

## The Multi-Model approach to rainfall–runoff modelling

ADAM M. WYATT & STEWART W. FRANKS

*School of Engineering, University of Newcastle, Callaghan 2308, New South Wales, Australia*  
[adam.wyatt@studentmail.newcastle.edu.au](mailto:adam.wyatt@studentmail.newcastle.edu.au)

**Abstract** This paper introduces the Multi-Model approach to rainfall–runoff modelling: a new lumped modelling method that incorporates numerous alternative process descriptions for the dominant processes within a catchment that affect the streamflow response to climatic forcing conditions. An extended GLUE approach is used to calibrate the numerous model structures contained, obtaining the best 100 parameter sets from two million (uniform) randomly sampled sets for each of the 45 model permutations contained. These model/parameter combinations were then used to produce prediction confidence limits for subsequent runoff predictions, including a calibration period, validation period and a synthetically determined extreme event period. Additionally, the ability of the calibration process to constrain the internal dynamics of models is investigated. The results indicate that calibration to simple runoff data alone is insufficient to constrain the saturation excess process description. Further potential conditioning of the models against saturated area is then investigated to refine the extreme event prediction uncertainty envelopes, showing that a single uncertain criteria limitation (on the extreme event flood peak saturated area) does little to improve the uncertainty envelopes. However, when multiple observations are utilized (on the flood peak and the recession saturated areas) the runoff generating process is sufficiently constrained to dramatically reduce the runoff prediction uncertainty envelopes for the extreme period, irrespective of model structure.

**Key words** extreme event; rainfall–runoff modelling; saturated area; uncertain conditioning

### INTRODUCTION

All environmental models are subject to three dominant sources of error that affect the accuracy of any predictions produced. These are input, parameter and model hypothesis uncertainty (Freer *et al.*, 1996; Kuczera & Franks, 2002; Sivapalan 2003). The modelling method described in this paper seeks to address the third source of modelling uncertainty, i.e. model hypothesis uncertainty. Environmental models consist of a series of mathematical functions that are simplifications of the very complex natural systems that are being modelled. Model hypothesis uncertainty arises when these simplifications do not reflect the natural system, therefore introducing an error into the model predictions. Indeed the huge spatial and temporal variation in natural systems present in catchments leads to the conclusions that the identification of a single model that is valid for all applications is impossible (Beven, 1989, 1993). As such, the development of rainfall–runoff models has typically been relatively catchment (and application) specific, often requiring manual modifications to a pre-existing model structure (e.g. Ambroise *et al.* (1996) demonstrated that three different

catchments required three different baseflow formulations within the TOPMODEL (e.g. Beven *et al.*, 1995 framework). The Multi-Model approach to rainfall–runoff modelling allows for the inclusion of numerous alternative process descriptions, as a means of incorporating the uncertainties that arise in rainfall–runoff modelling due to biases that are introduced into the runoff predictions through model structure uncertainty. By including many alternative process descriptions, runoff predictions are made using a broad range of modelling descriptions, weighted according to their ability to reproduce observed data series.

## MODEL OUTLINE

### Model structure

Multi-Model includes numerous alternative process descriptions for the principal components of most lumped rainfall runoff models: interception canopy, surface flow generation, root zone and deep storage zone. Figure 1, a schematic of Multi-Model shows how these processes are linked together, including the required inputs and outputs for each.

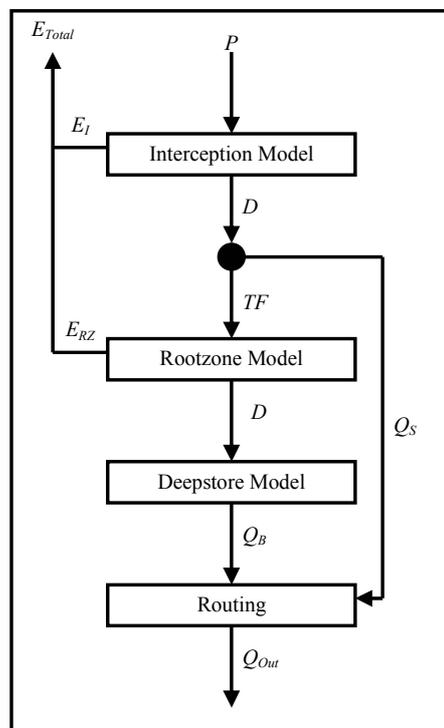


Fig. 1 Multi-Model schematic.

