A comparison of monthly runoff models for their performance under variable climate regimes within Australia

S Cherampatta Mana[a,g], **TJ Peterson**[b], **B Croke**[c,g], **T Iwanaga**[d], **TPurves**[f], **SJ Lade**[a,e,g]

[a] *Fenner School of Environment and Society, Australian National University, Canberra, 2600, Australia*
[b] *Department of Civil and Environmental Engineering, Monash University, Clayton Campus,3800, Australia*
[c] *Mathematical Sciences Institute, Australian National University, Canberra, 2600, Australia*
[d] *Australian Institute of Marine Science, Townsville,4810, Australia*
[e] *Stockholm Resilience Centre, Stockholm University, Stockholm, Sweden*
[f] *Icon Water Limited, Canberra ACT*
[g] *Institute for Water Futures, Australian National University, Canberra, 2600, Australia*

Email: sreelakshmi.cherampattamana@anu.edu.au

Abstract: Runoff models are generally used for simulating and predicting streamflow, often at a monthly timescale when used for water resource planning. Here, we evaluate the performance of WAPABA and ABCF monthly runoff models under dry and non-dry conditions for 4 inland catchments within the Australian Capital Territory (ACT) region and the neighbouring Queanbeyan River area in South-Eastern Australia (Figure 1). The ACT and Queanbeyan regions are both located within the Murrumbidgee River Catchment, which is part of the Murray-Darling Basin. The ACT's water supply needs are catered mainly by the Cotter, Queanbeyan, and Murrumbidgee Rivers. There are 4 reservoirs of interest: Bendora, Corin, Cotter, and Googong. The Murrumbidgee River divides the study area into western (3 catchments in the study) and eastern catchments (one catchment in the study). The eastern region of the study area has winter-dominated runoff regimes, while the western forested catchments exhibit higher runoff during spring. Land cover in the west is dominated by forests (> 50% of area), while the eastern catchment is predominantly agricultural land. We used the Bias-Corrected Nash-Sutcliff Efficiency metric and a Pareto trade-off parameter to choose the model parameters for different climatic conditions. Model performance was similarly evaluated using Bias-Corrected NSE. It was revealed that the monthly runoff model performance was better in non-dry periods than dry periods. This suggests the necessity of further studies on modifying the model structure to accurately model dry period flows. It was also found that the two runoff models displayed similar performance in all catchments except in the predominantly non-forested catchment. Here, WAPABA performed better with respect to Bias-Corrected NSE, adequately representing observed streamflow under drought conditions while the ABCF model underperformed. This study contributes to a nuanced interpretation of model performance under variable climatic scenarios and offers practical insights for water resource management and planning, for which monthly models are commonly used.

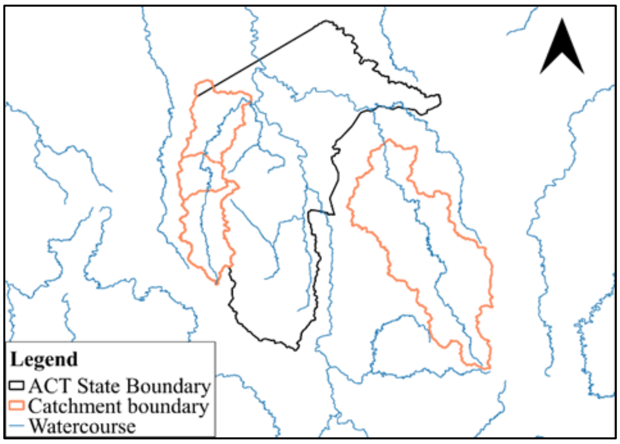

Figure 1 Study area map with catchments considered in the study. Three catchments, Corin, Bendora, and Cotter, fall in the west, while the Googong catchment falls in the East.

Keywords: *WAPABA, ABCF, runoff models, dry period, variable climate, streamflow*

1. INTRODUCTION

The behaviour of runoff models under a changing climate is an area of interest for water resource planning. Accurate estimation of streamflow is required for water availability forecasting during a changing climate. Droughts have been shown to shift runoff per unit rainfall (Saft et al. 2015) and models have been widely shown to inadequately simulate such periods (Fowler et al. 2020). This paper focuses on the performance of two monthly runoff models, ABCF and WAPABA, under dry and non-dry conditions, focused on the Australian Capital Territory and neighbouring Queanbeyan catchments.

