

# An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern Queensland

Ravinesh C Deo · Mehmet Şahin

Received: 15 May 2015 / Accepted: 6 January 2016 / Published online: 16 January 2016  
© Springer International Publishing Switzerland 2016

**Abstract** A predictive model for streamflow has practical implications for understanding the drought hydrology, environmental monitoring and agriculture, ecosystems and resource management. In this study, the state-of-art extreme learning machine (ELM) model was utilized to simulate the mean streamflow water level ( $Q_{WL}$ ) for three hydrological sites in eastern Queensland (Gowrie Creek, Albert, and Mary River). The performance of the ELM model was benchmarked with the artificial neural network (ANN) model. The ELM model was a fast computational method using single-layer feedforward neural networks and randomly determined hidden neurons that learns the historical patterns embedded in the input variables. A set of nine predictors with the month (to consider the seasonality of  $Q_{WL}$ ); rainfall; Southern Oscillation Index; Pacific Decadal Oscillation Index; ENSO Modoki Index; Indian Ocean Dipole Index; and Nino 3.0, Nino 3.4, and Nino 4.0 sea surface temperatures (SSTs) were utilized. A selection of variables was performed using cross correlation with  $Q_{WL}$ , yielding the best inputs defined by (month;  $P$ ; Nino 3.0 SST; Nino 4.0 SST; Southern Oscillation Index (SOI); ENSO Modoki Index (EMI)) for Gowrie

Creek, (month;  $P$ ; SOI; Pacific Decadal Oscillation (PDO); Indian Ocean Dipole (IOD); EMI) for Albert River, and by (month;  $P$ ; Nino 3.4 SST; Nino 4.0 SST; SOI; EMI) for Mary River site. A three-layer neuronal structure trialed with activation equations defined by sigmoid, logarithmic, tangent sigmoid, sine, hardlim, triangular, and radial basis was utilized, resulting in optimum ELM model with hard-limit function and architecture 6-106-1 (Gowrie Creek), 6-74-1 (Albert River), and 6-146-1 (Mary River). The alternative ELM and ANN models with two inputs (month and rainfall) and the ELM model with all nine inputs were also developed. The performance was evaluated using the mean absolute error (MAE), coefficient of determination ( $r^2$ ), Willmott's Index ( $d$ ), peak deviation ( $P_{dv}$ ), and Nash–Sutcliffe coefficient ( $E_{NS}$ ). The results verified that the ELM model as more accurate than the ANN model. Inputting the best input variables improved the performance of both models where optimum ELM yielded  $R^2 \approx (0.964, 0.957, \text{ and } 0.997)$ ,  $d \approx (0.968, 0.982, \text{ and } 0.986)$ , and  $MAE \approx (0.053, 0.023, \text{ and } 0.079)$  for Gowrie Creek, Albert River, and Mary River, respectively, and optimum ANN model yielded smaller  $R^2$  and  $d$  and larger simulation errors. When all inputs were utilized, simulations were consistently worse with  $R^2$  (0.732, 0.859, and 0.932 (Gowrie Creek),  $d$  (0.802, 0.876, and 0.903 (Albert River), and  $MAE$  (0.144, 0.049, and 0.222 (Mary River) although they were relatively better than using the month and rainfall as inputs. Also, with the best input combinations, the frequency of simulation errors fell in the smallest error bracket. Therefore, it can be ascertained that the ELM model offered an efficient approach for the streamflow

---

R. C. Deo (✉)  
School of Agricultural Computational and Environmental  
Sciences, International Centre of Applied Climate Science  
(ICACS), University of Southern Queensland, Springfield, QLD  
4300, Australia  
e-mail: ravinesh.deo@usq.edu.au

M. Şahin  
Department of Electrical and Electronics Engineering, Siirt  
University, 56100 Siirt, Turkey

simulation and, therefore, can be explored for its practicality in hydrological modeling.

**Keywords** Extreme learning machine · Streamflow prediction · Hydrological modeling

### Abbreviations

ANNs	Artificial neural networks
BOM	Bureau of Meteorology
DERM	Department of Environment
$D$	Willmott's Index of Agreement
$E$	Nash–Sutcliffe coefficient
EMI	ENSO Modoki Index
ELM	Extreme learning machine
$G_{\tan \text{ sig}}$	Tangent sigmoid function
$G_{\sin}$	Sine activation function
$G_{\text{hard-lim}}$	Hard-limit activation function
$G_{\text{rad bas}}$	Radial basis activation function
$G_{\text{tri bas}}$	Triangular basis function
$G_{\log \text{ sig}}$	Logarithmic sigmoid activation function
IOD	Indian Ocean Dipole
JISAO	Joint Institute of the Study of the Atmosphere and Ocean
$P$	Precipitation (or rainfall)
PE	Prediction error
PDO	Pacific Decadal Oscillation
$Q_{\text{WL pred}}$	Simulated streamflow water level
$Q_{\text{WL obs}}$	Observed streamflow water level
MAE	Mean absolute error
$R^2$	Coefficient of determination
SLFN	Single-layer feedforward neural network
SOI	Southern Oscillation Index
SST	Sea surface temperature

### Introduction

A versatile simulation model for streamflow water level can assist with the dissemination of lead time information on temporal evolution of local catchment hydrology and, therefore, is of paramount importance to water management, environmental monitoring, agriculture, ecosystems, and resource management (Ni et al. 2010). In Australia, streamflow has decreased since 1970 (IPCC 2001). As such, meteorological (rainfall-related) and hydrological (streamflow-related) drought

and its impact on water security are a significant cross-cutting issue. The Murray-Darling Basin, Australia's largest and economically sensitive river system that accounts for 70 % of irrigated crops and pastures, is likely experiencing decline in streamflow by 10–25 % by 2050 and 16–48 % by 2100, and little is known about future impacts on groundwater (Hennessy et al. 2007). Therefore, a prior knowledge of streamflow as a first-order integrator of rainfall conditions in a local catchment is appealing to the end users particularly for informed decision-making process (Chiew et al. 1998).

In Australia, the primary indicators of streamflow water level (e.g., rainfall or evaporation regime) are highly variable (Verdon et al. 2004) and respond rapidly to radiative perturbations (Deo et al. 2009; McAlpine et al. 2007; McAlpine et al. 2009). With a naturally variable climate influenced by El Nino Southern Oscillation, positive and negative aberrations in streamflow have pronounced impact on meteorological and hydrological disasters (Deo et al. 2009; Nicholls et al. 1997). However, streamflow is impacted by inter-annual variability of river flows (McMahon et al. 1992), which culminates added challenges in managing risks with irrigations, ecosystems, and marine life (Chiew et al. 1998). Although flood and drought are linked mainly to ENSO (Kiem and Franks 2004; Kiem et al. 2003), application of teleconnections driven by climate indices and sea surface temperatures (SSTs) for forecasting streamflow is an evolving area of interest.

The modeling of streamflow water levels should consider rainfall as an input as the variability of streamflow is influenced by surplus or deficit of rain water as a primary source of environmental flow. Therefore, the causal relationships between the current and the antecedent rainfall, river flow, and climate indices that affect these parameters should be employed for simulation of streamflow (Chiew and McMahon 2002; Chiew et al. 1998; Dettinger and Diaz 2000; Dettinger et al. 2000; Ouyang et al. 2014; Piechota et al. 2001; Simpson et al. 1993). Often, lagged relationship of streamflow with rainfall and climate indices such as Southern Oscillation Index, Indian Ocean Dipole, ENSO Modoki Index, and Pacific Decadal Oscillation Index is used (Zubair and Chandimala 2006). Hydrological droughts that are reflected by aberrations in streamflow are intrinsically linked to oceanic–atmospheric processes (Drosdowsky 1993; Kiem and Franks 2001; McBride and Nicholls 1983) and so are the Pacific Decadal Oscillation (PDO) (Power et al. 1999; Risbey et al. 2009) and Indian Ocean

Dipole (IOD) indices (Saji et al. 1999; Saji and Yamagata 2003). Also importantly, ENSO hydroclimatic links is known to foster below-normal streamflow discharge during El Niño and above normal during La Niña events. For example, a study found that sea surface temperatures (SSTs) can play a pivotal role in the prediction of January–March and April–June streamflow levels (Chiew and McMahon 2002). Therefore, the pivotal, yet crucial role of the inter-related inputs must be considered carefully for development of accurate and reliable streamflow models.

In general, data-driven models of streamflow that identify patterns and trends in historical observations use three primary approaches. First, historical changes in streamflow partitioned into training and testing datasets are utilized for predictive modeling as streamflow displays high degree of serial correlation (or persistence) arising from soil and groundwater storage and recharge that acts to delay the response of rainfall–runoff processes. This provides streamflow a memory of several months, which can represent the initial catchment hydrology to act as a source of prediction (Chiew et al. 1998). Second, concurrent or lagged correlations between rainfall and climate indices (McBride and Nicholls 1983) and streamflow (Chiew et al. 1998) can be utilized in a purely statistical sense. Third, probabilistic predictions are made, for example, by the Bureau of Meteorology with Bayesian joint probability where antecedent streamflow, rainfall, and climate indices are used as predictors to forecast streamflow (Robertson and Wang 2008; Wang and Robertson 2011; Wang et al. 2009). Recently, physically based approaches were also applied for monthly and seasonal forecasting of streamflow using rainfall–runoff models and historical dataset (Wang et al. 2011; Chowdhury et al. 2010).

The aforementioned methods have adopted statistical tools where artificial intelligence (AI) algorithms were used to extract predictive features embedded in historical input data. AI is a popular tool that is able to capture linear as well as non-linear relationships between streamflow and rainfall, climate indices, and related inputs (Asefa et al. 2006; Chang et al. 2002; Holland 1975; Koza 1992; Patterson 1998). It analyzes complex relationships between predictors and objective variable (Deo and Şahin 2015a; Deo et al. 2015; Salcedo-Sanz et al. 2015). Artificial neural network (ANN) is a popular model for simulation of hydroclimatic variables (Abbot and Marohasy 2012; Abbot and Marohasy 2014; Deo and Şahin 2015a; Masinde 2013; Nastos

et al. 2014; Ortiz-García et al. 2014; Ortiz-García et al. 2012; Shukla et al. 2011; Tran et al. 2011). Despite its widespread use, a bottleneck of the ANN is the iterative tuning of model parameters, slow response of gradient-based learning algorithm utilized by hidden neurons and low accuracy compared to more advanced algorithms (Acharya et al. 2013; Şahin et al. 2014). Alternative models based on relevance vector machine and multi-variate adaptive regression spline model have also been developed (Deo et al. 2015).

In this study, we applied the extreme learning machine (ELM) model for the simulation of streamflow in eastern Queensland. While its application for streamflow simulation is rarely found in literature for this particular region, the ELM model has successfully been applied elsewhere, for example in drought (Deo and Şahin 2015b) and evaporative loss (Deo et al. 2015) simulation, downscaling global climate model (Acharya et al. 2013), and solar and wind prediction (Şahin et al. 2014; Salcedo-Sanz et al. 2014). Despite its advantages over conventional data-driven model (e.g., ANN) and successful application in context of drought and evaporative modeling (Deo and Şahin 2015a; b; Deo et al. 2015; Samui and Dixon 2012), to our best knowledge, no study has tested the ELM for streamflow simulation in our region although the importance of streamflow as an environmental parameter (e.g., creek hydrology and ecosystems health) has been overstated (Chowdhury et al. 2010). As such, the adoption of ELM as an improved class of models in eastern Queensland is an original contribution of this research.

The novelty of our research is to develop and validate the utility of an ELM model for simulation of monthly streamflow water levels ( $Q_{WL}$ ) using predictors from set of nine variables for three hydrological catchments (Gowrie Creek: 18.44° S, 145.85° E; Mary River 25.95° E, 152.49° E, and Albert River 28.05° S, 153.05° E). The objective was to demonstrate the ability of the ELM for simulation of  $Q_{WL}$  using hydrometeorological data, climate indices (Southern Oscillation Index (SOI), IOD, PDO, and ENSO Modoki Index (EMI)), and SSTs (Nino 3.0; Nino 3.4, and Nino 4.0). A selection of inputs was performed using cross correlation analysis, and the sensitivity of developed ELM model for simulation of  $Q_{WL}$  based on single inputs (rainfall and periodicity) and all nine inputs was performed. In order to provide a benchmark, an ANN model with optimum input variables was also developed to elucidate the predictive accuracy of ELM model for simulation of  $Q_{WL}$ .

