How adequately do gauged tributaries represent the hydrological behaviour of ungauged tributaries in modelling large regulated catchments?

J.F. Costelloe [a], R. Adams [b] and A.W. Western [a]

[a] *Department of Infrastructure Engineering, The University of Melbourne, VIC 3010*
[b] *RMIT University, Melbourne, VIC, 3000*
Email: *jcost@unimelb.edu.au*

Abstract: The issue of how adequately gauged, unregulated tributaries represent the hydrological behaviour of ungauged, unregulated tributaries is rarely considered within the context of a single large regulated catchment. We show that in a catchment where the gauging of tributaries is biased towards perennial streams, this bias can result in overestimations of the water balance. The Broken River of south-eastern Australia is regulated and two gauged, unregulated tributaries comprise 49% of the catchment area downstream of the main storage. Both of these tributaries are largely perennial but short-term monitoring of other unregulated tributaries indicates that most of the remaining catchment downstream of the storage have intermittent to ephemeral flow regimes. Deterministic rainfall-runoff model parameters derived from calibration of the perennial tributaries were applied to ungauged intermittent tributaries and found to over-predict runoff, particularly during the low flow phase. Model performance was assessed in the ungauged tributaries using relatively short records (<18 months) of discharge data collected during a field campaign. The application of a probabilistic modelling approach (Generalised Likelihood Uncertainty Estimation, GLUE) to the gauged tributaries defined sets of behavioural model parameters with the median set better able to simulate the flow regime of the ungauged tributaries. The application of the GLUE approach to directly model data from the short record length of the ungauged tributaries required the development of "soft" model assessment criteria to identify behavioural datasets. The latter results showed some improvements in comparison to transferring behavioural datasets from gauged to ungauged tributaries but with wide ranges of uncertainty.
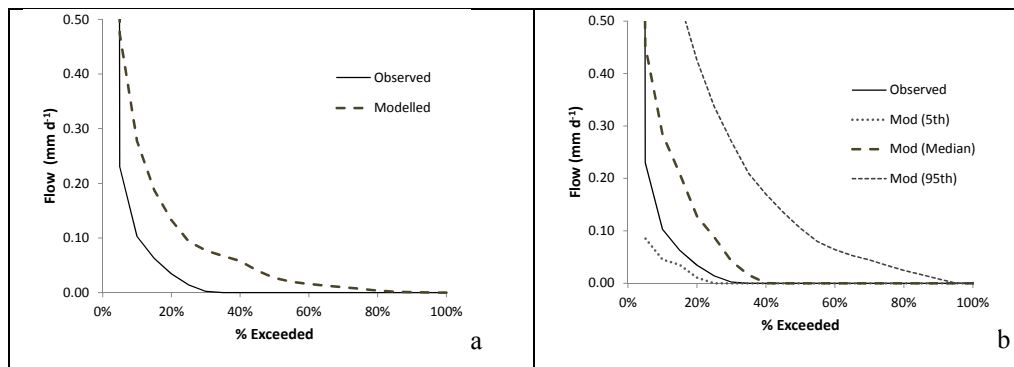
Figure. Flow duration curves for a previously ungauged, intermittent tributary showing observed versus modelled results from deterministic modelling (a) and probabilistic GLUE modelling (b).

Keywords: *Perennial, intermittent, flow regime, Generalised Likelihood Uncertainty Estimation*