The performance of runoff models in predicting streamflow across wet and dry periods has mostly been studied in daily or annual models (Bai et al. 2021; Fowler et al. 2016). For example, Vaze et al.(2010) showed that runoff models such as SIMHYD, IHACRES, Sacramento, and SMAG exhibit biased performance when models calibrated during wet periods were used to simulate dry period conditions. Conversely, models calibrated during dry periods tend to show less bias when simulating runoff during wetter periods (Vaze et al. 2010). Similar results were found in Tunisian catchments, where they specifically examined both precipitation and temperature and found that model transferability to wet and/or cold climates yielded good results. However, the model robustness dropped when there was a 25% or more decrease in precipitation and an increase in temperature (Dakhlaoui et al. 2017).

Monthly runoff models are more useful when the focus is on seasonal or annual streamflow volumes and are also less computationally intensive. Studies have shown that monthly models perform similarly to daily models calibrated on monthly flow data, hence suggesting that monthly time step models can be a valuable option(Wang et al. 2011). Monthly models are useful in water resource planning, decision making, and future forecasting, and have lighter data requirements as they require coarser-scale data (Clark et al. 2024). WAPABA and ABCF are two monthly models that differ in their model structure, with ABCF being a relatively new model and having shown good performance in different regions of Australia (Jaskierniak et al. 2019; Wang et al. 2011). WAPABA was compared against SIMHYD and AWBM models and found that WAPABA had a Nash Sutcliffe Efficiency (NSE) similar to SIMHYD, while WAPABA outperformed in terms of NSE and bias when compared with the AWBM model (Wang et al. 2011).

Few studies have explored the performance of runoff models on a monthly scale during dry and non-dry periods. While other research has focused on daily or annual streamflow models and their variability during dry periods, monthly models remain underexplored, even though they are primarily important for water resource planning. This paper explores the research question "In what ways do monthly runoff models WAPABA and ABCF vary in their performance across different climatic regimes and catchment vegetation types?". Rather than relying on standard split-sample testing, which may yield misleading results due to non-unique parameter identification, this study implements an approximate Pareto front approach that simultaneously optimises multiple performance criteria across varying climatic conditions to ensure robust model calibration (Fowler et al. 2016).

2. MATERIALS AND METHODS

2.1. Input data

Based on previously established data suitability for each catchment, either point rainfall recorded at each site or gridded rainfall data at a monthly timestep averaged over each catchment from the Bureau of Meteorology was used in the study (Peterson et al. 2020). The monthly inflow at the dam station was used for streamflow at each catchment. The (monthly) streamflow and evaporation data for this study were provided by Icon Water, the water service provider for the region, with the evapotranspiration (ET) calculated using Penman Monteith method. The volumetric streamflow in volumetric units (Million Litres) was then converted into streamflow depth using the corresponding catchment area. The data from 1980 to 2020 were used for the analysis.

2.2. Models

To understand the importance of runoff models and their varying performance during dry and non-dry periods, two monthly runoff models – WAPABA and ABCF – which differ in their model structure, were considered in this study. WAPABA is a water balance and partitioning model (Wang et al. 2011) that uses a Budyko type supply demand framework to partition rainfall into various fluxes. The ABCF model is a relatively new model that uses a conceptual storage approach (Jaskierniak et al. 2019), where runoff is estimated as the residual of rainfall after calculating all the losses and is routed through quick and slow flow components. A detailed description of each model is given in the following sections.

2.2.1. WAPABA model

WAPABA is a monthly runoff model that uses the Budyko framework consumption curves to partition available water into multiple components based on supply and demand within a catchment. The model streamflow calculation involves five model parameters.

