Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach

Article *in* Water Resources Research · January 2008

DOI: 10.1029/2006WR005205

Paul James Smith
waternumbers Limited

Keith Beven
Lancaster University

Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach

P. J. Smith,[1] K. J. Beven,[1] and J. A. Tawn[2]

Received 2 June 2006; revised 31 May 2007; accepted 6 July 2007; published 8 January 2008.

[1] In recent years, increasing computational power has been used to weight competing hydrological models in a Bayesian framework to improve predictive power. This may suggest that for a given measure of association with the observed data, one hydrological model is superior to another. However, careful analyses of the residuals of the model fit are required to propose further improvements to the model. In this paper we consider an alternative method of analyzing the shortcomings in a hydrological model. The hydrological model parameters are treated as varying in time. Simulation using a particle filter algorithm then reveals the parameter distribution needed at each time to reproduce the observed data. The resulting parameter, and the corresponding model state, distributions can be analyzed to propose improvements to the hydrological model. A demonstrative example is presented using rainfall-runoff data from the Leaf River, United States. This indicates that even when explicitly representing the uncertainty of the observed rainfall and discharge series, the technique shows shortcomings in the model structure.

Citation: Smith, P. J., K. J. Beven, and J. A. Tawn (2008), Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach, *Water Resour. Res.*, *44*, W01410, doi:10.1029/2006WR005205.

1. Introduction

[2] The hydrological modeling community has come to recognize three main sources of uncertainty in the modeling process which, though their precise definitions vary between modelers, are difficult to differentiate. The three sources of uncertainty can be titled observational, parameter and structural uncertainties. The observational uncertainty comes not only from the inaccuracy of the equipment used to make the observations but also from incommensurability arising from the differences in temporal and spatial scale from the variables defined in the model structure. Defined in such a way the observational error is directly linked to the model implementation. Parameter uncertainty relates to the spread of sets of parameters that, for a model structure, prove to have some level of consistency with the data and the modelers' prior belief. As such the parameter uncertainty is related both to the model structure and observational uncertainty. The structural uncertainty relates to the chosen representation of the system. That is for a given modeling purpose the structural uncertainty is all the classes of models that could be conceptualized and numerically implemented which have parameters that result in some level of consistency with the observed data and meet with the prior beliefs of the modeler. Note that in this definition the model structure is not only the processes represented, but also the representations chosen (e.g., St. Venant's or kinematic wave equation for channel routing) and the spatiotemporal scale of their application (i.e., grid size and time step). From this definition it can be quickly seen that, unless the types of structures conceptualized are constrained a priori, that the full evaluation of the structural uncertainty is an impossible task. In recent times low numbers of model structures have been compared [*Marshall et al.*, 2005; *Peters et al.*, 2003; *Ye et al.*, 2004], or combined to aid prediction [*Butts et al.*, 2004; *Georgakakos et al.*, 2004; *Marshall et al.*, 2006]. In doing this it may be possible to select one model structure out of those considered to be the most plausible. The most plausible model may still suffer from structural inadequacy, which is the inability to find a plausible set of model parameters that provide consistency with the observed data and the modeler's prior beliefs [see, e.g., *Peters et al.*, 2003].

[3] This paper proposes a technique to evaluate structural inadequacy as outlined above. First the observational errors along with plausible parameter and state spaces are defined [*Beven*, 2006]. Then, treating the parameters as evolving in time, a filtering algorithm (section 2) is used to determine the distribution of the parameters that best satisfy the joint requirements of consistency with the observed data and plausibility, at each time step. Analysis of the evolving sample from the parameter and state space, combined with the residual of the fit can then be used to reveal where the joint requirements are not met satisfactorily, indicating structural inadequacy. This allows an iterative process of model improvement [e.g., *Nash and Sutcliffe*, 1970; *Wagener et al.*, 2001] to be undertaken, with new models that remain consistent with the observational error definitions proposed and tested if necessary.

[4] The treatment of the model parameters as varying in time makes this approach distinct from methods of sequentially incorporating data to estimate a distribution of con-

[1]Department of Environmental Science, Lancaster University, Lancaster, UK.
[2]Department of Mathematics and Statistics, Lancaster University, Lancaster, UK.

0043-1397/08/2006WR005205$09.00

stant parameters such as GLUE [*Beven and Binley*, 1992] and BaRE [*Thiemann et al.*, 2001]. As with the DYNIA algorithm [*Wagener et al.*, 2003] regions of a time series with high information content may be highlighted, in this case through consideration of the variability in the parameter ranges.

[5] Comparison may be drawn to other published calibration algorithms particularly filtering algorithms related to the sequential incorporation of additional data. The most commonly applied filtering algorithms in hydrology are the extended and ensemble Kalman filter [*Evensen*, 1994], which are extension of the Kalman filter [*Kalman*, 1960] for nonlinear models. Some recent applications include updating states in a three dimensional hydrodynamic model [*Sorensen and Madsen*, 2004], parameter estimation in aquifer modeling [*Leng and Yeh*, 2003; *Yeh and Huang*, 2005] and state parameter estimation in rainfall runoff modeling [*Moradkhani et al.*, 2005b]. Alternative filtering techniques such as particle filters (PF) [*Kitagawa et al.*, 2001; *Moradkhani et al.*, 2005a], where the filtered distribution is approximated by sample, and the "Sequential Optimisation and Data Assimilation" (SODA) algorithm [*Vrugt et al.*, 2005] have also been applied in hydrological settings. In all the applications listed, with the exception of *Moradkhani et al.* [2005b], the model inadequacy and observational error of the inputs are absorbed into a state correction, which is not analyzed with the aim of improving the model structure, though it has been noted that "We are therefore left with the intuitively reasonable hypothesis that the computed time series of state updates should contain valuable information about model structural errors" [*Vrugt et al.*, 2005, p. 15].

[6] *Moradkhani et al.* [2005b] explicitly represent errors associated with the observed input and output of the model. However, rather then defining these on the basis of the perceived observational process, they are tuned so they are conditional upon the model. *Rajaram and Georgakakos* [1989] present a filtering algorithm which allows for the explicit representation of the observational uncertainties in parameter estimation. They also propose an extension to investigate the model inadequacy by considering the evolution of time varying parameters as outlined above. The computational algorithm presented by *Rajaram and Georgakakos* [1989] has two main limitations for the analysis of nonlinear hydrological models. The first of these is that the joint distribution of model states and parameters is summarized by its first two moments. This means that skew, or multimodality, of the distribution will remain undetected. The second limitation is that, as with the extended Kalman filter, a linear approximation of the hydrological model is constructed at each time step to evolve the joint distribution of model states and time varying parameters. For this to be effective the derivatives of the model functions with respect to the model states, parameters and input must exist and be well behaved. Parameters that define a threshold are an example where this assumption is violated. *Evensen* [1994] presents a more detailed analysis of this problem in proposing the ensemble Kalman filter. Given these limitations an alternative PF scheme is implemented.

[7] Section 2 outlines the PF algorithm used, its theoretical limits and presents an example where the extended or ensemble Kalman filter would provide misleading results. The main body of the work (section 3) proceeds to focus on the analysis of data from the Leaf River catchment located north of Collins, Mississippi, United States. This catchment, covering an area of 1944 km2, has been the subject of much study [*Boyle et al.*, 2000; *Vrugt et al.*, 2003] and also the basis of the HUGE PUB uncertainty workshop simulation study (IAHS-PUB Workshop, Lugano, Italy, http://www.es.lancs.ac.uk/hfdg/uncertainty_workshop/uncert_intro.htm).