1. INTRODUCTION

Hydrological modelling of large, regulated catchments (i.e. 1000 – 10,000 km^2) typically relies on calibration of conceptual rainfall-runoff models using data from gauged, unregulated sub-catchments (i.e. tributaries) and applying the models and calibrated parameter values to estimate streamflow from the remaining ungauged, unregulated sub-catchments (Sivapalan, 2003; Wagener and Montanari, 2011). The issue of how well the gauged tributaries represent the hydrological behaviour of ungauged tributaries is rarely considered within the context of a single, large, regulated catchment, although it has received considerable attention in the context of regionalisation studies (e.g. McIntyre et al., 2005; Reichl et al., 2009). Typically, the regionalisation studies compare catchment similarity using an extensive dataset of catchments distributed over large, regional (often sub-continental) scales. Within most water resource modelling of large catchments, the modeller would assume the 'spatial proximity' approach (c.f. Reichl et al., 2009) and would apply the parameter values from the nearest gauged tributary to ungauged tributaries, particularly in the absence of strong topographic (or other) gradients. The issue of representativeness of the gauged tributaries has important implications for the applicability of both model structure and calibrated parameter values, and for water resource management (Sivapalan, 2003; Wagener and Montanari, 2011).

In the Australian context, gauging stations are most common on rivers and streams that flow perennially due to their importance in contributing to the water resources of a catchment. A study of 830 Australian catchments with relatively natural flow regimes by Kennard et al. (2010) classified 63% as perennial or rarely ceasing to flow and these were clustered around the sub-humid to humid regions of eastern and southern Australia. The gauged intermittent to ephemeral streams were mostly confined to northern Australia and semi-arid inland areas. However, intermittent to ephemeral streams commonly occur to relatively high stream orders and represent significant percentages of larger catchments, particularly in the lowland reaches in sub-humid to semi-arid regions. Using models calibrated on gauged perennial tributaries to estimate unregulated streamflow may bias the water balance, and over-predict discharge from an intermittent-ephemeral tributary. In this study, we investigate the transferability of model structure and parameter values from gauged, unregulated tributaries of a large, regulated catchment (Broken River, southeastern Australia) to ungauged, unregulated tributaries of the same catchment with differing flow regime behaviour. This addresses the question of whether the flow regime of more perennial tributaries contains enough information to predict the flow regime of more intermittent tributaries and whether addition flow information is required.

2. THE STUDY CATCHMENT

The Broken River is a typical large, regulated sub-humid catchment of southeastern Australia. It has a catchment area of approximately 2500 km^2 (Figure 1) and is a major tributary of the Goulburn River. River flows in the Broken River have been regulated since 1967 by Lake Nillahcootie and the main-stem of the Broken River is gauged at five locations downstream of this reservoir (Figure 1). There are also two gauged, unregulated tributaries downstream of Lake Nillahcootie, Holland Creek (HC, area 460 km^2) and Lima Creek (LC, area 93 km^2). The total gauged area downstream of Lake Nillahcootie, not including main-stem gauges, comprises approximately 49% of this area. Mean annual rainfall for the catchment is 800 mm y^{-1}, the annual average areal potential evapotranspiration rate for the period 1961-2001 varied from 963 to 1190 mm y^{-1} and streamflow is winter-spring dominant (Australian Bureau of Meteorology data).

In order to determine the representativeness of the two gauged tributaries in the Broken, additional low-cost streamflow monitoring was installed at five unregulated Previously-Ungauged Tributaries (PUTs, see Figure 1). The PUT flow dataset comprised up to 17 months of observed flow data from the middle of 2009 (except Back Creek - operational in May 2008). Observed flow values were estimated from a rating curve comprising at least three points defining the relationship between observed stage and flow at a rated section at each site. Flow was measured manually in the streams using a combination of handheld acoustic Doppler velocity meter (FlowTracker, SonTek), floating acoustic Doppler current profiler (RiverSurveyor S5, SonTek) and propeller current meters (CMC20, Hydrological Services P/L). Stage (i.e. water level) was measured at each rated section using pressure transducers with integral loggers (Odyssey, Dataflow Systems) that were vented for atmospheric pressure compensation. Gaugings tended to be carried out during low flows, which was partly due to the severe south-eastern Australian drought that occurred in the first half of the sampling period. Therefore, uncertainly bounds were set on the observed flows (Q_{OBS}) at each site. Generally, above the maximum gauged discharge (Table 1) the upper bound of Q_{OBS} was assumed to approach infinity. Flows were converted to runoff depths (in mm d^{-1}) by dividing by the gauged subcatchment area (Table 1). The period that the PUTs ceased to flow (CTF), defined as observed flows below 0.001 mm d^{-1}, was also calculated (Table 1). This indicated that SC11 and 14 were flowing for more than 95% of the monitoring period, whereas, the other three PUTs ceased to flow for 35- 70% of the period, so can be classified as