The general form of the consumption curve used in the model is given by:

$$\frac{Consumption}{Demand} = F(Supply/Demand, \alpha) == 1 + \frac{Supply}{Demand} - [1 + (\frac{Supply}{Demand})^{\alpha}]^{1/\alpha}, \quad (1)$$

where α is the catchment consumption curve parameter. Here, the supply can be the total water available for evapotranspiration, while the demand is the potential evapotranspiration. The streamflow calculation is performed in five stages using five model parameters. The available rainfall is divided into water consumption and catchment yield, and thereafter, the yield is split into surface and baseflow. The total runoff is given by:

$$Q_t = Q_s(t) + Q_b(t), \quad (2)$$

where $Q_s(t)$ is surface runoff and $Q_b(t)$ is the baseflow at time t.

2.2.2. ABCF Model

The ABCF model is a 6-parameter model which uses a conceptual storage approach that accounts for water stress using catchment memory. The model defines a conceptual unconstrained soil store to model the catchment water balance.

$$S_{t+1} = S_t + P_t - AET_t - Q_s(t) - Q_b(t) \quad (3)$$

where S_t, P_t, AET_t are the soil store depth, rainfall and actual evapotranspiration at time t. The fluxes are calculated based on rainfall partitioned into actual evapotranspiration, and the residual is then routed as streamflow through quick and slow flow components. For calculating these fluxes, different model parameter values are used.

2.3. Model analysis

Both models were created for 4 water supply catchments based on dam location. The available data were divided into dry and non-dry periods, and the performance of each model during these different climate regimes was analysed based on objective function values derived from the Pareto front, which is detailed in the following section.

2.3.1. Selection of dry period

The dry period for the current study was defined based on the amount of rainfall received in each catchment. The algorithm adopted in this paper for dry period definition uses annual rainfall anomaly data relative to the annual mean (Saft et al. 2015). The rainfall anomaly time series was smoothed using the mean of a 3-year moving window to avoid single wet years. Years that showed a negative smoothed rainfall anomaly (less than -5%) in consecutive years (for at least 2 years) were initially identified as potential dry years. To refine the end of the dry year and minimise distortion from smoothing, the unsmoothed anomaly data from the final 3-year negative anomaly period were examined.

The dry periods were estimated separately for each catchment. Based on the results of the algorithm from all catchments, a common dry period was decided within this study for all the catchments. The dry periods shown by the algorithm were almost the same in every catchment, except for a difference in one or two years. As per these criteria, the years from 2001 to 2009 and 2018 to 2020 were decided as dry periods. This also coincided with the Millennium drought and Tinderbox drought within the region, respectively. The timeframe considered in this study is between 1980 and 2020. The drought of 1982-1983 was not considered as a dry period since that timeframe was not picked up in all the catchments. For uniformity, years identified as dry periods across all catchments were classified as dry periods in the analysis.

2.3.2. Analysis of models and Pareto curves

The data were divided into dry and non-dry years based on the algorithm detailed in the earlier section. The calibration is achieved through the Differential Evolution algorithm by optimising the objective function. The objective function used in this study is a meta-metric combining the Nash-Sutcliffe Efficiency and logarithmic function of bias (Viney, N.R. et al. 2009), referred to as Bias-Corrected NSE (BC-NSE) in this paper. Utilising efficiency and bias in the objective function provides a more comprehensive evaluation by accounting for both aspects rather than a single metric.

The BC-NSE, is given by:

$$F = NSE - 5 \times |\ln(1 + B)|^{2.5}, \quad (5)$$

where NSE is the transformed Nash-Sutcliffe Efficiency, computed by raising streamflow values to the power of 0.5 and B is the bias in the modelled streamflow. The transformation of streamflow in calculating NSE reduces the bias towards high flows.

To estimate the performance of each model during dry and non-dry periods the Pareto front was estimated using a weighted sum of objective functions corresponding to each period with the help of a trade-off parameter called β. Suppose the BC-NSE for dry and non-dry periods are denoted as F_{dry} and F_{nondry} respectively. Then each of these functions is given a weight as shown below:

$$F_{final} = \beta \times F_{dry} + (1 - \beta)F_{nondry} \tag{6}$$

The β values range between 0 and 1. For each of its values, the combined objective function F_{final} was optimised. After getting the optimal parameter set post calibration, the individual dry and non-dry period BC-NSE *(F)* was calculated. BC-NSE for each β plotted against dry and non-dry periods formed an approximate Pareto front. Pareto front is a useful tool to get parameter sets that allow a robust model performance during changing climate rather than the more widely adopted split-sample approach, which may lead to misleading conclusions (Fowler et al. 2016). From this Pareto front, the best trade-off value that gave a higher objective function value for both dry and non-dry periods was selected. This was done to identify parameter sets that can perform well during both dry and non-dry periods. The parameters corresponding to this beta value were considered a better trade-off between the dry and non-dry periods. Using the parameters of that set, the flow duration curve was obtained.

