26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

Quantitative evaluation of the influence of rainfall runoff model parameter uncertainty

Phillip Jordan [a], Matthew Scorah [a], Kaung Si (Kai) Thu [a], Michael Finger [b], Xiang Cheng [b], Katherine Szabo [a], Rory Nathan [c], Kate Austin [a]

[a] *HARC, Blackburn, Victoria, Australia*
[b] *Department of Energy, Environment and Climate Action, Melbourne, Victoria, Australia*
[c] *Emeritus Professor, University of Melbourne, Melbourne, Victoria, Australia*
Email: Phillip.Jordan@harc.com.au

Abstract: Eight case study catchments were selected to investigate the uncertainty in inflow predictions from rainfall runoff models in Northern Victoria. These eight models provided a representative mix of: GR4J and Sacramento models; intermittent and perennial catchments spread geographically across Northern Victoria; and the range of methods typically used for deriving observed flow inputs, including gauged flows, calculated flow balances for reaches and storage reservoirs. Bayesian uncertainty analysis, implemented via adaptive Differential Evolution Markov Chain with Snooker updater, was applied to assess the influence of parameter and input uncertainties on flow predictions. Results show GR4J models generally produced tighter uncertainty bounds and flow distributions that were more consistent with their calibrated model fits in perennial catchments, while Sacramento models exhibited challenges in representing low flow in intermittent catchments, likely due to both model structural limitations and observed data quality. Uncertainty increased significantly at the lower end of flow duration curves, especially in catchments with observed data derived from regression relationships or reach balances. Even well-calibrated models showed 5–20% uncertainty in mean annual flow, with greater uncertainty in drier or poorly gauged catchments. These findings highlight the value of explicitly accounting for inflow uncertainty in water resources modelling to better inform planning decisions, particularly for management approaches that rely on low-flow estimates. Integration of rainfall runoff model uncertainty with climate projection uncertainty can better span the range of possible future flow outcomes under climate change.

Keywords: *Rainfall runoff, Bayesian uncertainty analysis, GR4J, Sacramento, Monte Carlo Markov chain*

1. INTRODUCTION

More than 100 rainfall runoff models have been developed for catchments in the Goulburn Broken Campaspe Coliban and Loddon, Wimmera-Glenelg, Kiewa and Ovens systems. The existence of catchment specific rainfall runoff models covering Northern Victoria provides the opportunity to investigate climate change projections developed using these models as an alternative to the current approach of regionalisation of climate change projections from a more limited population of catchments with rainfall runoff models (Department of Environment Land Water and Planning, 2020; Potter et al., 2016). It is therefore useful to understand and potentially quantify the uncertainty in streamflow produced from the rainfall runoff models, GR4J (Perrin et al., 2003) and Sacramento (Burnash et al., 1973), that have been applied in recent water resources modelling in the Northern Victoria.

This paper and associated report (HARC, 2025) provide a quantitative assessment of the influence of uncertainty in inputs and rainfall runoff model parameters on inflows derived from rainfall runoff models that are typically used in water resources models applied in the Northern Victoria. Uncertainty introduced via rainfall runoff model structure is not explicitly addressed in this investigation, although some aspects of model structural uncertainty are revealed by the uncertainty analysis results. The investigation was conducted using rainfall runoff models for eight catchments, an equal mix of Sacramento and GR4J models with loss tables. The subset of selected models also represents the range of methods that have been applied to develop "observed" flows: streamflow gauges (3 catchments), inflows to reservoirs calculated from water balance (2), residual inflows for reaches between gauges (2) and transposed from a regression relationship with recorded flows at another gauge (1). The quantitative assessment also considered whether the uncertainty changes based on using the entire observed period of record for calibration or restricting the observed data record only to the particularly dry conditions observed during the Millennium Drought, which was defined as July 1997-June 2009 for all catchments in this study. Figure 1 shows the spatial distribution of the selected catchments across Northern Victoria.