ephemeral. Daily flow data were obtained for the gauged catchments, Holland Creek (ID 404207) and Lima Creek (ID 404208), from the Victorian Data Warehouse (http://www.vicwaterdata.net) for the period 1974 to 2008. Examination of the gauging records for the gauged tributaries indicated that both were near-perennial (see observed flow duration curves in Figure 2), although Holland Creek did cease to flow for 6.6% of the period that data were collected from the PUTs, and for about 5% over the full record covering the period 1974-2008.
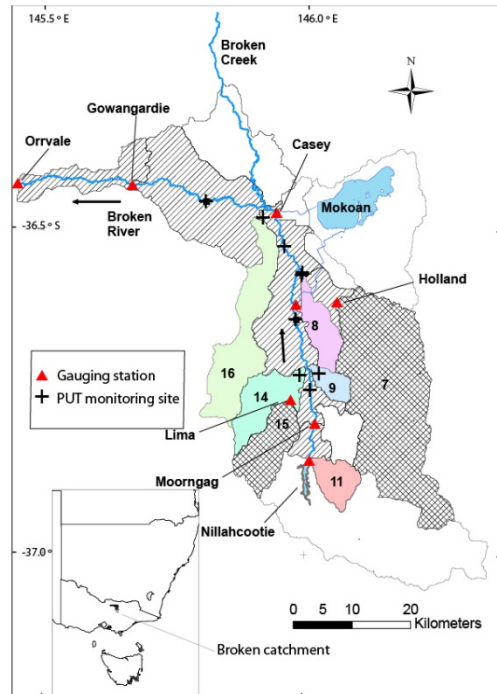

Figure 1. Broken River study reaches. The gauged, unregulated subcatchments (7 - Holland Creek, 15 – Lima North Creek) are shown with cross-hatching; unregulated PUTs are shown as shaded polygons (8 – Blind Creek, 9 – Castle Hill Creek, 11 – Back Creek, 14 – Lima North Creek, 16 – Five Mile Creek); unmonitored flanking subcatchments are shown with diagonal hatching; other unmonitored subcatchments are shown with no shading.

3. METHODS

In the first step (Section 3.1) the gauged tributaries were modelled using a deterministic approach. The parameter sets obtained from the calibration using the gauged flows were simply transferred to the ungauged tributaries (PUTs). This step assumed no prior knowledge of the behaviour of the ungauged tributaries except that obtained from the gauged tributaries (i.e. assuming they are representative of all ungauged tributaries in the Broken catchment). The second step (Section 3.2) explored the uncertainty around the parameter values, using the GLUE methodology on the gauged tributaries. The third step (Section 3.3) evaluated whether the application of the probabilistic GLUE model parameters obtained from the gauged tributaries (Step 2) to the PUTs could adequately represent the flow regime of the PUTs.