**Interception** The interception canopy models the influence of evaporation and canopy storage on the amount of water that reaches the land surface. Using the

observed rainfall ( $P$  [ $\text{LT}^{-1}$ ]) and potential evapotranspiration ( $E_P$  [ $\text{LT}^{-1}$ ]), canopy drainage ( $D$  [ $\text{LT}^{-1}$ ]) and evaporation losses ( $E_I$  [ $\text{LT}^{-1}$ ]) are computed, whilst accounting for changes in canopy storage ( $V_S$  [ $\text{L}$ ]). Three alternative computations are included: a “Null” model (i.e.  $V_S = 0$ ,  $D = P$  and  $E_I = 0$ ); a “Linear” model (i.e.  $D = \max[V_S - V_{CAP}, 0]$  &  $E_I = \min[E_P, E_P V_S / V_{CAP}]$ ); and a “Rutter” (e.g. Rutter *et al.*, 1971, 1975) model:

$$D = \begin{cases} 0 & \text{if } V_S - V_{CAP} \leq 0 \\ D_0 \exp[\beta(V_S - V_{CAP})] & \text{if } V_S - V_{CAP} > 0 \end{cases} \quad (1)$$

**Surface flow** Surface flow ( $Q_S$  [ $\text{LT}^{-1}$ ]) is calculated as the canopy drainage that falls on the saturated fraction ( $A_S$  [-]) of the catchment (i.e.  $Q_S = A_S D$ ), with the balance becoming throughfall ( $TF$  [ $\text{LT}^{-1}$ ]).

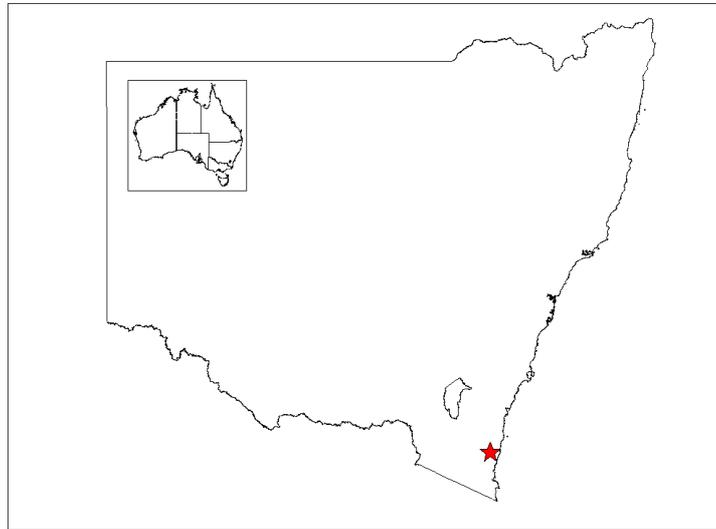
**Rootzone** The rootzone component models the influence of evapotranspiration losses ( $E_{RZ}$  [ $\text{LT}^{-1}$ ]) and storage effects ( $S_S$  [ $\text{L}$ ]) within the uppermost soil layers on the infiltration ( $I$  [ $\text{LT}^{-1}$ ]) to deep storage. Aside from the “Null” model ( $E_{RZ} = 0$ ,  $I = TF$  and  $S_S = 0$ ), four parametric models are included. Evapotranspiration losses are linearly proportional to storage depth with respect to either the storage capacity (i.e.  $E_{RZ} = \min[(E_P - E_I), (E_P - E_I) S_S / S_{CAP}]$ ) or storage threshold (i.e.  $E_{RZ} = \min[(E_P - E_I), (E_P - E_I) S_S / S_{THRES}]$ ). Infiltration is computed as either the storage excess (i.e.  $I = \max[S_S - S_{CAP}, 0]$ , where  $S_S^+ = S_S^- + TF - E_{RZ}^-$ ), or as a fixed fraction of  $TF$ , with any additional storage excess (i.e.  $I = I^0 + \alpha TF$ , (i.e. where  $I^0 = \max[S_S - S_{CAP}, 0]$  and  $S_S^+ = S_S^- + (1 - \alpha)TF - E_{RZ}^-$ ).