2. Hydrological State Space Models and Particle Filters

[8] Consider the vector $\mathbf{x}_t$ which contains both the states of the hydrological model (e.g., storage volumes) and the unknown, evolving, parameters of the model. For process-based hydrological models this appears intuitive since the unknown parameters are often believed to represent physical characteristics of the system which may not be constant over time. A state space model acting on $\mathbf{x}_t$ consists of two functions. The observation equation $\mathbf{y}_t \sim f_t(\cdot|\mathbf{x}_t, \phi)$ relates $\mathbf{y}_t$; the observations at time t; to the unobserved $\mathbf{x}_t$ with known (fixed) parameters ϕ. The Markovian state evolution equation $\mathbf{x}_t \sim g_t(\cdot|\mathbf{x}_{t-1}, \theta)$ relates $\mathbf{x}_t$ at two consecutive time steps, with known (fixed) parameters θ. At a time t the filtering density, which is the probability of the current values of the states and unknown parameters of the system given the observations up to that time ($\mathbf{y}_{1:t} = \{\mathbf{y}_1, \ldots, \mathbf{y}_t\}$), is denoted $\pi_t(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ where $\mathbf{x}_{0:t} = \{\mathbf{x}_0, \ldots, \mathbf{x}_t\}$. The conditional distribution of $\mathbf{x}_t$ given $\mathbf{x}_{0:t-1}$ and $\mathbf{y}_{1:t}$, denoted $\pi_t(\mathbf{x}_t|\mathbf{x}_{0:t-1}, \mathbf{y}_{1:t})$, is related to new observation $\mathbf{y}_t$ through the Markovian observation and evolutions equations. This allows $\pi_t(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ to be expressed in terms of $\pi_0(\mathbf{x}_0)$, the known prior distribution and $\mathbf{y}_{1:t}$. In particular,

$$\pi_t(\mathbf{x}_{0:t}|\mathbf{y}_{1:t}) = \frac{f_t(\mathbf{y}_t|\mathbf{x}_t,\phi)g_t(\mathbf{x}_t|\mathbf{x}_{t-1},\theta)}{p(y_t|y_{1:t-1})}\pi_{t-1}(\mathbf{x}_{0:t-1}|\mathbf{y}_{1:t-1}) \propto \pi_0(\mathbf{x}_0)\prod_{i=1}^{t} f_i(\mathbf{y}_i|\mathbf{x}_i,\phi)g_i(\mathbf{x}_i|\mathbf{x}_{i-1},\theta).$$

[9] Particle filters are a form of sequential Monte Carlo estimation [*Pitt and Shephard*, 1999]. They are derived to approximate $\pi_t(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ as a series of proper weighted samples, known as particles. A sample of particles $\{\mathbf{x}_{0:t}^{[j]}; j = 1, \ldots, m\}$ from $\pi_t(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ is termed proper if the particles have weights $\{w_t^{[j]}; j = 1, \ldots, m\}$ such that

$$\lim_{m\to\infty}\frac{\sum_{j=1}^{m} h\left(\mathbf{x}_{0:t}^{[j]}\right)w_t^{[j]}}{\sum_{j=1}^{m} w_t^{[j]}} = \mathrm{E}[h(\mathbf{x}_{0:t})]$$

for some function $h(\cdot)$ which maps the support of $\mathbf{x}_{0:t}$ onto a real space of constant dimension and is integrable with respect to $\pi_t(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$. In particular this allows $\pi_{t+1}(\mathbf{x}_{0:t+1}|\mathbf{y}_{1:t+1})$ to be approximated by

$$\pi_{t+1}\left(\mathbf{x}_{0:t+1}|\mathbf{y}_{1:t+1}\right) \approx \frac{\sum_{j=1}^{m} f_{t+1}\left(\mathbf{y}_{t+1}|\mathbf{x}_{t+1},\phi\right)g_{t+1}\left(\mathbf{x}_{t+1}|\mathbf{x}_t^{[j]},\theta\right)w_t^{[j]}}{\sum_{j=1}^{m} w_t^{[j]}}. \quad (1)$$

[10] The idea behind particle filtering is that if π_{t+1} $(\mathbf{x}_{0:t}|\mathbf{y}_{1:t+1})$ is close to that of π_t $(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ then a random sample from π_t $(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})$ can be reused to help construct a random sample from $\pi_{t\ +1}$ $(\mathbf{x}_{0:t+1}|\mathbf{y}_{1:t+1})$. This, hopefully, results in an improvement in computational efficiency. Several techniques have been proposed for sampling from (1) to generate a proper sample from π_{t+1} $(\mathbf{x}_{0:t+1}|\mathbf{y}_{1:t+1})$ [*Fearnhead*, 2002; *Guo et al.*, 2005; *Liu and Chen*, 1998; *Pitt and Shephard*, 1999]. These techniques have been based around constructing an approximation of π_{t+1} $(\mathbf{x}_{t+1}|\mathbf{x}_{0:t},\ \mathbf{y}_{1:t+1})$ under various conditions. One of the simplest and most general algorithms is that of sample importance resampling (SIR) [*Marshall*, 1956] where π_{t+1} $(\mathbf{x}_{t+1}|\mathbf{x}_{0:t},\ \mathbf{y}_{1:t+1})$ is approximated by g_{t+1} $(\mathbf{x}_{t+1}|\mathbf{x}_t,\ \theta)$. However, since the function g is a potentially highly nonlinear hydrological model, such an approximation may prove inefficient and weights of the resulting sample highly correlated, thus providing a poor sample of π_{t+1} $(\mathbf{x}_{0:t+1}|\mathbf{y}_{1:t+1})$. Here an adaptation of SIR proposed by Hürzeler and Künsch [1995] on the basis of the use of rejection sampling [*von Neumann*, 1951] is used. Within this adaptation particles are proposed using g_{t+1} $(\mathbf{x}_{t+1}|\mathbf{x}_t,\ \theta)$ then either rejected, or accepted as part of the sample from π_{t+1} $(\mathbf{x}_{0:t+1}|\mathbf{y}_{1:t+1})$ dependant upon the value of f_{t+1} $(\mathbf{y}_{t+1}|\mathbf{x}_{t+1},\ \phi)$. The advantage of this rejection sampling algorithm is that at any time step an independent sample of particles from π_{t+1} $(\mathbf{x}_{0:t+1},\ |\mathbf{y}_{1:t+1})$ is generated, with each particle having equal weight. The penalty forusing rejection sampling is that if g_{t+1} $(\mathbf{x}_{t+1}|\mathbf{x}_t,\ \theta)$ diverges greatly from $\pi_{t+1}(\mathbf{x}_{t+1}|\mathbf{x}_{1:t},\ \mathbf{y}_{1:t+1})$, as is likely to be the case if an evolution equation representing ignorance about the parameters is chosen; then many particles will be proposed and then rejected. The rejection sampling–based algorithms implementation can be summarized by the following three steps:

[11] 1. Select $J \in (1,\ \ldots,\ m)$ with $P\ (J = j) \propto w_t^{[j]}$.

[12] 2. Generate $\hat{\mathbf{x}}_{t+1}$ as a sample from g_{t+1} $(\mathbf{x}_{t+1}|\mathbf{x}_t^{[J]},\ \theta)$.