3.1. Deterministic modelling (Step 1)

The AWBM model (Boughton, 2004) was used and calibrated on the gauged tributaries (HC and LC) for the period 1988 to 2006 (1987 was used as a warm-up year) using a deterministic approach to obtain a single (optimal) parameter set from each subcatchment. The software package Rainfall Runoff Library was used for the model calibration. The pattern-search optimisation algorithm with multiple start points was used with the Nash-Sutcliffe Efficiency measure (NSE) as the primary objective function, calculated on the modelled (Q_{MOD}) and observed (Q_{OBS}) flows. The secondary objective function was to limit the mass balance error (MBE, where $MBE = 100*((Q_{OBS}-Q_{MOD})/Q_{OBS})$) to < 10%. The optimum parameter sets were transferred to the PUTs for estimating the flows using deterministic modelling. The set from the nearest of the two gauges was used (so LC parameters were used for SC9, SC11 and SC14, HC for the remainder). The modelled flows were then compared against observed data from the period 2009-2010, with particular attention to comparisons during low flow periods, using flow duration curves to qualitatively assess the model's performance at simulating the intermittent PUTs. An alternative version of the mass balance error, termed MBE[*] (Table 2), was also evaluated in this and subsequent steps involving the PUT data due to high flows not being acceptably constrained with the available rating data. MBE[*] was only calculated on pairs of flows where the observed flow was less than the maximum gauged flow (Q_{thresh}, Table 1).

3.2. GLUE simulation on gauged tributaries

In Step 2 we modified AWBM by introducing a loss term that represented the percolation to deep groundwater in each subcatchment (Q_{loss}), which enabled the model to better simulate intermittent flow patterns. The revised model, AWBM2, was used to simulate flows in the gauged tributaries using the GLUE methodology (Beven and Binley, 1992). The GLUE methodology provides a convenient framework to

estimate the uncertainty range of the predicted flows on each day, subject to the assumptions made when calculating the likelihood weights for each simulation. The NSE objective function was modified using an additional shaping parameter α (0-1) that raises the power of the NSE criteria to better weight the statistic for low flow periods (termed NSE^α, see Equation 1).

$$NSE^\alpha = 1 - \left[\frac{\sum SSE^\alpha}{\sum VAROBS^\alpha} \right] \qquad (1)$$

Where: SSE^α = Sum of squared error, based on transformed flow data $(Q_{OBS}^\alpha - Q_{MOD}^\alpha)^2$ and $VAROBS^\alpha$ = Variance of transformed observed flows Q_{OBS}^α. We also excluded high flows from the calibration using a filter (Q_{thresh}), since the focus on this study is on how well parameter sets from the gauged tributaries could replicate the low flow behaviour of the PUTs (value in Table 1). Behavioural parameter sets (total N_b) were then chosen on the basis of $MBE < \pm 10\%$ and NSE^α falling in the top 1% of all simulations (99th percentile). The α parameter was set to 0.5 for both HC and LC. Runs were carried out to test the sensitivity of the results to α. The final step in GLUE is to obtain a distribution of the predicted flows on each day, using the likelihood weights $L(\theta)_i$, as described in Beven (2009). In all, we generated 10,000 parameter realisations which were assessed for the period 1974-2008. This was the longest period with high quality flow data (few gaps) at both HC and LC. Parameter sets falling above the 99th percentile threshold of the NSE^α values were chosen, subject to the MBE constraint being met. Parameter values were constrained within the ranges suggested by Boughton (2004), apart from K_b, K_s and Q_{loss}. It was found that K_b should be constrained between 0.8 and 1.0 and always $> K_s$ (so K_s was constrained to be <0.8).

Table 1. Characteristics of Previously Ungauged Tributaries (PUTs)

Measure	Unit	SC8 (Blind Ck)	SC9 (Castle Ck)	SC11 (Back Ck)	SC14 (Lima Ck)	SC16 (5 Mile Ck)
Gauged Area	(km^2)	59.7	22.6	48.2	71.0	172.8
Q_{thresh}	(mm d^{-1})	2.62	0.09	2.18	1.12	0.32
Time Q> Q_{thresh}	%	1.8	1.8	14.6	2.2	13.2
Mean Q	(mm d^{-1})	0.22	0.02	0.39	0.31	0.26
% Time Zero Flow		70.0	35.8	1	2.6	60.3
Period of Record		3/7/2009-24/11/2010	3/7/2009-4/9/2010	29/5/2008-30/11/2010	2/7/2009-24/11/2010	4/7/2009-24/11/2010
% Missing Data		0	0	8.2	0	52.4

