involved, we outline a simple spatial regression model which may be useful in some hydrologic problems. See Marin *et al* [2002] for further details.

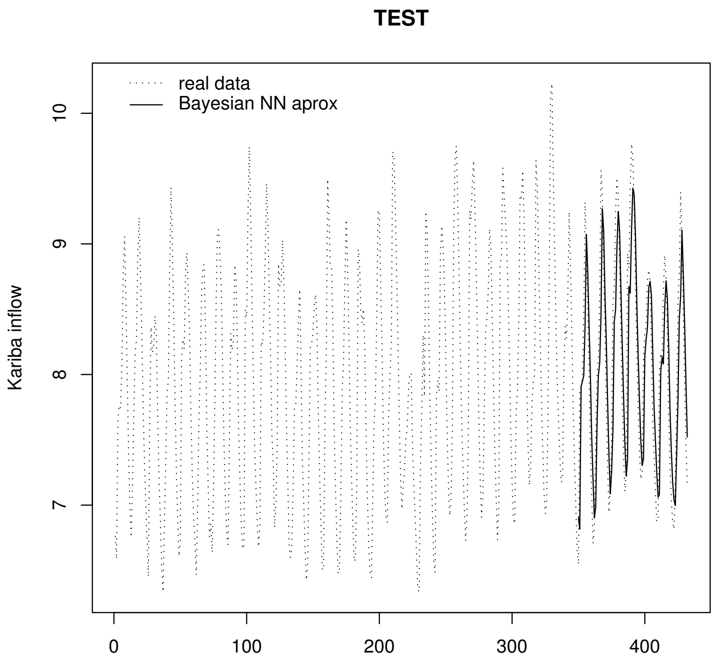

Figure 4. Time series (log-transformed) and predicted values for the test data set.

4.1. A model for spatial regression

We consider here a model to account for spatial variations in a variable of interest, say e.g. rainfall at various locations. The phenomenon in clustered mainly by location, for which we use mixtures. Diebolt and Robert [1994] provide a general framework for modelling with mixtures. An issue of importance is the number of terms in the mixture; Richardson and Green [1997] provide a solution for the case of mixtures of normals of variable size, based on reversible jump MCMC samplers.

Here we shall provide a general model to deal with multi-level mixture phenomena. We limit the discussion to two levels. We assume we have data $D = \{z_1, \ldots, z_N\}$ where $z_i = (x_i, y_i)$, with x_i describing the bivariate location of certain phenomena, whereas y_i will describe a (possibly) multivariate hydrologic phenomena, say rainfall at location x_i.

At the level of the observed data, we induce clustering by assuming a mixture of (bivariate) normal models with an unknown number of terms, M

$$z_i \sim \sum_{j=1}^{M} q_j \; N(\cdot|\mu_j, \Sigma_j), \quad i = 1, \ldots, N$$

implying a locally weighed mixture of normal linear regressions, with weights varying as a spatial process. Equivalently, the implied clustering may be formulated in terms of latent indicator variables r_i with

$$Pr(r_i = j) = q_j \quad \text{and} \quad p(z_i|r_i = j) = N(\mu_j, \Sigma_j).$$

The latent indicators r_i define clusters $\Gamma_j = \{i : \; r_i = j\}$.

At a second level of the nested cluster model, we need to deal with cluster locations μ_j and covariance matrices Σ_j, which we shall assume to be clustered by a similar process themselves

$$\theta_j = (\mu_j, \Sigma_j) \sim \sum_{k=1}^{L} p_k \, W\left[\cdot|\nu, (\nu\, S_k)^{-1}\right] \; N(\cdot|\beta_k, \rho\,\Sigma_j), \quad j = 1, \ldots, M.$$

where $W[\cdot|\cdot,\cdot]$ denotes Wishart distribution. We can rewrite this mixture, using latent indicators $s_j \in \{1, \ldots, L\}$, as

$$Pr(s_j = k) = p_k \quad \text{and} \quad p(\mu_j, \Sigma_j | s_j = k) = W\left[\Sigma_j^{-1}|\nu, (\nu\, S_k)^{-1}\right] \; N(\mu_j|\beta_k, \rho\,\Sigma_j).$$

As in the top level mixture, the latent indicators s_j define super-clusters $\Delta_k = \{j : \; s_j = k\}$.

