

Quantification of hydrological uncertainty in short lead time forecast of levels in frequently spilling reservoirs

Prafulla Pokhrel^a

^a *Entura/Hydro Tasmania*

Email: Prafulla.Pokhrel@entura.com.au

Abstract: Lake Gairdner is a frequently spilling hydropower diversion reservoir in Northern Tasmania. Comparison of forecast lake levels generated using a coupled inflow-storage model with observed lake levels suggest significant changes in the relationship between observed and forecast lake levels as the lake approaches and exceeds Full Supply Level (FSL). A wider spread (greater variance) in the relationship (between forecast and observed levels) is observed at lower lake levels and sharp converging scatter is observed above FSL. Importantly, further examination of the scatter reveals that observed and forecast levels are characterised by bimodal distributions.

A common approach to quantify uncertainty in hydrological forecasts is to model the relationship in a transformed bivariate Gaussian space. However, the peculiar shape of the scatter due the presence of FSL in Lake Gairdner rendered application of bivariate Gaussian unfeasible. This study proposed a post-processor designed to handle the change in dependence structure (relationship) and account for the ‘converging scatter’ at the upper tail. The post-processor served two purposes (1) quantify the hydrologic uncertainty and (2) improve the accuracy of short term lake level forecast in Lake Gairdner. It is envisaged that the method can be generalized for other frequently spilling reservoirs with similar conditions.

The post-processor was made up of 2-parameter copula functions. Marginal distributions were modelled using a number of unimodal distributions and mixtures of Gaussian distributions. The mixtures of Gaussian distributions were introduced to account for the bi-modal nature of the marginal distributions.

Separate post-processors were fitted for winter and summer to account for the seasonal operation of the reservoir levels. The best winter model consisted of BB1 copula and Gaussian Mixture marginal distributions and the summer model consisted of a T – copula and Weibull distributions. The post-processors were tested for their capacity to improve accuracy and quantify the uncertainty in the lake level forecast. The results showed that the post-processors consistently improved the accuracy of the forecast. The winter model generated reliable forecasts, and characterized the uncertainty reasonably well. However, the summer model resulted in some conditional bias and slightly over dispersive forecast distribution.

Keywords: *Post processor, hydrologic uncertainty, uncertainty quantification, short term forecast, Lake level prediction, copula, Gaussian mixture model*

1. INTRODUCTION

Hydro Tasmania (Hydro) manages 13500 GL (Robinson et al., 2015) of water impounded by 54 major dams located across Tasmania. The dam safety operations in a number of these reservoirs are informed by a forecast modelling system that provides forecasts of inflow and reservoir levels 7 days in to the future. The Hydro forecast system consists of a coupled inflow, storage routing and operation model (hereafter called storage model). The inflow model is driven by Bureau of Meteorology's Australian Digital Forecast Database (ADFD) rainfall product. The modelled inflow is routed through the storage using the level pool reservoir routing method. The storage model corrects for the observed lake level and generates forecasts 7 days into the future.

The forecast system uses a statistical post-processor for bias correction and estimation of 75 and 25 percentile rainfall forecast traces that serve as an estimate of uncertainty in the rainfall forecast. In the current form, however, the forecast system does not account for the combined uncertainty due to inadequacies in model structures (of inflow and storage models), errors in parameter estimation, and model initialization. These uncertainties, hereafter called hydrologic uncertainty, become increasingly important and constitute a significant proportion of reservoir level forecast uncertainty at shorter lead times when the reservoir inflows are dominated by the storage and release characteristics of the catchment.

Quantification of hydrological uncertainty generally involves modelling the joint distribution between the forecast and the observed variable. A common approach is to apply a transformation to normalize the joint distribution and use a bivariate (or multivariate) Gaussian distribution to model the dependence (Wang et al., 2009; Robertson et al., 2013a; Robertson et al., 2013b; Pokhrel, et al., 2013). In frequently spilling reservoirs however, the dependence structure between the model and forecast change significantly before and after the reservoir levels reach FSL. Importantly, the marginal distributions also demonstrate bi-modal characteristics, which make application of a transformed Gaussian multivariate distribution unfeasible.