### Theory of machine learning model

#### Extreme learning machine

The extreme learning machine (ELM) algorithm was developed by Huang et al. (2006). It is a state-of-art AI algorithm developed under single-layer feedforward neural network (SLFN) (Fig. 1) and since has widely been used in prediction problems (Acharya et al. 2013; Belayneh and Adamowski 2012; Deo and Şahin 2015b; Şahin et al. 2014). As the ELM is easy to use with no requirement for parameters to be tuned except the predefined network architecture, the algorithm avoids complications faced by the gradient-based algorithms (e.g., ANN) in terms of their relatively slow learning, difficulty with learning epochs, and the problems encountered by the local minima within the predictive data. Importantly, ELM learning process is faster compared to other conventional learning algorithms such as ANN or support vector machine (SVM) (Deo and Şahin 2015b; Rajesh and Prakash 2011).

In the ELM model, the training of input data is accomplished in time span of seconds and minutes, even for large dataset and even complex applications, which are not achievable using conventional techniques. Besides,

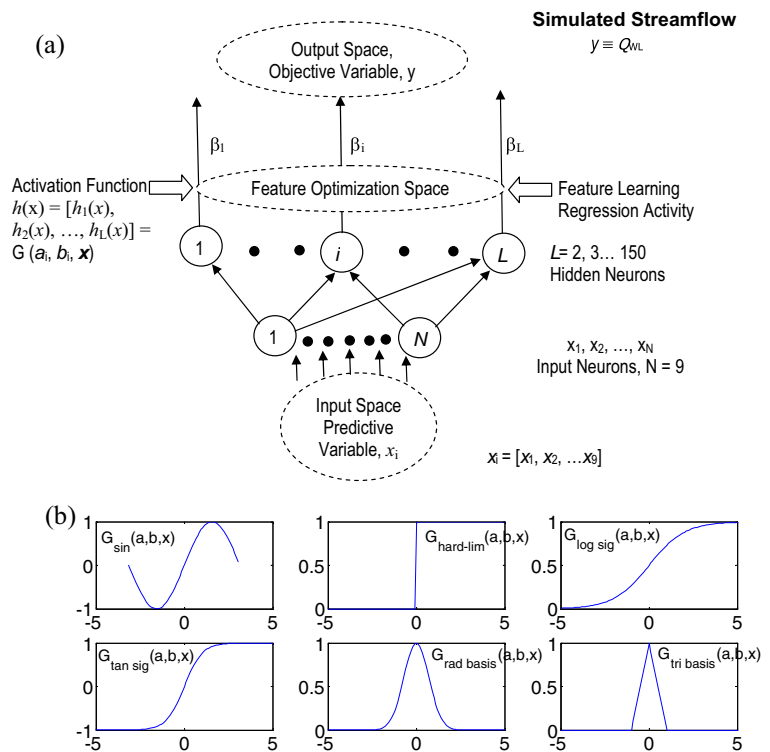
ELM model has good generalization performance compared to the ANN, SVM, and the singular value decomposition (SVD) algorithms in various classification and regression problems (Acharya et al. 2013; Deo and Şahin 2015b; Huang et al. 2015; Sánchez-Monedero et al. 2014). Consequently, the ELM is used as an ideal algorithm for forecasting atmospheric variables such as solar energy, air temperatures, and rainfall (Deo and Şahin 2015b; Leu and Adi 2011; Şahin 2012; Şahin et al. 2013; Şahin et al. 2014; Sánchez-Monedero et al. 2014; Wu and Chau 2010

In Fig. 1a, the schematic structure of ELM is exemplified. Consider a set of  $N$  distinct samples  $(x_i, y_i) \in R^n \times R^m$ ,  $n, m$  are the domains and  $i = 1, 2, \dots, N$ . The SLFNs with  $L$  hidden neurons and activation function,  $G(a_i, b_i, x)$  is described viz (Huang et al. 2006; Şahin et al. 2014):

$$f_L(x) = \sum_{i=1}^L \beta_i h_i(x) \tag{1}$$

where  $\beta = [\beta_1, \beta_2, \dots, \beta_L]^T$  is the output weight matrix between hidden layer of  $L$  nodes to the  $m > 1$  output nodes, and  $h(x) = [h_1(x), h_2(x), \dots, h_L(x)]$  is the non-linear mapping space, which comprised the output row vector with respect to input  $x$ . Note that  $h_i(x)$  is the output

**Fig. 1** a A schematic view of the ELM algorithm with input space, optimization space, hidden neuron parameters ( $a, b$ ), and output space. Inputs are defined by nine predictor variables ( $x$ ) while single output has the objective variable ( $y$ ). Hidden neurons have different type of computational nodes (Huang and Chen 2007). b A sample plot of activation functions,  $G(a, b, x)$ , tried in this study



of the  $i$ th hidden node output. In optimization space, ELM model analyses input data to “learn” the natural patterns or trends present in the time series in order to simulate the objective variable, which is intrinsically related to the input space. According to Huang and Chen (2007), the output function of hidden nodes may not be unique and different functions are used with various neuron arrangements to finalize the optimum model.

In this study, a number of  $h_i(x)$  (Eqs. (7–12)) were tried in terms of the following:

$$h_i(x) = G(a_i \cdot b_i \cdot x), \quad a_i \in R^N, b_i \in R, i = [1, 2, \dots, L] \tag{2}$$

where  $G(a, b, x)$  is the activation function for ELM model with hidden parameters  $(a, b)$ . The  $G(a, b, x)$  is a non-linear piecewise continuous function that satisfies the ELM universal approximation capability theorem (Huang and Chen 2007; Huang et al. 2006), and therefore, it may be optimized for the best unsupervised learning outcome. In the ELM model, the hidden node parameters  $(a, b)$  are randomly generated without any interference from the predictor data and may be deduced by continuous probability distribution instead of being explicitly trained. This leads to remarkable computational efficiency compared to traditional neural network (e.g., SVM or ANN) (Huang et al. 2006; Huang et al. 2015).

The optimum ELM model was deduced by trailing activation functions with several model runs. Figure 1b shows a sample of the plots for  $G(a, b, x)$ . In this study, the activation functions were defined by sine ( $G_{\text{sin}}$ ), hard-limit ( $G_{\text{hard-lim}}$ ), logarithmic sigmoid ( $G_{\text{log sig}}$ ), tangent sigmoid ( $G_{\text{tan sig}}$ ), radial basis ( $G_{\text{rad bias}}$ ), and triangular basis ( $G_{\text{tri basis}}$ ) equations viz

$$G_{\text{sin}}(a, b, x) = \sin(ax + b) \tag{7}$$

$$G_{\text{hard-lim}}(a, b, x) = \begin{cases} 1, & \text{if } ax + b < 0 \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

$$G_{\text{logsig}}(a, b, x) = \frac{1}{1 + \exp(-(ax + b))} \tag{9}$$

$$G_{\text{tansig}}(a, b, x) = \frac{2}{1 + \exp(-2(ax + b)) - 1} \tag{10}$$

$$G_{\text{rad bias}}(a, b, x) = \exp(-(ax + b)^2) \tag{11}$$

$$G_{\text{tribasis}}(a, b, x) = \begin{cases} 1 - |ax + b|, & \text{if } -1 \leq (ax + b) \leq 1 \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

The trialing of the activation functions was necessary to develop a robust model that considers the best equation(s) for extracting features within the input data to simulate  $Q_{\text{WL}}$ . Upon selecting the suitable  $G(a, b, x)$  and deducing weights for hidden and output layer ( $\beta$ ), simulation of  $Q_{\text{WL}}$  was performed by minimizing the approximation error

$$\min_{\beta \in R^L} \|H\beta - T\|^2 \tag{13}$$

where  $\|\cdot\|$  denotes the Frobenius norm and  $H$  is the hidden layer (randomized) matrix produced by the activation function, mathematically defined viz

$$H = \begin{bmatrix} h(x_1^L) \\ \vdots \\ h(x_N^L) \end{bmatrix} = \begin{bmatrix} h_1(x_1) \cdots \cdots h_L(x_1) \\ \vdots \\ h_1(x_N) \cdots \cdots h_L(x_N) \end{bmatrix} \tag{14}$$

Note that  $T$  is the target output matrix that in our case represents the simulated variable,  $y$  ( $\equiv Q_{\text{WL}}$ ):

$$T = \begin{bmatrix} t_1^L \\ \vdots \\ t_m^L \end{bmatrix} = \begin{bmatrix} t_{11} \cdots \cdots t_{1m} \\ \vdots \\ t_{N1} \cdots \cdots t_{Nm} \end{bmatrix} \tag{15}$$

Finally, the optimal solution to Eq. (13) is given by  $\beta^* = H^+ T$  (16)

where  $H^+$  is the Moore–Penrose generalized inverse of the hidden layer matrix  $H$ .

### Artificial neural network

In order to benchmark the performance of the ELM model, an artificial neural network (ANN) model was also developed. The ANN model is a computational paradigm composed of non-linear elements (neurons) operating in parallel and massively connected by networks characterized by different weights. A single neuron computes the sum of its inputs, adds a bias term, and drives the simulation results through a generally non-linear activation function to produce a single output termed as the activation level of the neuron. ANN models are specified by network topology, neuron characteristics, and training or learning rules (Lippman 1987) with inputs, output, and hidden layers with interconnections. The fundamental processing

unit is a neuron, which computes a weighted sum of its input signals,  $x_i$ , for  $i=0, 1, 2, \dots, N$ , hidden layers,  $w_{ij}$ , and then applies a non-linear activation function to produce an output signal  $u_j$  (Deo and Şahin 2015a; Şahin 2012).

A neuronal model consists of an externally applied bias,  $b_k$ , which has the effect of increasing or decreasing the net input of activation function depending on whether it is positive or negative. Mathematically, it is described by

$$y_k = \Phi(u_k + b_k), \text{ where } U_k = \sum_{j=1}^m W_k X_j \quad (17)$$

where  $x_1, x_2, \dots, x_m$  are the inputs signals;  $w_{k1}, w_{k2}, \dots, w_{km}$  are the synaptic weights of neuron  $k$ ;  $u_k$  is the linear combiner output due to input signals;  $b_k$  is the bias;  $\Phi(\cdot)$  is the activation function; and  $y_k$  is the output signal of the neuron. Bias  $b_k$  has the effect of applying an affine transformation to the output  $u_k$  of the linear combiner viz

$$v_k = u_k + b_k \quad (18)$$

In particular, depending on whether the  $b_k$  is positive or negative, the relationship between the induced local field or activation potential  $v_k$  of neuron  $k$  and linear combiner output  $u_k$  can be modified. Note that as a result of this affine transformation, the graph of  $v_k$  versus  $u_k$  no longer passes through the origin. The bias  $b_k$  is an external parameter of artificial neuron  $k$  (Deo and Şahin 2015a; Şahin 2012).

Equivalently, combinations of Eqs. (17) and (18) may be formulated as follows (Haykin 2010)

$$v_k = \sum_{j=0}^m w_{kj} x_j \quad (19)$$

$$y_k = \Phi(v_k) \quad (20)$$

The tangent sigmoid,  $\phi(x)$ , logarithmic sigmoid,  $\psi(x)$ , and linear,  $\chi(x)$ , transfer function are described as follows (Vogl et al. 1988)

$$\varphi(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (21)$$

$$\psi(x) = \frac{1}{1 + e^{-x}} \quad (22)$$

$$\chi(x) = \text{linear}(x) = x \quad (23)$$

Thus, Eqs. (21–23) were trialed in order to determine the best ANN model.

In this study, we have also optimized the ANN-based simulation models by utilizing a number of learning algorithms including the quasi-Newton, resilient, scaled conjugate gradient, Levenberg–Marquardt, conjugate gradient with Powell–Beale restarts, conjugate gradient with Fletcher–Reeves updates, conjugate gradient with Polak–Ribière updates, one-step secant and the gradient descent with momentum and adaptive learning rate, following earlier studies (Deo and Şahin 2015a; Şahin 2012). Therefore, based on the lowest mean square error, the optimum ANN model with most appropriate transfer function and the best machine learning algorithm was adopted for the final simulation of the monthly  $Q_{WL}$ .