The model is completed with priors on the unknown parameters, for the mixture sizes M and L, the mixture weights $q = (q_1, \ldots, q_M)$ and $p = (p_1, \ldots, p_L)$, and the second level mixture hyperparameters β_k and S_k.

Posterior simulation for mixture models is habitually implemented using the latent indicator variables (r_i, s_j), conditional on which, the problem is reduced to the corresponding model without the mixture. Conditional on the parameters of each term in the mixture, resampling the indicators only requires multinomial sampling. In our specific problem, assuming M and L fixed for now, the Gibbs sampling scheme would be

1. Sample from $r_i|M, \mu, \Sigma_j, D$ (multinomial).
2. Sample from $q|r, M, D$, (M-dimensional Dirichlet).
3. Sample from $\mu_j|r, \Sigma_j, s_j = k, \beta_k, D$, (normal).
4. Sample from $\Sigma_j|\mu_j, r, s_j = k, \beta_k, D$, (inverse Wishart).
5. Sample from $s_j|L, \beta_k, S_k, D$, (multinomial).
6. Sample from $p|s, L, D$, (k-dimensional Dirichlet)
7. Sample from $\beta_k|s, \mu, \Sigma_j, D$ (normal).
8. Sample from $S_k|s, \Sigma_j, D$, (inverse Wishart).

where $i = 1, \ldots, N, \; j = 1, \ldots, M, \; k = 1, \ldots, L$ and D designates dependence from data.

Keeping random the size of the mixture, as in our case with L and M, considerably complicates posterior simulation, and reversible jumpers or other methods like Stephens' [2002] should be used. Full details may be seen in Marin *et al* [2002].

5. Spatio-temporal issues

Space time issues have gained also popularity in recent years due to an abundance of environmental sciences applications, including hydrology. For example, we could be interested in forecasting rainfall through time and space over a region. Space-time data sets are often large and, therefore, require substantial computing resources to fit even simple models, which generally combine time series models with features of spatial statistics.

Here we briefly describe a powerful model in Stroud *et al* [2001], which they illustrate with space-time forecasting of rainfall. Let $y(x)$ be the observed variable (e.g. rainfall) at location x. The proposed spatial model is of the form

$$y(x) = \sum_{j=1}^{J} \pi_j(x) f_j'(x)\theta_j + v(x)$$

where $\sum_{j=1}^{J} \pi_j(x) f_j'(x)\theta_j = S(x,\theta)$ is a spatial mean function, which is a locally weighted mixture of linear regressions, where $f_j(x)$ is a set of known basis functions, θ_j is a vector of unknown random parameters, $\pi_j(x)$ is a non-negative weighting kernel, J is the number of components and $v(x)$ is a Gaussian noise process.

An observation of the process is $y = (y(x_1), \ldots, y(x_n))$ at n locations $x_1, ..., x_n$. Then, if we define $\pi_j = (\pi_j(x_1), \ldots, \pi_j(x_n))$ and $F_j = (f_j(x_1), \ldots, f_j(x_n))$, we may rewrite the model as

$$Y = F\theta + v.$$

To introduce time effects, we let the parameters evolve through time and consequently consider the DLM

$$\begin{aligned} y_t &= F_t\theta_t + v_t,\ v_t \sim N(0, V_t) \\ \theta_t &= G_t\theta_{t-1} + w_t,\ w_t \sim N(0, W_t) \end{aligned}$$

completed with a prior on the initial state.

The approach proposed is computationally efficient, allowing for on-line implementation and full posterior inference. As new data are collected, the model provides updated estimates of the mean and uncertainty field as well as predictions at any desired location. This is critical in problems like monitoring rainfall, where fast algorithms are necessary to issue warnings in a timely manner.