The objective of this study was (1) assess the dependence between the observed and forecast lake levels for Lake Gairdner, a frequently spilling reservoir operated by Hydro Tasmania and (2) propose a statistical post-processor aimed at improving forecast accuracy and quantifying hydrologic uncertainty. While this study specifically focused on Lake Gairdner, it is envisaged that the method can be generalized for other frequently spilling reservoirs with similar conditions.

2. CATCHMENT AND MODELS

Lake Gairdner was formed by impounding the Lea River and Iris River by the Wilmot Dam. The catchment covers 161 km² of forested area. The reservoir holds 7.4 GL of water at the full supply level (FSL). The Wilmot Dam has a free flowing spillway and is prone to frequent spills. Releases from Lake Gairdner are diverted to an adjacent catchment via a 32 MW capacity power station. The reservoir releases are governed by fixed seasonal release targets, based on an economic operation level (EOL). Generally, the reservoir is maintained at higher levels during the summer (October –March) and at lower levels during winter (April –September). However, in winter, large inflows result in larger reservoir level fluctuations and frequent spills.

The inflow to Lake Gairdner was modelled by a modified, 2 tap, version of the Australian Water Balance Model (AWBM) (Boughton, 2004). The model was calibrated to reservoir inflow data derived from lake water balance, taking into account the changes in storage and the outflows from the dam. The storage was modelled by a level pool routing method and the reservoir operation model, based on the target seasonal EOL.

The forecast model was used to generate a long time series of forecast level. The model was driven by the observed rainfall, and measured power station discharges, essentially excluding uncertainties due to rainfall forecast and errors in power station operation from the analysis. The model was run at 6 hourly time-step 28 times with a starting lag of 6 lag hours, thus creating 28 time series with forecast lead-times 6 hours to 7 days.

3. COMPARISON OF OBSERVED AND MODELLED SCATTER

Figure 1 shows the scatter between the observed and forecast reservoir levels for winter and summer seasons at lead times of 7 days. The graphs indicate a strong relationship between the forecast and the observed lake levels; however in both cases the forecasts tend to underestimate the observed level. More importantly, the spread of the scatter is large, and almost constant, for levels lower than the FSL (= 472.44 m), but variance

decreases sharply above the FSL. This behavior is not unique to Lake Gairdner, In fact we have observed a similar spread in Lake Margaret and other frequently spilling reservoirs of Hydro Tasmania.

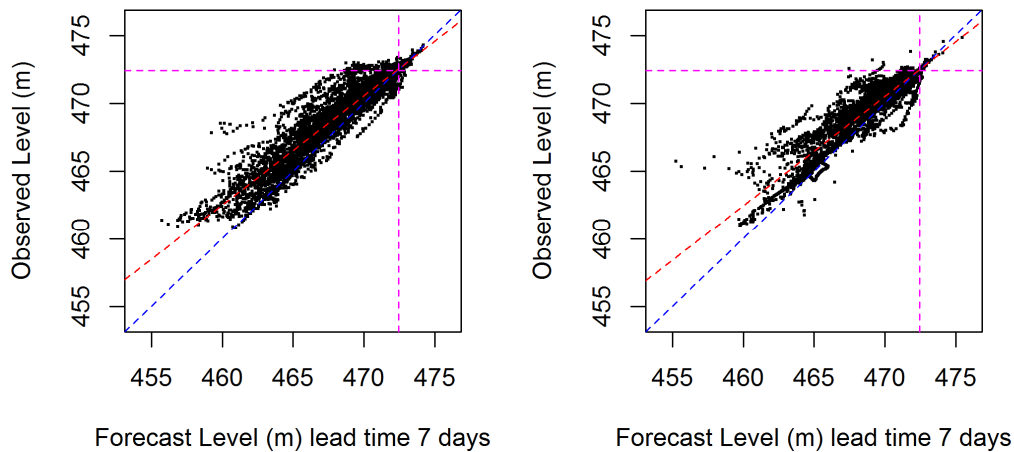


Figure 1. Forecast and the observed level in winter (left) and summer (right) [magenta = FSL, red = linear fit, blue = 1:1]

There could be two reasons for this behaviour (1) errors in the storage model do not have significant impact after the level crosses the FSL, as any inflow to the reservoir directly results in a spill (2) when the reservoir is on spill, the storage above the FSL increases sharply and for any model runs initialized to level above FSL the differences in the level become smaller.