3. RESULTS

Figure 2 shows the approximate Pareto front for the Bias-Corrected NSE values for all the catchments within the study area. The WAPABA model consistently outperforms the ABCF model with higher values of BC-NSE for 3 of the 4 catchments-Bendora, Cotter, and Googong assessed (Figure 2). However, Figure 2B shows that at the Corin catchment, the ABCF model has higher Bias-Corrected NSE values for the dry period when compared to other catchments. It is to be noted that the dominance of the WAPABA model in terms of BC-NSE values in the majority of catchments is only by a small margin between 0 and 0.1, except for the Googong catchment. In the Googong catchment, the WAPABA model considerably exceeds its performance with respect to ABCF during both dry and non-dry periods. Another fact that can be inferred from Figure 2 is that both the ABCF and WAPABA models perform better during the non-dry period compared to the dry period. The difference in the objective function values between dry and non-dry periods varies in the range between 0.1 and 0.25, approximately, on average. This indicates that the model did not replicate dry period conditions as accurately as it did for the non-dry period in both models. This finding was, however, contradicted in Googong, where the WAPABA model performed better in both dry and non-dry periods.

Based on the Pareto front, which consists of different combinations of dry and non-dry periods with different weightage, a trade-off value(β) was selected that maximised both dry and non-dry bias-corrected NSE, as specified in the Methods section. The trade-off value for each catchment is highlighted in Figure 2 for both models. The parameter set corresponding to this trade-off value was used to calculate the Bias-corrected NSE values for each climatic period and was summarised in Table 1 for all catchments. As indicated by the values in Table 1, the performance was slightly better for the non-dry period, suggesting that models lack an accurate estimation of dry weather flow when compared to non-dry flows. This was not the case for the Googong catchment, where the ABCF model failed to adequately simulate streamflow under both dry and non-dry climatic conditions. In contrast, the WAPABA model demonstrated clearly superior performance within this catchment.

The models were also evaluated using the flow duration curve for modelled and simulated values of streamflow. The parameter values obtained for the optimal beta were used to generate these plots for each model of each catchment. Figure 3 presents the flow duration curve for all 4 catchments using the parameters at the optimal weight for WAPABA and ABCF for a period of 1980-2020, dry and non-dry conditions. The figure reveals that for each catchment, both models exhibit broadly similar behaviour in simulating low and high flows across different climatic conditions or irrespective of the time period. Overall, the WAPABA model was better at representing low flows more accurately. The WAPABA models were able to simulate the low flows at par with their observed flow. The ABCF model had difficulty in modelling the low flows. While the model underestimated low flows within Bendora, Corin, and Googong (except during the dry period), it overestimated low flows in Cotter and during the dry period within Googong. In case of high flows, the WAPABA model was able to simulate more accurately compared to ABCF in all the climatic conditions across Bendora, Corin, and Googong, with underestimation in the Cotter catchment. ABCF was better at predicting the high flow conditions compared to its ability to simulate low flow conditions.

Table 1 Bias-Corrected NSE values across all catchments during different climate regime for two models

	ABCF			WAPABA		
Catchment	Optimal beta	BC-NSE		Optimal beta	BC-NSE	
		Dry	Non-dry		Dry	Non-dry
Bendora	0.5	0.68	0.8	0.4	0.7	0.82
Corin	0.3	0.72	0.8	0.7	0.7	0.8
Cotter	0.3	0.52	0.54	0.3	0.59	0.69
Googong	0.6	0.5	0.46	0.3	0.72	0.78