**Deep storage** The deep storage model produces the baseflow ( $Q_B$  [ $\text{LT}^{-1}$ ]) and saturated area ( $A_S$  [-]). Two alternative models are included: a “Linear” model which uses simple linear functions to relate deep storage depth to the saturated area (i.e.  $A_S = \min[S_D \gamma / S_{D_{MAX}}, 1]$ ) and baseflow (i.e.  $Q_B = k_B S_D$ ), and; the “VIC” (e.g. Wood *et al.*, 1992), which uses a nonlinear equation to relate saturated area to deep storage (i.e.  $A_S = 1 - (1 - S_D / S_{MAX})^{\beta/(\beta+1)}$ ). In cases where the “Null” rootzone model is in use, evapotranspiration is computed from the deep storage (i.e.  $E_D = (E_P - E_I) (1 - (1 - S_D / S_{D_{max}})^{1/\beta_e})$ )

## Structure management

The resulting range of model permutations is managed using a “look-up” table which assigns a unique model identification number ( $MI$ ) to each model permutation. For each model component ( $i$ ), a component option ( $j_i$ ) is selected from the number of possibilities ( $N_i$ ), which combine according to equation (2) below to produce the unique  $MI$ .

$$I = 1 + \sum_{i=1}^M b_i (j_i - 1), \text{ where } b_i = \prod_{k=0}^{i-1} N_k \text{ and } 1 \leq j_i \leq N_i \quad (2)$$



**Fig. 2** Location of Brogo River (represented by the star), NSW, Australia

**Table 1** Physical and hydrological characteristics of the Brogo River.

Characteristic	Value	Unit
Catchment area (total forested)	460 (430)	km <sup>2</sup>
Latitude	36.54	Degrees south
Longitude	149.83	Degrees east
Mean annual rainfall	906 (416.76)	mm ( $\times 10^6$ m <sup>3</sup> )
Mean annual runoff	291 (133.86)	mm ( $\times 10^6$ m <sup>3</sup> )
Runoff coefficient	0.32	–
Mean annual potential evapotranspiration	1078 (495.88)	mm ( $\times 10^6$ m <sup>3</sup> )

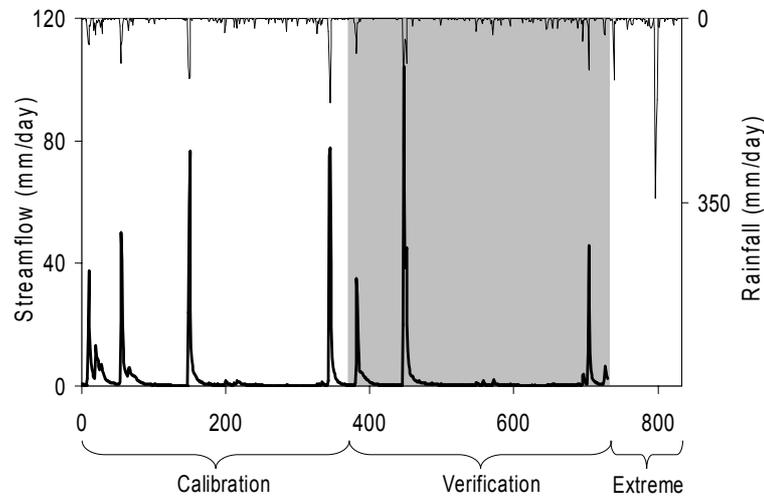
## STUDY SITE AND DATA

### Catchment location

This study is focused on the Brogo River catchment, North Brogo, situated in the southeastern corner of New South Wales, Australia (see Fig. 2). Table 1 summarizes some physical and hydrological characteristics of the study catchment.