© MSSANZ and The Authors, CC-BY

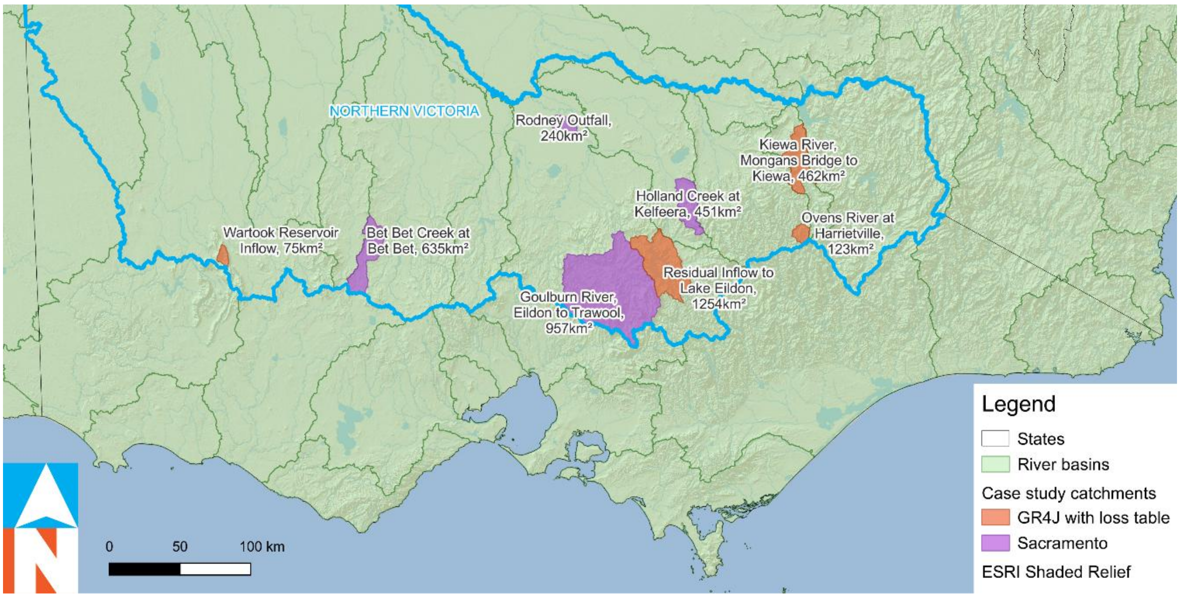

Figure 1 Location of rainfall runoff model catchments chosen for uncertainty analysis

2. METHOD

2.1. Bayesian uncertainty analysis implementation

Bayesian uncertainty analysis was applied to assess the uncertainty in rainfall runoff model predictions. Uncertainty analysis was undertaken using the BATEA (Bayesian Total Error Analysis) method (Kavetski et al., 2006a, 2006b; Renard et al., 2010). A brief summary of the method and equations is provided here but for further details on implementation, reference should be made to Kavetski et al. (2006a) and Kavetski et al. (2006b).

Rainfall runoff models are complicated transfer functions between a time series of inputs, such as rainfall and PET, and a time series of modelled outputs. Mathematically this can be represented by:

$$Q = f(\theta, X)$$

where Q are modelled flow outputs from the model, f represents the rainfall runoff model, θ are the parameters of the rainfall runoff model and X are the time series of inputs, i.e. rainfall and PET.

The observed time series of flows contain observational errors. For observed flows from a single gauge, the observational errors are due to application of a rating curve to convert observed water levels to a flow rate. For observed flows obtained by reach or storage balance calculations, the observational errors will contain contributions from the gauged inflow and outflows and (for storage balances) from rainfall and evaporation from the reservoir. The observed flows therefore are the sum of the expected value of the modelled flow at the given time step and the observational error:

$$\tilde{Q} = E[Q|\theta, \tilde{X}] + \varepsilon_Q$$

where $\tilde{Q}$ are the observed time series of flows, $\tilde{X}$ are the observed time series of inputs (usually observed rainfall at gauges, observed PET) and ε_Q are the errors in observing the flows from data collected.