3.3. Simulating PUTs using gauged tributary behavioural parameters (Step 3)

The capacity of the gauged tributary parameter sets, defined by the GLUE methodology, to model flow in the ungauged tributaries was then tested against the PUT datasets, assuming no prior knowledge of the hydrological response of the PUTs themselves. Observed flow data were defined with upper and lower uncertainty bounds and flows above the maximum gauged discharge were also excluded from the analysis. We ran the N_b behavioural parameter sets, obtained by running AWBM2 on HC and, on the five PUTs in turn, and predicted the flows Q_{MODi} from each set i on each PUT ($i=1,N_b$). The results then comprised an array of N_b sets of modelled flow time-series j (Q_{MODj}). For each PUT we extracted the 5th (Q_{MOD5}) and 95th (Q_{MOD95}) percentiles of the Q_{MOD} data on each day, from the time series array. For Step 3 we assumed that each simulation had an equal likelihood (no formal likelihood function was defined). To assess the fraction of time that the observed flows lie within the 5th and 95th percentile of the modelled flows on each day, a containment metric B, proposed by Ewen et al. (2006), was used in this and the subsequent PUT GLUE analyses (Equation 2).

$$B = \frac{\sum days\ with\ Q_{OBS(LB)} \leq Q_{MOD5}\ \&\ Q_{OBS(UB)} \geq Q_{MOD95}}{N} \qquad (2)$$

Where N = the total number of days of flow data, and the subscripts LB and UB indicate the lower and upper bounds respectively of Q_{OBS}. A value of unity for B indicates that the observed flows always lie within the modelled bounds. Periods of zero or missing Q_{OBS} data were excluded from this analysis.

4. RESULTS

4.1. Deterministic modelling of the gauged catchments (Step 1)

The AWBM model performance in the deterministic simulation of the gauged tributaries produced acceptable results as shown by flow duration curves comparisons (Figure 2) and performance metrics in

Table 2. Model parameters were similar for both catchments, except the *BFI* which was much higher in LC (0.80) compared to HC (0.49), indicating a higher contribution of baseflow to total flow in LC (Table 2). Total storage (*C*) differed between the catchments by only 4% which was surprising as the historical HC data indicated that it ceased to flow for about 5% of the period. The modelled flow duration curves showed a good fit with the observed data but the model was unable to simulate periods of zero flow, and tended to compensate for over-predicting the low flows by under-predicting the higher flows in both catchments.

We then transferred the Lima Creek and Holland Creek AWBM parameter sets to the PUTs and estimated the flows from each with no prior knowledge of the PUT observed data and no recalibration of the model. It was found that the deterministic AWBM simulations over-predicted flows from the PUTs, as shown by the large negative MBE* values with the exception of SC8 (Table 2), and flow duration curves (Figure 3). The PUT with the lowest mass balance error and best hydrograph fit was Lima North (SC14) with a MBE* <30%. In the case of Lima North, the flow regime (cease to flow for only 2.6% of the monitored period – Table 1) is similar to that of the neighbouring Lima Creek (SC15) gauged catchment (compare flow-duration curves in Figures 2b and 3b). In the case of Blind Creek (SC 8), this tributary had the highest gauged discharge and so had the best constrained MBE over the entire flow range (MBE = +38.6% compared to negative MBE values for other PUTs). The deterministic model simulated the high flows reasonably well for this tributary (Figure 3a). However, the observed flow duration curve (Figure 3a) indicates that this creek ceased to flow for about 70% of the total record but AWBM failed to reproduce this and generally poorly simulated low flow periods. NSE values were poor with tributaries Back Creek (SC11, 0.45) and Blind Creek (SC8, 0.42) showing the highest values (Table 2). Any modelling study of the PUTs should reject these simulations as not being able to reproduce the observed data.
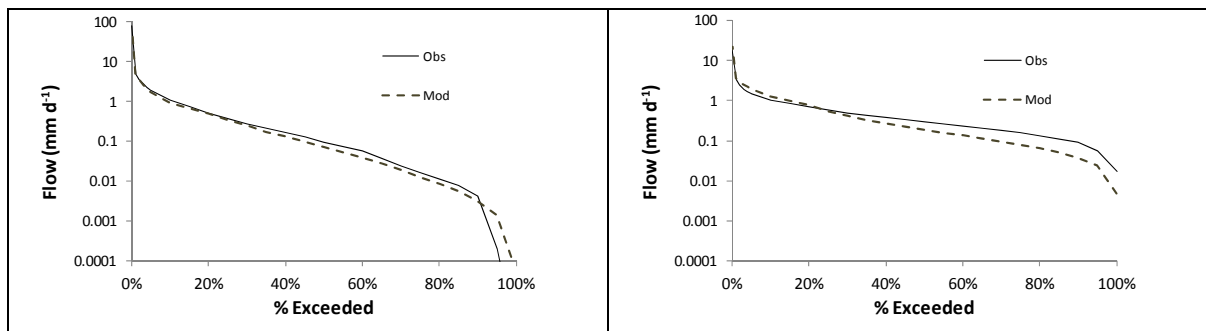