[13] 3. Accept $(\mathbf{x}_{0:t}^{[J]},\ \hat{\mathbf{x}}_{t+1})$ as a member of

$$\left\{\mathbf{x}_{0:t+1}^{[i]}\ i = 1,\ldots,m\right\}$$

with probability

$$p = \frac{f_{t+1}\left(\mathbf{y}_{t+1}|\hat{\mathbf{x}}_{t+1},\phi\right)}{c_{t+1}},$$

where

$$c_{t+1} = \sup_{\mathbf{x}_{t+1}}\left(f_{t+1}\left(\mathbf{y}_{t+1}|\mathbf{x}_{t+1},\phi\right)\right).$$

[14] The number of particles controls both the computational time and accuracy of the approximation (1). The number of particles was selected so that the computational time remained practical, while reasonable accuracy, relative to introducing more particles, of the approximation (1) was kept. The relative accuracy was determined by a low number of short simulations and is not reported.

[15] To reinforce the methodology outlined above let us consider a nonlinear example presented by *Doucet et al.* [2001]. For $t \geq 1$ the Markovian evolution equation for an unknown state at u_t is given by

$$u_t = \frac{1}{2}u_{t-1} + 25\frac{u_{t-1}}{1+u_{t-1}^2} + 8\cos(1.2t) + v_t$$

and the observation equation relating u_t to the observed value y_t by

$$y_t = \frac{u_t^2}{20} + z_t$$

where v_t and z_t are mutually independent zero mean Gaussian noise with known variances σ_v^2 and σ_z^2 respectively. In this example $\theta = \sigma_v^2$ and $\phi = \sigma_z^2$. The observation equation is given by

$$f_t(y_t|u_t,\phi) = d\left(\sigma_z^{-1}\left(y_t - \frac{u_t^2}{20}\right)\right),$$

where $d\ (\cdot)$ denote the density function of a standardized normal distribution. The supremum of f_t $(y_t|u_t,\ \phi)$ with respect to u_t occurs when $u_t^2 = 20\ y_t$ with $c_t = d\ (0)$. The prior distribution of u_0 is given by $u_0 \sim N\ (0,\ \sigma_0^2)$. The three summary steps for the application of the rejection sampling algorithm given above are now

[16] 1. Select $J \in (1,\ \ldots,\ m)$ with $P\ (J = j) \propto w_t^{[j]}$.

[17] 2. Generate $\hat{v}_t$ as a realization from $v_t \sim N\ (0,\ \sigma_v^2)$ and calculate

$$\hat{u}_t = \frac{1}{2}u_{t-1}^{[J]} + 25\frac{u_{t-1}^{[J]}}{1+u_{t-1}^{[J]\ 2}} + 8\cos(1.2t) + \hat{v}_t.$$

[18] 3. Accept $(u_{0:t}^{[J]},\ \hat{u}_{t+1})$ as a member of

$$\left\{u_{0:t+1}^{[i]}\ i = 1,\ldots,m\right\}$$

with probability

$$p = \frac{d\left(\sigma_z^{-1}\left(y_t - \frac{\hat{u}_t^2}{20}\right)\right)}{d(0)}.$$

[19] Figure 1 shows the filtered density π_t $(u_t|v_{1:t})$ for an example simulation of the model with $\sigma_v^2 = \sigma_0^2 = 10$, $\sigma_w^2 = 1$ and $m = 1000$. Note that for some values of t the filtered distribution is bimodal. This feature would not be revealed by an application of the extended or ensemble Kalman filter which would provide only the first two moments of the filtered distribution.

3. Leaf River Case Study

[20] This section of the paper focuses on the application of the methodology to detect structural inadequacy in three models, of increasing complexity, when fitted to 10 years (a) of daily data from the Leaf River catchment. This humid catchment is located north of Collins, Mississippi, United States, and covers an area of 1944 km^2. It has been the subject of much study [*Boyle et al.*, 2000; *Thiemann et al.*, 2001; *Vrugt et al.*, 2003] and also the basis of the HUGE PUB uncertainty workshop simulation study (see http://www.es.lancs.ac.uk/hfdg/uncertainty_workshop/uncert_intro.htm). The data selected for use are the daily precipitation, potential evapotranspiration and discharge observations for the period 29 May 1952 to 30 September 1962. This period was selected to allow the reader to make direct

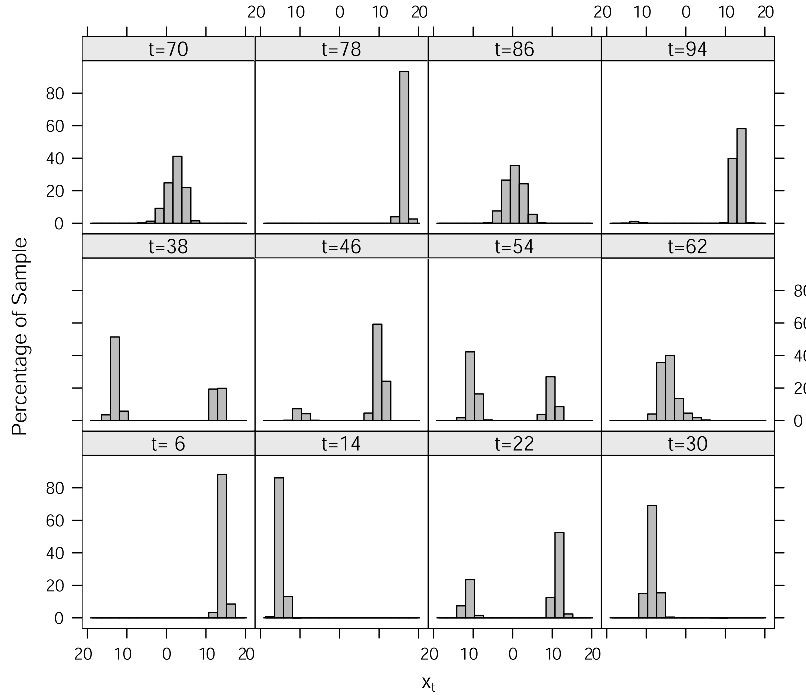

Figure 1. Summary of the filtered distribution of u_t for the example presented in section 2.1. Note the bimodal nature of the distribution which is not detectable from the second-order moments.

comparison to other studies [*Boyle et al.*, 2000; *Moradkhani et al.*, 2005a, 2005b; *Thiemann et al.*, 2001; *Vrugt et al.*, 2003, 2005] if desired.

[21] In the following work the error on the observed discharge $\tilde{q}_t$ at time t, was defined as an independent normal distribution on the ratio of the observed and modelled discharge constrained onto positive values, that is,

$$P(\tilde{q}_t|q_t) = \begin{cases} \frac{1}{\sqrt{2\pi}}\sigma_q^{-1}q_t^{-1}\exp\left(-\frac{1}{2\sigma_q^{-2}}\left(\frac{\tilde{q}_t}{q_t}-1\right)^2\right) & q_t > 0 \\ 0 & q_t \leq 0. \end{cases}$$

[22] The assumption that the error in the determination of discharge is considered proportional to magnitude stems from analysis of rating curves [e.g., *Petersen-Overleir*, 2004]. The errors in the determination of discharge are treated as uncorrelated on the presumption that they are caused by an independent random noise rather then systematic errors in the observational process, for example drift in the depth recording device or bias in the fit of the rating curve for certain ranges of stage.

[23] A similar error structure was chosen for the observed precipitation $\tilde{r}_t$, with the additional condition that if no precipitation was observed then it was assumed that none had fallen, that is,

$$P(\tilde{r}_t|r_t > 0) = \begin{cases} \frac{1}{\sqrt{2\pi}}\sigma_r^{-1}r_t^{-1}\exp\left(-\frac{1}{2\sigma_r^{-2}}\left(\frac{\tilde{r}_t}{r_t}-1\right)^2\right) & \tilde{r}_t > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$P(\tilde{r}_t|r_t = 0) = \begin{cases} 1 & \tilde{r}_t = 0 \\ 0 & \text{otherwise.} \end{cases}$$

[24] The increase in variance of the distribution as the magnitude of r_t the catchment average rainfall increases is motivated by the belief that the characterization of r_t by a network of rain gauges is less accurate as r_t increases. The condition that if no precipitation was observed then it was assumed that none had fallen is motivated that belief that the rain gauge network is dense enough to observe every event that will induce a detectable catchment response. The assumption of temporal independence presumes that the weather systems causing rainfall alter over a daily time period in such a way that the error in the determination of the r_t are not correlated.