Application of Bayes theorem yields the following posterior distribution of model parameters and errors in the model inputs (see Kavetski et al., 2006a):

$$p(\theta|\tilde{X}, \tilde{Q}) = \frac{p(\tilde{Q}|\theta, \tilde{X})p(\theta)}{p(\tilde{Q}|\tilde{X})}$$

where p is the probability. As explained in Kavetski et al. (2006a), it is only necessary to assess the relative likelihood of a set of parameters compared to any other set of parameters. Therefore, the likelihood function, L is given by:

$$L(\theta|\tilde{X}, \tilde{Q}) \propto L(\tilde{Q}|\theta, \tilde{X})p(\theta)$$

In other words, a model run that produces less error in modelled flows, relative to the observed flows had a higher likelihood of contributing to the output parameters than a model run that produces more error in modelled flows relative to the observed flows.

© MSSANZ and The Authors, CC-BY

Observational errors in inputs can be broken down into observational errors in the mean value of rainfall across the catchment and the mean value of PET across the catchment:

$$L\left(\theta|\tilde{R},\tilde{E},\tilde{Q}\right) \propto L\left(\tilde{Q}|\theta,\tilde{R},\tilde{E}\right)p(\theta)$$

where $\tilde{R},\tilde{E},\tilde{Q}$ are observed values of rainfall, PET and flow, respectively.

Bayesian uncertainty analysis was implemented using a Monte Carlo Markov Chain (MCMC) approach. MCMC approximates the posterior distribution of each parameter by iterative sampling. This involved millions of model runs, each with iterative sampling of the parameter values. To efficiently explore the parameter space and overcome potential convergence issues, MCMC was implemented using adaptive Differential Evolution Markov Chain with Snooker updater (DE-MCzS). Unlike traditional MCMC methods that rely on a single chain, differential evolution Markov chains maintain a population of chains that evolve collectively. This population-based approach allows chains to share information with candidate proposals informed by the states of other chains, thereby enhancing the algorithm's ability to navigate complex posterior landscapes. The DE-MCzS algorithm incorporates a snooker updater to further improve mixing and avoid local convergence traps (Cheng et al., 2014; ter Braak and Vrugt, 2008). The snooker updater generates proposals orthogonal to the chain's current trajectory. It leverages line projections to focus exploration on less-visited regions of the parameter space. The combination of differential evolution strategies and adaptive updates makes DE-MCzS effective and robust in Bayesian inference problems that are high-dimensional or involve multimodal distributions.

2.2. Prior parameter distributions

Parameter sets were obtained from the 35 GR4J and 59 Sacramento models that had been calibrated to catchments in the Northern Victoria. These parameter sets provide the basis for sampling the prior uncertainty distributions for each of the parameters of the rainfall runoff models. Whilst some of the prior parameter distributions well approximated by normal (Gaussian) distributions, the distributions fitted to the previously calibrated models for some of the parameters are strongly skewed. Skewed parameter distributions were transformed via a Box-Cox transformation, so that the prior probabilities were then generated from and assessed against distributions that were well approximated by a normal distribution.

2.3. Errors in climate inputs

Kavetski et al. (2006a) presents an approach whereby errors in catchment average rainfall depth are modelled separately for each rainfall event and did not allow for errors in catchment average PET. Applying separate scaling factors to each rainfall event was tested in the analysis for this study. However, due to the much longer observed data records used in the current study, it was found that the number prior parameters increased so markedly that the run times and number of model runs required to undertake the uncertainty analysis infeasible. Therefore, the uncertainty in climate data was represented by two factors: the error in catchment average mean rainfall across the entire model run and the error in catchment average mean PET across the entire model run. Uniformed priors were used for both mean rainfall and PET error, with the allowable range for each between 0.5 and 2.0 (i.e. mean rainfall or PET can vary, on average, between half and double the observed value). These are wide priors and the uncertainty analysis algorithm always converged posterior mean rainfall and PET ratios that were much closer to 1.

2.4. Errors in gauged flows

For gauged flows, the dominant source of uncertainty is in estimation of gauged flow from gauged water level using a rating curve. Rating curves are typically updated as new information is obtained by hydrographic rating of flows and levels at individual points in time.