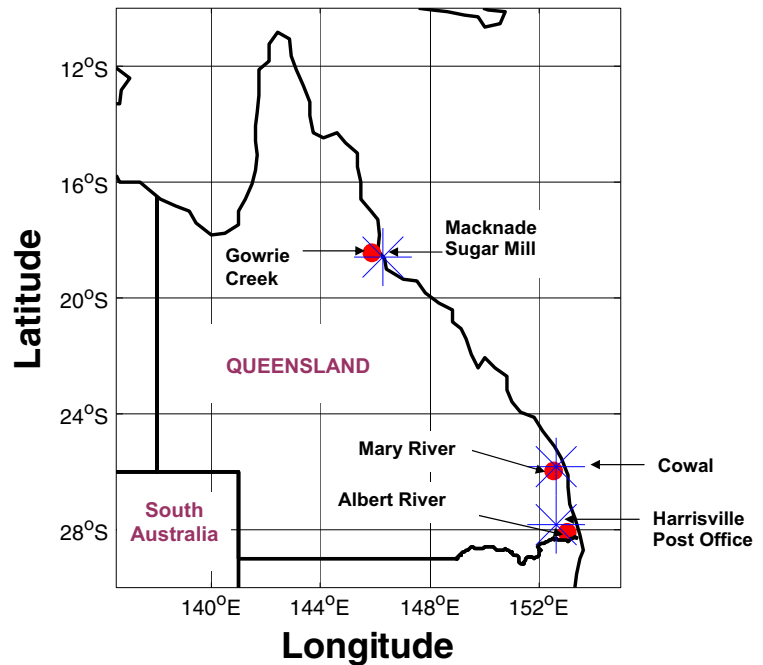
### Materials and method

#### Study area and model dataset

In order to develop a simulation model based on the ELM algorithm, this study has utilized eight predictor time series (plus the month as additional indicator of periodicity) where the rainfall ( $P$ ) was the primary meteorological input and streamflow water level ( $Q_{WL}$ ) was the objective (or simulated) output. Figure 2 plots a geographic map of the present study site, namely Gowrie Creek (Abergowrie), Mary River (Miva Road), and Albert River (Lumeah number 2). Table 1 lists characteristics of the hydrological stations used for simulation of  $Q_{WL}$ . Shown also are the closest meteorological stations that provided the matching  $P$  data for simulation of  $Q_{WL}$ .

The monthly values of  $Q_{WL}$  data were acquired from the Water Monitoring Data Portal (Dept of Environment & Resource Management), <http://watermonitoring.dnrm.qld.gov.au/host.htm> (DNRM 2014). Since reliable monthly  $P$  time series were not available for the streamflow sites, measurements of rainfall from the closest meteorological site (Macknade Sugar Mill, 44.96 km from Gowrie Creek; Cowal 0.99 km from Mary River; Harrisville 45.89 km from Albert River) were extracted from Bureau of Meteorology website (<http://www.bom.gov.au/climate/data-services/>) (Haylock and Nicholls 2000). Also, despite the availability of the  $Q_{WL}$  data from 1953 for Gowrie Creek and Albert River and from 1910 for Mary River, a number of quality control indicators had showed that the earlier period has

**Fig. 2** The hydrological sites from the Queensland Water Monitoring Portal (red circles) and the Australian Bureau of Meteorology's high-quality rainfall sites (blue asterisks). The approximate distance between the hydrological and meteorological stations is shown in Table 1



**Table 1** The characteristics of hydrological (streamflow) and meteorological (rainfall) stations for study sites in eastern Queensland

Description	Gowrie Creek (Abergowrie)	Mary River (Miva Road)	Albert River (Lumeah number 2)
Hydrological stations: streamflow water level, $Q_{WL}$			
QWM Portal Number	116008B	138001A	145101D
Latitude/longitude	18.44° S/145.85° E	25.95° S/152.49° E	28.05° S/153.05° E
Site commence	10 Jan 1953	1 Jan 1910	10 Jan 1953
Zero gauge	33.308	16.866	79.934
Control	Sand gravel	Gravel	Control weir
Maximum gauged level	2.68	17.3	7.84
Maximum gauge date	3 Oct 1983	27 Apr 1989	26 Jan 1974
Distance from stream mouth	6.9 km	126 km	75.5 km
Catchment area	124 km <sup>2</sup>	4755 km <sup>2</sup>	169 km <sup>2</sup>
Gauging	195 gauging between 11 Aug 1954 and 10 Dec 2013	543 gauging between 30 Apr 1909 and 19 Mar 2014	318 gauging between 16 Mar 1954 and 01 Nov 2013
Description	Macknade Sugar Mill	Cowal	Harrisville Post Office
Meteorological stations: rainfall, $P$			
BOM station ID	32032	40013	40094
Latitude/longitude	18.58° S/146.25° E	25.95° S/152.50° E	27.81° S/152.67° E
Site commence	01 May 1890	01 Sept 1894	01 Jan 1896
Site ceased	No	No	No
Elevation	124 m	61 m	55 m
Corresponding distance from the hydrological station	44.96 km	0.99 km	45.89 km

significantly fragmented time series and, therefore, was considered unreliable.

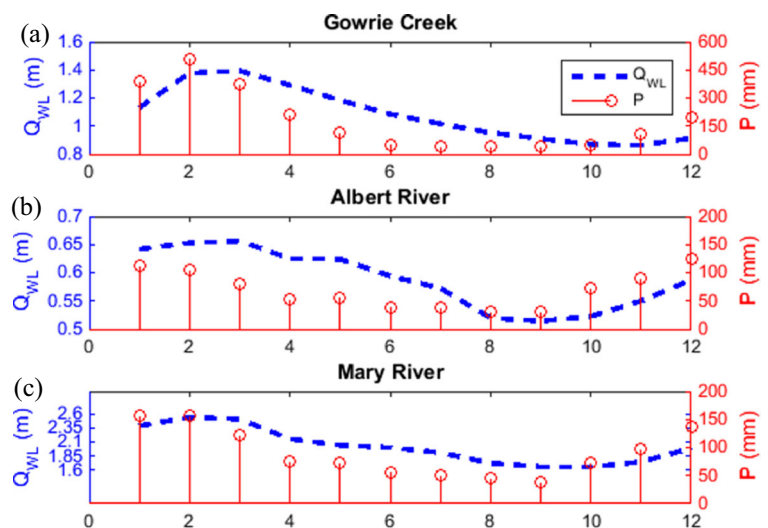
Consequently, the monthly  $P$  data for the period 1960–2012 were used as the predictor variables for both the ELM and the ANN models. Any missing  $P$  and  $Q_{WL}$  in this period were recovered from calendar averaged values within the monthly hydrological period. Figure 3a–c compares the monthly pattern of  $Q_{WL}$  for Gowrie Creek, Mary River, and Albert River and the corresponding  $P$  for Macknade Sugar Mill, Cowal, and Harrisville. For all hydrological stations, there was good correspondence with their monthly cycle of rainfall. The overall trends showed generally high streamflow (and also rainfall) values from the month of January–April and lower from August–November. Also, since the  $P$  data were acquired some distance away from actual hydrological site and that the hydrological peaks can be slightly delayed compared to the rainfall cycle (Brodie et al. 2008), there was a subtle but notable lag between the peak value of  $Q_{WL}$  and the corresponding  $P$  time series as the recharge component of rainfall may not be returned as streamflow water immediately.

In addition to the monthly  $P$  data used as the primary variable for simulation of the mean monthly  $Q_{WL}$ , training datasets utilized the monthly variation in the climate mode indices (SOI, PDO, EMI, and IOD) and SSTs (Nino 3.0 5° N–5° S, 150° W–90° W; Nino 3.4 5° N–5° S, 170° W–120° W; Nino 4.0 5° N–5° S, 160° W–150° W) (Table 2). The SOI and IOD data were acquired from Australian Bureau of Meteorology (Trenberth 1984), and the PDO index was acquired from the Joint

Institute of the Study of Atmosphere and Ocean (Mantua et al. 1997; Zhang et al. 1997). In its original form, the SOI was calculated by the Troup's method using the differences in mean sea-level pressure between Tahiti and Darwin, while the PDO index was created using the UKMO Historical SST dataset for 1900–1981, Reynolds Optimally Interpolated SST (Morid et al. 2007) for January 1982–Dec 2001, and OI.v2 SST fields from January 2002 onward.

In this study, we employed the monthly EMI as a predictor variable, which represented the first two dominant modes of variation in the EOF analysis of SST anomalies area-averaged over (165° E–140° W, 10° S–10° N), (110° W–70° W, 15° S–5° N), and (125° E–145° E, 10° S–20° N), respectively. Data were acquired from Japanese Agency for Marine–Earth Science and Technology (<https://www.jamstec.go.jp/frcgc/research/d1/iod/e/index.html>). It is imperative to note that the EMI reflects coupled ocean–atmosphere interactions in the Pacific Ocean (Ashok et al. 2003, 2007; Weng et al. 2007). Several studies have shown that the EMI is prominent for identifying teleconnection pattern arising from tropical Pacific (Cai and Cowan 2009; Taschetto 2009). Therefore, considering the impacts of large-scale climate drivers and SSTs on rainfall (and streamflow) variability in Australia, the climate mode indices used in the ELM and ANN model development were considered appropriate, as they have also been employed in previous models (Abbot and Marohasy 2012; Abbot and Marohasy 2014; Deo and Şahin 2015a, b; Mekanik et al. 2013; Morid et al. 2007

**Fig. 3** The pattern of monthly cycle of streamflow,  $Q_{WL}$  (m), and precipitation,  $P$  (mm), for the period 1960–2012. Note that the  $P$  data are from nearest meteorological station specified in Table 1





**Table 2** The cross correlation coefficients ( $r_{\text{cross}}$ ) computed for input ( $x$ ) and objective variable ( $y \equiv Q_{\text{WL}}$ ) at zero lag

All predictor (input) variables	Gowrie Creek	Albert River	Mary River
Rainfall: $P$	<i>0.648</i>	<i>0.306</i>	<i>0.718</i>
Nino 3.0 SST (5° N–5° S, 150° W–90° W): Nino 3.0 SST	<i>0.341</i>	0.046	0.044
Nino 3.4 SST (5° N–5° S, 170° W–120° W): Nino 3.4 SST	0.151	–0.099	–0.201
Nino 4.0 SST (5° N–5° S, 160° E–150° W): Nino 4.0 SST	–0.318	–0.099	–0.367
Southern Oscillation Index (SOI)	<i>0.215</i>	<i>0.141</i>	<i>0.277</i>
Pacific Decadal Oscillation Index (PDO)	–0.015	<i>0.174</i>	–0.105
Indian Ocean Dipole Index (IOD)	0.025	<i>0.134</i>	0.049
ENSO Modoki Index (EMI)	–0.314	–0.237	–0.328

The statistically significant  $r_{\text{cross}}$  at 95 % confidence interval is shown in italics

Development of predictive model and performance assessment

All experiments were performed under MATLAB programming platform on a Pentium 4, 2.93GHz CPU. Table 3 details the ELM and its comparative (ANN) model used for simulation of  $Q_{\text{WL}}$ . As a first step, the 53 years of data (1960–2012) were partitioned in two segments: one for training phase (1960–2005) where monthly rainfall, climate indices, and SSTs were used as predictor variables and the other set had testing phase (2006–2012) where the simulated  $Q_{\text{WL}}$  was validated with site-specific measurements. The training dataset was used to develop the appropriate ELM and ANN models. After training the networks, a weight matrix was obtained for each parameter in consideration and applied to the independent input parameters in “test” set. The simulations were compared with the observed streamflow.

Crucial for any robust predictive model is the consideration of the pertinent factors such as the proper selection of input variables, training of these inputs while ensuring minimum level of over-fitting, selecting the best training algorithm and activation function, ensuring smallest generalization error, and using the best performance metrics (Maier and Dandy 2000; Maier et al. 2010; Tiwari and Adamowski 2013). Unlike previous studies that have utilized a prescribed set of predictor variables (e.g., rainfall, temperature, climate index, etc.) without a prior selection process (Deo and Şahin 2015a, b; Salcedo-Sanz et al. 2015), in this study, we have deduced the input signal(s) by the correlation with  $Q_{\text{WL}}$  in order to develop a robust model.

As there is no “rule of thumb” to determine significant inputs (Adamowski 2008; Tiwari and Adamowski

2013), the input signal patterns were analyzed by correlation statistics using a cross correlation function (Adamowski et al. 2012). The cross correlation function measured the statistical similarity between inputs ( $x$ ) and shifted (lagged) copies of  $Q_{\text{WL}}$  as a function of lag. For a discrete signal, correlation of time series,  $x_i = (x_1, x_2 \dots x_{M-1})$  and  $y = (y_1, y_2 \dots y_{N-1})$ , is

$$\varphi_{xy,k} = \sum_{j=\max(0,k)}^{\min(M-1+k,N-1)} x_{j-k}, \quad k = -(M+1), \dots, 0, \dots, (N-1) \quad (24)$$

and the normalized correlation coefficient,  $r_{\text{cross}}$ , is defined as

$$r_{\text{cross}}(t) = \frac{\varphi_{xy}(t)}{\sqrt{\varphi_{xx}(0) \varphi_{yy}(0)}} \quad (25)$$

such that the quantity  $r_{\text{cross}}(t)$  will vary between –1 and 1. A value  $r_{\text{cross}}(t) = 1$  indicates that at alignment  $t$ , the two time series have the same exact shape (although amplitudes may be different) while –1 indicates the same shape with opposite sign. However, when  $r_{\text{cross}}(t) = 0$ , the two signals are uncorrelated but if  $r_{\text{cross}}(t) \geq 0.70$ , a good match is evident between  $x$  (inputs) and  $y$  ( $Q_{\text{WL}}$ ).