[25] The variance parameters $\phi = \{\sigma_r^2, \sigma_q^2\}$ were set such that the probability of $\tilde{q}_t$ or $\tilde{r}_t$ lying in the region $q_t \pm 20\%$ or $r_t \pm 20\%$ respectively was 0.95. The specification of the standard deviation for the flow is taken to be conservative compared to values reported in the literature [e.g., *Whalley et al.*, 2001]. The observed value of evapotranspiration at time t, $\tilde{e}_t$ was used as to model input. The state of the system $\mathbf{x}_t$ contains the volumes of the stores $\mathbf{s}_t$ and parameters $\mathbf{p}_t$ of the hydrological model as well as r_t. The evolution $g_{t+1}(\mathbf{x}_{t+1}|\mathbf{x}_t, \theta)$ is a two stage process. The first stage is the proposal of $\Delta_{t+1} = (r_{t+1}, \mathbf{p}_{t+1})$ from a reflective random walk in the plausible parameter bounds ($\Delta^{\min}$, $\Delta^{\max}$), that is

$$\Delta_{t+1} \sim TMVN\left(\Delta_t, \Gamma, \Delta^{\min}, \Delta^{\max}\right),$$

where Γ is the covariance matrix of the Markovian evolution and

$$TMVN\left(\Delta_t, \Gamma, \Delta^{\min}, \Delta^{\max}\right)$$

is the truncation of the multivariate normal distribution MVN (Δ_t, Γ) by the limits ($\Delta^{\min}$, $\Delta^{\max}$). In the general

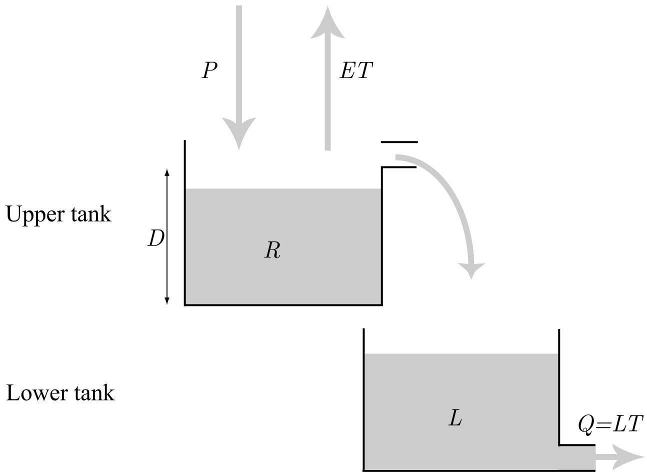

Figure 2. Schematic of the twin tank hydrological model. D is the maximum depth of the upper tank, with R being the upper tank storage. L is the storage in the lower tank, and T is the reciprocal of the lower tank time constant. P is the precipitation, and ET is the evapotranspiration.

notation this gives $\theta = \{\Gamma, \Delta^{\min}, \Delta^{\max}\}$. The second stage is to calculate $\mathbf{s}_{t+1}$ by applying the hydrological model to $\mathbf{s}_t$ using Δ_{t+1}.

3.1. Two Tank Model

[26] The first model considered is a twin tank rainfall runoff model (Figure 2). The upper tank representing the upper soil has storage depth D_t [mm] which when exceeded spills to the lower routing store. The computational sequence for the upper store (at each time step) results in the spill to the lower store occurring before evapotranspiration is accounted for, as such the store acts in a similar way to the root zone store used in many TOPMODEL applications [*Beven et al.*, 1995]. The lower store is a simple linear routing store with a single delay parameter T_t [d^{-1}] and unlimited depth. The volume stores of the system are $\mathbf{s}_t = \{R_t, L_t\}$ and the parameters $\mathbf{p}_t = \{D_t, T_t\}$ where D_t was constrained to lie between 10mm and 300 mm, with these values being based upon previous TOPMODEL studies [*Beven and Freer*, 2001]. The value of T_t is physically limited between zero and one, and was constrained as such. The covariance matrix of the evolution equation was specified as

$$\Gamma = \begin{bmatrix} \lambda_r^2 & 0 & 0 \\ 0 & \lambda_D^2 & 0 \\ 0 & 0 & \lambda_T^2 \end{bmatrix},$$

giving a series independent reflective random walks with the implication that any correlation between the parameters in the output is the result of the conditioning provided by the observed data. The variances λ_D^2 and λ_T^2 were selected such that the standard deviation were a fifth of the parameter ranges in an attempt to allow a free movement of the parameters while curtailing switching from extreme positions in the parameter ranges. The evolution of D_t was further constrained so that spill to the routing tank could not occur unless there was rainfall. The variance λ_r^2 was set to 14400 mm^2, so the standard deviation was approximately the value of the maximum change in the observed rainfall series between consecutive time steps. This choice represents a position of ignorance as to the evolution of the rainfall pattern.

[27] The SMC algorithm was run using 15000 particles. The parameters of the initial particles were sampled from uniform distributions on the parameter ranges. The stores were initialized using the first day with no observed rainfall. The lower tank storage value was then calculated using a flow value sampled from the observational error distribution and the particles delay parameter value. The initial storage in the upper tank was sampled uniformly between zero and the particles' upper soil storage depth value.

[28] Results show the poor suitability of this model for the catchment in question. Figure 3 shows that the storage in the lower tank increases throughout the simulation period. Since this is physically unrealistic it suggests a structural failing in the model resulting in poor water balance. There are several possible causes for this including a poor representation of the evapotranspiration process by the root zone tank, failure to represent some subsurface outflow across the catchment boundary or incorrect definition of the catchment boundaries.

3.2. HyMOD

[29] HyMOD is a conceptually simple hydrological which can be perceived as Hammerstein type model featuring a nonlinear filter followed by a linear component. HyMOD, along with the Leaf River data set has been used by several authors for the testing of calibration strategies [*Boyle et al.*, 2000; *Moradkhani et al.*, 2005a, 2005b; *Vrugt et al.*, 2003; *Wagener et al.*, 2001, 2003]. The nonlinear filter component of HyMOD is based on the rainfall excess model of *Moore* [1985]. The output of this filter is routed through a series of linear stores representing slow and fast flow pathways. Figure 4 shows a schematic of the model. The rainfall excess model used assumes the soil moisture capacity at time t, c_t, varies across the catchment. The cumulative fractional area with capacity less then or equal to c_t is expressed, in terms of the maximum storage capacity in the catchment (C_t) and a summary of the degree of spatial variability (b_t), by the formula

$$F(c_t) = 1 - \left(1 - \frac{c_t}{C_t}\right)^{b_t}.$$

[30] The output from the rainfall excess filter is split using a factor α between a quick and slow flow pathway. The quick pathway is three in series tanks each with fractional discharge F_t at time t. The slow pathway contains a single tank with fractional discharge K_t at time t. The parameter vector at time t is then $\mathbf{p}_t = \{C_t, b_t, F_t, K_t, \alpha_t\}$. As with the twin tank the covariance matrix of the evolutionary equation was specified to give a series of independent reflective random walks, with the variances being chosen using a similar methodology. An additional condition is placed upon the evolution of C_t and b_t by limiting the proposed values so that they cannot result in maximum storage volume of the rainfall excess filter which is less then the current stored volume. This implies that all the discharge is the result of excess rainfall caused by precipitation and not as a result of the release of water caused by a shrinking