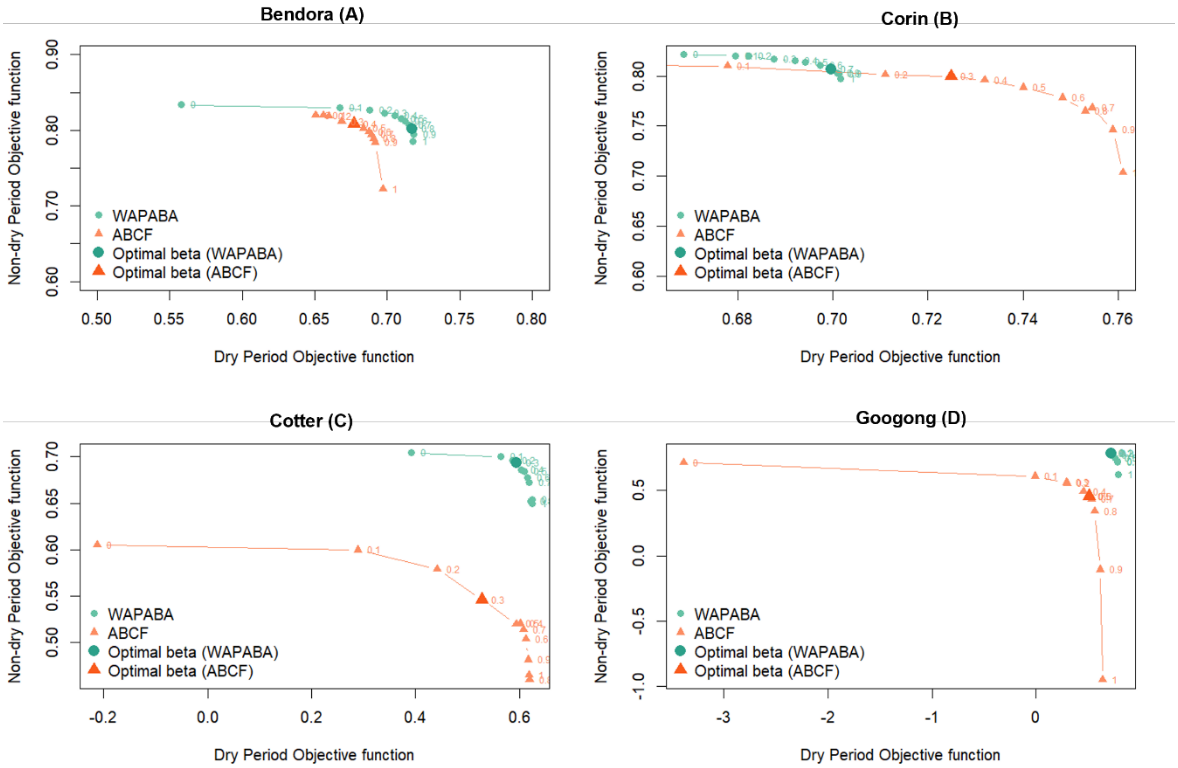

Figure 2 Pareto front for 4 catchments in the study for ABCF and WAPABA models. The graph shows the Bias-Corrected NSE values for dry and non-dry period when different beta values were used. The beta values corresponding to each data point is marked in the graph. Beta refers to the weight given to dry period. The optimal beta is highlighted for each model in the figure.