Figure 2. Flow duration curves from deterministic modelling (Step 1): (a) Holland Creek (b) Lima Creek.

4.2. GLUE simulations of gauged tributaries (Step 2)

The uncertainty range around the modelling of the gauged tributaries was then assessed with GLUE simulations using AWBM2. The model performance measures are shown in Table 2. A total of 453 parameter sets were selected as behavioural out of the 10,000 GLUE runs. NSE^{α} values from the runs on LC were usually worse than HC (Table 2) and in fact only 200 behavioural sets were obtained from LC. Over the simulation period the MBE values averaged out so that periods of over-prediction were compensated by periods of under-prediction, ensuring that the final error lay within the pre-set upper and lower bounds of ±10%. The containment statistic also varied between HC (*B*=0.86) and LC (*B*=0.51), with the latter result indicating that the observed flow for LC only lay within the modelled upper (95th) and lower bounds (5th) for just over half the time. We also attempted to apply the same methodology to modelling flow from the PUTs using AWBM2. The GLUE methodology was applied to each PUT in turn using the 15-17 months of flow data measured from 2009-10 and the same parameter bounds as for the gauged tributaries. With the exception of Lima Creek (SC 14) it was not possible to obtain behavioural parameter sets (i.e. passing both the performance criteria) using this method.

4.3. Simulation of PUTs using behavioural datasets from gauged tributaries (Step 3)

The results obtained from applying the GLUE-generated HC and LC behavioural parameters to run AWBM2 on the PUTs are shown in Table 2. The containment statistic *B* was surprisingly low (<0.25) for all tributaries indicating a wide distribution of modelled flow values tended to lie outside the upper and lower bounds of the observed flows. However, the median MBE* measure showed an improvement compared to the application of deterministic parameter sets to the PUTs (Table 2). This modest improvement is further illustrated with example flow duration curves for two tributaries (Blind Creek and Lima North Creek, Figure 3c-d) shown in comparison with Step 1 deterministic simulations on the same tributaries (Figure 3a-b). Results for Blind