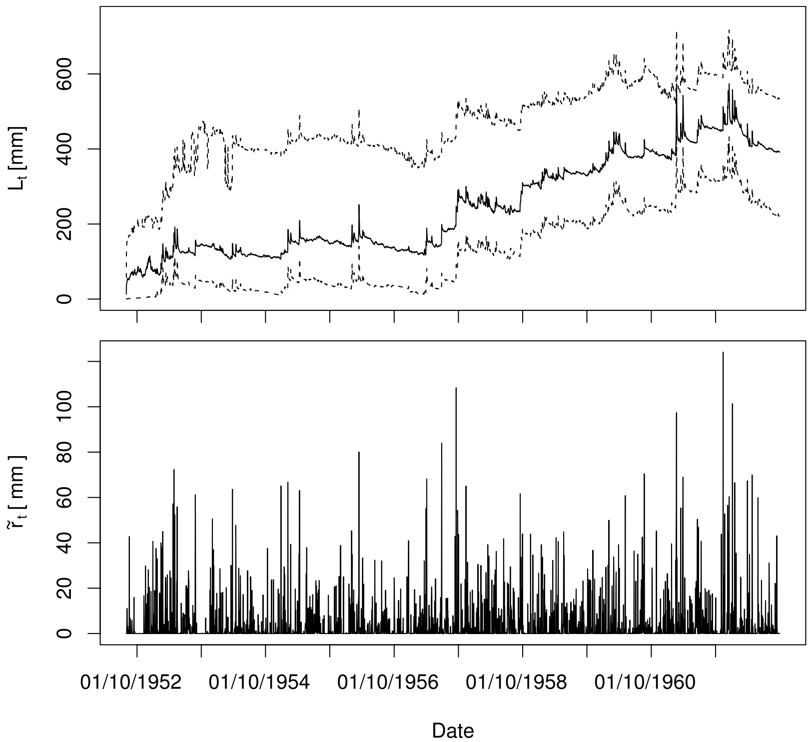

Figure 3. (top) Fluctuation in storage for the lower tank L_t of the twin tank model throughout the simulation period. Dotted lines represent the 95th and 5th percentiles of the sampled particles, and the solid line represents the sample mean. (bottom) Corresponding observed rainfall.

of the catchments storage potential. Table 1 shows the minimum and maximum parameter values. These were selected on the basis of previous studies [*Vrugt et al.*, 2003].

[31] The simulation result of the HyMOD model indicates that the structure of the model fails to reproduce the observed data, given the parameter ranges in Table 1 and the error assumptions outlined at the start of this section. Figure 5 shows that after a period with no observed rainfall in September and October 1952 the discharge is not adequately reproduced. Investigation of the states of the model indicate that the routing tanks contain inadequate water to reproduce the observed discharge. This indicates that the inability to recharge the routing tanks without rainfall is an inadequacy in the model structure that requires correction. Performing a simulation starting in October 1952 indicates (Figure 6 with the corresponding observed rainfall and flow in Figure 7) that apart from limited periods (e.g., late November 1952) where the model inadequacy outlined above is again revealed, the model structure may be an appropriate starting point for a representation of the system.

[32] Figure 8 shows the evolving parameter distributions for the first year of this simulation. During November and December 1952 the distribution of the parameters (C_t and b_t) controlling the volume of rainfall excess filter become more defined as the catchment "wets up" during a period of rainfall which does not produce a significant response in the

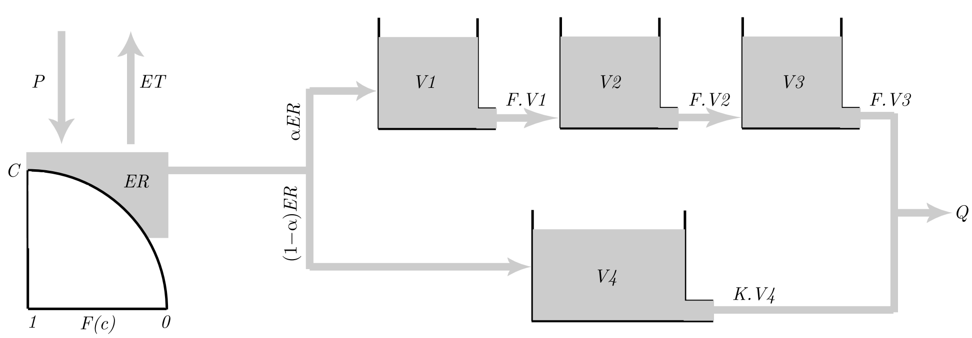

Figure 4. Schematic of the HyMOD hydrological model. Excess rainfall *ER* is generated conditional to the current moisture content of the catchment. The storage capacity distribution function $F(c)$ is related to the maximum storage capacity (C_t) and a shape parameter (b_t). The excess rainfall is routed with respect to parameter α_t either through the three linear reservoirs of the quick pathway, which has fraction discharge F_t, or through the single linear reservoir of the slow pathway, with residence time K_t. *P* is the precipitation, and *ET* is the evapotranspiration.

Table 1. Selected Parameter Bounds for HyMOD and PDM Models

Parameter	HyMOD	PDM
C_t, mm	100–500	100–500
b_t	0.1–2	0.1–2
α_t	0.4–0.99	na[a]
K_t, d^{-1}	0.001–0.1	na
F_t, d^{-1}	0.1–0.99	na
G_t	na	0–0.2
ks_t, d^{-1}	na	0.2–0.99
kf_t, d^{-1}	na	ks_t–0.99

[a]NA means not applicable.

discharge. The routing tank parameters, particularly F_t, display a fluctuating pattern which appears to correspond to the response of the catchment to rainfall. A similar pattern is witnessed to a lesser extent in α_t. The distribution of α_t is not conditioned heavily away from the prior for many time steps. This may indicate that if the slow and fast flow pathways were separated a clearer relationship between the routing tank parameters and the states of the system, or input, could be derived. Section 3.3 proposes two alterations to the HyMOD structure which separates the slow and fast flow pathways and allow recharge of the slow routing tank without rainfall.

3.3. Probability Distributed Moisture Model

[33] To attempt to overcome the issues relating to the performance of HyMOD during the dry period in September and October 1952 an alternative to the rainfall excess filter used in HyMOD is considered. This representation is contained within the probability distributed moisture (PDM) model proposed by Moore [*Moore*, 1985; *Moore and Clarke*, 1981] considered by Senbeta [*Senbeta et al.*, 1999] and the basis of several modeling studies [*Moore*, 2007; *Moore and Bell*, 2002; *Lamb*, 1999; *Cameron et al.*, 2002]. A representation of the PDM model used is shown in Figure 9. The moisture storage component (MSC) is identical to that used in HyMOD, with rainfall excess is generated in a similar fashion. This rainfall excess, considered the fast response, is routed through a linear tank with fractional discharge kf_t. The method of generation of slow flow in the PDM model, by drainage from the MSC, is

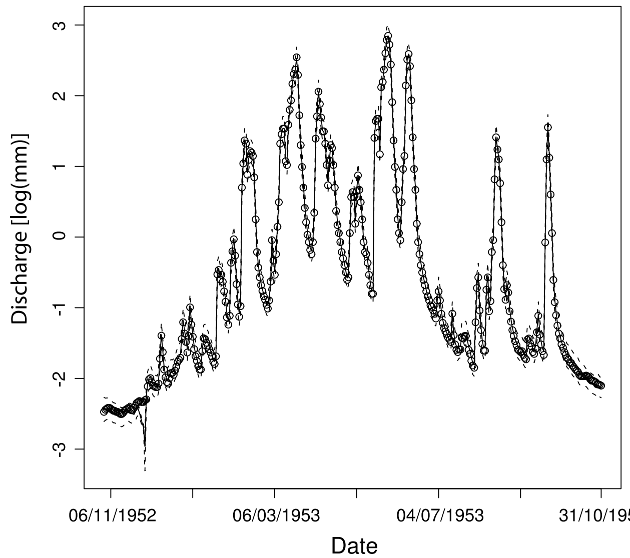

Figure 6. Traces of the discharge for the HyMOD simulation starting in October 1952. Dashed lines represent the 95th and 5th percentiles of the modeled values, and the solid lines represent the mean. The points represent the observed values. Note the brief periods in November 1952 when the model is unable to reproduce the discharge because of insufficient water in the routing tanks.