### Observed data series

One of the common applications of rainfall–runoff models is the prediction of responses to extreme events. However, the validity of such predictions is uncertain since this requires the models to be applied to forcing conditions that are well beyond the observed limits used for calibration (e.g. Kalma *et al.*, 2001; Seibert, 2003). To test the performance of the model performance outside the forcing conditions, a limited data period is used for calibration and preliminary verification. A two year period of area-corrected observed precipitation, streamflow and potential evapotranspiration data (recorded at daily time steps) is selected and divided into calibration and verification



**Fig 3** Observed Rainfall & Runoff series, with appended (synthetically derived) Extreme event period.

periods of one year each. An additional 100 day period of observed data was appended to this series, containing the 1 in 100 year, 72 h duration design storm for the catchment (calculated according to the method described in *Australian Rainfall and Runoff*, Chapters 2 and 3, Pilgrim *et al.*, 1987) to test the behaviour of the models beyond the observed limits of the forcing data (see Fig. 3).

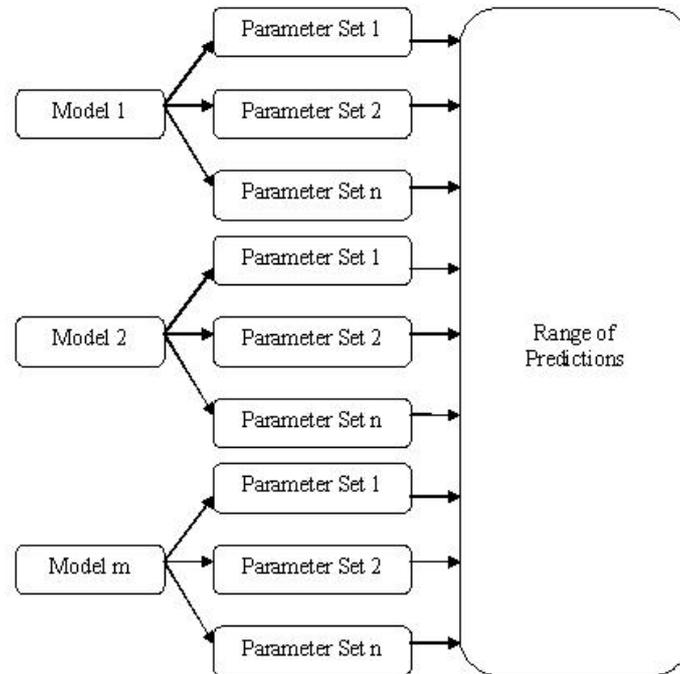
### Glue calibration

The Generalized Likelihood Uncertainty Estimation scheme (e.g. Beven & Binley, 1992; Beven, 1993; Romanowicz *et al.* 1994; Freer *et al.*, 1996; Franks *et al.*, 1998) provides a means of conditioning the parameter distributions of environmental models and generate prediction uncertainty envelopes that incorporate parameter uncertainties. When combined with the Multi-Model method, a means of producing runoff predictions that combine both parameter and model structure uncertainty is achieved. Figure 4 below shows how the multiple parameter-sets and model structures interact to produce the range of runoff predictions.

For each of the 45 model permutations contained within Multi-Model, the best 100 parameter sets from two million randomly sampled parameter sets are stored (resulting in a total of 90 million GLUE realizations, with 4500 parameter samples and model structures being used for uncertainty envelope estimates). The performance of each parameter set is assessed and compared for the calibration period using the Sum of Squared Errors (SSE).

### Confidence limits

The efficiencies that are calculated during the calibration period are used to compute the 95 percentile confidence limits for the modelled runoff series. To do this, the raw