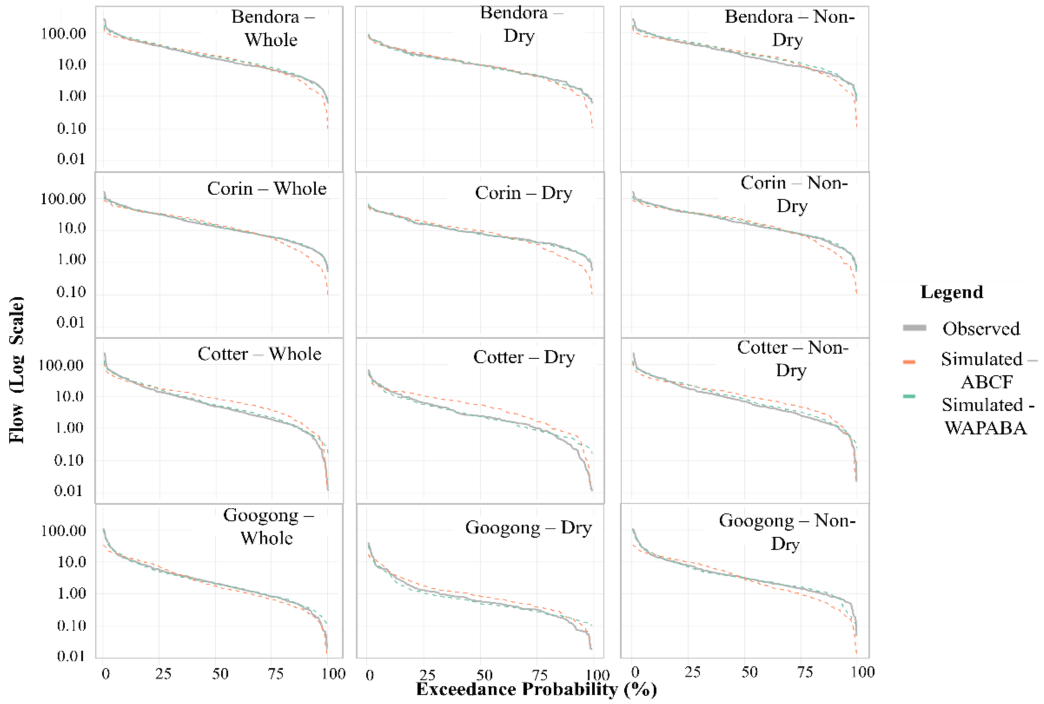

Figure 3 Flow duration curves(FDC) for all 4 catchments. Column 1,2 and 3 shows the FDC for the entire period, dry and non-dry period, respectively. Each image contains the FDC for observed flow and simulated flow using ABCF and WAPABA models.

4. DISCUSSION AND CONCLUSION

Runoff simulated by two monthly models – WAPABA and ABCF – was compared during dry and non-dry periods by applying an approximate Pareto front approach that optimises performance criteria across varying climatic conditions. This approach was adopted to mitigate issues with split sample testing approaches, which were demonstrated to provide misleading conclusions (Fowler et al. 2016). Both monthly runoff models were found to consistently underperform during the dry period compared to the non-dry period. These findings are consistent with previous research investigating model performance conducted at both annual and daily time scales, despite the use of separate calibration methodology (Pareto front-based compared to split sample tests), which highlighted the difficulties in using models calibrated under different climatic conditions (Dakhlaoui et al. 2017; Vaze et al. 2010).

The two models adopted for this study showed only minimal differences in terms of BC-NSE across half of the catchments tested. The approximate Pareto fronts indicate that the spread of objective function values between WAPABA and ABCF was narrow in the Corin and Bendora catchments, suggesting minimal performance differences. In contrast, the Googong and Cotter catchments exhibited a wider separation between the bias-corrected NSE value range, with WAPABA showing clear Pareto dominance, particularly in Googong. Although the approximate Pareto front indicated slightly higher performance for WAPABA, the difference was negligible. When the parameters for optimal beta were used (Table 1), both models yielded similar results despite their difference in model structure. This indicates that given the catchment hydroclimatic conditions, the simulated flow outputs from both models align similarly with observed streamflow. The similarity in model performance observed here contrasts with findings from other studies, where conceptual models exhibited notable differences in their simulated streamflow values (Fowler et al., 2016). This study also opens the opportunity to evaluate physically based models under similar catchment conditions, as physical models have an upper hand in modelling streamflow closer to the observed values (Deb and Kiem 2020).

Another aspect of the runoff models considered in the study is predicting the high and low flows. The model parameters for the optimal value, as shown in the Pareto front, were used to develop a flow duration curve for the entire period (1980-2020), for the dry and non-dry periods. The results show that WAPABA models are good at simulating both high and low flow conditions. The ABCF model exhibits limited accuracy in simulating low flow conditions. This implies that the usage of ABCF models for estimating low streamflow requires careful consideration, and the model requires alterations to accommodate accurate low flow estimations in these catchments or with similar catchment conditions. It is worth mentioning that the results can be different in hydrologically and geographically different scenarios.