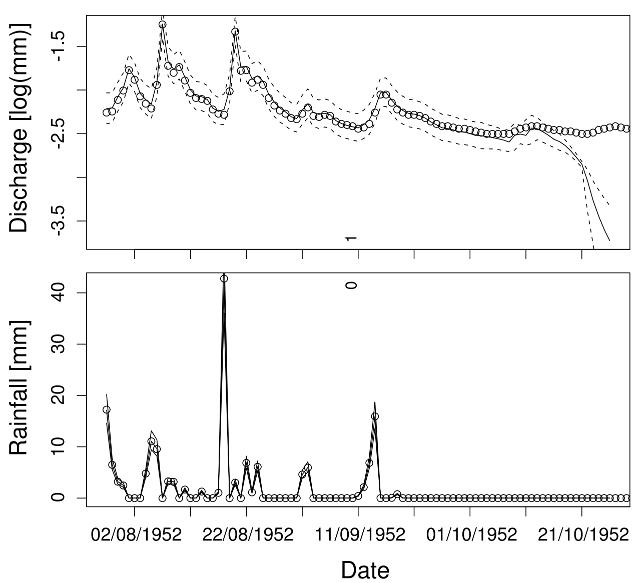

Figure 5. Traces of the discharge and rainfall for the HyMOD simulation from August to October 1952. Dashed lines represent the 95th and 5th percentiles of the modeled values, and the solid lines represent the mean. The points represent the observed values. The decline in discharge witnessed at the end of the period results from insufficient water being present in the routing tanks to generate adequate discharge.

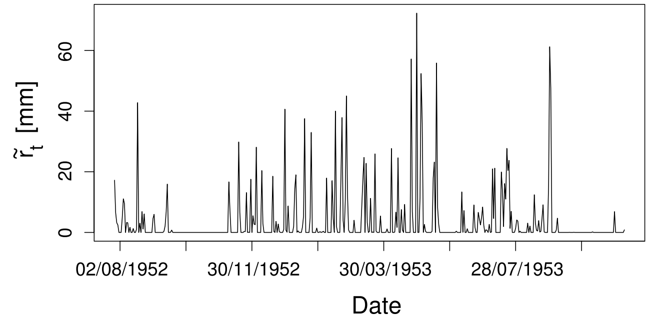

Figure 7. Observed precipitation and discharge for the first 365 d of the Leaf River simulation period.

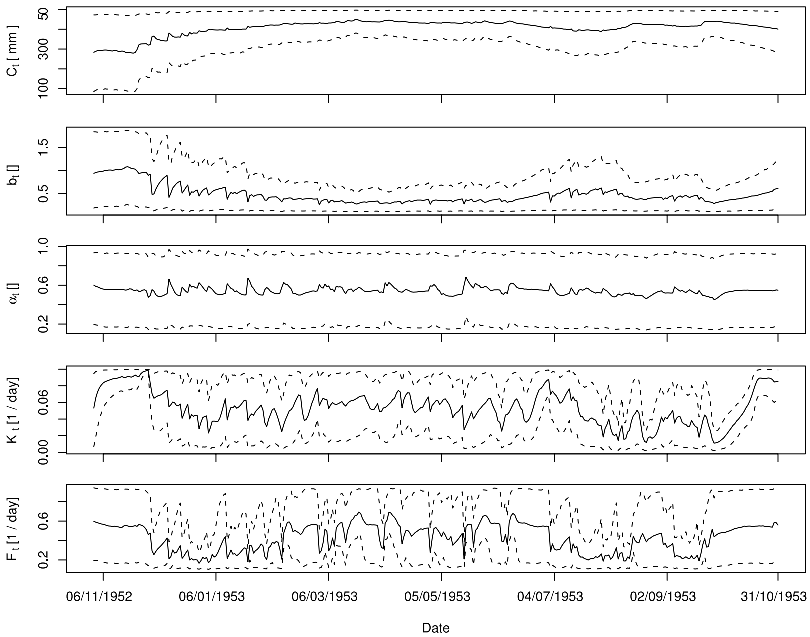

Figure 8. Traces of the five HyMOD model parameters during the first 365 d of the simulation. Dashed lines represent the 95th and 5th percentiles, and the solid lines represent the mean.

different to that used in HyMOD. This volume drained is expressed as a fraction (G_t) of the volume stored in the MSC and is routed through a linear tank with fractional discharge ks_t. The range of the parameter values, $\mathbf{p}_t = \{C_t, b_t, G_t, ks_t, kf_t\}$, selected is shown in Table 1. As with the analysis of HyMOD the values of C_t and b_t considered are limited so that fast response cannot be generated by shrinkage of the MSC.

[34] The simulated results show this model to be a more suitable representation; however, there is evidence that the model cannot adequately reproduce a large storm event at the end of the 10 a period. The states of the model show no long term trend, with the storage volume in the MSC fluctuating more slowly than the stored volume in the routing tank. This indicates that the MSC is one possible method of representing the interaction of the evapotranpiration, precipitation and the soil moisture for this data set. As with the HyMOD model patterns can be observed in the evolving parameter distributions; to illustrate these the first 365 d of the simulation will be considered. Traces of the parameters are shown in Figures 10 and 11 with the observed data shown in Figure 7. As the dry period between September and October 1952, already noted in the HyMOD evaluation, progresses the value of the G_t, ks_t and kf_t increase. The result of this is that the proportion of the water held within the model that is discharged increased throughout the period. This is consistent with the belief that the catchment is "drying out" during this time and some minimum base flow has been reached. Other than during this dry period, the distribution of G_t is similar to that which would be expected from the prior distribution (i.e., uniform on the parameter range) indicating that this parameter is not identifiable within the current model configuration. The dry period, and a period of low precipitation during June 1953, reveals little information about C_t and b_t.

[35] November and early December 1952 represent a period of precipitative activity and low catchment response. During this period C_t and b_t tend to the upper and lower end of there ranges respectively, thus maximizing the storage available in the SMC. These distributions stabilize during the first half of 1953 where the response of the catchment to rainfall is most notable in the discharge record. This pattern is similar to that witnessed for the parameters of the effective rainfall filter in HyMOD. In the case of the PDM model the parameters C_t and b_t appear less well identified then for HyMOD perhaps because of the inclusion of the slow flow generation mechanism. The fluctuation in the parameter distribution over shorter time periods, particularly evident for b_t, suggest that the parameters are most identifiable when the catchment is driven by precipitation.

