23 A Bayesian View of Rainfall–Runoff Modelling: Alternatives for parameter estimation, model comparison and hierarchical model development

ASHISH SHARMA, LUCY MARSHALL & DAVID NOTT

INTRODUCTION

The period 2003–2012 was recently named as the International Association for Hydrological Sciences (IAHS) Decade for Predictions in Ungauged Basins (PUB) (Sivapalan *et al*., 2003), and a major initiative was launched to promote development of strategies offering improvements on the current knowledge in this important area of research. The PUB initiative is one amongst many such efforts that have attempted to draw attention to a problem that has been possibly the biggest challenge in hydrology. While many advances have been made, predicting flow in ungauged catchments remains as difficult as before.

The traditional approach to PUB involves selecting a hydrological model that works well across catchments, and calibrating its parameters as functions of relevant catchment attributes (area, slope, river length). Once developed, it can be applied in an ungauged catchment using only its attributes and the incident rainfall record. Examples of such approaches are many (see Pilgrim, 1997; Campbell & Bates, 2001, and the references therein), and their limitations are few as long as three basic assumptions are satisfied. These are: (i) the hydrological model is applicable for all flow regimes (for example, both rising and falling limbs of the hydrograph, both dry and wet antecedent conditions); (ii) the model is applicable across the many catchments it is intended for use in; and (iii) the parameter regionalization is appropriate, or, the catchment attributes selected to formulate the regionalization model are capable at representing the variability in model parameters across a range of catchments. Satisfying these assumptions is not simple, with violations known to lead to a multitude of errors resulting in incorrect designs and poor management decisions.

We present here a Bayesian view of rainfall–runoff model development and comparison, with suggestions on their usefulness for predictions in ungauged basins. We then proceed to present a new class of rainfall–runoff models (called a Hierarchical Mixture of Experts or HME) where the model is dynamic and switches between alternate states in a probabilistic fashion depending on catchment antecedent conditions. We emphasize that such a model where a “soft” transition from one state to another is allowed, is difficult to accommodate in a non-Bayesian framework, and highlights one of the many advantages of using Bayesian methods. These methods are becoming an increasingly popular means for assessing parameter and model uncertainty. Bayesian inference addresses the issue of uncertainty by formulating a probability distribution (the posterior distribution) on the model unknowns (parameters and model outputs) that describe the uncertainty after taking into account the observed data. While recent years have seen an increase in applications of Bayesian methods to rainfall runoff modelling and parameter estimation (Kuczera, 1983; Romanowicz *et al.*, 1994; Freer *et al.*, 1996; Kuczera & Mroczkowski, 1998; Kuczera & Parent, 1998; Bates & Campbell, 2001; Theimann *et al.*, 2001), the use of these techniques has traditionally been limited due to computational difficulties. The advent of Markov Chain Monte Carlo (MCMC) methods has helped address some of these computational difficulties. MCMC schemes explore the posterior distribution by generating a random process where the stationary distribution is the posterior distribution of the parameters. It is no surprise that MCMC methods have been adopted widely in hydrological studies in recent years (Romanowicz *et al.*, 1994; Kuczera & Parent, 1998; Bates & Campbell, 2001; Campbell & Bates, 2001; Marshall *et al.*, 2004).

What follows is a brief background on Bayesian methods and MCMC sampling, after which a Bayesian basis for model comparison that is computationally viable and naturally accounts for model complexity is described. A new rainfall–runoff modelling framework, the Hierarchical Mixtures of Experts (HME), is presented next, and applied to simulate daily streamflow in selected Australian catchments of varying attributes. We conclude our discussion with a summary of the main points raised and suggestions on how the proposed HME rainfall–runoff modelling framework can be used for predictions in ungauged basins.

BAYESIAN METHODS AND MCMC SAMPLING

Bayesian methods provide a framework within which pre-existing knowledge about the parameters of a model can be combined with observed data and the model output. This results in a probability distribution on the parameter space (called the posterior distribution), that summarizes the uncertainty in the parameters ascertained based on the combination of pre-existing (or prior) knowledge and the sampled data values. The complications and uncertainties in accurately determining the parameters in conceptual rainfall–runoff models have led to the development of schemes for assessment of parameter uncertainty in a Bayesian framework.

Let the observed flow for a catchment at time t be Q_t. A rainfall–runoff model can then be cast as:

$$Q_t = f(x_t;\theta) + \varepsilon_t \tag{1}$$

where $f(x_t;\theta)$ is the corresponding model output for time t, x_t is the set of model inputs at time t (such as rainfall and evapotranspiration), θ is the set of unknown parameter values and ε_t is an error term.