6. Hydrological extreme value analysis

Extreme value modelling of environmental processes is a standard practice for the design of many large scale constructions, say reservoirs. The main problem here is that at the most extreme levels data are scarce, see Figure 1, whereas for design purposes it is the behavior at these levels which is of greatest interest. Scarcity of data makes Bayesian methods even more relevant, when expert information may be used to take into account physical information based on the understanding of the corresponding processes. The most interesting application, from an hydrology point of view, is the prediction of extreme rainfall. Here we shall briefly outline ideas in Coles and Tawn [1996] and Coles [2001], who include additional references.

Assume $y_1, y_2, \ldots, y_n$ are observations (rainfall) with common distribution. Our interest lies in estimating the upper tail behavior of such distribution. We assume that for large thresholds u, the sequence $y_1, y_2, \ldots, y_n$, viewed on (u, ∞) is, approximately, a non-homogeneous Poisson process with intensity function

$$\lambda(y) = \frac{1}{\sigma} \max\left\{ \left(1 + \psi\frac{y-\mu}{\sigma}\right)^{-(\psi+1)/\psi}, 0 \right\},$$

the parameters being μ, σ, ψ, resulting in the distribution function

$$Pr\left(\max(y_1, \ldots, y_n) \le y | \mu, \sigma, \psi\right) = \exp - \left[\max\left\{ \left(1 + \psi\frac{y-\mu}{\sigma}\right)^{-(\psi+1)/\psi}, 0 \right\} \right]$$

corresponding to a generalized extreme value distribution.

In extreme value analysis, data tends to be sparse and knowledge of the expert hydrologist may be used to supplement information, based on general understanding of the physical generation of rainfall, and a specific knowledge of the rainfall characteristics within the vicinity of the particular site. Note that the $1-p$ quantile of the maximum distribution is given by

$$q_p = \mu + \sigma[\{-\log(1-p)\}^{-\psi} - 1]/\psi$$

Values of q_p for small p are design parameters in various applications. The idea would be to elicit prior information in terms of $(q_{p_1}, q_{p_2}, q_{p_3})$ for, specified, small values of $p_1 > p_2 > p_3$. Given the order constraint $q_{p_1} < q_{p_2} < q_{p_3}$, we may alternatively work with the differences

$$\begin{aligned} q_1 &= q_{p_1} - 0 \\ q_2 &= q_{p_2} - q_{p_1} \\ q_3 &= q_{p_3} - q_{p_2} \end{aligned} \tag{2}$$

and assume

$$q_i \sim gamma(\alpha_i, \beta_i), \quad i = 1, 2, 3. \tag{3}$$

To determine α_i and β_i, $i = 1, 2, 3$, the experts are asked for estimates of, e.g., the median and .90 quantile of each q_i and then the corresponding equations are solved.

From the prior specification given by equations (2) and (3), the joint prior distribution for the q_{p_i} is obtained as

$$f(q_{p_1}, q_{p_2}, q_{p_3}) \propto q_{p_1}^{\alpha_1 - 1} \exp\left(-\beta_1 q_{p_1}\right) \prod_{i=2}^{3} (q_{p_i} - q_{p_{i-1}})^{\alpha_i - 1} \exp\left\{-\beta_i (q_{p_i} - q_{p_{i-1}})\right\}$$

which multiplied by the likelihood

$$L(\mu, \sigma, \psi | y) = \exp\left(-\frac{n}{N}\Lambda[u, \infty)\right) \prod_{i=1}^{n} \lambda(y_i)$$

where

$$\Lambda[u, \infty) = \int_u^{\infty} \lambda(y) \mathrm{d}y.$$

gives the posterior $p(\mu, \sigma, \psi | y)$, up to a proportionality constant. Explicit analytical calculation of the marginal posterior distributions is intractable, however direct simulation based on a hybrid MCMC sampler is straightforward. See Coles and Tawn [1996] for further details. This may be used, e.g., to forecast the maximum rainfall over a future period of L years.

7. Decision making issues: Reservoir operations

The previous sections showed models which are useful in dealing with time-space features which are frequent in hydrology problems. We emphasize now decision making issues, with reference to reservoir operations. It is worth mentioning that the field of reservoir operations has been a motor for the development of novel optimization methods, see Yeh [1985] and Yakowitz [1982] for reviews.