Subsequently, a set of input combinations were deduced by analyzing  $r_{\text{cross}}$  of each variable with  $Q_{\text{WL}}$ . Figure 4 plots correlogram of  $Q_{\text{WL}}$  with inputs (i.e., rainfall, SSTs, and climate indices). In this plot, the statistically significant  $r_{\text{cross}}$  at 95 % confidence levels is located outside the blue lines. Evidently, the level of correlation of the  $P$  data with  $Q_{\text{WL}}$  acquired the highest magnitude for all stations at zero lag ( $r_{\text{cross}} \approx 0.306$ – $0.718$ ) which was also statistically significant (Table 2) (note that the cross correlation of streamflow with inputs

**Table 3** The ELM and ANN models with best input combinations as per Table 2

Station	Input combinations, $x$	ELM model	ANN model
Gowrie Creek	$x = [\text{month}, P, \text{Nino 3.0 SST}, \text{Nino 4.0 SST}, \text{SOI}, \text{EMI}]$	Number of layers: 3 Input neurons: 6 Hidden neurons: $\{2 \dots 150\}$ Output neuron: 1 ( $Q_{WL}$ ) Activation functions: sig; sin; <i>hardlim</i> ; tribas; radbas; logsig; tansig Learning rule: ELM for SLFN  Architecture: <i>6-106-1</i>	Number of layers: 3 Input neurons: 6 Hidden neurons: $\{2, 4 \dots 50\}$ Output neuron: 1 ( $Q_{WL}$ ) Transfer function: <i>tansig</i> , logsig  Learning algorithm: trainbfg, trainrp, trainscg, <i>trainlm</i> , traincgb, traincgf, traincgp, trainoss, traingdx  Architecture: <i>6-28-1</i>
Albert River	$x = [\text{month}, P, \text{SOI}, \text{PDO}, \text{IOD}, \text{EMI}]$	Number of layers: 3 input neurons: 6 hidden neurons: $\{2 \dots 150\}$ Output neuron: 1 ( $Q_{WL}$ ) Activation functions: sig; sin; <i>hardlim</i> ; tribas; radbas; logsig; tansig Learning rule: ELM for SLFN Architecture: <i>6-74-1</i>	Number of layers: 3 Input neurons: 6 Hidden neurons: $\{2, 4 \dots 50\}$ Output neuron: 1 ( $Q_{WL}$ ) Transfer function: <i>tansig</i> , <i>logsig</i> Learning algorithm: trainbfg, trainrp, trainscg, <i>trainlm</i> , traincgb, traincgf, traincgp, trainoss, traingdx  Architecture: <i>6-18-1</i>
Mary River	$x = [\text{month}, P, \text{Nino 3.4 SST}, \text{Nino 4.0 SST}, \text{SOI}, \text{EMI}]$	Number of layers: 3 Input neurons: 6 Hidden neurons: $\{2 \dots 150\}$ Output neuron: 1 ( $Q_{WL}$ ) Activation functions: sig; sin; <i>hardlim</i> ; tribas; radbas; logsig; tansig Learning rule: ELM for SLFN Architecture: <i>6-146-1</i>	Number of layers: 3 Input neurons: 6 Hidden neurons: $\{2, 4 \dots 50\}$ Output neuron: 1 ( $Q_{WL}$ ) Transfer function: <i>tansig</i> , <i>logsig</i> Learning algorithm: trainbfg, trainrp, trainscg, <i>trainlm</i> , traincgb, traincgf, traincgp, trainoss, traingdx  Architecture: <i>6-20-1</i>

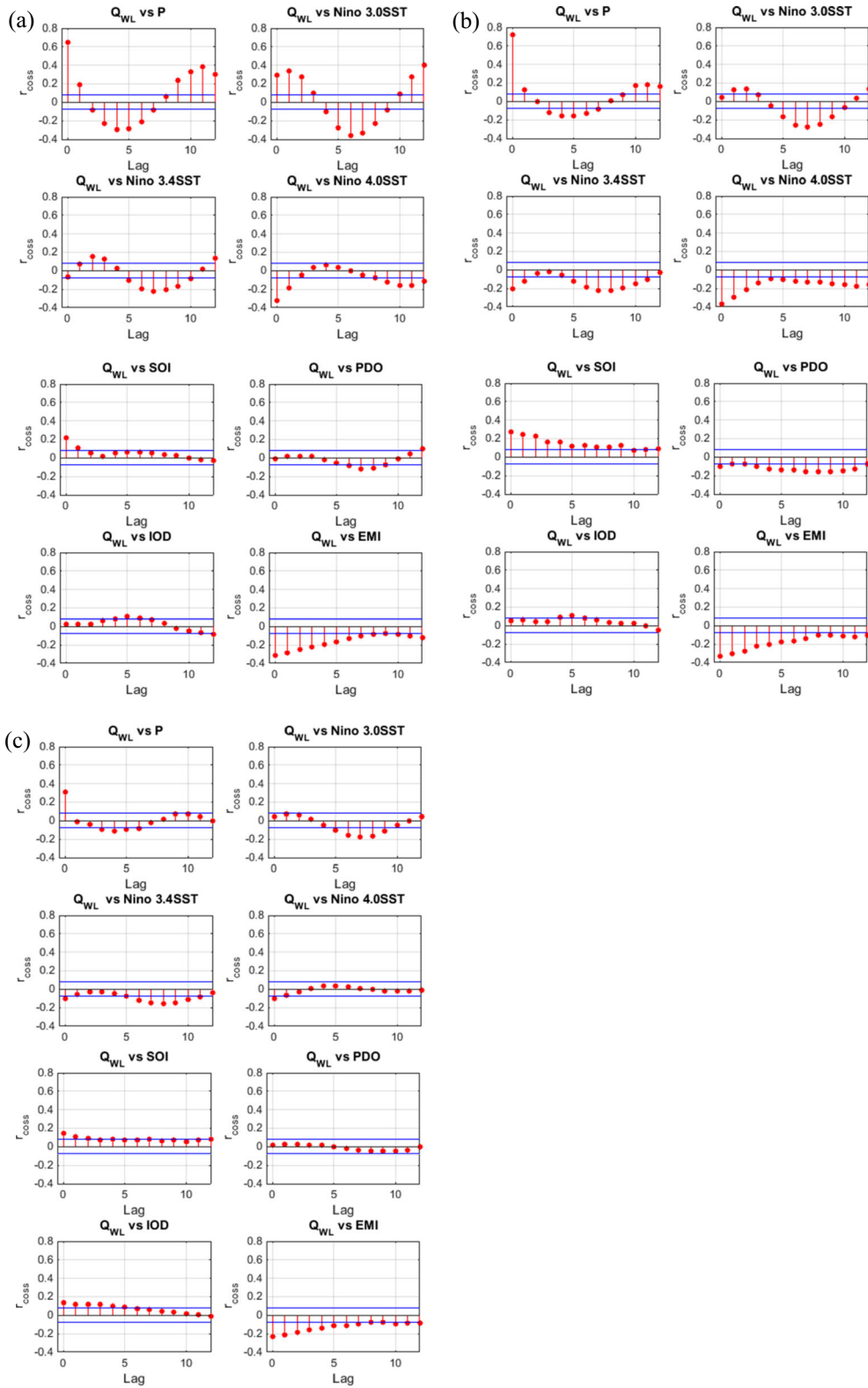
Acronyms for the ELM model functions are *hardlim* hard limit, *tribas* triangular basis, *radbas* radial basis, *logsig* log-sigmoid, *tansig* hyperbolic tangent sigmoid. Acronyms for the ANN model functions are *trainbfg* BFGS quasi-Newton, *trainrp* resilient, *trainscg* scaled conjugate gradient, *trainlm* Levenberg–Marquardt, *traincgb* conjugate gradient BP with Powell–Beale restarts, *traincgf* conjugate gradient BP with Fletcher–Reeves update, *trainoss* one-step secant, *trainbr* Bayesian regulation, *traingdx* gradient descent with momentum and adaptive learning. The optimum (best) functions used in the ELM and ANN models are in italics

was also significant at various other lags; however, in this study, only the zero lag has been used as the  $r_{cross}$  attained the highest magnitude). Except for the EMI and SOI which had statistically significant  $r_{cross}$  from  $(-0.237)$  to  $(-0.328)$  and  $(0.241)$  to  $(0.277)$ , respectively, for all three stations, cross correlation of  $Q_{WL}$  with the other input variables varied quite significantly. That is, for Gowrie Creek, a statistically significant value of  $r_{cross} \approx 0.341$  was obtained for  $Q_{WL}$  with Nino 3.0 SST, but a very low statistically insignificant value was obtained for Albert and Mary River. Furthermore, Nino 3.4 SST correlated well with  $Q_{WL}$  for Mary River ( $r_{cross} \approx -0.201$ ), but it was poorly correlated with

streamflow measurements for Gowrie Creek and Albert River. For the PDO and IOD indices, the correlation with  $Q_{WL}$  was statistically significant only for Albert Creek station ( $r_{cross} \approx 0.134-0.174$ ) (Table 2).

Taken together, discernible differences in cross correlations of streamflow with prescribed inputs showed

**Fig. 4** A set of correlogram of streamflow water levels ( $Q_{WL}$ ) and its predictor variable as the rainfall ( $P$ ), sea surface temperatures (Nino 3.0 SST, Nino 3.4 SST; Nino 4.0 SST), Southern Oscillation Index (SOI), Pacific Decadal Oscillation Index (PDO), Indian Ocean Dipole (IOD), and ENSO Modoki Index (EMI). **a** Albert River, **b** Gowrie Creek, and **c** Mary River. Statistically significant cross correlation coefficients ( $r_{cross}$ ) at 95 % confidence level are outside the *blue lines*



the significance of selecting input combinations rather than using all variables. Consequently, in this study, three sets of ELM models were developed: (1) an optimum model with unique combinations of inputs ( $x$ ) where  $x$ =[month,  $P$ , Nino 3.0 SST, Nino 4.0 SST, SOI, EMI] for Gowrie Creek station,  $x$ =[month,  $P$ , SOI, PDO, IOD, EMI] for Albert River station, and  $x$ =[month,  $P$ , Nino 3.4 SST, Nino 4.0 SST, SOI, EMI] for Mary River station, (2) trial ELM model 1 with the month and rainfall as inputs, and (3) trial ELM model 2 nine input variables. The latter were utilized to check the response of ELM algorithm for simulating streamflow when no feature selection process was incorporated. It is also imperative to note that a follow-up study could incorporate appropriate lagged signals with statistically significant cross correlations.

A three-layer neuron arrangement with input space (where input variables were fed in), learning space (where activation or transfer functions were applied to formulate the best neuronal arrangements with lowest mean square error), and output space (where streamflow was simulated as objective variable) was designed. Table 3 provides details of the ELM and ANN models. For both cases, the maximum number of input neurons ( $x$ ) was 9 (denoted as  $x_1, x_2, x_3 \dots, x_9$ ), where one neuron was assigned for month (to account for periodicity), one neuron for meteorological inputs (rainfall), four neurons for climate indices (SOI, IOD, EMI, and PDO), and remainder three neurons for sea surface temperatures (Nino SSTs). In the learning space, the activation functions defined by Eqs.(7–12) were applied to simulate the  $Q_{WL}$  based on the sine, log-sigmoid, hyperbolic tangent sigmoid, radial bias, triangular bias, hyperbolic tangent sigmoid, and the hard-limit equations.

In each trial, the number of nodes in hidden layer was increased gradually by an interval of five with different activation functions tried (Eq. (7–12)) and subsequent adjustments of the model’s parameters were performed to deduce the optimum ELM architecture (Table 3). Then, the nearly optimal node determined by lowest mean square error, which was intrinsically disparate for each site, was adopted. For ANN model, two sets of simulations were performed: (1) with optimum combination of predictor variables and (2) with the month and rainfall as the prescribed inputs. The simulations were compared with an equivalent ELM model. The tuning of hidden layer for the ANN model was performed using 5 to 50 neurons with different machine learning algorithms. Table 3 shows the architecture of

ELM and ANN models used for simulation of monthly streamflow.