For the Googong catchment, the WAPABA model outperformed ABCF in both dry and non-dry periods based on performance metrics. Moreover, the model performance was similar in both periods for WAPABA. This difference in model behaviour is likely influenced by several factors such as vegetation type, rainfall variability, and catchment streamflow response, with these factors acting individually or interactively. The Googong catchment is dominated by agricultural lands in contrast to other catchments, which are dominated by forested regions. The similar model performance in dry and non-dry periods for WAPABA in the Googong catchment indicates that the model is capable of accurately simulating in a changing climate in an agricultural catchment. The result is specific to Googong, and the current study only has one catchment with reduced forest cover. Hence, to generalise this finding, it requires further exploration in multiple non-forested catchments under changing climatic conditions.

Overall, this study investigated the performance of monthly runoff models under dry and non-dry climatic conditions. The results from this paper indicate that overall, both WAPABA and ABCF models estimate streamflow less precisely during the dry period when compared to the non-dry period. This underperformance in dry periods highlights the necessity for careful application of these models for future water resource forecasting. Since water supply forecasting and decision-making processes are reliant on monthly models, it is essential to incorporate changes in the model structure to improve model accuracy during dry periods. Additionally, the results suggest that the extent of forest cover is likely to impact the model performance, given the results of this study, which showed the WAPABA model performing better on catchments with less woody cover under a changing climate. This study provides valuable insights for water resource planners and hydrologists, aiding in the selection of appropriate models for a changing climate scenario.

REFERENCES

Bai P, Liu X and Xie J (2021) ‘Simulating runoff under changing climatic conditions: A comparison of the long short-term memory network with two conceptual hydrologic models’, *Journal of Hydrology*, 592:125779.

Clark SR, Lerat J, Perraud J-M and Fitch P (2024) ‘Deep learning for monthly rainfall–runoff modelling: a large-sample comparison with conceptual models across Australia’, *Hydrology and Earth System Sciences*, 28(5):1191–1213.

Dakhlaoui H, Ruelland D, Tramblay Y and Bargaoui Z (2017) ‘Evaluating the robustness of conceptual rainfall-runoff models under climate variability in northern Tunisia’, *Journal of Hydrology*, 550:201–217.

Deb P and Kiem AS (2020) ‘Evaluation of rainfall–runoff model performance under non-stationary hydroclimatic conditions’, *Hydrological Sciences Journal*, 65(10):1667–1684.

Fowler K, Knoben W, Peel M, Peterson T, Ryu D, Saft M, Seo K and Western A (2020) ‘Many Commonly Used Rainfall-Runoff Models Lack Long, Slow Dynamics: Implications for Runoff Projections’, *Water Resources Research*, 56(5):e2019WR025286.

Fowler KJA, Peel MC, Western AW, Zhang L and Peterson TJ (2016) ‘Simulating runoff under changing climatic conditions: Revisiting an apparent deficiency of conceptual rainfall-runoff models’, *Water Resources Research*, 52(3):1820–1846.

Jaskierniak D, Kuczera G, Benyon RG, Haydon S and Lane PNJ (2019) ‘Top-down seasonal streamflow model with spatiotemporal forest sapwood area’, *Journal of Hydrology*, 568:372–384.

Peterson TJ, Wasko C, Saft M and Peel MC (2020) ‘AWAPer: An R package for area weighted catchment daily meteorological data anywhere within Australia’, *Hydrological Processes*, 34(5):1301–1306.

Saft M, Western AW, Zhang L, Peel MC and Potter NJ (2015) ‘The influence of multiyear drought on the annual rainfall-runoff relationship: An Australian perspective’, *Water Resources Research*, 51(4):2444–2463.

Vaze J, Post DA, Chiew FHS, Perraud J-M, Viney NR and Teng J (2010) ‘Climate non-stationarity – Validity of calibrated rainfall–runoff models for use in climate change studies’, *Journal of Hydrology*, 394(3–4):447–457.

Viney, N.R., , J. Perraud, , J. Vaze, , F.H.S. Chiew, , D.A. Post, and A. Yang (2009) ‘The usefulness of bias constraints in model calibration for regionalisation to ungauged catchments.

Wang QJ, Pagano TC, Zhou SL, Hapuarachchi HAP, Zhang L and Robertson DE (2011) ‘Monthly versus daily water balance models in simulating monthly runoff’, *Journal of Hydrology*, 404(3–4):166–175.