In winter, also noticeable was the cluster of points (Figure 1) above the FSL which are created due to the fact that the model underestimates the volume of flow. This is further highlighted by the histogram of marginal distributions of observed (Figure 2) and forecast lake levels (not shown) that also demonstrate that the marginal are bi-modal.

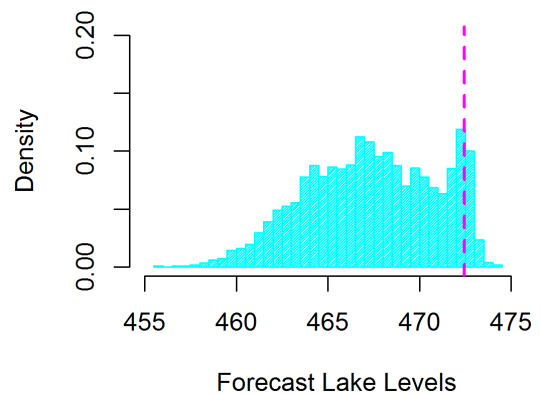


Figure 2. Distribution of the forecast and the observed levels [magenta line = FSL]

4. STATISTICAL MODEL

In this study, the changes in dependence between the lake levels (before and after the FSL) have been modelled by a 2 parameter copula. The 2 parameter copulas accommodate more than one type of dependence, and can specifically account for the ‘converging scatter’ observed at the upper tail of the distribution. The multimodality in the distributions has been modelled by mixtures of Gaussian distributions. Separate models were set up for each lead-time. However, the equations, the results and examples used in this paper correspond to the lead time of 7 days.

A copula provides a flexible way to construct a multivariate distribution from a number of independent marginal distributions. In this paper a very brief introduction to copulas is provided, readers are referred to Nelson (1999) for more details. Based on Sklar’s theorem (Sklar, 1959), the joint cumulative distribution and joint density between the observation and the forecast are given by

$$H(o, s) = C[G(o), F(s)] = C(u_o, u_s) \quad \& \quad h(o, s) = c(u_o, u_s)g(o)f(s) \quad \text{eqn. 1 \& 2}$$

Where, s = model forecast, o = observed value; $H(o, s)$ = joint distribution between s and o , $F(\cdot)$, $G[\cdot]$ are the marginal distribution of forecast and observation; $u_s, u_o \in [0, 1]$ are uniformly distributed variates, such that $u_s = F(s), u_o = G(o)$; $C(\cdot, \cdot)$ = copula distribution, $c(\cdot, \cdot)$ = copula density, $f(s)$ and $g(o)$ are the marginal densities.

A number of copula are available that describe different types of dependence structure, for this study “Rotated Tawn Copula” (Schepsmeier et al., 2012), “T-Copula” (Nelson, 1999, Madadgar et al., 2014) and “BB1” (Genest & Favre, 2007) were found to be best suited to describe the relationship between the model and forecast levels. Equation 3 describes a BB1 copula that has been used extensively in this study.

$$C(u_o, u_s) = \left[1 + \left((u_o^{-\theta} - 1)^\delta + (u_s^{-\theta} - 1)^\delta \right)^{\frac{1}{\theta}} \right]^{-1} \quad \text{eqn. 3}$$

Where, θ and δ are the parameters that define the lower and upper tail dependence respectively

Given the joint distribution (equation 1 & 2), the predictive density conditioned on the model forecast s at time t is given by:

$$h(o_t | s_t) = \frac{h(o_t, s_t)}{f(s_t)} = \frac{c(u_{o,t}, u_{s,t}) f(s_t) g(o_t)}{f(s_t)} = c(u_{o,t}, u_{s,t}) g(o_t) \quad \text{eqn. 4}$$

Using the properties of the conditional distribution, the predictive distribution and the expectation are obtained as below.

$$H(o_t | s_t) = \int_{o_{min}}^{o_t} h(o | s_t) d(o) \quad \& \quad E[o_t | s_t] = \int_{o_{min}}^{o_{max}} o \times h(o_t | s_t) d(o) \quad \text{eqn. 5 \& 6}$$