Based on simulated and observed  $Q_{WL}$ , model performance was assessed using mean absolute error (MAE), coefficient of determination ( $R^2$ ) (Paulescu et al. 2011; Ulgen and Hepbasli 2002), Willmott’s Index ( $d$ ) (Acharya et al. 2013; Willmott 1981; Willmott 1982), and the Nash–Sutcliffe coefficient ( $E$ ) (Krause et al. 2005; Nash and Sutcliffe 1970) viz the following mathematical equations:

$$MAE = \frac{1}{N} \sum_{i=1}^n \left| (Q_{WL_{pred,i}} - Q_{WL_{obs,i}})_t \right| \tag{26}$$

$$R^2 = \left( \frac{\sum_{i=1}^n (Q_{WL_{obs,i}} - \bar{Q}_{WL_{obs}})(Q_{WL_{pred,i}} - \bar{Q}_{WL_{pred}})}{\sqrt{\sum_{i=1}^n (Q_{WL_{obs,i}} - \bar{Q}_{WL_{obs}})^2} \sqrt{\sum_{i=1}^n (Q_{WL_{pred,i}} - \bar{Q}_{WL_{pred}})^2}} \right)^2 \tag{27}$$

$$d = 1 - \left[ \frac{\sum_{i=1}^N (Q_{WL_{obs,i}} - Q_{WL_{pred,i}})^2}{\sum_{i=1}^N (|Q'_{WL_{pred,i}} - \bar{Q}_{WL_{obs}}| - |Q'_{WL_{obs,i}} - \bar{Q}_{WL_{obs}}|)^2} \right], \quad 0 \leq d \leq 1 \tag{28}$$

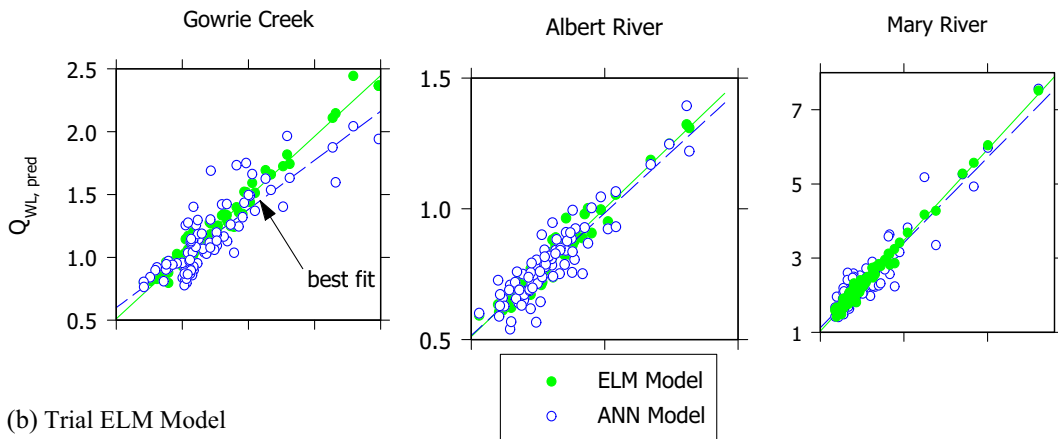
$$E = 1 - \left[ \frac{\sum_{i=1}^N (Q_{WL_{obs,i}} - Q_{WL_{pred,i}})^2}{\sum_{i=1}^N (Q_{WL_{obs,i}} - \bar{Q}_{WL_{obs}})^2} \right], \quad 0 \leq E \leq 1 \tag{29}$$

where  $Q_{WL_{p_i}}$  and  $Q_{WL_{o_i}}$  were the  $i$ th monthly value of simulated and observed  $Q_{WL}$  in the test period  $t$ , respectively,  $i$  is the month of the test data,  $\bar{Q}_{-WL_{obs}}$  and  $\bar{Q}_{-WL_p}$  are the overall mean parameters, and  $N$  (=84) is the length (number of samples in the test set) for period  $t$  (2006 to 2012).

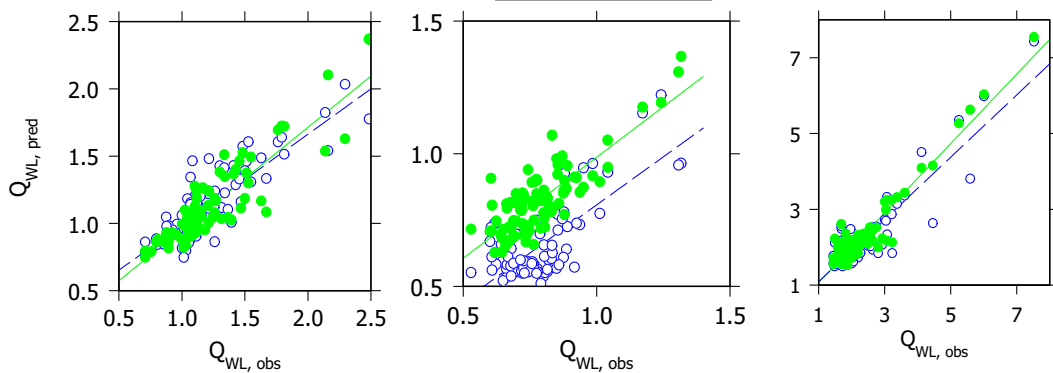
### Results and discussion

The relationship between simulation and observed streamflow water levels obtained from the ELM and ANN models with optimum input combinations (Fig. 4; Tables 2 and 3) is shown in Fig. 5a. Also shown are the simulation results for trial ELM model using

(a) Optimum ELM Model



(b) Trial ELM Model



**Fig. 5** A scatter plot of simulated and observed streamflow ( $Q_{WL}$ ) in test period (2006–2012) with best fit line of the form  $Q_{pred} = mQ_{obs} + C$ . **a** Optimum ELM model with best input

combinations specified in Table 3. **b** Trial ELM model with rainfall ( $P$ ) and corresponding month as input variables

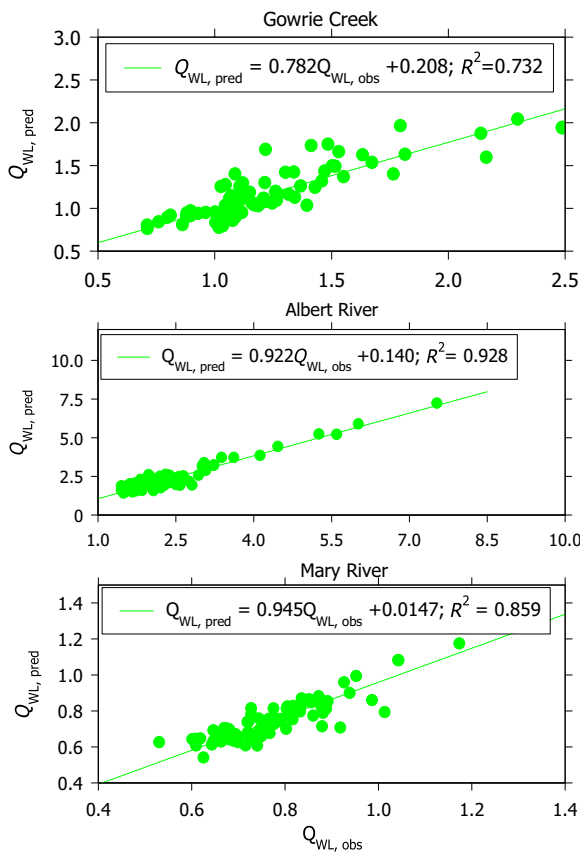
month and rainfall as the predictor dataset (Fig. 5b). Undoubtedly, the predictive ability of the ELM and the ANN models varied greatly between the different stations and the input combinations used to simulate the monthly  $Q_{WL}$ . In fact, for all three stations considered, the ELM model appeared to be accurate than the ANN model since its level of scatter was the lowest. However, when optimum input combinations for each station (Table 3) were used, the degree of scatter was lower for both the ELM and ANN models compared to the simulations when only rainfall and corresponding month were inputted into the model.

Among the three stations tested, the simulation for Mary River for both input cases was the best whereas simulations for Gowrie Creek were the least accurate. Interestingly, when only the rainfall and the month were considered as inputs for Albert River site (Fig. 5b), the scatterplot appeared to be consistently shifted to higher simulated  $Q_{WL}$  for ELM model. Although the actual

cause of this is currently unclear, the lesser scatter for the ELM model showed its better performance even with less number of predictor datum points.

In order to check the performance of the ELM model with nine input variables, Fig. 6 shows the simulated and observed mean streamflow water levels for the trial ELM model. The square of the correlation coefficient ( $R^2$ ) and the best-fit equations are shown for each station. Also, Table 4 lists the statistical metrics of regression analysis performed between predicted streamflow ( $Q_{WL, pred}$ ) and observed streamflow ( $Q_{WL, obs}$ ) by ELM and ANN models, assessed by a simple regression equation,  $Q_{WL, pred} = mQ_{WL, obs} + C$ . Note that regression coefficient ( $r$ ) and maximum deviation (maxDev) of simulations from observed index with gradient ( $m$ ) of linear regression plot, regression coefficient ( $r$ ), and the  $y$ -intercept ( $C$ ) are included.

Comparing Fig. 6 with Fig. 5, it was noticeable that the inclusion of irrelevant input variables seems to



**Fig. 6** A scatter plot of simulated and observed mean streamflow ( $Q_{WL}$ ) for the trial ELM model with all nine inputs,  $x = [\text{month}; P; \text{Nino } 3.0 \text{ SST}; \text{Nino } 3.4 \text{ SST}; \text{Nino } 4.0 \text{ SST}; \text{SOI}; \text{PDO}; \text{IOD}; \text{EMI}]$  in test period (2006–2012). For each station, the square of correlation coefficient ( $R^2$ ) and best-fit equation have been shown

deteriorate the performance of the ELM model. For example, for the case of Gowrie Creek, the gradient and correlation coefficients were 0.969 and 0.964, respectively, for the optimum model with six inputs, but with the inclusion of nine inputs in the trial ELM model, there was a reduction in  $m$  and  $r$  by about 19.3 and 24.1 %, respectively. Likewise, the  $y$ -intercept (whose ideal zero value is expected to show the best fit of observed and simulated water levels) increased from 0.003 to 0.203 when all nine inputs were incorporated into the trial ELM model. Likewise, the maximum deviation increased from 0.188 to 0.523. However, it was imperative to note that the trial ELM model with only rainfall and the corresponding month yielded better results compared to the trial ELM model with all nine input variables.

According to the results of Fig. 6 and Table 4, the inclusion of irrelevant inputs into the ELM model inadvertently deteriorated the accuracy of the simulated

$Q_{WL}$ . Also, based on the scatterplot, it was evident that the optimum ELM model simulations for Mary River yielded the best results (with  $m \approx 0.979$  and  $r \approx 0.990$ ). However, simulations by the optimum ANN model provided the best results for Albert River ( $m \approx 0.936$ ,  $r \approx 0.830$ ) followed by Mary River ( $m \approx 0.917$ ,  $r \approx 0.892$ ) and the worst for Gowrie Creek ( $m \approx 0.729$ ,  $r \approx 0.732$ ). Notwithstanding this, the performance indicators remained smaller in magnitude compared those of the optimum ELM model (Table 4), thus demonstrating the superiority of the ELM over ANN model for monthly streamflow simulation.

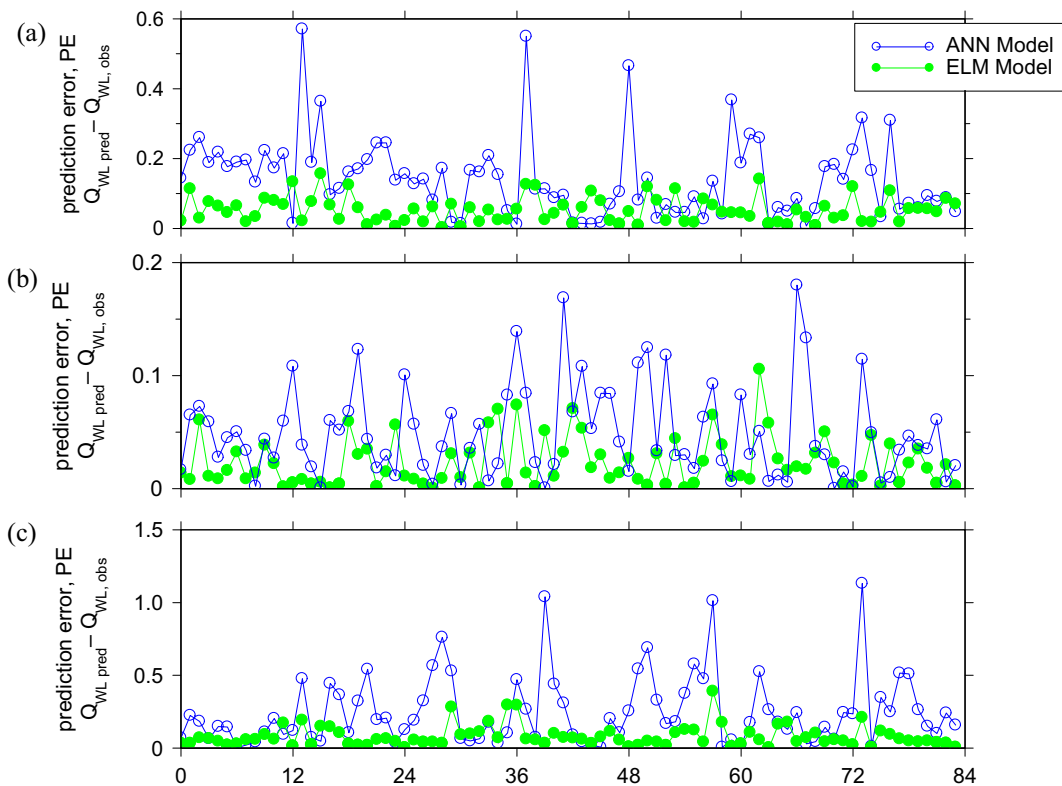
In addition to inspecting the visual agreement between simulated and observed  $Q_{WL}$ , it was important to assess the actual model prediction error (PE) for each month where the quantity  $PE = Q_{WL, pred} - Q_{WL, obs}$  was calculated over the test period. Figure 7 shows the absolute value of PE for the optimum ELM model. For comparison, the results of the optimum ANN model were also included. There was unambiguous evidence that the simulated value of  $Q_{WL}$  was in much better agreement with the observed values of  $Q_{WL}$  for the optimum ELM model, as this model yielded much smaller error values compared to the optimum ANN model. For all three stations, the performance of the ELM model outweighed the performance of the ANN, as it resulted in significantly large values of PE.