Prior to considering the observed data, the current knowledge about the parameter set is summarized in a distribution $p(\theta)$ (called the prior distribution). The decision to use a particular prior should be based on any available knowledge about the parameter. This may include comparing the function of the parameter within the model to the function of similar parameters within established models. Estimated optimum values of parameters from previous studies may be helpful for establishing the prior, and constraints on the model parameters should also be taken into account.

The posterior distribution $P(\theta \mid Q)$ of the parameter set may be found through the application of Bayes' theorem:

$$P(\theta \mid Q) = \frac{p(\theta)P(Q \mid \theta)}{P(Q)} \tag{2}$$

where $P(\theta \mid Q)$ is the likelihood function summarizing the model for the data given the parameters and $P(Q)$ is a proportionality constant. The posterior distribution assumes a shape similar to the prior when available data are limited. For large sample sizes, on the other hand, the posterior is influenced more by the data than by the assumed prior distribution. One of the reasons Bayesian methods are so attractive for use in hydrological settings is that they can incorporate prior "expert" knowledge through the assumed prior distribution, which is then combined with actual data to lead us to the final posterior distribution.

Consider, for instance, the prior and posterior distribution in Fig. 23.1 for the parameter K (representing the baseflow recession rate) of the Australian Water Balance Model (Boughton, 2004), using 11 years of daily runoff data from the 52 km^2 Bass River catchment in Australia. Note that the posterior distribution converges to a range that has a relatively low prior probability of occurrence, again highlighting the usefulness of Bayesian techniques in allowing observed information to control the end results even if prior knowledge indicates otherwise. Note also that the same would not have been the result were the sample significantly shorter than the 11 years of data used. A posterior distribution not overly different to the prior would have been the likely outcome as the information in the data would not have been sufficient to influence the information contained in the assumed prior distribution.

While the conceptual elegance of the Bayesian framework is undeniable, the computational effort required to implement Bayesian procedures has hindered their application. This effort is often due to the difficulty in evaluating moments and other statistics of interest from the posterior distribution in equation (2), which is a multi-dimensional function expressed in terms of the unknown model parameters θ, and can

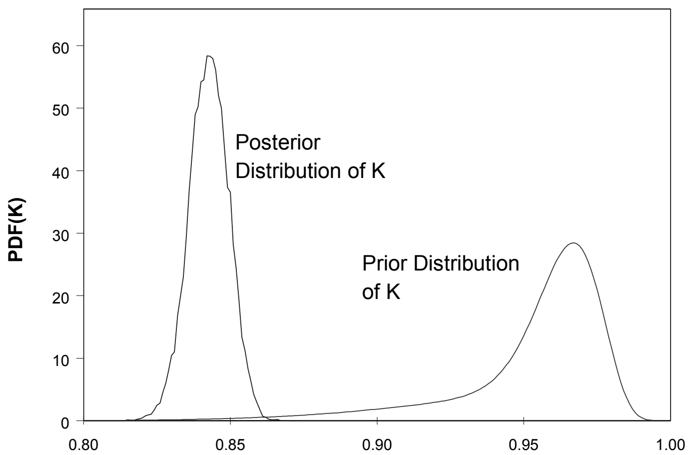

Fig. 23.1 *Posterior distribution and prior distribution for the K parameter of the Australian Water Balance Model.*

assume a form that is not always easy to integrate. Markov Chain Monte Carlo (MCMC) sampling is a relatively recent addition to the Bayesian statistical literature and provides a simple and elegant way around the computational difficulties noted above. The aim of MCMC sampling is to generate samples of the parameter values from the posterior distribution by simulating a random process that has the posterior distribution as its stationary distribution. While there exist many different MCMC sampling algorithms, the general form of commonly used algorithms is as follows:

1. Start the simulation at time $t = 0$ with an arbitrarily chosen initial parameter vector $\theta = \theta_0$.
2. (a) Generate a proposed value θ^* for θ from a proposal distribution depending on the current value θ_t of θ.
   (b) Compute an acceptance probability α that controls whether the proposed value θ^* is accepted or not.
   (c) Accept $\theta_{t+1} = \theta^*$ with probability α. Otherwise $\theta_{t+1} = \theta_t$.
3. Increment time t and repeat step 2.