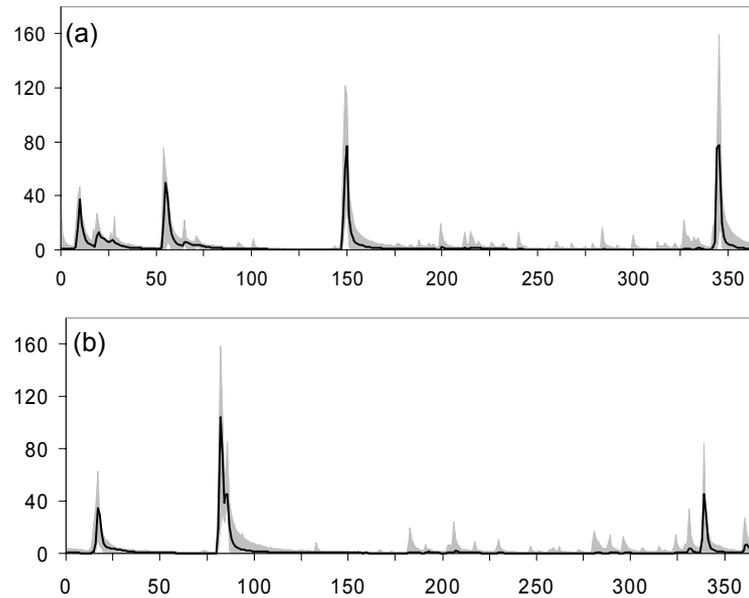
**Fig. 4** Schematic of the Extended GLUE method employed.

SSE values for each model / parameter set combination ( $n = 1$  to 4500) is converted to a relative likelihood ( $L_n$  [-]) according to  $L_n \propto (1/SSE_n)^k / \sum_{n=1}^{4500} (1/SSE_n)^k$ . These relative likelihoods are then used to compute the 95 percentile confidence limits for the runoff predictions.

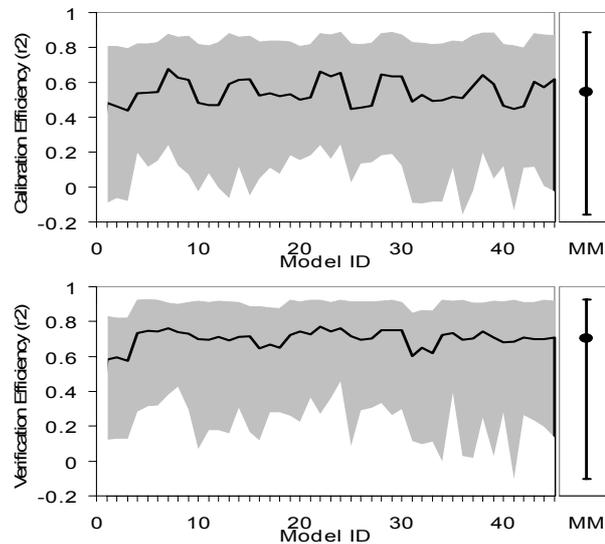
## MODEL SIMULATIONS

### Calibration series

The calibration period performance of Multi-Model (across all model permutations and associated parameter sets) is shown below in Fig. 5(a). Importantly, the 95 percentile confidence limits of the modelled streamflows contain the entire observed series, including peaks and low flow periods, indicating that across the spectrum of model structures, the complete dynamic of the system (at least within the calibration period) is modelled. The upper confidence limit seems to indicate an overestimation of the low flow periods. This is possibly a result of parameter sets which are poorly targeted at low-flow periods (since the *SSE* places little weight on low-flow periods); some model structures that are unable to handle extended periods of low-flows; errors within the observed rainfall series (which will be amplified during such periods), or a combination of the three.



**Fig. 5** 95% confidence limits (shaded region) on the modelled runoff and observed time series (solid line) for (a) the Calibration, and (b) Validation periods.



**Fig 6** Summary of the calibration and verification Nash-Sutcliffe efficiencies for each model structure and overall Multi-Model

Figure 6 provides a summary of the performance of each model contained within the Multi-Model framework, highlighting the wide range of average performances between the “best” and “worst” of the individual model structures. Although not shown here, the uncertainty limits that are observed for the “best” model very closely resemble those of the Multi-Model, whereas the uncertainty limits for the “worst” model fail to contain any of the observed low-flow period.