Creek (SC8) indicated that the flow duration curve representing the median modelled flow improved upon the deterministic modelling in being able to better predict the cease-to-flow period but generally over-predicted the flow during periods when the creek was observed to be flowing. The median modelled flow duration curve for Lima North Creek (SC14), the most perennial of the PUTs, was quite similar in shape to the deterministic flow duration curve. These results indicate the median of the GLUE behavioural parameter sets trained on the gauged tributaries were able to improve the modelling of the PUTs, particularly those with an ephemeral flow regime, in comparison to using the deterministic parameter sets from the gauged tributaries.
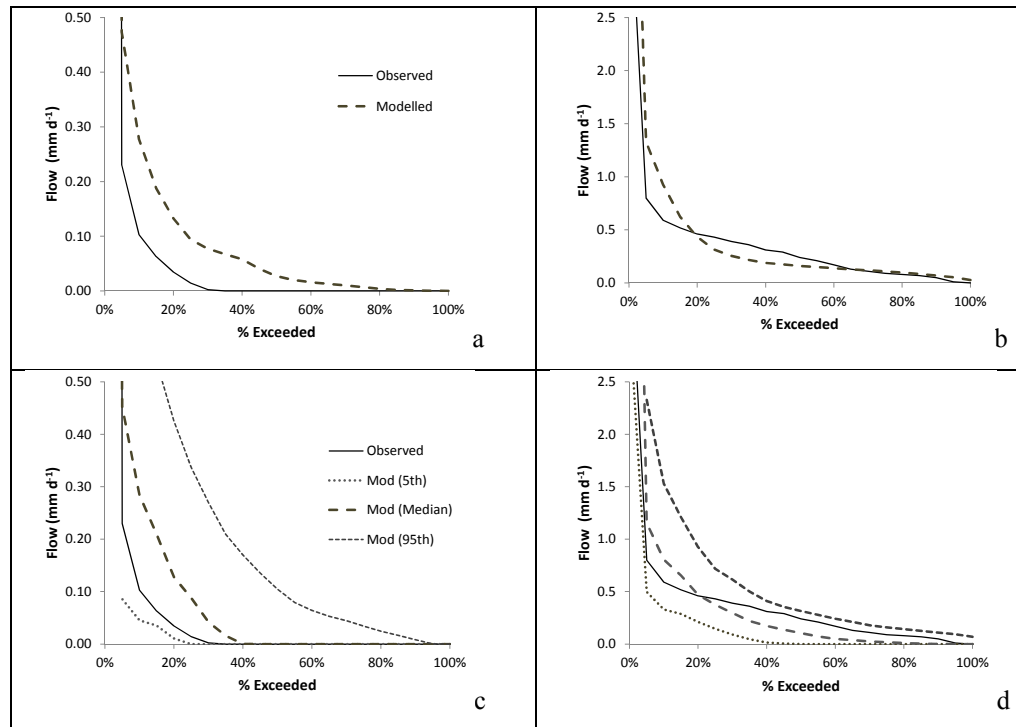

Figure 3. Flow duration curves from SC8 (a,c) and SC14 (b,d) from Step 1 deterministic modelling (a,b) and Step 3 GLUE modelling (c,d).

Table 2. Model performance measures for both deterministic and GLUE modelling for the gauged (HC, LC) and ungauged (SC8, SC9, SC11, SC14, SC16) tributaries.

Performance Measures	HC	LC	SC8	SC 9	SC11	SC14	SC16
NSE (Step 1) (-)	0.86	0.67					
Mean NSE$^{\alpha}$ (Step 2) (-)	0.76	0.61					
MBE (Step 1) (%)	8.8	-0.32					
Mean MBE (Step 2) (%)	0.97	-0.15					
MBE range (Step 2) (%)	-9.97-10.0	-9.96 - 9.98					
NSE (Step 1)			0.42	<0	0.45	0.01	0.10
MBE (Step 1) mm y^{-1}			+33	-68	-19	-15	-18
MBE (Step 1) %			+39	-1104	-13	-13	-18
MBE* (Step 1) %			+251	-2650	-88	-30	-133
Containment (*B*) (Step 3)			0.04	0.05	0.23	0.05	0.1
MBE*(Step 3) Median %			+19.3	-1500	-56	+12.3	-116

5.   DISCUSSION AND CONCLUSIONS

Constrained resources mean that large catchments are unlikely to have all significant, unregulated tributaries gauged. As a result, the issue of the representativeness of the hydrological behaviour of the gauged tributaries compared to the remaining ungauged, unregulated tributaries is important for water resource modelling and management. This is particularly important during the low flow phases when water resource managers need to determine reservoir releases required to fulfil environmental flow requirements. The flow duration curves for the gauged tributaries (Figure 2) show perennial (Lima Creek) to perennial-intermittent behaviour (Holland Creek) compared to the much wider range of behaviour observed in the PUTs (Figure 3). For