The traditionally difficult step of evaluating moments and other statistics from the posterior distribution becomes relatively simple once a trace of sampled parameter values has been generated. However, using sampled parameter traces for statistical evaluation assumes that the traces are representative of the estimated posterior distribution, and that for a given starting value θ_0 for the model parameters and proposal distribution, the sequence θ_t has converged to its stationary distribution over the length of the simulation. Typically, a "burn-in" period of initial iterates of the sequence θ_t must be discarded where values are influenced by the initial value θ_0 and are not typical of the posterior distribution. Deciding on the length of the burn-in sequence is one difficult issue in implementation of an MCMC scheme. The choice of proposal is also crucial in obtaining a good MCMC algorithm. If only very small changes are proposed to the current state, then exploration of the posterior distribution can be very slow. On the other hand, if very large changes are proposed to the current parameter values this may result in values inconsistent with the posterior distribution, resulting in turn in a low acceptance rate for proposals and once again slow exploration of the posterior distribution. The problem of choosing good proposal distributions becomes even more difficult in the situation where two or more of the parameters are highly interdependent. Readers are referred to Gelman *et al*. (1995) and Gilks *et al*. (1996) for more detailed discussion of the issues involved in devising good MCMC sampling algorithms.

Adaptive Metropolis and associated advantages

The Adaptive Metropolis (AM) algorithm (Haario *et al*., 2001) is a variation on the conventional MCMC algorithm outlined above. It is designed to avoid the necessity of specifying the proposal density, instead adopting an approach where the proposal is assumed to be a multivariate normal with a covariance that is estimated based on the sampled parameter values. Hence, at time t the proposal is specified with mean given by the current value, $N(\theta_t, C_t)$ where C_t is the proposal covariance. The covariance C_t has a fixed value for the first iterations and is updated after some minimum number t_0 of iterations has elapsed:

$$C_t = \begin{cases} C_0 & t \le t_0 \\ s_d Cov(\theta_0, \cdots, \theta_{t-1}) + s_d \varepsilon I_d & t > t_0 \end{cases} \tag{3}$$

where C_0 is the initial fixed covariance, ε is a small parameter chosen to ensure C_t does not become singular; and s_d is a scaling parameter to ensure reasonable acceptance rates of the proposed states. As a basic guideline, Haario *et al*. (2001) suggest choosing s_d for a model of dimension d as $(2.4)^2/d$.

The calculation of the proposal covariance requires the definition of an arbitrary initial covariance, C_0. To automate the process, the initial covariance is set at the covariance of the parameters under their prior distributions. The steps involved in implementing the AM algorithm then become:

1. Initialize $t = 0$ and set C_0, the covariance matrix for the parameters under the prior distributions.
2. (a) Select C_t for the current iteration number t using equation (3).
   (b) Generate a proposed value θ^* for θ where $\theta^* \sim N(\theta_t, C_t)$.
   (c) Calculate the acceptance probability, α, of the proposed value:
   $$\alpha = \min\left\{1, \frac{p(y \mid \theta^*) p(\theta^*)}{p(y \mid \theta_t) p(\theta_t)}\right\} \tag{4}$$
   where $p(y \mid \theta)$ is the likelihood function, $p(\theta)$ is the prior distribution of θ.
   (d) Generate $u \sim \underline{U}[0,1]$
   (e) If $u < \alpha$, accept $\theta_{t+1} = \theta^*$, otherwise set $\theta_{t+1} = \theta_t$.
3. Repeat step 2.

For parameter values that were proposed outside the parameter constraints, the acceptance probability is set to zero.

A major advantage of the AM approach is that the entire parameter set is updated simultaneously. This reduces computation time and complexity, and is especially of use when the parameters are highly interdependent. This is in contrast to most classical algorithms where separate proposal distributions need to be specified for each sampled parameter, the complexity of this specification increasing when the parameters are sampled as a block. Marshall *et al.* (2004) compares the performance of the Adaptive Metropolis algorithm in the context of rainfall–runoff modelling, with other sampling algorithms used routinely in hydrology, and find that the Adaptive Metropolis algorithm offers significant advantages over the others in speed and in the ease with which it can be implemented. Readers are referred to Marshall *et al*. (2004) for further details, limitations, and suggestions for use with high dimensional models.

BAYESIAN MODEL SELECTION

The Bayesian approach to model selection explicitly describes the issues of model uncertainty by describing parameter and model uncertainty in a probabilistic sense. In comparing two models, the traditional approach requires calculation of the Bayes factor, which represents the odds of one model *versus* the other after observing the data assuming equal prior model probabilities. Say we wish to choose a model from the set of models $M = \{M_1, \ldots M_n\}$, given data y for implementing the model. Let $p(M_i)$ be the prior probability of model M_i, and θ_i be the set of uncertain model parameters corresponding to model M_i. The traditional approach to Bayesian model selection proceeds by pair-wise comparison of the models through their posterior probability ratio:

$$\frac{P(M_1 \mid y)}{P(M_2 \mid y)} = \frac{P(M_1)}{P(M_2)} \times \frac{m(y \mid M_1)}{m(y \mid M_2)} \tag{5}$$