Where, o_{min} , o_{max} are the lower and upper possible values of Lake levels

The marginal distributions were modelled using mixtures of Gaussian distributions as well as unimodal distributions including; normal, logistic, Weibull, log normal, exponential and Gamma. To prevent the model from “over-fitting”, to a large number of Gaussian mixtures, the selection of marginals was guided by Bayesian Information Criteria (BIC) (Bhat & Kumar, 2010). The BIC penalizes the goodness of fit of the models by taking into account the length of data (used for model fitting) and the complexity of model. Equation 7 and 8 describe form of Gaussian mixture models used in this study.

$$g(o | \theta^o) = \sum_{i=1}^M w_i^o \phi_i^o(o) \quad \& \quad f(s | \theta^s) = \sum_{j=1}^N w_j^s \phi_j^s(s) \quad \text{eqn. 7 \& 8}$$

Where, $g(.|.)$ and $f(.|.)$ = Gaussian mixture marginal corresponding to the observed and forecast level; superscripts o and s denote observed and forecast. $\theta = (w, \mu, \sigma)$ = vector of parameters θ with w = individual Gaussian weights that sum up to 1, μ and σ = are the parameters of each of the Gaussian distributions (ϕ); $i = 1 \dots M$ and $j = 1 \dots N$ correspond to number of Gaussian distribution used for o and s .

5. MODEL CALIBRATION

The statistical models were calibrated and validated using a leave-2-years out cross validation approach. Where the models were calibrated for all the data (from 2006 to 2015) except two years and the forecast performance verified for the left out years. Separate models were fitted for summer and winter seasons to account for seasonal operation of the reservoirs.

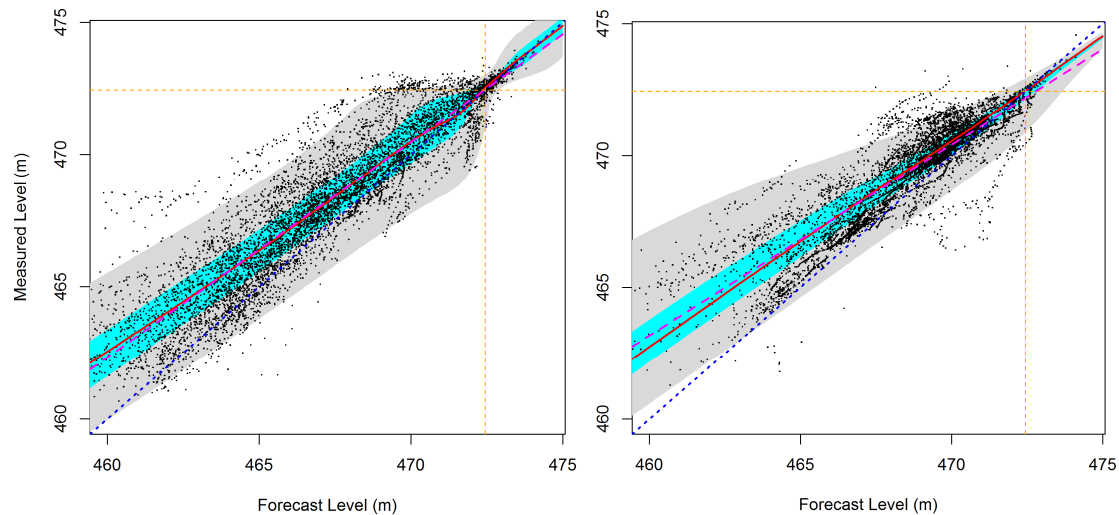


Figure 3. Statistical model fit to forecast and observed levels for winter (left) and summer (right) seasons [grey shade = 95 % uncertainty bounds, cyan = 50 % uncertainty bound, blue dotted line = 1:1, magenta dotted line = linear fit to the scatter, red line = mean forecast]

For each calibration run, the marginals and copula were fitted separately. The copula was fitted by maximizing the likelihood [$L(\varphi | o, s) = c(o, s | \varphi) \times g(o)$; where φ = copula parameters]. The R package “CDVine” (Brechmann & Schepsmeier, 2013) was used to setup and fit the copula. The GMM marginals

were fitted by Expectation Maximization (EM), while other distributions were fitted by maximum likelihood. The R package “mixtools” (Benaglia et al., 2009) was used to model and fit the GMMs. For each distribution BIC was calculated, and a simpler model was preferred in cases where the BIC values were close. The models fitted to the calibration period inclusive of all years except 2006 and 2007 are shown in Figure 3.