Rating errors are likely to vary with flow magnitude. McMahon et al. (2025) proposed a standardised approach for assessing uncertainty in flow estimation from rating curves, which they applied to 459 gauging stations across Australia. Following the McMahon et al. (2025) approach, errors in flow gauging were represented by applying the following equations:

$$\varepsilon_Q = (1+\epsilon)\tilde{Q}$$

where the proportional error in gauged flow, ϵ, is given by an independently generated normally distributed error for each rating table period, p, as follows:

$$\epsilon = Max\left[-\epsilon_{Max}, Min\left[\epsilon_{Max}, N\left(0,\sigma_{Q,p}\right)\right]\right]$$

and where the standard deviation of the rating error varies with flow magnitude:

© MSSANZ and The Authors, CC-BY

$$\sigma_{Q,p} = \begin{cases} \sigma_{QL,p}, \tilde{Q} \le \tilde{Q}_{Median} \\ \sigma_{QU,p}, \tilde{Q} > \tilde{Q}_{Median} \; and \; \tilde{Q} \le \tilde{Q}_{Max\ Rated} \\ \sigma_{QU,p} \dfrac{\tilde{Q}}{\tilde{Q}_{Max\ Rated}}, \tilde{Q} > \tilde{Q}_{Max\ Rated} \end{cases}$$

where $\sigma_{QU,p}$ and $\sigma_{QL,p}$ are the RMS percentage errors in rated flows greater than and less than the median gauged flow respectively for the particular rating table period; and ϵ_{Max} is the maximum proportional error in gauging any individual flow rate, which was set at 50%.

2.5. Likelihood function

Differences in daily flow rates between the observed and modelled flows can be characterised by a probability distribution. Most of the rainfall runoff models used in Victoria have been optimised using the Standard Deviation Exceedance and Bias (SDEB) objective function. SDEB minimises a weighted sum of the differences in the square root of the flows and the sum of the differences in the ranked square root of the flows multiplied by the absolute mean flow bias. An analogous likelihood function for the SDEB, L_{SDEB} was calculated from:

$$L_{SDEB} = \left[\prod_{t=1}^{T} \frac{0.1}{\sqrt{2\pi\sigma_{MS}}} e^{-\frac{\left(\sqrt{Q}-\sqrt{\tilde{Q}}\right)^2}{2\sigma_{MS}^2}} + \prod_{R=1}^{T} \frac{0.9}{\sqrt{2\pi\sigma_{MS}}} e^{-\frac{\left(\sqrt{Q_R}-\sqrt{\tilde{Q}_R}\right)^2}{2\sigma_{MS}^2}} \right] \times \frac{1}{\sqrt{2\pi\sigma_B}} e^{-\frac{\left(\sum_{t=1}^{T} Q - \sum_{t=1}^{T} \tilde{Q}\right)^2}{2\sigma_B^2 \left(\sum_{t=1}^{T} \tilde{Q}\right)^2}}$$

where Q_R and $\widetilde{Q_R}$ are the gauged and modelled flows, with each ranked from highest to lowest and σ_{MS} is the RMSE in the square-roots of the gauged and modelled daily flows derived using the calibrated parameter set and σ_B is the estimated standard deviation of bias in mean annual flow from calibrated models.

3. RESULTS AND DISCUSSION

3.1. Posterior distributions of flow duration curves for observed flow data periods

Figure 2 shows flow duration curves for a selection of the case study catchments, to demonstrate some of the key insights. Flow duration curves for all eight case study catchments, for the full period of available gauged flow data and for the Millennium Drought only (July 1997-June 2009) are provided in HARC (2025).

When considering the posterior flow duration curves produced by the GR4J with loss model for the Ovens River at Harrietville, there was an excellent fit to the gauged flow data for both the full period of gauged flow (1988-2021) and the Millennium Drought (see (a) and (b) of Figure 2). In both cases, the gauge flow duration curve sits within the 90% confidence limits when considering parameter uncertainty. The width of the confidence interval in the flow duration curves gradually increase at lower flows. The driest ~1% of days of gauged flow during the Millennium Drought had flows below 10 ML/d and there were no gauged flow days below 1 ML/d. However, the posterior distributions show that there is a 5% chance that the 97th percentile flow will be less than 1 ML/d and 5% chance that the 76th percentile flow will be less than 10 ML/d.