7.1. Sequential optimization for water resources management

The aim of the control policy is to determine at every discrete moment of time (for instance, once a month) controls $a_{t+1}, \ldots, a_{t+k}$, that is, volumes of water to be released, where k is the planning horizon and t is the current time. Usually, we distinguish types of releases associated with various operational purposes, e.g. for hydro-power generation, irrigation, flood control, spill, ... so that $a_t = (a_t^1, a_t^2, \ldots, a_t^m)$, where a_t^l denotes the volume of water released for purpose l at time t.

Information about the inflow process is given in the form of a predictive density $h(y_{t+1}, \ldots, y_{t+k} | D_t)$ based on the analysis of historical data records, possibly with models described in Section 3.. A preference model u evaluates the consequences $c(a, y)$ associated with releasing a when the inflows are y. Storage values at time t are denoted s_t. An evaluation of the final state of the reservoir is given through a function

Φ. Then, at time t, the reservoir management planning problem consists of finding controls $a_{t+1}, \ldots, a_{t+k}$ maximizing the expected utility

$$\int \left(\sum_{j=1}^{k} u(c(a_{t+j}, y_{t+j})) + \Phi(s_{t+k+1}) \right) h(y_{t+1}, \ldots, y_{t+k}|D_t) \mathrm{d}y_{t+1} \ldots \mathrm{d}y_{t+k} \quad (4)$$

taking into account the dynamics of the reservoir system, constraints over controls and reservoir storages. Typical constraints would include bounds on various types of releases, bounds on maximum and minimum allowed reservoir storages, and continuity conditions relating storages from previous to consecutive periods of time given inflows, releases and evaporation according to equation:

$$s_{t+1} = s_t - e_t + y_t - \sum_l a_t^l,$$

where e_t is the evaporated volume.

Should we apply the above framework to large reservoirs, we would require a 36 month, or longer, planning horizon. The corresponding maximization problem becomes unmanageable for such long planning horizons. Note also that the solution of such problem requires, for the evaluation of each control, the solution of a high dimensional integral. Additionally, the uncertainty about the inflow process rapidly propagates through the considered time horizon.

For problems solved over shorter horizons, stochastic programming provides appropriate computational and/or approximation schemes, see Birge and Louveaux [1997] for a review, and Carlin *et al* [1997] for alternative approaches based on forward simulation. White [1993] reviews several solution procedures including linear programming, policy improving algorithms and value improving algorithms. Rogers *et al* [1991] use aggregation-disaggregation techniques to deal with large scale algorithms. Many other methods aim at mitigating the dimensionality curse, either by taking advantage of the problem structure, including search methods, decomposition methods or methods to approximate the value function, e.g, with neural networks, as in Bertsekas and Tsitsiklis [1996]. We describe here an alternative method. Another possibility is to use the augmented simulation method extended to sequential problem, as in Virto [2002].

7.2. A perturbated myopic approach

An alternative strategy based on a *reference trajectory* may be adopted. This strategy assumes availability of a *reference* storage level for each period. Then, the problem can be formulated as

$$max \int (u(c(a_{t+1}, y_{t+1})) + \delta(s_{t+1}, s^*_{t+1}))h(y_{t+1}|D_t)\mathrm{d}y_{t+1}, \quad (5)$$

where $\delta(s_{t+1}, s^*_{t+1})$ represents the deviation of the final storage s_{t+1} from the *reference* storage s^*_{t+1}. Intuitively speaking, if reference storages are defined in such a way so as to account for the dynamic aspects of the problem, we would not loose too much with this modified *myopic* approach.

To compute the *reference trajectory*, we use a deterministic version of problem (4), with inflows fixed at their predictive expected values $\bar{y}_{t+j}$. We use the same dynamic equations and constraints on storages and controls, and select an initial storage s_0. The objective function to be maximized is then

$$\sum_{j=1}^{k} u(c(a_{t+j}, \bar{y}_{t+j})) - \rho\,(s_{t+k+1} - s_0)^2. \quad (6)$$

where ρ is a weight for the *reference trajectory*. The optimal solution of (6) provides a *reference trajectory*. The (deterministic) dynamic programming problem (5) may be solved using discrete dynamic programming.