where $P(M_i \mid y)$ is the posterior probability of model M_i, $P(M_i)$ is the prior probability of model M_i, and $m(y \mid M_i)$ is the model's marginal likelihood. The second term on the right hand side of equation (5) is known as the Bayes factor. The marginal likelihood $m(y \mid M_i)$ is defined through the integral:

$$m(y \mid M_i) = \int f(y \mid \theta_i, M_i)\, p(\theta_i \mid M_i) \mathrm{d}\theta \tag{6}$$

over the parameter space, which must usually be estimated numerically. The difficulty in estimating the marginal likelihood in (6) lies in the fact that the integration of the likelihood function is performed over the assumed prior distribution and not the posterior distribution. Were that not the case, the integral could have been approximated as an average of the likelihood values sampled via the MCMC runs. To get around this difficulty, Chib & Jeliazkov (2001) proposed an approximation which involves first identifying the parameter set that maximizes the posterior distribution, then using the local probability density about this optimal set as the basis for approximating the value of the marginal likelihood. Readers are referred to Marshall *et al.* (2005) for details on the approach and an illustration of how the method is a more stable and suitable choice for a generic model comparison problem.

There are other Bayesian approaches which analytically approximate the marginal likelihood. The BIC criterion of Schwarz (1978) is an asymptotic approximation to the Bayes factor. The criterion holds that the model log-marginal likelihood is approximately –0.5 BIC, where (if N is the size of the sample):

$$\text{BIC} = -2(\text{log maximized likelihood}) + (\log N)(\text{number of parameters}) \tag{7}$$

The BIC and other associated approximations have the distinct advantage of simplicity over the Bayes factor, since they do not require numerical methods for approximating the integral in the marginal likelihood. In cases where the prior information is small relative to the information provided by the data, the BIC can yield results consistent with estimating the Bayes factor (Kass & Raftery, 1995). The BIC also accounts for overfitting by tending to favour simple models, as the criterion penalizes according to the number of parameters in a model. However, there are inherent risks in using the BIC alone. The method requires some regularity conditions on the model and can differ from the use of Bayes factors if the available data are limited. An application of the marginal likelihood, the BIC, and the commonly used Nash-Sutcliffe coefficient of efficiency for comparing across alternate modelling configurations in both synthetic and real settings is presented in Marshall *et al.* (2004). Results confirm the inadequacy of the Nash-Sutcliffe coefficient to select simpler models if a more complex alternative exists, and the efficiency of both the marginal likelihood and the BIC in overcoming this limitation. For this reason the BIC is used as the basis for the model comparison in the results presented in the next section. Note that while explicit computation of the Bayes factor has not been performed in the results that follow, if each candidate model is assumed to have an equal prior probability of selection, the Bayes factor can be computed as the ratio $\mathrm{e}^{2BIC_1} / \mathrm{e}^{2BIC_2}$, where BIC_1 and BIC_2 are the respective BIC estimates for models 1 and 2. It should also be noted that if expert knowledge favours one model over another, the above ratio should be multiplied with the ratio of the respective model prior probabilities to arrive at a more meaningful estimate of the Bayes factor.

HIERARCHICAL MIXTURES OF EXPERTS

A Hierarchical Mixtures of Experts (HME) model (Jordan & Jacobs, 1994; Peng *et al.*, 1996) aims to combine the output from two or more models in a probabilistic sense. Consider the single-level, two-component HME in Fig. 23.2. Incident rainfall, evapotranspiration (ET) and the parameters that describe the framework serve as inputs to each HME component model (denoted Models 1 and 2 in Fig. 23.2). The output from each HME component model is an ensemble of streamflow hydrographs that aims to represent the uncertainty in the parameters used. As each HME component model could be any suitable rainfall–runoff model, rigidity in the model structure limits representation of the varied flow mechanisms present. This is addressed by combining the HME component outputs in a probabilistic sense using a *gating* function (often specified to be a logistic regression model) that estimates the probability with which each component model may be used. If, say, this probability equals 0.3 for component 1 for a given time step, 30% of the ensemble of HME hydrographs would originate from HME component 1, and the remaining 70% from HME component 2. This probability would depend on the state of the predictors used in formulating the gating function, and would change from one time step to the next.

Estimating the probability of selecting each HME component model is central to the effectiveness of the proposed HME framework. An intuitive choice for the predictors used to estimate this probability is the antecedent soil moisture state, the time of the year, or a combination of the two. This choice, however, needs to be made through an elaborate comparison that considers the various mechanisms that result in a "switch" in the flow generation process. Use of a model comparison approach such as the Bayes factor provides a sensible basis for deciding the predictors to be used.