In Table 5, we assess the overall model PE, standard deviation ( $\sigma$ ), and the number of datum points within the  $\pm (0-1)\sigma$ ,  $\pm (1-2)\sigma$ ,  $\pm (2-3)\sigma$ , and  $>\pm 3.0\sigma$  for the optimum ELM and ANN models. Indeed, the metrics showed that for all stations considered, the optimum ELM model was more accurate with  $|PE| \approx 0.0278-0.0793$  compared to the optimum ANN model with  $|PE| \approx 0.0489-0.2493$ . It was also noteworthy that the range of standard deviations for the ELM model simulations was smaller ( $\sigma \approx 0.0218-0.0725$ ) relative to the larger range of 0.0407–0.2381 (for the ANN model). This suggested that the optimum ELM model was not only more accurate but also stable and, therefore, exhibited lower fluctuations in the simulated value of  $Q_{WL}$ . Another interesting observation can be made by checking the number of error values within the  $\pm (0-1)$ ,  $\pm (1-2)$ ,  $\pm (2-3)$ , and  $\pm (>) 3$  standard deviations. When the PE was considered within  $\pm (>) 3\sigma$ , a maximum of 1 datum point was located within this error range for the ELM model whereas for the ANN model, up to 3 data points within the test period had model errors located within  $\pm (>) 3\sigma$ . Consequently, this also

**Table 4** Statistical metrics of regression analysis for simulated ( $Q_{WL\ pred}$ ) and observed streamflow ( $Q_{WL\ obs}$ ) based on regression equation:  $Q_{WL\ pred} = mQ_{WL\ obs} + C$

Regression parameter	Optimum ELM model (selected input combinations)			Optimum ANN		
	Gowrie Creek	Albert River	Mary River	Gowrie Creek	Albert River	Mary River
$m$	0.969	0.979	0.979	0.728	0.936	0.917
$R^2$	0.964	0.957	0.990	0.732	0.830	0.892
maxDev	0.188	0.101	0.377	0.523	0.180	1.170
$C$	0.003	0.022	0.055	0.208	0.048	0.207
	Trial ELM (with inputs, $x = [month, P]$ )			Trial ANN (with inputs, $x = [month, P]$ )		
$m$	0.761	0.761	0.916	0.674	0.725	0.823
$R^2$	0.781	0.685	0.910	0.723	0.571	0.823
maxDev	0.385	0.219	1.010	0.415	0.236	1.32
$C$	0.193	0.226	0.156	0.315	0.019	0.261
	Trial ELM (with nine inputs, no prior selection)					
$m$	0.782	0.922	0.945			
$R^2$	0.732	0.928	0.859			
maxDev	0.523	0.821	0.181			
$C$	0.208	0.140	0.015			

The regression coefficient ( $R$ ) and maximum deviation (maxDev) of simulations are shown.  $m$  is the gradient of linear regression plot;  $R$  is the regression coefficient; maxDev is the maximum deviation of simulated  $Q_{WL\ obs}$  from observed variable;  $C$  is the  $y$ -intercept



**Fig. 7** The prediction error ( $PE = Q_{WL\ pred} - Q_{WL\ obs}$ ) using optimum ELM model shown as the difference between observed ( $Q_{WL\ obs}$ ) and simulated streamflow ( $Q_{WL\ pred}$ ) for **a** Gowrie Creek, **b** Albert River, and **c** Mary River in test period (2006–2012)

**Table 5** An assessment of model prediction error, |PE|, standard deviation,  $\sigma$ , and number of datum points in  $\pm (0-1)\sigma$ ,  $\pm (1-2)\sigma$ ,  $\pm (2-3)\sigma$ , and  $>\pm 3.0\sigma$  for optimum models

Station	Mean error,  PE  (m)	Standard deviation, $\sigma$	Number of points			
			$\pm (0-1)\sigma$	$\pm (1-2)\sigma$	$\pm (2-3)\sigma$	$>\pm 3\sigma$
Optimum ELM model						
Gowrie Creek	0.0528	0.0369	62	19	3	0
Albert River	0.0278	0.0218	66	14	3	1
Mary River	0.0793	0.0725	69	11	3	1
Optimum ANN model						
Gowrie Creek	0.1437	0.1137	62	19	1	2
Albert River	0.0489	0.0407	58	22	3	1
Mary River	0.2493	0.2381	68	12	1	3

All units are in meters (m)

concur with the deduction that the optimum ELM model was highly accurate compared to the optimum ANN model.

In Table 6, we show the model performance metrics defined by Eqs. (17–20) computed within the test period (2006–2012). Note that performance metrics for the optimum ELM model with six best input combinations, trial ELM models with only rainfall and month as input variable, and the trial ELM model with all nine inputs have been shown. For comparison purposes, the optimum ANN (best inputs) and trial ANN model (all nine inputs) results have also been tabulated. In interpreting these performance metrics, one must be mindful that the magnitude of MAE for the best model is expected to be

as small as possible in order to demonstrate that the simulated streamflow exhibits the lowest deviation from the observed values. As MAE is considered to be less sensitive to extreme values in the simulated data than the root mean square error (Fox 1981), we have ignored the latter metric in our analysis. Also, for best model with most reliable simulations, the magnitude of  $R^2$  which was determined from a scatter plot of observed and simulated  $Q_{WL}$  is expected to be close to unity, and  $d$  and  $E$  should be unity for a perfect fit (Krause et al. 2005). In model assessment, a disadvantage of the  $R^2$  and  $E_{NS}$  arises from the fact that the differences of observed and predicted streamflow parameters use the squared values. Consequently, large errors in time series

**Table 6** The model performance based on coefficient of determination ( $R^2$ ), Willmott’s Index ( $d$ ), Nash–Sutcliffe coefficient ( $E_{NS}$ ), peak percentage deviation ( $P_{dv}$ ), and mean absolute error (MAE) in the test period (2006–2012)

Station	ELM					ANN				
	$R^2$	$d$	$E_{NS}$	$P_{dv}$ (%)	MAE (m)	$R^2$	$d$	$E_{NS}$	$P_{dv}$ (%)	MAE (m)
Optimum model										
Gowrie Creek	0.964	0.968	0.963	1.993	0.053	0.732	0.802	0.698	18.080	0.144
Albert River	0.957	0.962	0.955	-0.091	0.023	0.830	0.863	0.816	-5.527	0.049
Mary River	0.990	0.986	0.989	0.372	0.079	0.892	0.855	0.891	-0.254	0.249
Trial model with only rainfall and month as inputs										
Gowrie Creek	0.781	0.802	0.691	4.918	0.128	0.723	0.773	0.656	18.389	0.149
Albert River	0.685	0.537	0.617	-3.503	0.072	0.571	0.588	-0.312	7.500	0.145
Mary River	0.910	0.870	0.909	0.016	0.227	0.823	0.776	0.800	1.555	0.312
Trial model all nine inputs (no prior selection)										
Gowrie Creek	0.732	0.809	0.698	18.080	0.144					
Albert River	0.859	0.876	0.812	2.358	0.049					
Mary River	0.928	0.903	0.926	4.178	0.222					

The optimum model was selected based on input combinations (x) and compared with an equivalent ANN model



can be overestimated whereas small values can be neglected (Legates and McCabe 1999). This insensitivity was overcome using Willmott's Index,  $d$  (Willmott 1981), where the ratio of mean square error and potential error was considered for ELM model assessment instead of the squared differences between simulated and observed parameters (Willmott 1984).

The statistical metrics (Table 6) provided undisputed evidence that the optimum ELM was highly accurate compared to the optimum ANN model. In comparison with the ANN model, the ELM model yielded correlation statistics that were very high ( $R^2 \approx 0.957\text{--}0.994$ ) and so were the magnitudes of Willmott's Index and Nash–Sutcliffe efficiency ( $d \approx 0.968\text{--}0.986$  and  $E_{NS} \approx 0.955\text{--}0.989$ ). Importantly, the peak percentage deviation for the ELM model was also quite low ( $P_{dv} \approx -0.091\text{--}1.993$ ) compared to  $P_{dv} \approx -0.254\text{--}18.080$  for the ANN model. This showed that the ELM model generated much reliable results compared to the ANN, which also concurred emphatically with relatively lower magnitudes of the MAE for the former model.

In terms of the simulation of  $Q_{WL}$  from the trial models with only rainfall and month as inputs, the ELM yielded better results than the ANN, albeit its overall performance was worse compared to the optimum ELM with selected inputs. For example, simulations for the trial ELM model with rainfall and month as input yielded  $R^2$  that was lower by 18.98 % (Gowrie Creek), 28.42 % (Albert River), and 8.08 % (Mary River), and the magnitudes of  $d$  were by 17.15, 44.18, and 11.77 %, respectively. Likewise, the peak percentage deviation and MAE values were significantly higher when trial ELM model was simulated with only rainfall and month as the input variable. Similar observations were made when the ELM model was executed with all nine inputs rather than the selected optimum input combinations.

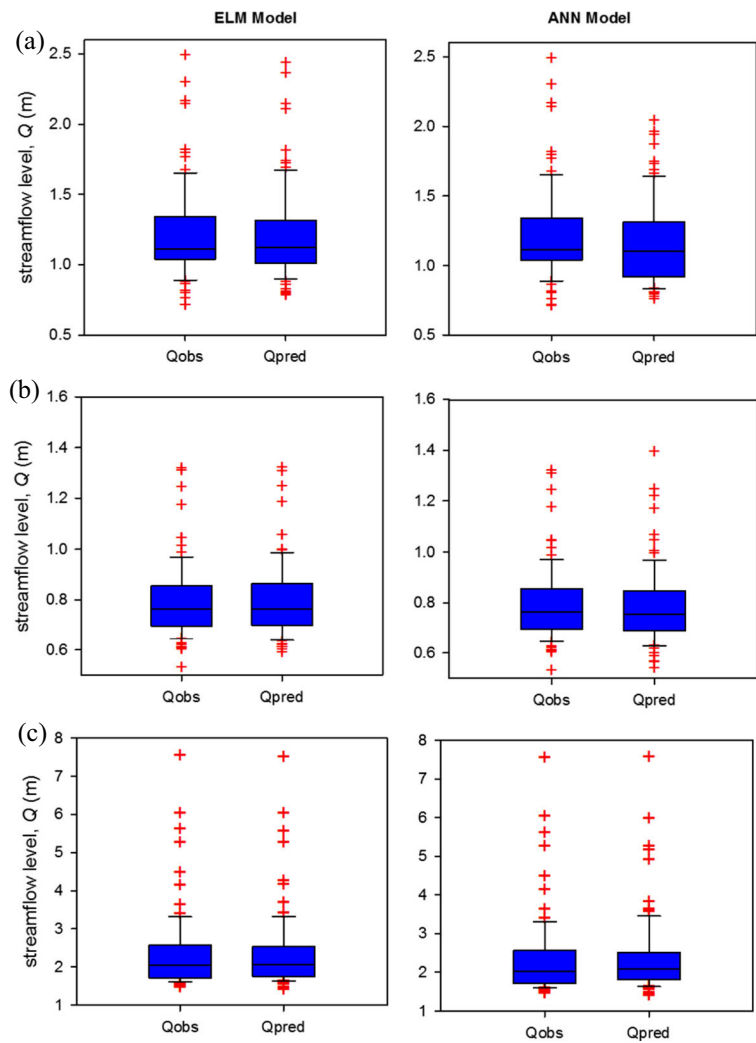
It was noteworthy that for all models tested, the overall simulation accuracy over the test period represented by Eqs. (18–20) had better performance for Mary River compared to the other two stations, although PE represented by Eq. (17) was higher. That is, for Mary River, the  $r^2$ ,  $d$ , and  $E_{NS}$  registered approximately 0.990, 0.986, and 0.989, respectively, but the MAE was 0.079 m (optimum ELM) compared to 0.964, 0.968, and 0.963 (Gowrie Creek) and 0.957, 0.962, and 0.955 (Albert River). This was also confirmed by simulations obtained from the optimum ANN as well as the trial

ELM and ANN models. As confirmed later (Figs. 8 and 9), the distribution of simulated  $Q_{WL}$  was closer to the observed and the magnitude of statistically significant cross correlation coefficients was larger for Mary River (Table 2; Fig. 4). Taken together, it is conclusive that among the other stations, simulations for Mary River were more accurate than Gowrie Creek and Albert River. This was perhaps attributable to the closer distance (which was  $<0.99$  km) of rainfall station (Cowal) to the (hydrological) station (Mary River) where simulations were performed. By contrast, rainfall stations used for Gowrie Creek and Albert River were inadvertently 45 km away (Table 1) and, therefore, were unresponsive to changes in hydrological flow. It was perhaps this reason, and others, that the simulations for Mary River yielded the best agreement with observed  $Q_{WL}$ .