The width of the confidence intervals in the flow duration curves for Holland Creek at Kelfeera increase markedly for flows below 10 ML/d (see Figure 2(c)). The gauged flow sits well above the 90% confidence interval from the model for the Millennium Drought period for the driest 80% of days. It may be that the structure of the Sacramento model has difficulty in representing low flows in this intermittent catchment. It was also challenging to fit low flows in the intermittent Bet Bet Creek catchment with a Sacramento rainfall runoff model (see Figure 2(d)). Fitting the uncertainty analysis only to Millennium Drought data slightly improved the flow duration curve representation for the wettest 10% of days. The gauged flow record shows a larger proportion of days between 1 ML/d and 15 ML/d than is covered by the 90% posterior confidence limits.

Figure 2(e) shows that there was a good fit to the flow duration curve of water balance derived inflows to Lake Wartook across all gauged inflows greater than about 3 ML/d. However, posterior confidence intervals on the flow duration curve from the GR4J model become very wide for the driest 50% of days. The posterior 90% confidence intervals on proportion of days below 1 ML/d covers between 0% and 33% of days. Figure 2(f) shows that there was reasonable fit of the uncertainty in GR4J modelled inflows to the inflows derived water balance on storage calculations for the ungauged inflows to Lake Eildon at moderate and higher flows (greater than ~50 ML/d). These flows produce most of the inflow volume from this catchment to Lake Eildon. It was also difficult to quantitatively estimate the error in each of the individual daily values of inflow from the storage water balance, as this was calculated as the difference between the change in volume of Lake Eildon, minus releases, minus the sum of inflows from seven gauged upstream catchments. The posterior confidence intervals

© MSSANZ and The Authors, CC-BY

on the modelled flow duration curve of Eildon ungauged inflows become very wide for the driest 30% of days. The 90% confidence intervals on proportion of days below 10 ML/d covers between 0% and 20% of days.

Figure 2(g) shows that the uncertainty analysis in the Sacramento model parameters for the Eildon to Trawool reach failed to represent the distribution of daily flows calculated from the reach balance. Even allowing for uncertainty in the Sacramento model parameter values, the posterior confidence limits on flow duration curves are generally not wide enough to capture the flows derived from the reach balance model. This could indicate difficulties in accurately deriving observed data from the reach water balance in this reach, or it could indicate structural difficulties with the Sacramento rainfall runoff model in representing rainfall runoff processes in this reach. Figure 2(h) shows that the uncertainty analysis provides an excellent representation of the flow duration curve for full period of the reach water balance data between Mongans Bridge and Kiewa.

3.2. Posterior distributions of mean annual flow

Figure 3 shows the interquartile and 90% confidence intervals in mean annual flow for each of the catchments due to rainfall runoff model parameter uncertainty, for the post-1975 period and the Millennium Drought. Figure 3(a) shows that for the 1975–present period, 90% confidence intervals on mean annual flow remain within ±7% for five catchments: Holland Creek, Lake Eildon inflow, Lake Wartook inflow, the Eildon-Trawool reach, and the Mongans Bridge–Kiewa reach. In contrast, Bet Bet Creek and the Rodney Outfall, both intermittent and the smallest catchments by absolute volume of runoff, display intervals around ±30%. While parameter uncertainty is present in all catchments, its impact on mean annual flow uncertainty is larger in catchments with low base flows.

Figure 3(b) shows that the rainfall runoff models for many of the catchments produce biased estimates of mean annual flow for the Millennium Drought, when uncertainty analysis is undertaken with the entire period of observed data. The width of the confidence intervals for posterior mean annual flow also increases. Both issues could be due to changes in catchment response during the long period of the Millennium Drought, which were not adequately represented within the structure of the Sacramento and GR4J models. This problem could be mitigated in 6 of the 8 catchments by calibrating the rainfall runoff models for the Millennium Drought only to observed data during the drought. However, there was still considerable bias and uncertainty in the mean annual flow estimates for Bet Bet Creek and the Rodney Outfall, even when uncertainty analysis was undertaken only using Millennium Drought data.