Specifying the form of the gating function to be used is the other important aspect of developing an effective HME model. A simplistic choice for the gating function is a logistic function, but more complex choices involving nonlinear or nonparametric functional terms (such as a Generalised Additive Model, GAM; Hastie *et al.*, 2001)

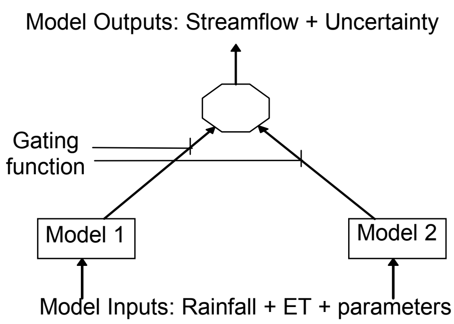

Fig. 23.2 *A single-level two-component Hierarchical Mixture of Experts model.*

offer greater functionality. An exhaustive model comparison should be performed in order to identify the appropriate gating function. The use of Bayes factors as the basis for this comparison is again recommended.

Application of 2-component HME for rainfall–runoff modelling

A single-level 2-component HME, similar to that in Fig. 23.2, was applied to simulate daily streamflow from 10 catchments of varying sizes and attributes, selected from different parts of Australia. Table 23.1 provides details on the rainfall–runoff data for each of the 10 catchments (Peel *et al.*, 2001), referred to as catchments A–J for brevity in the rest of this document. For simplicity, the two component models were both specified to be the reduced three-parameter Australian Water Balance Model (AWBM; Boughton, 2004), a conceptual rainfall–runoff model consisting of three inter-connected surface stores. The AWBM has been used widely to simulate daily runoff in Australia, using daily rainfall and daily evapotranspiration as the two hydrological inputs. For the purpose of this presentation, a simplified version of the AWBM was used (Fig. 23.3) (Boughton, 2004), the end model needing the specification of only three parameters (K, BFI and S in Fig. 23.3) that are calibrated using daily rainfall–runoff records. The "switch" from one component to the other was accomplished through a simplified gating function, a fixed transition probability that depends on the model state at the previous time step. The performance of the model is documented in the last two columns of Table 23.1. It is easy to observe the advantages in using the HME formulation in comparison to the use of a single AWBM. One should note specifically the improvements in terms of the BIC, the Bayesian measure of model performance that accounts for the added complexity in the HME configuration. These results are all the more encouraging given the considerable simplification used in the experiment design (a naïve predictor variable choice, a transition probability based gating function, and the three-parameter AWBM as the HME component model).

The optimal parameter values for both the AWBM and the HME, along with a segment of the observed and modelled hydrographs for catchment J are presented in Table 23.2 and Fig. 23.4, respectively. Note that the dominant component (conditional probability greater than 0.5) is indicated as M1 or M2 in the HME results. Note care

Table 23.1 *Application of AWBM and HME to selected catchments. Under location: (xx–C) denotes coastal; MDB denoted Murray-Darling basin. Catchment J results are presented in detail later. Refer to Peel et al. (2001) for additional details on the data used.*

	Station number	Location	Area (km^2)	Annual rainfall (mm)	Annual runoff (mm)	AWBM: Nash coeff.	AWBM: BIC	HME: Nash coeff.	HME: BIC
A	117002	Qld (NE-C)	260	1195	311	0.74	-495	0.78	239
B	136202	Qld (NE-C)	640	898	84	0.58	1345	0.67	2332
C	204025	NSW (SE-C)	135	1793	755	0.81	-2273	0.83	-1808
D	410033	NSW (MDB)	1891	882	134	0.41	-4677	0.64	2290
E	239519	SA (Gulf)	1130	607	26	0.51	9277	0.67	12173
F	410730	ACT (MDB)	148	1139	318	0.66	-6825	0.73	-3687
G	601001	WA (SW-C)	1610	390	3	0.52	19941	0.59	20723
H	8210007	NT (N-C)	260	1357	409	0.76	-3033	0.80	-2338
I	314207	TAS	500	1775	1014	0.81	-14881	0.82	-14496
J	*205014*	*NSW(SE-C)*	*51*	*2036*	*1114*	*0.79*	*-3751*	*0.81*	*-3379*

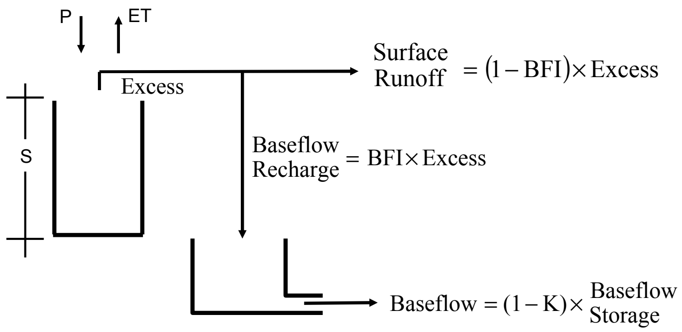

Fig. 23.3 *A simplified AWBM with three parameters: S (surface store capacity), K (recession) and BFI (base-flow index). Note that the surface storage S consists of three sub-stores whose sizes are expressed as functions of S.*