A boxplot with the distribution of simulated and observed streamflow in the prescribed test period is shown in Fig. 8. Note that the whiskers in boxplot were used to show the extremities of predicted and observed streamflow properties using their respective quartile values where the lower end of each boxplot was at the lower quartile,  $p_{25}$  (25th percentile), and upper end for the upper quartile,  $p_{75}$  (75th percentile), and the second quartile,  $p_{50}$  (50th percentile), was the median. Two horizontal whiskers were extended from the top and bottom of the box, where the bottom whisker extends from  $p_{25}$  to the smallest non-outlier in the dataset, whereas the other one goes from  $p_{75}$  to the largest non-outlier. Table 6 summarizes the statistical properties for simulated and observed streamflow.

According to Fig. 8, the distribution of simulated streamflow was close to the observed values for all three stations when the optimum ELM model was utilized. This confirmed that the ELM model was more accurate in simulating the monthly  $Q_{WL}$ . However, when the optimum ANN model was used, there was a very significant discrepancy between the distribution of  $Q_{WL}$  pred and  $Q_{WL}$  obs especially for Gowrie Creek. Quite clearly, the extreme values of streamflow were poorly represented in the ANN simulation model for this site. In fact, the simulated  $Q_{WL}$  pred was highly underpredicted as most of the simulated values were lower than the observed. For the case of Albert River, the optimum ANN model was more responsive and, therefore, yielded more accurate simulations despite the presence of a notable outlier in  $Q_{WL}$  pred. The most accurate streamflow simulations were obtained for Mary River both by optimum ELM and ANN models.

**Fig. 8** Boxplot of distributions of predicted ( $Q_{obs}$ ) and simulation streamflow ( $Q_{pred}$ ) for **a** Gowrie Creek, **b** Albert River, and **c** Mary River. *Whiskers* are used to show extremities of predicted and observed flow properties

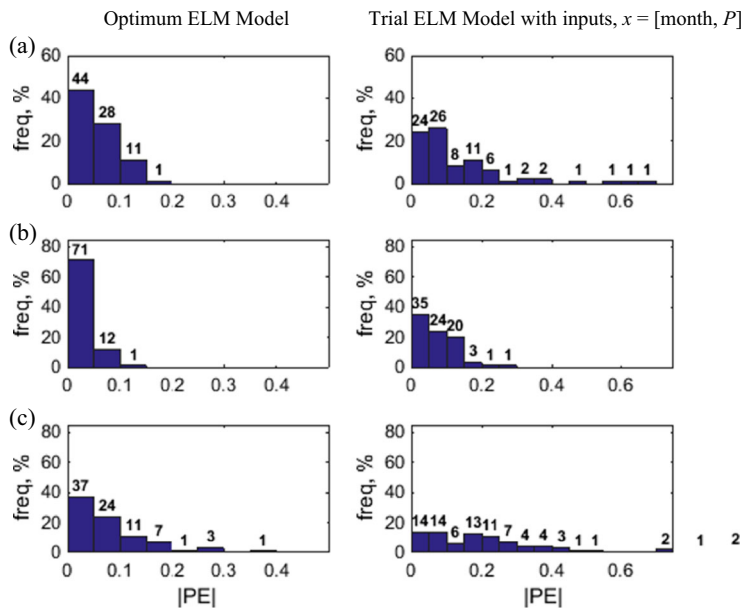


In order to inspect the model simulation skill more closely, the differences in distribution of error statistics in terms of maximum, minimum, lower quartile ( $Q_{25}$ ), median ( $Q_{50}$ ), upper quartile ( $Q_{75}$ ), range, skewness, and flatness of streamflow for the optimum ELM and the optimum ANN models are shown in Table 7. Also shown are the error statistics of the trial ELM and trial ANN models with month and rainfall as the input variables. Again, these statistics confirmed more accurate simulations for the ELM compared to the ANN model for both the optimum and the dual input (rainfall and month) case. In terms of the maximum simulated and observed streamflow, the optimum ELM model yielded a value of 0.16 m whereas the optimum ANN model had 0.57 m for Gowrie Creek. Likewise, for Albert River, the optimum ELM model generated a difference of

0.07 m whereas the ANN model had 0.18 m, and for Mary River, it was 0.39 versus 1.13 m. Similar deduction was made when the minimum, lower quartile, median, upper quartile, and the range of simulated and observed  $Q_{WL}$  were analyzed for the optimum ELM and ANN models, as well as the trial ELM and ANN models. This demonstrated the efficacy of the ELM model for reliable simulation of extreme values of streamflow compared to the ANN model.

Figure 9 shows a histogram of the relative frequency of prediction error (PE) in various error brackets for the optimum ELM compared to the trial ELM model with only month and rainfall ( $P$ ) as the inputs. Also shown are numbers on each bar representing the actual percentage values for the particular error bin. When the results for the optimum ELM model (left) were considered,

**Fig. 9** A histogram of relative frequency of absolute prediction error (PE) (m) in error bracket for the optimum ELM model compared with the trial ELM model with month and rainfall (*P*) as inputs for **a** Gowrie Creek, **b** Albert River, and **c**, **d** Mary River. The numbers on each bar show the actual count of months in test period



almost 44 % of simulations were in  $\pm 0.05$ -m bin width (Gowrie Creek) whereas that for Albert River was 71 % and Mary River was 37 %. This confirmed that the optimum ELM model yielded the most simulation of

streamflow for Albert River, which was also consistent with earlier results (Figs. 5 and 6). It was also noticeable that the next error bin ( $\pm (0.05-0.10)$ m) was only 12 % and ( $\pm (0.10-0.15)$ m) was only 1 % compared to about

**Table 7** Difference in distribution of error statistics in terms of maximum, minimum, lower quartile ( $Q_{25}$ ), median ( $Q_{50}$ ), upper quartile ( $Q_{75}$ ), range, skewness, and flatness of streamflow in the test period

Error statistic	Optimum ELM model			Optimum ANN model		
	Gowrie Creek	Albert River	Mary River	Gowrie Creek	Albert River	Mary River
Maximum	0.16	0.07	0.39	0.57	0.18	1.13
Minimum	-0.14	-0.11	-0.29	-0.47	-0.14	-1.04
$p_{25}$	-0.03	-0.02	-0.06	-0.06	-0.04	-0.20
$p_{50}$	0.02	0.00	-0.02	0.06	-0.01	0.02
$p_{75}$	0.05	0.01	0.05	0.17	0.04	0.16
Range	0.30	0.18	0.69	1.04	0.32	2.17
Skewness	-0.22	-0.13	0.74	-0.03	0.40	0.32
Flatness	-0.21	1.14	2.43	1.20	0.33	1.90
	Trial ELM model (inputs, $x = [\text{month}, P]$ )			Trial ANN model (inputs, $x = [\text{month}, P]$ )		
Maximum	0.67	0.12	1.12	0.72	0.36	1.85
Minimum	-0.17	-0.30	-0.90	-0.38	-0.12	-0.85
$p_{25}$	0.00	-0.10	-0.16	-0.05	0.05	-0.05
$p_{50}$	0.07	-0.03	0.03	0.10	0.15	0.12
$p_{75}$	0.17	0.02	0.21	0.18	0.20	0.30
Range	0.84	0.42	2.02	1.09	0.47	2.70
Skewness	1.48	-0.42	0.50	0.61	-0.21	1.33
Flatness	3.15	0.04	2.42	1.90	-0.17	4.44

All units are in meters (m)

28 and 11 % (Gowrie Creek) and 24 and 11 % (Mary River), respectively.

Based on this analysis, it is axiomatic to conclude that ELM model simulations for Albert River generated the smaller PE when compared with the other two stations. In fact, a closer assessment showed that for the case of Mary River, the optimum ELM model yielded relatively small yet notable error values in larger error bins (cumulative frequency was approximately 5.0 % for error of  $\pm (0.2\text{--}0.4\text{ m})$ ), which concurred with earlier the results (Fig. 7c). It is therefore conclusive that the trial ELM model with only the month and rainfall as input variables was more inaccurate compared to the optimum ELM model (Figs. 8 and 9, right).

For all stations considered, the frequency distribution of model simulation errors for trial models (with unselected inputs and with only rainfall and month as input) was inherently located in larger error bins compared to the optimum ELM model where errors were concentrated in smaller bin. This was also the true for the comparative trial ANN model with nine inputs (not shown here). Therefore, an ostensible deduction is made: that streamflow modeling requires selection of the most relevant input variables by a careful and robust assessment of statistical dependence of input(s) and target (objective) variable (Table 2; Fig. 4). This is necessary for accurate modeling since relationship between inputs and outputs needs to be better identified and irrelevant variables whose features are not useful for the simulation of the objective variable are eliminated (Tiwari and Adamowski 2013). Consequently, proper selection of variables tends to achieve better simulation accuracy compared to models with no selection of variables, as was evident for the case of trial ELM and ANN models developed in this study.

## Summary and conclusion

In the twenty-first century where environmental challenges are exacerbated by the underlying consequences of climate shift, the appraisal of streamflow using efficient models as a useful stratagem for applications in environmental monitoring, adaptive water resources planning, sustainable agriculture, and ecosystem management. In this study, we exemplified the utility of ELM model for simulating streamflow in eastern Queensland and results were validated with the ANN model. In order to develop the ELM model, hydrological data were

utilized for Gowrie Creek (18.44° S; 145.85° E), Albert (28.05° S; 153.05° E), and Mary River (25.95° S; 152.49° E) for the period 1960–2012. The rainfall data were acquired from weather stations, namely Macknade Sugar Mill (18.58° S; 146.25° E), Harrisville Post Office (27.81° S; 152.67° E), and Cowal (25.95° S; 152.50° E) with climate indices (SOI, PDO, IOD, EMI), and SST (Nino 3.4 SST, Nino 3.4 SST, and Nino 4.0 SST). In order to determine significant inputs, cross correlation of streamflow with input series was performed.

Based on the statistically significant correlations ( $r_{\text{cross}}$ ), three sets of models were developed that included a trial ELM model with month (to consider periodicity) and eight input variables, a trial ELM model with only month and rainfall as input, and an optimum ELM model with selection of best inputs for each station. All predictive models were trained using data for the period 1960–2005, tested over 2006–2012, and the model's performance was assessed using statistical measures to verify its ability to simulate the streamflow. The findings are enumerated as follows.

- A cross correlation analysis of observed streamflow with each of the eight inputs revealed a statistically significant dependence on rainfall ( $r_{\text{cross}} \approx 0.306\text{--}0.718$ ), SOI ( $r_{\text{cross}} \approx 0.141\text{--}0.215$ ), and EMI ( $r_{\text{cross}} \approx -0.237$  to  $-0.314$ ) for all stations. However, cross correlation of streamflow with time series of PDO, EMI, IOD, and SSTs was different for different hydrological stations.
- For ELM model, trial and error experiments were performed using activation functions (sigmoid, sine, hard-limit, triangular basis, radial basis, logarithmic sigmoid, and tangent sigmoid functions) and randomly assigned weights and incrementally varying hidden neurons (2 to 150). Therefore, the optimum ELM was designed with hard-limit activation equation and neuronal architecture of 6-106-1 (Gowrie Creek), 6-74-1 (Albert River), and 6-146-1 (Mary River) corresponding to its input–hidden–output arrangements.
- The comparative ANN model was trialed using tangent and logarithmic sigmoid transfer functions with several learning algorithms. This resulted in best ANN model with Levenberg–Marquardt algorithm and a neuronal architecture of 6-28-1 (Gowrie Creek), 6-18-1 (Albert River), and 6-20-1 (Mary River).