4. SUMMARY AND CONCLUSIONS

This study analysed posterior distributions of specific flow metric, mean annual flow over various time windows and daily percentiles on flow duration curves. Examining these metrics highlighted which aspects of flow regimes were most and least certain given parameter uncertainty. This approach clarified whether models could reliably capture extremes, low-flow thresholds, or overall system behaviour.

GR4J models with loss tables produced consistently good fits in four catchments: Ovens at Harrietville, Mongans to Kiewa reach, and inflows to Lake Eildon and Lake Wartook. These catchments were perennial or near-perennial, and the posterior loss functions enveloped the prior calibration loss functions. As a result, posterior confidence intervals on both mean annual flow and daily flow duration percentiles remained relatively tight across most of the flow range. This robustness suggested that GR4J with loss table calibration yielded dependable estimates of key flow statistics in catchments with sustained base flow.

In the four catchments modeled using Sacramento: Bet Bet Creek, Holland Creek, Rodney Outfall and the Eildon-Trawool reach, uncertainty analyses struggled to generate reliable flow duration curves. Three of these were intermittent, and the fourth depended on a reach water balance rather than direct gauge data. The Sacramento structure appeared to have difficulty representing low flows in settings with sporadic base flow or compounded observational errors. Consequently, posterior confidence intervals for daily flows at the dry end of the distribution were wide, and these intervals likely underestimated the true uncertainty when model structural error was also considered.

Uncertainty in flow estimates grew markedly as flows decreased, regardless of model or catchment. The transition point, where confidence intervals began to balloon, varied between systems and model structures. Catchments with flows derived from storage balances or regressions exhibited particularly large uncertainty in low-flow estimates compared to those with continuous gauge records. Water resources modellers should treat low-flow projections from both Sacramento and GR4J with caution, especially in systems where low flow metrics are important drivers of system response.