Table 23.2 *Optimal parameter values for AWBM and HME models for catchment J.*

AWBM			HME Component 1			HME Component 2		
K	*BFI*	*S*	*K*	*BFI*	*S*	*K*	*BFI*	*S*
0.96	0.53	64	0.97	0.71	53	0.96	0.44	79

fully the implications of the parameter values for each HME component state in reference to the AWBM structure illustrated in Fig. 23.3. A high *BFI* (HME component 1) implies more rain gets stored in the baseflow storage, or, a small fraction of the rainfall enters the stream as direct runoff. A low *BFI* (HME component 2) implies the reverse. Also note the mechanism that forces each component to become dominant (Fig. 23.4). The "slow-flow" process (HME component 1) generally dominates when a rainfall event

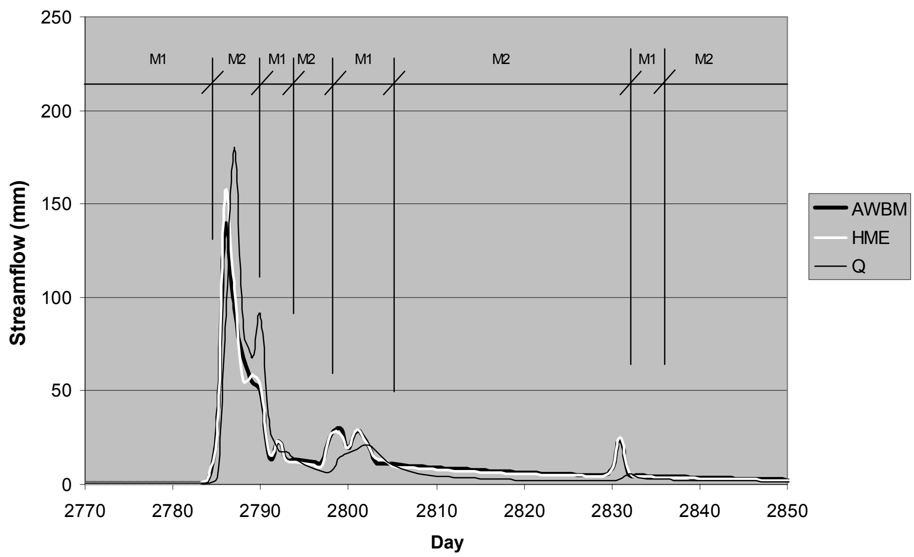

Fig. 23.4 *Observed and simulated hydrographs for catchment J. M1 and M2 represent the dominant HME component for each time step.*

has occurred, while the "quick-flow" process takes over towards the recession limb of the hydrograph. While these results are based on simplistic predictor and gating function choices, they do illustrate clearly that there exist two distinct states in the HME model.

While the above results pertain to a single catchment, the analysis was conducted for a total of 10 Australian catchments, all representing varying climatic, topographic and hydrological regimes. It was notable that in all the catchments analysed, the two dominant states corresponded to the "quick-flow" and "slow-flow" mechanisms that were noticed in the detailed results for catchment J in Table 23.2. Hence the differences were not related to the dissimilarity in the two states, but the residence times and the overall probability with which each state existed. This brings us to the next question: How should such a multiple-state HME modelling framework be used for prediction in ungauged basins?

IMPLICATIONS FOR PUB

The HME results in Tables 23.1 and 23.2 suggest that there are two distinct states in the flow mechanism. The application across a range of catchments indicated that these states correspond to a slow-flow and a quick-flow mechanism in each of the catchments investigated. The main difference across the catchments appeared to be the transition or the switching mechanism between the two states, with differences in both the residence time for each state, and the frequency with which a state might be selected. As a two-component HME requires more than twice the number of parameters compared to either component model, a traditional regionalization of the parameters for use in ungauged catchments is both onerous and unstable. Hence extension of the HME framework through a traditional regionalization approach (such as a Bayesian regionalization approach by Campbell & Bates, 2001) is not an attractive option to pursue.