7.3. Decision support systems

In many cases, a decision making problem with important enough consequences has to be solved repeatedly through time as in, e.g., reservoir operations. In such case, we may encode the knowledge obtained in a Decision Support System (DSS), which is a computer system that supports the decision making process, helping decision makers to explore the implications of their judgments in order to make decisions based on understanding of the underlying assumptions, available information and consequences of plausible decisions.

One example is BayRes (see Salewicz *et al*, [2002]) which is a decision support system for reservoir operations, supporting all phases of the decision making process, including:

- a module to load historical and new data from their sources, typically a text file, into the system,
- a module to build a forecasting model, facilitating the construction of the model, say for example, specification of the DLM process or data analysis,
- a module to build a preference model, including the computation of a *reference trajectory*,
- an optimizer to solve problems of type (5),
- several sensitivity analysis tools.

The features of the application and its operation are described and controlled through several windows. From the main menu, you can launch one of five different modules depending on the stage you are in the process: loading data, building the forecasting model, building the preference model, optimizing or making sensitivity analysis. As this is not a sequential process, each module can be launched at a time with no order predefined except, of course, that optimizing and sensitivity tools require a previous model to work on it.

As an example, the preference module has the following steps: 1) Introducing the number of attributes and their characteristics (maximum value, minimum value,...), 2) Obtaining some values of the utility function through lotteries for each attribute, 3) Fitting the different utilities function to a concave-convex or convex-concave family depending on the monotonicity properties, 4) Obtaining the weights of each utility function for the general additive multiattribute utility function, 5) Construction of the general multiattribute additive utility function.

As a result of the optimization module, BayRes shows different charts for the control policy in the next k periods. BayRes provides capabilities to modify the suggested control policy as external input arrives.

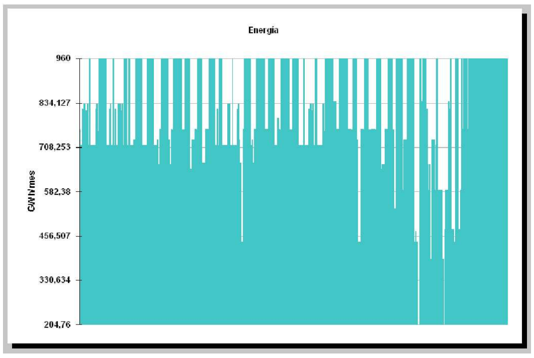

Figure 5. Example of a control policy obtained after the optimization process.

For example, its forecasting module allows for interventions, illustrating a principle of *management by exception* mentioned above. Examples would include a sudden rainfall, a big release from an upstream reservoir, or the detection of a wet period. In such cases, the system is open to external (user-initiated and user-performed) interventions, typically by inclusion of additional subjective information.

8. Synthetic hydrology

As we have mentioned, another purpose of statistical analysis within water resources management is synthetic hydrology. Once we have identified and estimated a model for a hydrologic phenomenon of interest, we may use it to simulate sequences of observations that mimic the phenomenon behavior for engineering design or analysis. Note however, that the standard method of i) estimating model parameters; ii) plugging the estimates into the model; and iii) using the estimated model for prediction or simulation; typically greatly underestimates uncertainty in predictions, since the uncertainty in the model parameters is not taken into account. In exchange, in Bayesian forecasting the output is the entire predictive distribution, not just summaries. Thus, we can use this distribution for any purposes taking expectations or simulating the future.

To wit, suppose we use a simulation model to evaluate some engineering design. There will typically be some design parameters λ and some (random) input parameters θ, some fixed costs $f(\lambda)$ and variable costs $g(\lambda, \theta)$ and we aim at minimizing (in λ)

$$f(\lambda) + E(g(\lambda, \theta)).$$

The standard approach would estimate $\widehat{\theta}$ and use simulation to estimate

$$f(\lambda) + E(g(\lambda, \widehat{\theta})),$$

ignoring uncertainty about θ. Alternatively, we could have a sample $\{\theta_i\}$ from the posterior distribution of θ and based our cost estimate on the sample

$$\{f(\lambda) + E(g(\lambda, \theta_i))\},$$

acknowledging uncertainty about θ. Further details on the impact of Bayesian methods on Monte Carlo and discrete event simulation may be seen in Palomo and Rios Insua [2003].