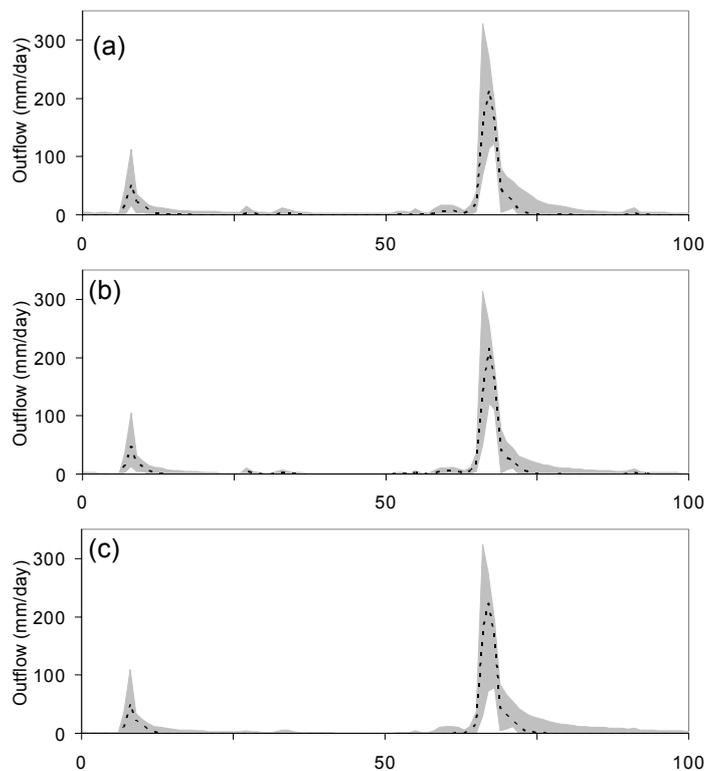
### Verification limits

The performance within the validation period (Fig. 5(b)) shows the performance of Multi-Model outside the calibration period. Like the calibration period performance, the observed runoff series is well contained within the 95% confidence limits produced by Multi-Model.

Again Fig. 6 shows the wide range of performances for the various models contained within Multi-Model. The uncertainty limits computed for the validation period are congruous with those computed for the calibration period, with a further decay of the uncertainty limits of the “worst” model, failing to contain parts of the flood recession and overestimates the low flow periods (for the same reasons that are mentioned above).

### Extreme event performance

Figure 7(a) shows the Multi-Model performance for the extreme event, well above the forcing ranges observed during both the calibration and validation periods. Note that the uncertainty estimates for the extreme peak flow range from  $\sim 120 \text{ mm day}^{-1}$  to  $340 \text{ mm day}^{-1}$ . The “best” and “worst” of the individual models contained within Multi-Model for the extreme period produce uncertainty limits that are similar to those achieved with the Multi-Model approach (Fig. 7(b) and 7(c), respectively). This is particularly interesting as it demonstrates that irrespective of model structure, extrapolation to extreme events seems highly uncertain.



**Fig. 7** 95% confidence limits (shaded region) and median (dashed line) for the extreme event performance using (a) Multi-Model; (b) the “Best” model; and (c) the “Worst” model according to calibration efficiencies.

## EXTREME EVENT PREDICTION REFINEMENT USING UNCERTAIN CONDITIONING

### Uncertain conditioning – Saturated area

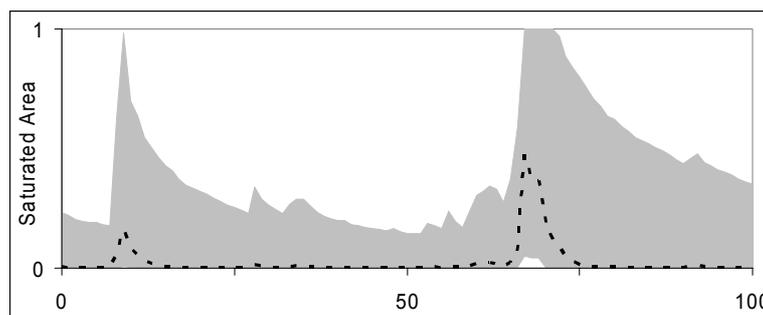
The wide uncertainty limits observed during the extreme event period highlight the shortcomings of “runoff-only” based calibration common for most calibration strategies. The wide modelled saturated area uncertainty envelopes for the extreme event (shown in Fig. 8) show clearly that the inability of “runoff-only” calibration techniques to constrain the internal variables of conceptual rainfall–runoff models. These wide uncertainties provide an opportunity to introduce additional information to describe the saturated areas at selected points in the times series, in order to “differentiate between previously acceptable models, rejecting many as non-behavioral” (Franks *et al.*, 1998). In the absence of real saturated area measurements, reasonable estimates are used.