Costelloe *et al.,* How adequately do gauged tributaries represent the hydrological behaviour of ungauged tributaries in modelling large regulated catchments?

instance, three of the PUTs showed highly intermittent behaviour (e.g. Blind, Five Mile and Castle Creeks) and flowed for <50% of the time. The other two PUTs (Back Creek and Lima Creek) showed perennial-intermittent behaviour (Table 1) similar to the gauged tributaries. This observed range in flow regime behaviour occurred in sub-catchments with similar size, mid-catchment position and close proximity (see Figure 1). It is difficult to determine how widespread is the issue of representativeness of gauged tributaries compared to neighbouring ungauged tributaries in other large catchments of southeastern Australia due to the inherent lack of information on ungauged tributaries. The issue of representativeness in the context of limited gauging station data enters a Catch-22 dilemma. How can water resource managers determine if representativeness is resulting in unacceptable errors in the water balance without additional information on the hydrological behaviour of ungauged tributaries?

The results presented show that relatively short periods of low-cost monitoring data on ungauged tributaries provide sufficient information (in the context of the continuum between perennial to ephemeral flow regimes) to improve understanding of tributary behaviour and identify the limitations of deterministic modelling but only had moderate success in further constraining probabilistic modelling, as the bounds of the predicted flows were quite wide. It was necessary to make a structural modification to an existing hydrological model to simulate ephemeral subcatchments. In our study it was necessary to develop rating curves from scratch using limited flow gauging carried out between 2009 and 2010. The resulting uncertainty bounds around the rating were poorly constrained for high flows. Moreover, about 75% of this time period encompassed a very dry spell (the end of the 1997-2010 south-eastern Australian drought), so model parameters derived from this flow record may not be truly representative of the flow regime from these subcatchments in wetter climatic periods. However, in terms of fine-tuning the water balance and management of regulated catchments with substantial unregulated tributaries, an increase in short to medium term, low cost monitoring can be beneficial, particularly if gaugings are able to constrain the upper flow range.

ACKNOWLEDGMENTS

REFERENCES

Beven, K. (2009). *Environmental modelling: an uncertain future?: An introduction to techniques for uncertainty estimation in environmental prediction*. Routledge, Abingdon, Oxon (UK).

Beven, K.J., and A.M. Binley (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*, 6, 279-298.

Boughton, W. (2004). The Australian water balance model. *Environmental Modelling and Software*, 19, 943-956.

Ewen, J., G. O'Donnell, A. Burton, and E. O'Connell (2006). Errors and uncertainty in physically-based rainfall runoff modelling of catchment change effects. *Journal of Hydrology*, 330, 641-650.

Kennard, M.J., B.J. Pusey, J.D. Olden, S.J. Mackay, J.L. Stein, and N. Marsh (2010). Classification of natural flow regimes in Australia to support environmental flow management. *Freshwater Biology*, 55, 171-193.

McIntyre, N., H. Lee, H.S. Wheater, A. Young, and T. Wagener (2005). Ensemble prediction of runoff in ungauged watersheds. *Water Resources Research*, 41, W12434.

Reichl, J., A.W. Western, N. McIntyre, and F.H.S. Chiew (2009). Optimization of a similarity measure for estimating ungauged streamflow. *Water Resources Research*, 45. W10423.

Sivapalan, M. (2003). Prediction in ungauged basins: A grand challenge for theoretical hydrology, *Hydrological Processes*, 17, 3163–3170.

Wagener, T., and A. Montanari (2011). Convergence of approaches toward reducing uncertainty in predictions in ungauged basins. *Water Resources Research*, 47, W06301.