9. Discussion and conclusions

The intrinsically uncertain nature of hydrological processes makes Bayesian analysis natural within this field, whenever statistical problems are considered. We have tried to convey the relevance of such studies through various examples, covering wide areas of interest in hydrology including modelling and forecasting time–space features, decision making issues relating reservoir operations and simulation models to help engineering design or analysis.

Space prevents from dealing with many other relevant topics. For example, hydrologists have been traditionally interested in changepoint detection problems, to study bursty phenomena like waterflow of rivers, e.g. in order to prevent floods. Vellekoop and Clark [2002] provide a recent Bayesian analysis of the problem. We have illustrated decision making issues with reservoir operations, however the ideas are relevant in many other problems; for example, Sansó and Müller [1998] use the augmented simulation model to redesign a rainfall monitoring network. Water metering problems have been studied from a Bayesian perspective in Pasanisi *et al* [2002]. Finally, recall that we have stressed water quantity issues; some Bayesian references in water quality include Dilks *et al* [1992], Solow and Gaines [1995] and Dominici *et al* [1997].

Acknowledgement. Research supported by grants from CICYT, CAM, URJC and the Iberdrola Foundation. We are grateful to discussions with K. Salewicz, J. Berger, P. Müller and J. Martin.

References

[1] Beran, J. (1994). *Statistical Methods for long memory processes*, Chapman & Hall.

[2] Berger, J. and Ríos Insua, D. (1998). Recent developments in Bayesian Inference, with applications in Hydrology, in Parent *et al.* (eds.) *Statistical and Bayesian Methods in Hydrology*, UNESCO Press.

[3] Bertolino, F. and Racugno, W. (1996). *Bayesian Model Comparison*, Pittagora, Edittrice.

[4] Bertsekas, D. and Tsitsiklis, J. (1996). *Neuro-Dynamic Programing*, Athena Scientific.

[5] Birge and Louveaux, (1997). *Introduction to Stochastic Programming*, Springer.

[6] Bras, R. and Rodriguez Iturbe, I. (1993). *Random functions and Hydrology*, Dover.

[7] Carlin, B., Kadane, J., Gelfand, A. (1997). Approaches for optimal sequential decision analysis in clinical trials. *Tech.Rep.* U. Minnesota.

[8] Clarke, R. (1994). *Statistical modelling in Hydrology*, Wiley.

[9] Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer.

[10] Coles, S. and Tawn, J. (1996). A Bayesian analysis of extreme rainfall data. *Applied Statistics*, **45**, 463–478.

[11] Diebolt, J. and Robert, C. (1994). Estimation of finite mixture distributions through Bayesian sampling. *J. Roy. Statist. Soc. B*, **56**, 163–175.

[12] Dilks, D., Canale, R. and Meier, P. (1992). Development of Bayesian Monte Carlo techniques for water quality model uncertainty. *Ecological Modeling*, **62**, 149–162.

[13] Dominici, F., Parmigiani, G., Reckhow, K. and Wolpert, R. (1997). Combining information from related regressions. *Journal of Agricultural Biological and Environmental Statistics*, **2**, 313–332.

[14] French, S. and Ríos Insua, D. (2000). *Statistical Decision Theory*, Arnold.

[15] Green, P. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82**, 711–732.

[16] Helsel, D. and Hirsch, R. (1992). *Statistical methods in water resources*, Elsevier.

[17] Hurst, H. (1951). Long term storage capacities of reservoirs, *Trans. A.S.C.E.*, **116**, 776–808.

[18] Mandelbrot, B. and Wallis, J.(1968). Noah, Joseph and Operational Hydrology. *Water Resources Research*, **5**, 909–18.