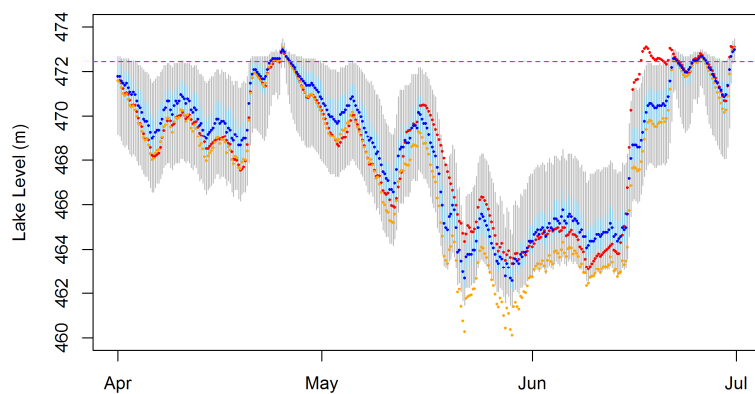
The winter model was fitted to a two parameter BB1 Copula with 2 mixtures of GMM marginals. While the BB1 copula provided a best fit, other 2 parameter copula including T and Rotated Twan copula (Schepsmeier et al., 2012) also provided very good representation of the dependence. Higher numbers of Gaussian mixtures provided closer fit to the marginals, with slightly lower BIC values, but the number of Gaussian mixtures were kept at 2 to simplify the model. Figure 3 shows that the 95% uncertainty bounds cover the scatter well. The winter model represents the convergence of the scatter above the FSL and wider uncertainty before the FSL.

The summer model (Figure 3) was fitted using “T-copula”, with Weibull marginal. For the summer models the use of a larger number of GMM models did provide some improvement over using Weibull marginals, however the differences in BIC values were small and a simpler model was preferred.

6. FORECAST PERFORMANCE VERIFICATION

The performance of the statistical model (hereafter called post-processor or PP) was assessed for each of 2 years not used for calibration. The ability of the PP to improve the accuracy of the forecast, and consistently represent the distribution of the observed lake level was evaluated.

Accuracy: The ability of the PP to improve accuracy was evaluated through visual inspection of the PP forecast mean (\bar{s}_{pp}) and the observed (\bar{o}), volume bias ($Bias = \frac{\bar{o} - \bar{s}_{pp}}{\bar{o}} \times 100$) and the ratio of root mean



squared error ($RR = \frac{\sqrt{\sum(o - s_{pp})^2}}{\sqrt{\sum(o - s_{raw})^2}}$) between the PP and the deterministic (Raw) forecast.

Figure 4. Observed and forecast level generated by the winter PP model for year 2015, grey = 95% uncertainty bound, light blue 50% uncertainty bound, blue dots = mean, red = observed, orange dots raw model forecast

Figure 4 illustrates typical outputs generated by the PP. In general, the visual inspection of the PP forecast shows that the mean PP is closer to the observed and improves the raw forecast. This was consistent for all the validation period. Similarly, Figure 5 shows % Bias and RR calculated for the summer and winter model. Each point in the plot represents Bias or RR calculated for the 2 year cross

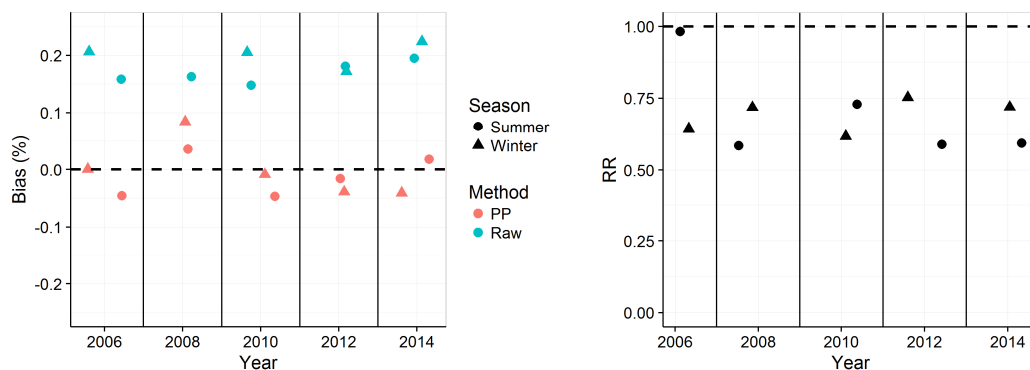


Figure 5. Bias (left plot) and RR (right) calculated for winter and summer models.

validation period. Values of RR less than 1 and lower value of Bias indicate improvement over the raw. In general the Bias (%) calculated for the PP are closer to the zero value compared to raw and the RR are less than 1, indicating that the post-processed forecast are more accurate than the raw.