### Case 1: Single uncertain constraint

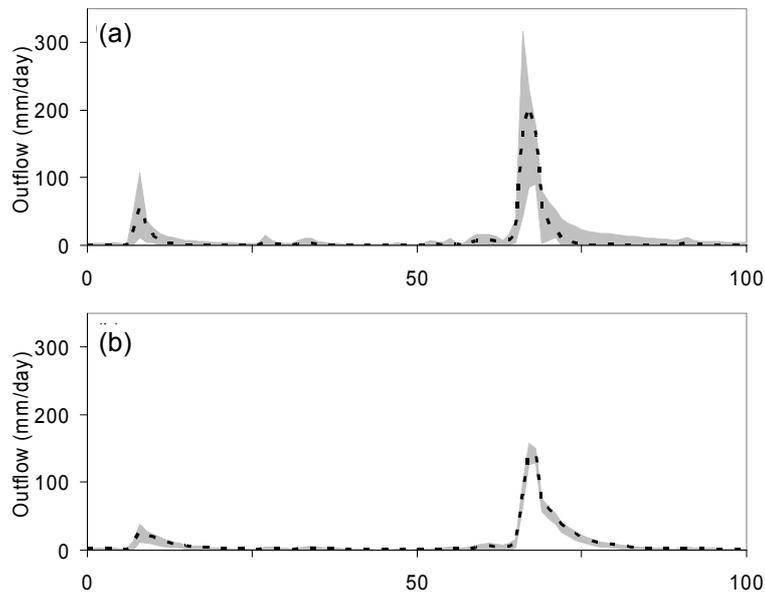
The first implementation of uncertain conditioning applies a single uncertain constraint on the saturated area at the flood peak (day 68 in the time series). An upper limit of 30% and a lower limit is 25%. This reduction in saturated area uncertainty did not however translate into a reduction in the uncertainty estimate envelopes for the predicted runoffs (Figure 9(a)), which changed very little when compared to the original extreme event period. These results are consistent with Blazkova & Beven (2003).

### Case 2: Two uncertain constraints

The second implementation of uncertain condition of the extreme event predictions added an additional uncertain condition to the saturated area, four days after the extreme flood peak (i.e. day 72 in the time series), with upper and lower saturated area limits of 25 and 20%, respectively. The resulting runoff prediction confidence limits for the extreme period (Fig. 9(b)) are found to be dramatically narrower than previous runoff predictions, resulting in a very well defined extreme event runoff prediction.



**Fig. 8** 95% confidence limits (shaded) and median (dashed line) for the modelled saturated area during the Extreme event using Multi-Model.



**Fig. 9** 95% uncertainty limits (shaded) and median (dotted line) for the extreme event predictions using (a) a single uncertainty constraint, and (b) two uncertain constraints on saturated area.

## CONCLUSIONS

The Multi-Model approach presented here provides a means to make runoff predictions which, when combined with the GLUE methodology, produce confidence limits that include uncertainty effects arising from both parameter uncertainty (by using multiple parameter samples) and model uncertainty (by using multiple models equally, regardless of complexity). The extended GLUE approach allows the numerous runoff predictions (based on different model structures and different parameters) to be treated equally, regardless of model complexity, since the model behaviour that is required beyond the calibration observation limits in forcing conditions is unknown, meaning that any method that weights the model complexity and calibration performance together (e.g. Bayes factors) are only valid when applied to forcing conditions within the range of hydrological conditions observed during the calibration period.

The usefulness of combining uncertain conditioning estimates for the internal variables that are used within Multi-Model was also shown to be an effective means of constraining the runoff predictions from what were very wide confidence limits (resulting from the wide range of models used and the range of associated parameter sets, being applied well outside of the limits of the observed forcing conditions within the calibration period). Franks *et al.* (1998) showed valuable constraint of TOPMODEL simulations when conditioned on saturated area during a single storm. However, the constraint appears largely as a function of conditioning only a single, predetermined model. As has been shown here, when multiple alternative and possibly divergent model representations are utilized, a single constraint is insufficient to constrain the dynamics of these representations, indicating that multiple observations of saturated

area may be required to improve the realism of model simulations. Above all, it must be acknowledged that calibration to runoff data alone is insufficient to constrain the key flow generating mechanisms in lumped rainfall–runoff models.