[19] Marin, J. M., Müller, P. and Ríos Insua, D. (2002). Spatial clustering with multi-level mixture models, *Tech. Rep.* URJC.

[20] Menchero, A., Montes, R., Müller, P. and Ríos Insua, D. (2002). Bayesian Neural networks for nonlinear autoregression, *Tech. Rep.* URJC.

[21] Muster, H. and Bardossy, A. (1998). Rainfall forecasts for flood management in River Basins. in Parent *et al.* (eds) *Statistical and Bayesian Methods in Hydrology*, UNESCO Press.

[22] Palomo, J. and Ríos Insua, D. (2003). Bayesian discrete-event simulation, with application to workflow modeling. *Tech. Rep.*, URJC.

[23] Parent, E., Hubert, P., Bobée, B. and Miguel, J. (1997). *Statistical and Bayesian Methods in Hydrological Sciences*. UNESCO.

[24] Pasanisi, A., Parent, E., Arnac, P. and Paquet, F. (2002). Describing the Ageing of Water Meters in Vivendi Water Distribution Networks by Dynamic State Model. *Bayesian Statistics 7*, Tenerife.

[25] Petris, G. (1997). *Bayesian analysis of long memory time series*, Ph.D. Thesis. ISDS, Duke University.

[26] Petris, G. and West, M. (1998). Bayesian time series modelling and prediction with long-range dependence. *Tec. Rep.* ISDS Duke University.

[27] Richardson, S. and Green, P.J. (1997). On Bayesian analysis of mixtures with an unknown number of components. *J. Roy. Statist. Soc. Ser. B*, **59**, 731–792.

[28] Ríos Insua, D., Salewicz, A., Müller, P. and Bielza, C. (1997). Bayesian methods in reservoir operations: The Zambezi River Case. French and Smith eds. *Bayesian methods in reservoir operations*. Arnold.

[29] Ríos Insua, D. and Salewicz, A. (1995). The operation of Lake Kariba. *J. of Multicriteria Decision Analysis*, **4**, 203–222.

[30] Rogers, D. F., Plate, R. D., Wong, R. T. and Evans, J. R. (1991). Aggregation and Disaggregation techniques and Methodology in Optimization, *Oper. Res.*, **39**, 553–582.

[31] Salewicz, K., Ríos Insua, D., Palomo, J. and Nakayama, M. (2002). Building the Bridge Between Reservoir Management and Decision Analysis. *Tech. Rep.* URJC.

[32] Sansó, B. and Müller, P. (1998). Redesigning a network of rainfall stations In Case Studies in Bayesian Statistics, vol. **IV**, 383–394, Springer-Verlag.

[33] Solow, A. and Gaines, A. (1995). An empirical Bayes approach to monitoring water quality, *Environmetrics*, **6**, 1–5.

[34] Stephens, M. (2002). Bayesian analysis of mixture models with an unknown number of components: an alternative to reversible jump methods, *J. Royal Statistical Society*.

[35] Stroud, J., Müller, P. and Sansó, B. (2001). Dynamic models for spatiotemporal data, *J. Roy. Statist. Soc. Ser. B*, **63**, 673–689.

[36] Vellekoop, M. and Clark, J. (2002). A nonlinear filtering approach to changepoint detection problems: Direct and Differential-Geometric Methods. *Tec. Rep.* U. Twente.

[37] Virto, M. (2002). *Métodos Montecarlo en análisis de decisiones*. Ph.D. Thesis. ETSI Industriales. UNED.

[38] West, M. and Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. Springer.

[39] White, D.J. (1993). *Markov Decision Processes*, Wiley.

[40] Yakowitz, S. (1982). Dynamic programming applications in water resources, *Water Resources Research*, **18**, 673–696.

[41] Yeh, W. (1985). Reservoir Management and Operation Models: A State-of-the-Art Review. *Water Resources Research* **21**(12), 1797–1818.

D. Rios Insua, R. Montes Díez, J. Palomo
Statistics and Decision Sciences Group
Universidad Rey Juan Carlos
Calle Tulipán s/n, Móstoles, 28933 Madrid, Spain