[36] The evolution of the storage volume of the slow routing tank shows a seasonal pattern (Figure 11), with the volume stored being higher during the winter and spring months. The storage volume of the fast flow routing tank is

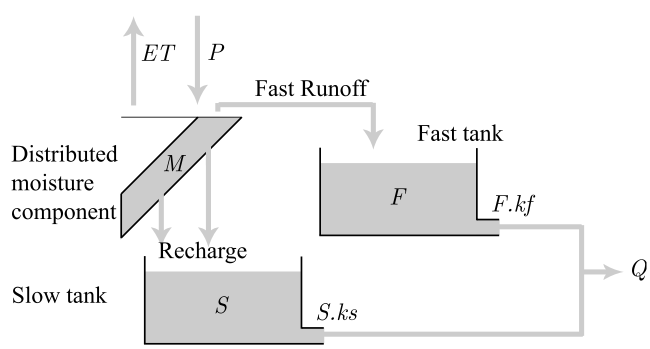

Figure 9. Schematic diagram of the complex PDM-based model. M represents the multiple storage capacities. S is the storage in the slow tank, which has fractional drainage ks. F is the storage in the fast tank, which has fractional drainage kf. P is the precipitation, and ET is the evapotranspiration.

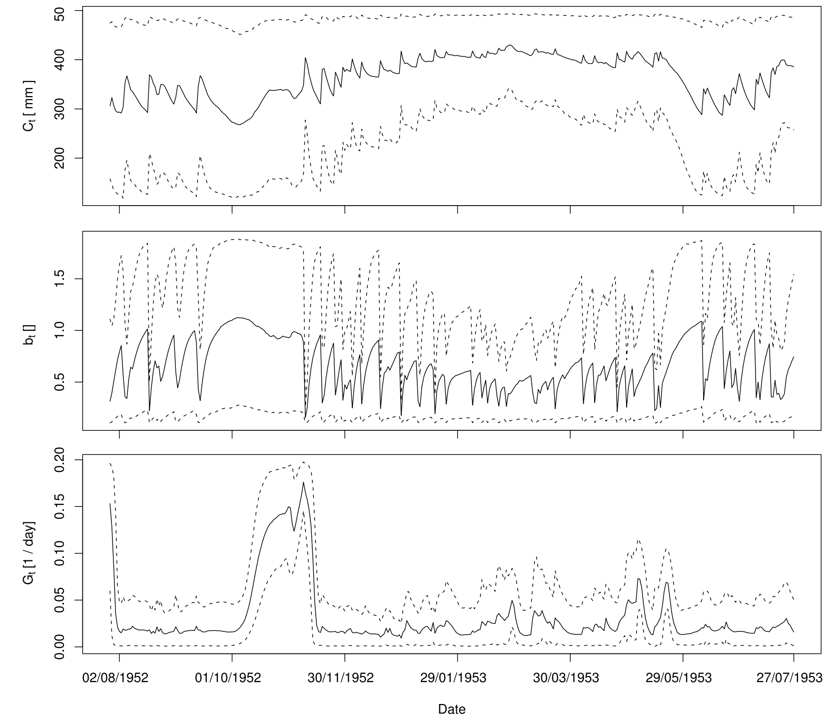

Figure 10. Traces of the three parameters of the SMC for the PDM model during the first 365 d of the simulation. Dashed lines represent the 95th and 5th percentiles, and the solid lines represent the mean.

clearly related to the rainfall pattern. The routing parameters (*ks* and *kf*) are most accurately defined (i.e., with least variability) at time steps where the model is attempting to replicate the catchment's response to rainfall. This suggests that these "driven" periods contain the most information about these parameters. No relationship was found between the routing parameters specific aspects of the flow or rainfall pattern during events (e.g., peaks). This may be indicative of the more complex underlying processes being approximated in the model.

4. Conclusions

[37] The results presented in the case study show that the simple hydrological models used fail to capture the full dynamics present in the data. This is despite the fact that

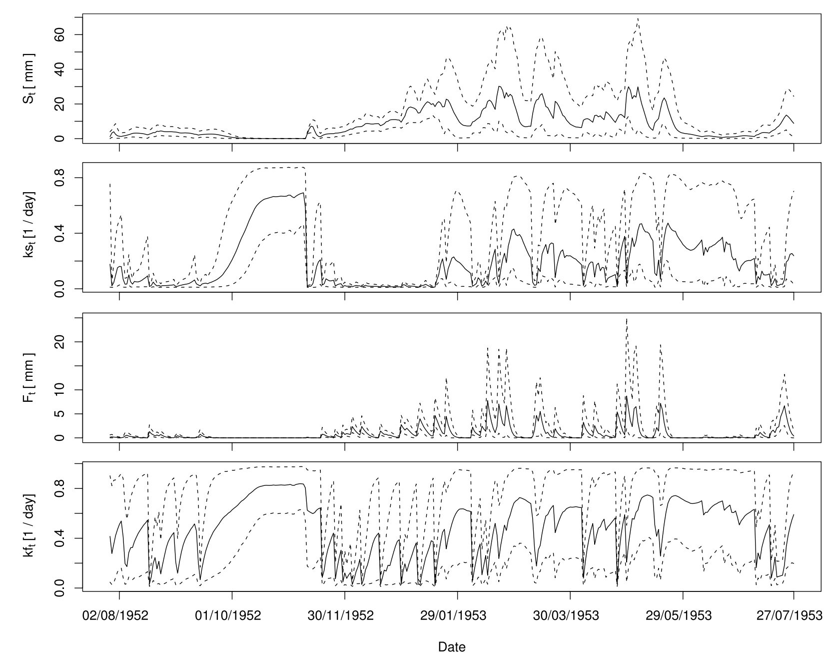

Figure 11. Traces of the two routing tank storage volumes and fractional drainages during the first 365 d of the simulation for the PDM model. Dashed lines represent the 95th and 5th percentiles, and the solid lines represent the mean.

some of the models used have been shown to characterize the dominant modes of the system [*Thiemann et al.*, 2001; *Wagener et al.*, 2003]. The case study also shows the iterative testing and improving of model structure that can be made on the road to finding an ''adequate'' (section 1) model. The detailed investigation of the evolution of the parameters and states of the model encouraged by this type of analysis reveals the shortcoming in the process representations. It may also allow some of the more general modeling assumptions, such as the closure and limits of the system, to be queried.

[38] The ability to undertake this detailed analysis comes at a cost, the need to accurately specify the error distributions of the observed data, both input and output series, at each time step. The undertaking of this, which is not discussed in detail within the paper, forces the modeler to clarify their belief of observational errors thus making it a considered part of the modeling process and opens it up to sensible questioning. This clear specification limits the ability for the ''error model'' to compensate for failings in the physical model so adding clarity and realism to the interpretation of the effectiveness of the process representations. The computational examples using the particle-filtering methodology show that such a using a detailed error representation is not unreasonable given modern computing capabilities. However, if computational approximations, such as those discussed in section 1 with regards to the work of *Rajaram and Georgakakos* [1989], are deemed expedient then more efficient computation schemes based on the Kalman filter may be derived.

[39] **Acknowledgments.** The authors have benefited from the support of the NERC long-term research program ''Uncertainty in Environmental Modelling'' (NER/L/2001/00658).

References

Beven, K. (2006), A manifesto for the equifinality thesis, *J. Hydrol.*, *320*(1–2), 18–36.

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

Beven, K., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *J. Hydrol.*, *249*(1–4), 11–29.

Beven, K., R. Lamb, P. Quinn, R. Romanowicz, and J. Freer (1995), TOPMODEL, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 629–668, Water Resour. Publ., Highlands Ranch, Colo.