Our suggestion for extending the HME framework to PUB works off the assumption that a probabilistic combination of a finite number of HME component models is sufficient to represent the runoff response across a variety of catchments. In other words, we assume that, instead of changing model parameters from one catchment to another, one can change the switching mechanism or the gating function that specifies the transition probability with which each component model is to be used at a given time step. Based on the results presented in Table 23.1, we hypothesize that the parameters of the HME gating function vary from catchment to catchment, while the parameters of each HME component are reasonably stable (or can be expressed as functions of bulk catchment attributes in a reasonable stable manner). Consequently, we feel that if the HME framework were to be extended to ungauged catchments, regionalization of only the HME gating function would be necessary.

It must be noted that, depending on the HME component model used, the model parameters may need to be specified in terms of specific catchment attributes (such as the catchment area). While this is a form of regionalization, we expect this to be accomplished more through the physics rather than statistics. It must also be noted that the two-component HME illustrated in Fig. 23.2 and used in developing the results for our presentation, is likely to be insufficient at representing the various hydrological mechanisms that need to be simulated. The optimal number of components and levels (layers) (if needed) of the HME would need to be ascertained through elaborate comparisons using a rationale such as the Bayes factor.

CONCLUSIONS

Two frequent applications in hydrology are design flood estimation and water resources management. The vast majority of catchments where these are needed are ungauged. Hence prediction in ungauged catchments is amongst the more significant needs of applied hydrology, and an area that requires major advances to bridge the limitations of current methods. The two most strident criticisms of existing alternatives for PUB are:

- the requirement that a single model be used across the many catchments it is intended for use in, and
- that the parameter regionalization be appropriate, or, the catchment attributes selected to formulate the regionalization model be capable at representing the variability in model parameters across a range of catchments.

In addition to these issues, a generic criticism of most rainfall–runoff models is the limited applicability of the hydrological model across all flow regimes that exist. Such regimes could include, for example, both rising and falling limbs of the hydrograph, both dry and wet antecedent conditions, or both saturation excess and infiltration excess flow mechanisms. It is difficult to find a rainfall–runoff model that can be specified to simulate such variability, let alone apply it in catchments based on crude relationships to estimate the handful of parameters it involves.

The Bayesian hierarchical approach presented in earlier sections aims to do away with the following two assumptions implicit in traditional alternatives for PUB:

(i) using a single model that is applicable through all states of the flow regime, and
(ii) ensuring that the same model applies across all catchments it is intended for use in.

The use of a "switching" mechanism whereby the model is specified to exist in more states than one, enables applications across a range of catchment antecedent conditions. The use of multiple component models representing each hydrological state, and the reduced need for regionalization of the resulting parameters, ensures that varied rainfall–runoff models may be used as the HME components, as long as the main hydrological mechanism simulated through that component is appropriately represented. While it is possible that some of the parameters that describe each HME component model may need some level of regionalization, we feel the focus of the regionalization will be confined to the HME gating function that specifies the probabilities and residence times associated with each HME component. Hence it will be possible to apply the HME framework to an ungauged catchment by specifying the form of the HME gating function through the regionalized relationships, and then simulating the streamflow as an ensemble of hydrographs proportionally simulated from each HME component model depending on the transition probability estimated at each time step.

It should be mentioned that the specification of the HME model outlined above is accomplished because of the flexibility that the Bayesian framework offers. Bayesian methods provide us with a simple basis for sampling plausible parameter values and establishing the associated posterior distributions. Consequently, they enable an effective basis for establishing the confidence limits associated with model outputs, and enable also a thorough basis for comparing across alternate models based on their respective predictive performances. While Bayesian methods were difficult to implement without developing exacting analytical solutions, the advent of Markov Chain Monte Carlo (MCMC) sampling has simplified their applicability considerably. A recent study by Marshall *et al*. (2004) presents the basis for MCMC sampling and a comparison between alternate sampling strategies, their pros and cons, and an assessment of their applicability when used in the context of rainfall–runoff modelling.

It should also be mentioned that while Bayesian methods are becoming increasingly popular as a tool to use to investigate parameter and model uncertainty, their use requires assuming the nature of the error distribution that the model produces. This error distribution is the basis for the likelihood function the posterior probability distributions are derived from. Characterization of the error is both important and imperative at the successful specification of the model. The common approach to formulating the likelihood is to assume model errors as independent and normally distributed. There are two problems in this. Firstly, model errors are seldom independent with high serial correlations due to consistent over/under estimation of the falling limb of the hydrograph. Secondly, resulting parameters are often suboptimal for the applications (flood estimation, simulation for water management) the model is to be used for. While there are alternatives which take into account the serial correlation in residuals (Bates & Campbell, 2001) or consider multiplicative errors in the rainfall inputs (Kavetski *et al.*, 2002), they require estimation of more parameters needed to represent the additional characteristics of interest. A multi-objective formulation using Bayesian loss functions, such as proposed by Vrugt *et al*. (2003) or Madsen (2003), enables consideration of the multiple criterion that denote a model's goodness of fit. While there is the difficulty of establishing the relative importance associated with each of the criteria considered, these functions are recommended as the model likelihood in both parameter assessment or model comparison applications.