## REFERENCES

- Ambrose, B., Beven, K. & Freer, J. (1996) Toward a generalization of the topmodel concepts: topographic indices of hydrologic similarity. *Water Resour. Res.* **32**(7), 2135–2145.
- Beven, K. (1989) Changing ideas in hydrology—the case of physically-based models. *J. Hydrol.* **105**, 157–172.
- Beven, K. & Binley, A. (1992) The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes* **6**, 279–298.
- Beven, K. J. (1993) Prophecy, reality and uncertainty in distributed hydrological modelling. *Adv. Water Resour.* **16**, 41–51.
- Beven, K., Lamb, R., Quinn, P., Romanowicz, R. & Freer, J. (1995) Topmodel. In: *Computer Models of Watershed Hydrology* (ed. by V. Singh), 627–668. Water Resources Publications, Fort Collins, Colorado, USA.
- Blazkova, S., Beven, K. J. & Kulasova A. (2002) On constraining TOPMODEL hydrograph simulations using partial saturated area information. *Hydrol. Processes* **16**, 441–458.
- Chappell, N. A., Franks, S. W. & Larenus, J. (1998) Multi-scale permeability estimation for a tropical catchment. *Hydrol. Processes* **12**, 1507–1523.
- Franks, S. W., Gineste, P., Beven, K. J. & Merot, P. (1998) On constraining the predictions of a distributed model: The incorporation of fuzzy estimates of saturated areas into the calibration process. *Water Resour. Res.* **34**(4), 787–797.
- Freer, J., Beven, K. & Ambrose, B. (1996) Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. *Water Resour. Res.* **32**(7), 2161–2173.
- Kalma, J. D., Franks, S. W. & van den Hurk, B. (2001) On the representation of land fluxes for atmospheric modelling, *Met. Atmos. Phys.* **76**, 53–67.
- Kuczera, G. & Franks, S. W. (2002) Testing hydrologic models: fortification or falsification? In: *Mathematical Modelling of Large Watershed Hydrology* (ed. by V. P. Singh & D. K. Frevert), 141–185. Water Resources Publications, Fort Collins, Colorado, USA.
- Pilgrim, D. H., Kennedy, M. R., Rowbottom, I. A., Cordery, I., Canterford, R. P. & Turner, L. H. (1987) Temporal patterns of rainfall bursts. In: *Australian Rainfall and Runoff: A Guide to Flood Estimation*. Institution of Engineers, Australia.
- Romanowicz, R., Beven, K. J. & Tawn, J. (1996) Bayesian calibration of flood inundation models. In: *Floodplain Processes* (ed. by M. G. Anderson, D. E. Walling & P. D. Bates), 333–360. John Wiley, New York, USA.
- Rutter, A. J., Kershaw, K. A., Robins, P. C. & Motron, A. J. (1971) A predictive model of rainfall interception in forests. I. Derivation of the model from observations in a plantation of Corsican pine. *Agric. Met.* **9**, 367–384.
- Rutter, A. J., Morton, A. & Robins, P. C. (1975) A predictive model of rainfall interception in forests. II. Generalization of the model and comparison with observations in some coniferous and hardwood stands. *J. Appl. Ecol.* **12**, 367–380.
- Seibert, J. (2003) Reliability of models predictions outside calibration conditions. *Nordic Hydrol.* **34**(5), 477–492.
- Sivapalan, M. (2003) Process complexity and hillslope scale, process simplicity at the watershed scale: is there a connection? *Hydrol. Processes* **17**, 1037–1041.
- Spear, R. C. & Hornberger, G. M. (1980) Eutrophication in peel inlet. 2, Identification of critical uncertainties via generalised sensitivity analysis. *Water Res.* **14**, 43–49.
- Wood, E. F., Lettenmaier, D. P. & Zartarian, V. G. (1992) A land–surface hydrology parameterization with subgrid variability for general circulation models. *J. Geophys. Res.* **97**(D3), 2717–2728.