Forecast reliability: Reliability refers to the ability of the PP forecast probability distribution to consistently represent the observed frequency for the events of interest. In this study the ability of the PP probability distribution $[1 - H(\tau|s_t)]$ to consistently represent lake levels greater than a threshold ($\tau = 469$) was

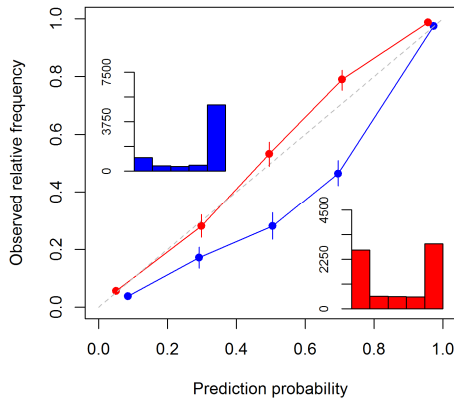


Figure 4. Forecast reliability diagrams for lake levels > 469 m, [red winter model, blue summer; the bar plots in the inset show number of samples in each bins, vertical lines represent 95 % boot strapped confidence intervals]

assessed. The threshold value was deemed to be important as it was close to FSL and covered events that lie on either side of the FSL, and covered the change in the dependence.

Figure 4 shows the reliability diagram generated by pooling the forecast probabilities over the entire cross validation period (2006 – 2015), binning them and plotting against the corresponding observed frequency.

The perfectly reliable forecast results in points close to the 1:1 line. The winter model (red line) generates forecast that is close to the 1:1 line indicating that the forecast distribution is reliable for the assessed events. The summer model results in a deviation from the 1:1 line, suggesting some conditional bias, and possibility that the PP might have slightly overestimated the forecast probabilities.

The ability of the PP to reliably represent the forecast over the entire range of values is assessed with PIT $[H(obs_t|s_t); \text{where } obs = \text{observed level}]$ histograms. An unbiased forecast distribution should have a uniform distribution (shown by red line) in Figure 5. In general the winter model produces relatively reliable forecasts; the summer model indicates the forecast might be slightly over dispersive (variance over estimated).

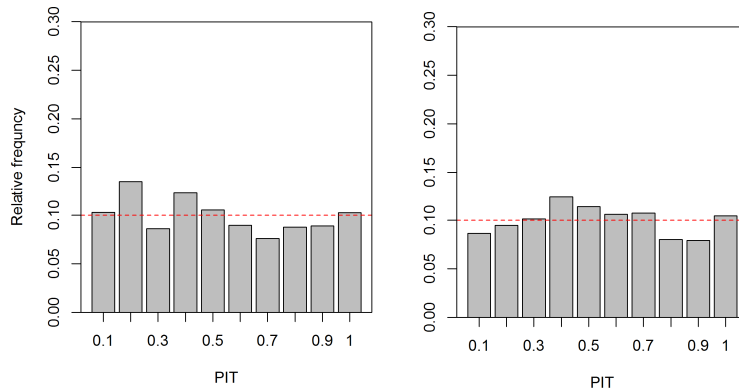


Figure 5. PIT histograms for the winter model (left) and the summer model (right)

7. CONCLUSION

The dependence between the raw forecasts (at lead times of 7 days) and observed lake level in Lake Gairdner, a frequently spilling reservoir in Northern Tasmania, demonstrated a wider spread at lower lake levels and converging scatter above FSL. Analysis of the observed lake levels also demonstrated that the marginal distribution was multi-modal. This study proposed a post-processor that can handle more than one type of dependence and

account for the ‘converging scatter’ at the upper tail.