References

Bates, B. C. & Campbell, E. P. (2001) A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall–runoff modeling. *Water Resour. Res*. **37**, 937–947.

Boughton, W. C. (2004) The Australian Water Management Model. *Environmental Modelling & Software* **19**, 943–956.

Campbell, E. P. & Bates, B. C. (2001) Regionalization of rainfall–runoff model parameters using Markov chain Monte Carlo samples. *Water Resour. Res*. **37**, 731–739.

Chib, S. & Jeliazkov, I. (2001) Marginal likelihood from the Metropolis-Hastings output. *J. Am. Statist. Assoc.* **96**, 270–281.

Freer, J., Beven, K. J. & Ambroise, B. (1996) Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. *Water Resour. Res*. **32**, 2161–2173.

Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. (1995) Bayesian Data Analysis. Chapman and Hall, Suffolk, UK.

Gilks, W. R., Richardson, S. & Spiegelhalter, D. J. (eds) (1996) *Markov Chain Monte Carlo in Practice*. Chapman and Hall/CRC, New York, USA.

Haario, H., Saksmen, E. & Tamminen, J. (2001) An Adaptive Metropolis Algorithm. *Bernoulli* **7**, 223–242.

Hastie, T., Tibshirani, R. & Friedman, J. (2001) *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer, New York, USA.

Jordan, M. I. & Jacobs, R. A. (1994) Hierarchical mixtures of experts and the EM algorithm. *Neural Computation* **6**, 181–214.

Kass, R. E. & Raftery, A. E. (1995) Bayes factors. *J. Am. Statist. Assoc.* **90**, 773–795.

Kavetski, D., Franks, S. W. & Kuczera, G. (2002) Confronting input uncertainty in environmental modelling. In: *Calibration of Watershed Models* (ed. by Q. Duan, H. V. Gupta, S. Sorooshian, A. N. Rousseau & R. Turcotte), 49–68. American Geophysical Union, USA.

Kuczera, G. (1983) Improved parameter inference in catchment models. 1. Evaluating parameter uncertainty. *Water Resour. Res*. **19**, 1151–1162.

Kuczera, G. & Mroczkowski, M. (1998) Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. *Water Resour. Res.* **34**, 1481–1489.

Kuczera, G. & Parent, E. (1998) Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *J. Hydrol.*, 69–85.

Madsen, H. (2003) Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives. *Adv. Water Resour.* **26**, 205–216.

Marshall, L., Nott, D. J. & Sharma, A. (2004) A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling. *Water Resour. Res.* **40**, 10.1029/2001WR002501.

Marshall, L., Nott, D. J. & Sharma, A. (2005) Hydrological model selection: a Bayesian alternative. *Water Resour. Res.* (in press).

Peel, M. C., Chiew, F. H. S., Western, A. W. & McMahon, T. A. (2001) Extension of unimpaired monthly streamflow data and regionalisation of parameter values to estimate streamflow in ungauged catchments. *CRC for Catchment Hydrology*, Melbourne, Australia.

Peng, F., Jacobs, R. A. & Tanner, M. A. (1996) Bayesian inference in Mixtures-of-Experts and Hierarchical Mixtures-of-Experts models with application to speech recognition. *J. Am. Statist. Assoc.* **91**, 953–960.

Pilgrim, D. H. (1997) Estimation of peak flows for small to medium-sized rural catchments. In: *Australian Rainfall and Runoff* (ed. by D. H. Pilgrim), IV-1 to IV-21. Institution of Engineers, Pilgrim, Australia.

Romanowicz, R., Beven, K. & Tawn, J. A. (1994) Evaluation of predictive uncertainty in nonlinear hydrological models using a Bayesian approach. In: *Statistics for the Environment 2: Water Related Issues (ed. by* B. Barnett & J. F. Turkman), 297–317. John Wiley, Chichester, UK.

Schwarz, G. (1978) Estimating the dimension of a model. *Ann. Statist.* **6**, 461–464.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiodo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S. & Zehe, E. (2003) IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrol. Sci. J.* **48**(6), 857–880.

Theimann, M., Trosset, M., Gupta, H. & Sorooshian, S. (2001) Bayesian recursive parameter estimation for hydrologic models. *Water Resour. Res.* **37**, 2521–2535.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W. & Sorooshian, S. (2003) Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resour. Res.* **39**, 10.1029/2001WR001214.