

# **A Bayesian Joint Probability Post-Processor for Reducing Errors and Quantifying Uncertainty in Monthly Streamflow Predictions**

**Prafulla Pokhrel, David E. Robertson and Q.J. Wang**

We would like to thank the reviewer for assessment on the paper and constructive suggestions/comments. We have revised the manuscript, incorporating most of the comments made by the reviewers. To help in the assessment of our revision, each reviewer's comments and our specific response to those comments are included.

## **Response to comments made by Reviewer # 1:**

**General comment:** The manuscript investigates the post-processing of monthly streamflow simulations generated by a hydrological model (WAPABA). A Bayesian joint probability modelling approach is used for the post-processing, and applied to 18 catchments in eastern Australia. This work builds on recent works of the authors (the post-processing technique and hydrological model). The overall motivation of the work is scientifically sound; the manuscript is well organized and written. This is a good example of integrating different modeling components for a direct application. Therefore I recommend the manuscript to be accepted after some minor comments (see below) have been addressed by the authors.

**Response:** We thank the reviewer for the positive assessment on the paper and the comment that the paper is scientifically sound, well organised and well written.

### **Comment 1: Prediction vs. simulations**

The authors use the word prediction throughout the text, including the title and abstract. My first impression after reading the title and abstract was that the methods were applied to streamflow predictions/forecasts, i.e with a certain lead time. In my opinion, the word prediction is associated with something made “in advance”. The authors use “simulation” to refer to the raw model output and prediction to the post-processed streamflow, however the post-processing is not adding “lead time” to the streamflow simulations, and this can be misleading. Therefore I would suggest changing “predictions” to “simulations”, including in the title and abstract.

**Response:** We appreciate the reviewer’s concern on the use of ‘prediction’ and agree that predictions and forecast are made with certain lead times. However, we would like to avoid using the word ‘simulation’ in the context of this paper. The main reason is that we feel that the use of word model simulation does not adequately include all the aspects of post-processing including, for example, real-time output updating using streamflow observations.

However, to address reviewer’s comment we now include a tight definition of the word ‘prediction’ at the beginning of section 3. We believe a clear definition of the word ‘prediction’ in the context of this study eliminates any ambiguity regarding the use of ‘prediction’ vs. ‘simulation’. In the paper we define ‘prediction’ as “one step time step ahead forecast of streamflow, under perfect rainfall forecast”.

The first paragraph of section 3 now reads -

“In each catchment, we calibrate parameters of a hydrologic water balance model and generate streamflow predictions. Hereinafter, in the context of this study, we define prediction as one time step ahead forecast of streamflow, under perfect rainfall forecast. The ‘raw’ deterministic streamflow predictions generated by the model contain errors that are unreconciled during calibration process. The BJP post-processor aims to reduce such errors and quantify uncertainty. This section describes the process of generating streamflow predictions, using a hydrologic model and their subsequent post-processing.”

#### **Comment 2: Methods description**

The generation of streamflow simulation (section 3.1) and statistical post-processing (section 3.2) are much resumed, building on some recent work developed by the authors. To understand the model and statistical post-processing we need to read the previous papers. For example, I only understood the meaning of the parameter vector ( $\theta$ ) in eq. 1. after reading Wang et al (2009). I do not have access to the journal describing the hydrological model (Wang et al 2011), and it is very unclear how the calibration was performed, or the exact meaning of “scalarized multi-objective measure” (line 12, 11204). Without understanding this, a question comes to my mind: how sensitive is the post-processing to the calibration? I suggest that the authors include a more detailed description of the hydrological model, especially the calibration, and also of the post-processing. This could be included as appendix, but it is not mandatory, and I leave that decision to the authors and editor consideration.

**Response:** We have added more information as well as an equation that will provide readers some additional information about the BJP modelling approach. We have included following text in section 3.2 of the manuscript -

“The posterior distribution of the parameters  $p(\theta|Y_{OBS})$ , including mean, variance and transformation parameters for each variable and a correlation matrix for the multivariate normal distribution, is estimated using a Bayesian inference (equation 1).

$$p(\theta|Y_{OBS}) \propto p(\theta) \cdot p(Y_{OBS}|\theta) \quad (1)$$

where  $Y_{OBS}$  contains the historical data of both predictor  $y(1)$  and predictand  $y(2)$  variables used for model inference, and  $\theta$  is the parameter vector.  $p(\theta)$  is the prior distribution of the parameters of the multivariate normal distribution, representing any information available before the use of historical data  $Y_{OBS}$ .  $p(Y_{OBS}|\theta)$  is the likelihood function defining the probability of observing the historical data given the model and the parameter sets. The posterior parameter distribution is approximated by 1000 sets of parameters sampled using a MCMC method.”

In addition, we also include added description of the multi-objective measure. The modified part of section 3.1 in the manuscript now reads –

“ We maximise a scalarized multi-objective measure consisting of a uniformly weighted average of the Nash-Sutcliffe efficiency (NS) coefficient (Nash and Sutcliffe, 1970), the NS of log transformed flows, the Pearson correlation coefficient and a symmetric measure of bias. The NS is an ‘observed-variance-normalized mean squared error’ measure that emphasises large errors, often occurring during large events. The NS of log transformed flow emphasises errors occurring during low flow events. The Pearson correlation measures the co-variability of the simulated and the observed. The

symmetrical measure of bias evaluates the match between average simulation and average observation (Wang et al., 2011)."

**Comment 3:** Result of method C : including WAPABA lagged simulations The negligible impact of including the WAPABA lagged simulations in the postprocessing is interesting. Could this be related to the model design? small size of the catchments ? monthly time-scale ? It would be interesting to see this point further discussed.

**Response:** What we are saying is that inclusion of the third predictor does not lead to additional improvement over the two (WAPABA prediction + Lag-1 streamflow). We have added following text to section 4.1.4 to make this point clearer. The added text in the manuscript reads as -

"This suggests that, two predictors in the BJP post-processors (WAPABA prediction and Lag-1 streamflow observation) are able to capture all information about the residual error structure from the training data, thus making contributions from an additional predictor redundant."