© MSSANZ and The Authors, CC-BY

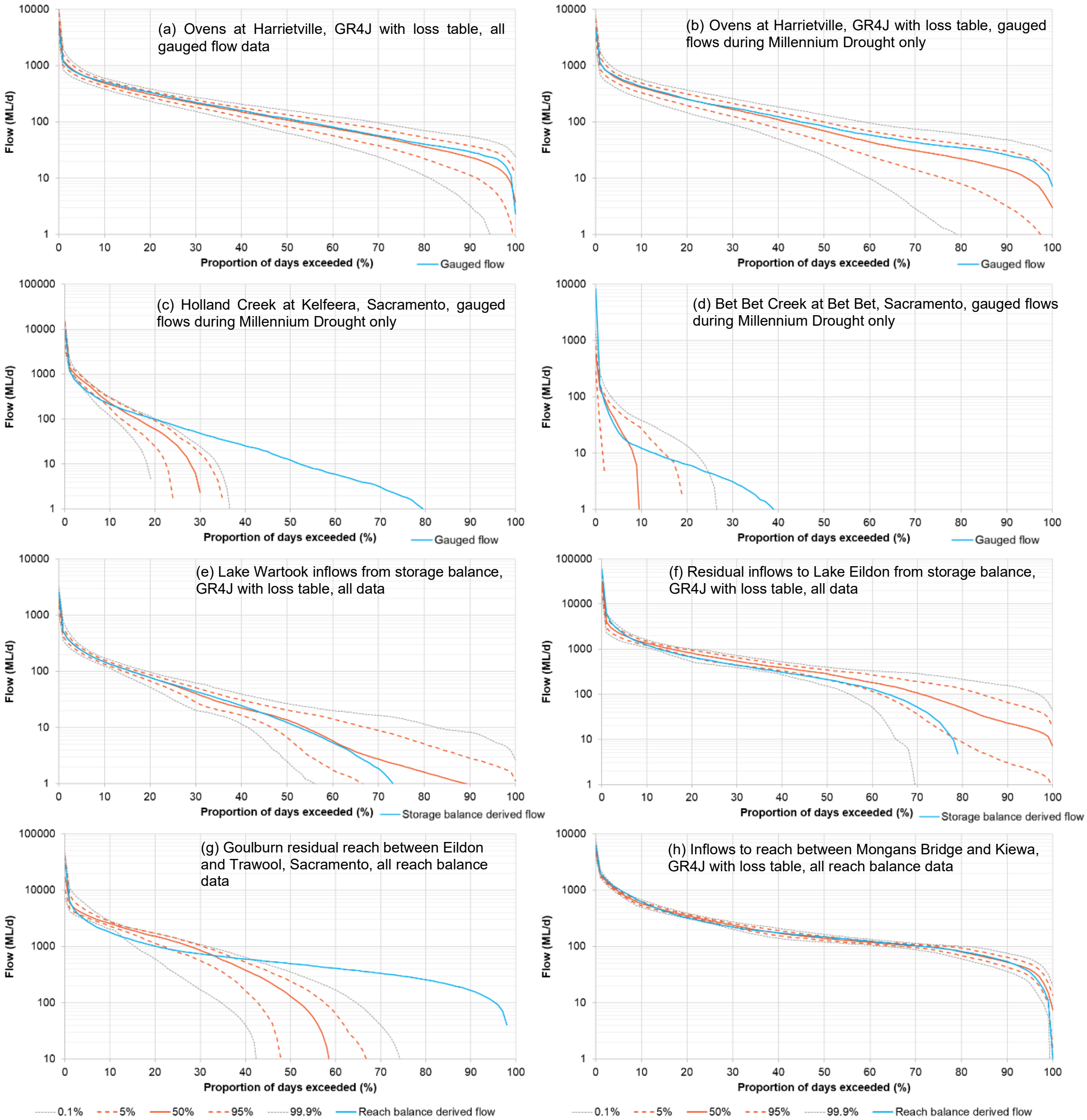

Figure 2 Observed flow duration curves compared to posterior modelled flow duration curves with parameter uncertainty: (a) Ovens at Harrietville, GR4J with loss table, all data; (b) Ovens at Harrietville, GR4J with loss table, Millennium Drought; (c) Holland Creek at Kelfeera, Sacramento, Millennium Drought; (d) Bet Bet Creek at Bet Bet, Sacramento, Millennium Drought; (e) Lake Wartook inflows from storage balance, GR4J with loss table, all data; (f) Residual inflows to Lake Eildon from storage balance, GR4J with loss table, all data; (g) Goulburn residual reach between Eildon and Trawool, Sacramento, all data; (h) Inflows to reach between Mongans Bridge and Kiewa, GR4J with loss table, all data

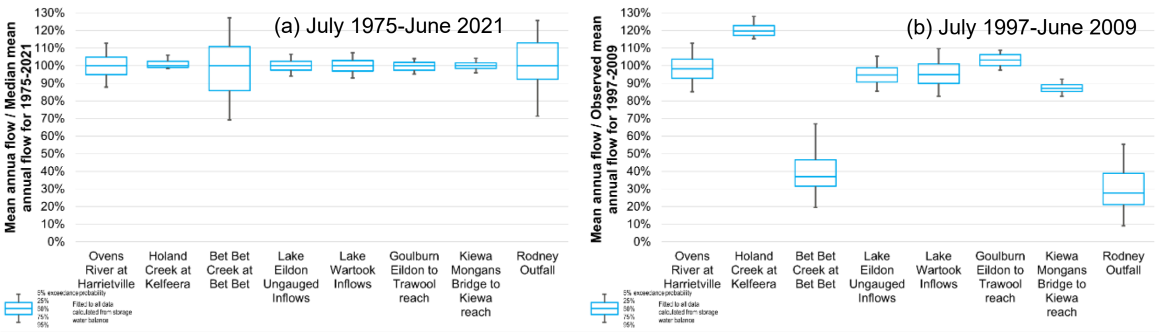

Figure 3 Posterior distributions of the ratio of mean annual flow to the posterior median of mean annual flow for each of the catchments: (a) July 1975-June 2021; (b) Millennium Drought July 1997-June 2009