The post-processor is a bivariate model consisting of a 2-parameter copula function. It uses mixtures of Gaussian distributions to handle multimodality shown by the observed level. Separate post-processors were established over the winter and summer seasons to account for the seasonal operation of the reservoir levels. The winter model was fitted with Gaussian Mixture marginals and the summer model with Weibull distributions. The post-processors were tested for their capacity to improve accuracy and quantify the uncertainty in the Lake level forecast. The results showed that the post-processors improved the accuracy of the forecast. The winter model generated reliable forecast, and characterized the uncertainty reasonably well. However the summer model resulted in some conditional bias and slightly over dispersive forecast distributions.

Pokhrel P., Quantification of hydrological uncertainty in short lead time forecast of levels in frequently spilling reservoirs

ACKNOWLEDGEMENTS

The author wishes to acknowledge Cassandra Blazely and Neil Smith (Hydro Tasmania) for their review of the paper and support in granting the permission to present the outcome of the project. The author also wishes to acknowledge Kim Robinson (now at WMA Water) for his review of the work during initial stage of the project and David Fuller, Bill Cohen, Mo Azmi, Jayson Peterson and Alex Wylie (Entura) for their constructive feedbacks and review of the work. The author also wishes to thank James Bennett and the two anonymous reviewers for their constructive feedbacks that helped to improve the paper.

REFERENCES

- Benaglia, T., Chauveau, D., Hunter, D. R., & Young, D. S. (2009). mixtools: An R package for analyzing mixture models. *Journal of Statistical Software*, 32(1), 1–29. <https://doi.org/10.18637/jss.v032.i06>
- Bhat, H., & Kumar, N. (2010). On the derivation of the Bayesian Information Criterion. *School of Natural Sciences, University of ...*, (August), 1–4. Retrieved from <http://nscs00.ucmerced.edu/~nkumar4/BhatKumarBIC.pdf>
- Boughton, W. (2004). The Australian water balance model. *Environmental Modelling & Software*, 19(10), 943–956. <https://doi.org/10.1016/j.envsoft.2003.10.007>
- Brechmann, E. C., & Schepsmeier, U. (2013). Modeling Dependence with C- and D-Vine Copulas: The R Package CDVine. *Journal of Statistical Software*, 52(3), 1–27. <https://doi.org/10.18637/jss.v052.i03>
- Genest, C., & Favre, A.-C. (2007). Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask. *Journal of Hydrologic Engineering*, 12(4), 347–368. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2007\)12:4\(347\)](https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347))
- Madadgar, S., Moradkhani, H., & Garen, D. (2014). Towards improved post-processing of hydrologic forecast ensembles. *Hydrological Processes*, 28(1), 104–122. <https://doi.org/10.1002/hyp.9562>
- Nelson, R. B. (1999). *An Introduction to Copulas*. New York: Springer: New York.
- Pokhrel, P., Robertson, D. E., & Wang, Q. J. (2013). A Bayesian joint probability post-processor for reducing errors and quantifying uncertainty in monthly streamflow predictions. *Hydrology and Earth System Sciences*, 17(2), 795–804. <https://doi.org/10.5194/hess-17-795-2013>
- Robertson, D. E., Pokhrel, P., & Wang, Q. J. (2013a). Improving statistical forecasts of seasonal streamflows using hydrological model output. *Hydrology and Earth System Sciences*, 17(2), 579–593. <https://doi.org/10.5194/hess-17-579-2013>
- Robertson, D. E., Shrestha, D. L., & Wang, Q. J. (2013b). Post-processing rainfall forecasts from numerical weather prediction models for short-term streamflow forecasting. *Hydrology and Earth System Sciences*, 17(9), 3587–3603. <https://doi.org/10.5194/hess-17-3587-2013>
- Robinson, K., Blundy, S., & Parkyn R. (2015). Implementing Australian Digital Forecast Database (ADFD) forecasts in Hydro Tasmania's Dynamic Inflow Forecasting System. In *HWRs*. Hobart: Engineers Australia.
- Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., Nagler, T., & Erhardt, T. (2012). VineCopula: Statistical inference of vine copulas. *R Package Version*, 1.
- Sklar, A. (1959). Fonctions de Répartition à n Dimensions et Leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris: Paris*, 8, 229–231.
- Wang, Q. J., Robertson, D. E., & Chiew, F. H. S. (2009). A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites. *Water Resources Research*, 45(5), 1–18. <https://doi.org/10.1029/2008WR007355>