Boyle, D. P., H. V. Gupta, and S. Sorooshian (2000), Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, *Water Resour. Res.*, *36*(12), 3663–3674.

Butts, M. B., J. T. Payne, M. Kristensen, and H. Madsen (2004), An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation, *J. Hydrol.*, *298*(1–4), 242–266.

Cameron, D., P. Kneale, and L. See (2002), An evaluation of a traditional and a neural net modelling approach to flood forecasting for an upland catchment, *Hydrol. Processes*, *16*(5), 1033–1046.

Doucet, A., N. de Freitus, and N. Gordon (2001), An introduction to sequential Monte Carlo methods, in *Sequential Monte Carlo Methods in Practice*, edited by A. Doucet, N. de Freitus, and N. Gordon, pp. 3–13, Springer, New York.

Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, *99*(C5), 10,143–10,162.

Fearnhead, P. (2002), Markov chain Monte Carlo, sufficient statistics, and particle filters, *J. Comput. Graphical Stat.*, *11*(4), 848–862.

Georgakakos, K. P., D. J. Seo, H. Gupta, J. Schaake, and M. B. Butts (2004), Towards the characterization of streamflow simulation uncertainty through multimodel ensembles, *J. Hydrol.*, *298*(1–4), 222–241.

Guo, D., X. D. Wang, and R. Chen (2005), New sequential Monte Carlo methods for nonlinear dynamic systems, *Stat. Comput.*, *15*(2), 135–147.

Hürzeler, M., and H. R. Künsch (1995), Monte Carlo approximations for general state space models, *Res. Rep. 73*, ETH Zurich, Zurich, Switzerland. (Available at http://e-collection.ethbib.ethz.ch/ecol-pool/incoll/incoll_126.pdf)

Kalman, R. (1960), A new approach to linear filtering and prediction problems, *J. Basic Eng.*, *82*(2), 35–45.

Kitagawa, G., T. Takanami, and N. Matsumoto (2001), Signal extraction problems in seismology, *Int. Stat. Rev.*, *69*(1), 129–152.

Lamb, R. (1999), Calibration of a conceptual rainfall-runoff model for flood frequency estimation by continuous simulation, *Water Resour. Res.*, *35*(10), 3103–3114.

Leng, C. H., and H. D. Yeh (2003), Aquifer parameter identification using the extended Kalman filter, *Water Resour. Res.*, *39*(3), 1062, doi:10.1029/2001WR000840.

Liu, J. S., and R. Chen (1998), Sequential Monte Carlo methods for dynamic systems, *J. Am. Stat. Assoc.*, *93*(443), 1032–1044.

Marshall, A. (1956), The use of multi-stage sampling schemes in Monte Carlo computations, in *Symposium on Monte Carlo Methods*, edited by M. Meyer, pp. 123–140, Wiley, New York.

Marshall, L., D. Nott, and A. Sharma (2005), Hydrological model selection: A Bayesian alternative, *Water Resour. Res.*, *41*, W10422, doi:10.1029/2004WR003719.

Marshall, L., A. Sharma, and D. Nott (2006), Modeling the catchment via mixtures: Issues of model specification and validation, *Water Resour. Res.*, *42*, W11409, doi:10.1029/2005WR004613.

Moore, R. J. (1985), The probability-distributed principle and runoff production at point and basin scales, *Hydrol. Sci. J.*, *30*(2), 273–297.

Moore, R. J. (2007), The PDM rainfall-runoff model, *Hydrol. Earth Syst. Sci.*, *11*(3), 483–499.

Moore, R. J., and V. A. Bell (2002), Incorporation of groundwater losses and well level data in rainfall-runoff models illustrated using the PDM, *Hydrol. Earth Syst. Sci.*, *6*(1), 25–38.

Moore, R. J., and R. T. Clarke (1981), A distribution function approach to rainfall runoff modeling, *Water Resour. Res.*, *17*(5), 1367–1382.

Moradkhani, H., K.-L. Hsu, H. Gupta, and S. Sorooshian (2005a), Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, *Water Resour. Res.*, *41*, W05012, doi:10.1029/2004WR003604.

Moradkhani, H., S. Sorooshian, H. V. Gupta, and P. R. Houser (2005b), Dual state-parameter estimation of hydrological models using ensemble Kalman filter, *Adv. Water Resour.*, *28*(2), 135–147.

Nash, J. E., and J. Sutcliffe (1970), River flow forecasting through conceptual models part I: A discussion of principles, *J. Hydrol.*, *10*(3), 282–290.

Peters, N. E., J. Freer, and K. Beven (2003), Modelling hydrologic responses in a small forested catchment (Panola Mountain, Georgia, USA): A comparison of the original and a new dynamic TOPMODEL, *Hydrol. Processes*, *17*(2), 345–362.

Petersen-Overleir, A. (2004), Accounting for heteroscedasticity in rating curve estimates, *J. Hydrol.*, *292*(1–4), 173–181.

Pitt, M. K., and N. Shephard (1999), Filtering via simulation: Auxiliary particle filters, *J. Am. Stat. Assoc.*, *94*(446), 590–599.

Rajaram, H., and K. P. Georgakakos (1989), Recursive parameter estimation of hydrologic models, *Water Resour. Res.*, *25*(2), 281–294.

Senbeta, D. A., A. Y. Shamseldin, and K. M. O'Connor (1999), Modification of the probability-distributed interacting storage capacity model, *J. Hydrol.*, *224*(3–4), 149–168.

Sorensen, J. V. T., and H. Madsen (2004), Data assimilation in hydrodynamic modelling: On the treatment of non-linearity and bias, *Stochastic Environ. Res. Risk Assess.*, *18*(4), 228–244.

Thiemann, M., M. Trosset, H. Gupta, and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, *Water Resour. Res.*, *37*(10), 2521–2535.

von Neumann, J. (1951), Various techniques used in connection with random digits, in *Monte Carlo Method*, *Natl. Bur. of Stand. Appl. Math. Ser.*, vol. 12, p. 36–38, Natl. Bur. of Stand., Gaithersburg, Md.

Vrugt, J. A., H. V. Gupta, L. A. Bastidas, W. Bouten, and S. Sorooshian (2003), Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.*, *39*(8), 1214, doi:10.1029/2002WR001746.

Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten (2005), Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res.*, *41*, W01017, doi:10.1029/2004WR003059.

Wagener, T., D. P. Boyle, M. J. Lees, H. S. Wheater, H. V. Gupta, and S. Sorooshian (2001), A framework for development and application of hydrological models, *Hydrol. Earth Syst. Sci.*, *5*(1), 13–26.

Wagener, T., N. McIntyre, M. J. Lees, H. S. Wheater, and H. V. Gupta (2003), Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis, *Hydrol. Processes*, *17*(2), 455–476.

Whalley, N., R. S. Iredale, and A. F. Clare (2001), Reliability and uncertainty in flow measurement techniques—Some current thinking, *Phys. Chem. Earth, Part C*, *26*(10–12), 743–749.

Ye, M., S. P. Neuman, and P. D. Meyer (2004), Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff, *Water Resour. Res.*, *40*, W05113, doi:10.1029/2003WR002557.

Yeh, H. D., and Y. C. Huang (2005), Parameter estimation for leaky aquifers using the extended Kalman filter, and considering model and data measurement uncertainties, *J. Hydrol.*, *302*(1–4), 28–45.

K. J. Beven and P. J. Smith, Department of Environmental Science, University of Lancaster, Lancaster LA1 4YQ, UK. (p.j.smith@lancs.ac.uk)

J. A. Tawn, Department of Mathematics and Statistics, University of Lancaster, Lancaster LA1 4YF, UK.