© MSSANZ and The Authors, CC-BY

Posterior uncertainty in the flow duration curve increases with decreasing flow. There were large increases in uncertainty in the flow duration curve at the drier end of the flow distribution in all catchments, although the specific flow rate or daily flow percentile at which the uncertainty starts to rapidly increase varied between the models and catchments. A calibrated rainfall runoff model that appears to produce an accurate representation of the driest end of the flow duration curve is likely to be providing a false sense of the precision with which low flows can be modelled using Sacramento or GR4J rainfall runoff models. The uncertainty in estimation of low flows may also be greater in catchments where the observed inflows are derived from storage or reach balances than in catchments where the observed flows are obtained at a single streamflow gauge. BATEA may also be useful for diagnosing issues with model structure, particularly for representing low flows.

The uncertainty analyses provided in this paper were undertaken across only eight catchments, representing four Sacramento models and four GR4J models with loss functions. Those eight catchments had a range of observed data types. It is recommended that in future a similar Bayesian uncertainty analysis is extended to a much larger set of models and catchments, so that it can be more readily determined how general the observations and conclusions are.

Consideration should be given to how uncertainty in inputs to Source models propagate through the analysis to influence the uncertainty in the conclusions made using those models. Even where an inflow input is derived using a rainfall runoff model that has apparently very good performance during calibration and it is fitted to a good record of reliable gauge data, the 90% uncertainty range on mean annual flow is typically between 5% and 20% of the mean gauged flow. The confidence interval increases markedly in intermittent catchments, for catchments where there is greater uncertainty in the observed data (e.g. it is derived from a regression) and/or poor calibration performance. Water resources modellers should therefore consider how the rainfall runoff model has been applied and how the resulting time series of flows may influence outputs generated by water resources system models, such as Source. For systems with large reservoirs, the uncertainties in the mean annual or monthly flows are likely to be most influential and well calibrated rainfall runoff models seem to produce relatively robust estimates of those. However, the rainfall runoff models for several catchments had larger uncertainties in low flows, which should be considered if low flow occurrence and timing are important outputs from models. Integration of rainfall runoff model uncertainty with climate projection uncertainty can better span the range of possible future flow outcomes under climate change.

5. REFERENCES

Burnash, R.J.C., Ferral, R.L., McGuire, R.A., 1973. A generalized streamflow simulation system: Conceptual modeling for digital computers. US Department of Commerce, National Weather Service, and State of California

Cheng, L., AghaKouchak, A., Gilleland, E., Katz, R.W., 2014. Non-stationary extreme value analysis in a changing climate. Clim Change 127, 353–369. https://doi.org/https://doi.org/10.1007/s10584-014-1254-5

Department of Environment Land Water and Planning, 2020. Guidelines for assessing the impact of climate change on water availability in Victoria. DELWP, Melbourne.

HARC, 2025. Quantitative evaluation of rainfall runoff model parameter uncertainty.

Kavetski, D., Kuczera, G., Franks, S.W., 2006a. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. Water Resour Res 42. https://doi.org/10.1029/2005WR004368

Kavetski, D., Kuczera, G., Franks, S.W., 2006b. Bayesian analysis of input uncertainty in hydrological modeling: 2. Application. Water Resour Res 42. https://doi.org/10.1029/2005WR004376

McMahon, T.A., Peel, M.C., Amirthanathan, G.E., 2025. Assessing rating curve uncertainty. Hydrological Sciences Journal. https://doi.org/10.1080/02626667.2024.2442736

Perrin, C., Michel, C., Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. J Hydrol (Amst) 279, 275–289. https://doi.org/https://doi.org/10.1016/S0022-1694(03)00225-7

Potter, N.J., Chiew, F.H., Zheng, H., Ekström, M., Zhang, L., 2016. Hydroclimate projections for Victoria at 2040 and 2065.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., Franks, S.W., 2010. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. Water Resour Res 46. https://doi.org/10.1029/2009WR008328

ter Braak, C.J.F., Vrugt, J.A., 2008. Differential evolution Markov chain with snooker updater and fewer chains. Stat Comput 18, 435–446.

© MSSANZ and The Authors, CC-BY