- For all stations considered for simulation of mean monthly streamflow, the optimum ELM outperformed the ANN model. For Gowrie Creek, performance metrics yielded a value of  $R^2 \approx 0.964$ ,  $d \approx 0.968$ , and  $E_{NS} \approx 0.963$  (ELM) compared to 0.732, 0.802, and 0.698 (ANN), while for Albert River, a value of  $R^2 \approx 0.957$ ,  $d \approx 0.962$ , and  $E_{NS} \approx 0.955$  (ELM) was obtained compared to 0.830, 0.863, and 0.816 (ANN). Similarly, a value of  $R^2 \approx 0.990$ ,  $d \approx 0.986$ , and  $E_{NS} \approx 0.989$  (ELM) was obtained compared to 0.892, 0.855, and 0.891 (ANN).
- Given correlation of streamflow with rainfall for all stations, trial models using rainfall and month as inputs yielded good performance for the ELM compared to the ANN model. However, simulation accuracy of both models was worse compared to the optimum models with selected input. When all nine inputs were incorporated into the trial ELM model without selection of variables, the performance deteriorated quite significantly.
- An analysis of the relative frequency of simulation errors in different error brackets showed that the trial ELM with month and rainfall as inputs resulted in a higher frequency for the wider error brackets ( $> \pm 0.10$  m) compared to the optimum ELM model that produced the highest frequency of all errors in the larger  $\pm (0-0.05)$  m bracket.

Based on our findings, it is ascertained that the ELM model with selected input variables has a good ability to simulate streamflow water level and its performance is comparatively better than the ANN model. It is therefore advocated that the ELM algorithm is a useful utility for environmental monitoring where predictive modeling of parameters (e.g., streamflow) is required. Finally, the ELM-based streamflow model can be employed to implicate or analyze trends in hydrological parameters (e.g., rainfall), assessing the viability of irrigation systems, river flows, or lakes as well as sustainable use of water for agriculture and developing flood and drought response strategies.

**Acknowledgments** The data were acquired from the Australian Bureau of Meteorology, Joint Institute of the Study of Atmosphere and Ocean (JISAO), and Japanese Agency for Marine-Earth Science and Technology. Dr RC Deo was supported by the Academic Division Researcher Activation Incentive Scheme (RAIS; July–September 2015) grant and the Australian Government Endeavor Executive Fellowship (2015) to Dr R.C. Deo to collaborate with

Dr M Sahin (Turkey). We thank two reviewers and the Editor for their comments that improved the overall clarity of this paper.

## References

- Abbot, J., & Marohasy, J. (2012). Application of artificial neural networks to rainfall forecasting in Queensland. *Australia Advances in Atmospheric Sciences*, 29, 717–730.
- Abbot, J., & Marohasy, J. (2014). Input selection and optimisation for monthly rainfall forecasting in Queensland. *Australia, using artificial neural networks Atmospheric Research*, 138, 166–178. doi:10.1016/j.atmosres.2013.11.002.
- Acharya N, Shrivastava NA, Panigrahi B, Mohanty U (2013) Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: an application of extreme learning machine *Climate Dynamics*:1–8
- Adamowski J, Fung Chan H, Prasher SO, Ozga-Zielinski B, Sliusarieva A (2012) Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada *Water Resources Research* 48
- Adamowski, J. F. (2008). Development of a short-term river flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis. *Journal of Hydrology*, 353, 247–266.
- Asefa, T., Kemblowski, M., McKee, M., & Khalil, A. (2006). Multi-time scale stream flow predictions: the support vector machines approach. *Journal of Hydrology*, 318, 7–16.
- Ashok, K., Guan, Z., Yamagata, T. (2003). Influence of the Indian ocean dipole on the Australian winter rainfall. *Geophysical Research Letters*, 30. doi:10.1029/2003GL017926.
- Ashok, K., Behera, S. K., Rao, S. A., Weng, H., Yamagata, T. (2007). El Niño Modoki and its possible teleconnection. *Journal of Geophysical Research: Oceans*, (1978–2012) 112. doi:10.1029/2006JC003798.
- Belayneh A, Adamowski J (2012) Standard precipitation index drought forecasting using neural networks, wavelet neural networks, and support vector regression *Applied Computational Intelligence and Soft Computing* 2012:6 doi:10.1155/2012/794061.
- Brodie, R. S., Hostetler, S., & Slatter, E. (2008). Comparison of daily percentiles of streamflow and rainfall to investigate stream–aquifer connectivity. *Journal of hydrology*, 349, 56–67.
- Cai W, Cowan T (2009) La Niña Modoki impacts Australia autumn rainfall variability *Geophysical Research Letters* 36
- Chang, F., Chang, L. C., & Huang, H. L. (2002). Real time recurrent learning neural network for stream flow forecasting. *Hydrological Processes*, 16, 2577–2588.
- Chowdhury R, Gardner T, Gardiner R, Chong M, Tonks M, Begbie D, Wakem S Catchment hydrology modelling for stormwater harvesting study in SEQ: from instrumentation to simulation. In: Science Forum, 2010

- Chiew, F. H., & McMahon, T. A. (2002). Modelling the impacts of climate change on Australian streamflow. *Hydrological Processes*, *16*, 1235–1245.
- Chiew, F. H., Piechota, T. C., Dracup, J. A., & McMahon, T. A. (1998). El Niño/Southern Oscillation and Australian rainfall, streamflow and drought: links and potential for forecasting. *Journal of Hydrology*, *204*, 138–149.
- Deo, R. C., & Şahin, M. (2015a). Application of the Artificial Neural Network model for prediction of monthly Standardized Precipitation and Evapotranspiration Index using hydrometeorological parameters and climate indices in eastern. *Australia Atmospheric Research*, *161–162*, 65–81.
- Deo, R. C., & Şahin, M. (2015b). Application of the extreme learning machine algorithm for the prediction of monthly Effective Drought Index in eastern. *Australia Atmospheric Research*, *153*, 512–525. doi:10.1016/j.atmosres.2013.11.002.
- Deo RC, Samui P, Kim D (2015) Estimation of monthly evaporative loss using relevance vector machine, extreme learning machine and multivariate adaptive regression spline models Stochastic Environmental Research and Risk Assessment:1–16
- Deo RC, Syktus J, McAlpine C, Lawrence P, McGowan H, Phinn SR (2009) Impact of historical land cover change on daily indices of climate extremes including droughts in eastern Australia Geophysical Research Letters *36*
- Dettinger, M. D., & Diaz, H. F. (2000). Global characteristics of stream flow seasonality and variability. *Journal of Hydrometeorology*, *1*, 289–310.
- Dettinger, M. D., Cayan, D. R., McCabe, G. J., Marengo, J. A. (2000). Multiscale streamflow variability associated with El Niño/Southern oscillation. Cambridge: Cambridge University Press.
- DNRM (2014) Establishing a new water monitoring site (WM65). version 1.0. Brisbane Qld: State of Queensland (Department of Natural Resources and Mines), Service Delivery.
- Drosowsky, W. (1993). An analysis of Australian seasonal rainfall anomalies: 1950–1987. II: temporal variability and teleconnection patterns. *International Journal of Climatology*, *13*, 111–149.
- Fox, D. G. (1981). Judging air quality model performance. *Bulletin of the American Meteorological Society*, *62*, 599–609.
- Haykin, S. (2010). *Neural networks: a comprehensive foundation*, 1994 Mc Millan. New: Jersey.
- Haylock, M., & Nicholls, N. (2000). Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910–1998. *International Journal of Climatology*, *20*, 1533–1541.
- Hennessy K et al. (2007) Australia and New Zealand in ML Parry, OF Canziana, JP Palitkof, PJ van der Linder, and CE Hanson, editors. Climate change 2007: impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge
- Holland JH (1975) Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. U Michigan Press
- Huang, G.-B., & Chen, L. (2007). Convex incremental extreme learning machine. *Neurocomputing*, *70*, 3056–3062.
- Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2006). Extreme learning machine: theory and applications. *Neurocomputing*, *70*, 489–501.
- Huang, G., Huang, G.-B., Song, S., & You, K. (2015). Trends in extreme learning machines. *A review Neural Networks*, *61*, 32–48.
- IPCC (2001) The scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change In: Houghton JT, Ding, Y., Griggs, D. J., Noguer, M., Van der Linden, P. J., Dai, X., Maskell, K. and Johnson, C. A. (Eds.) (ed). Cambridge University Press, Cambridge and New York
- Kiem, A. S., & Franks, S. W. (2001). On the identification of ENSO-induced rainfall and runoff variability. *A Comparison of Methods and Indices Hydrological Sciences Journal*, *46*, 715–727.
- Kiem, A. S., & Franks, S. W. (2004). Multi-decadal variability of drought risk, eastern Australia. *Hydrological Processes*, *18*, 2039–2050.
- Kiem AS, Franks SW, Kuczera G (2003) Multi-decadal variability of flood risk Geophysical Research Letters *30*
- Koza JR (1992) Genetic programming: on the programming of computers by means of natural selection vol 1. MIT press
- Krause, P., Boyle, D., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, *5*, 89–97.
- Legates, D. R., & McCabe, G. J. (1999). Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, *35*, 233–241.
- Leu S-S, Adi TJW (2011) Probabilistic prediction of tunnel geology using a Hybrid Neural-HMM Engineering Applications of Artificial Intelligence *24*:658–665
- Lippman R (1987) An introduction to computing with neural nets IEEE ASSP Magazine *4*:4–22
- Maier HR, Dandy GC (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications Environmental modelling & software *15*:101–124
- Maier, H. R., Jain, A., Dandy, G. C., & Sudheer, K. P. (2010). Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions. *Environmental Modelling & Software*, *25*, 891–909.
- Mantua NJ, Hare SR, Zhang Y, Wallace JM, Francis RC (1997) A Pacific interdecadal climate oscillation with impacts on salmon production Bulletin of the American Meteorological Society *78*:1069–1079
- Masinde M (2013) Artificial neural networks models for predicting effective drought index: factoring effects of rainfall variability Mitigation and Adaptation Strategies for Global Change:1–24
- McAlpine C, Syktus J, Deo R, Lawrence P, McGowan H, Watterson I, Phinn S (2007) Modeling the impact of historical land cover change on Australia’s regional climate Geophysical Research Letters *34*
- McAlpine, C., Syktus, J., Ryan, J., Deo, R., McKeon, G., McGowan, H., & Phinn, S. (2009). A continent under stress: interactions, feedbacks and risks associated with impact of modified land cover on Australia’s climate. *Global Change Biology*, *15*, 2206–2223.

- McBride JL, Nicholls N (1983) Seasonal relationships between Australian rainfall and the Southern Oscillation Monthly Weather Review 111:1998–2004
- McMahon TA, Finlayson B, Haines A, Srikanthan R (1992) Global runoff: continental comparisons of annual flows and peak discharges. Catena Verlag
- Mekanik, F., Imteaz, M., Gato-Trinidad, S., & Elmahdi, A. (2013). Multiple regression and artificial neural network for long-term rainfall forecasting using large scale climate modes. *Journal of Hydrology*, 503, 11–21.
- Morid, S., Smakhtin, V., & Bagherzadeh, K. (2007). Drought forecasting using artificial neural networks and time series of drought indices. *International Journal of Climatology*, 27, 2103–2111.
- Nash, J., & Sutcliffe, J. (1970). River flow forecasting through conceptual models part I—a discussion of principles. *Journal of Hydrology*, 10, 282–290.
- Nastos, P., Paliatatos, A., Koukouletsos, K., Larissi, I., & Moustiris, K. (2014). Artificial neural networks modeling for forecasting the maximum daily total precipitation at Athens. *Greece Atmospheric Research*, 144, 141–150.
- Ni, Q., Wang, L., Ye, R., Yang, F., & Sivakumar, M. (2010). Evolutionary modeling for streamflow forecasting with minimal datasets: a case study in the West Malian River. *China Environmental Engineering Science*, 27, 377–385.
- Nicholls, N., Drosowsky, W., & Lavery, B. (1997). Australian rainfall variability and change. *Weather*, 52, 66–72.
- Ortiz-García, E., Salcedo-Sanz, S., & Casanova-Mateo, C. (2014). Accurate precipitation prediction with support vector classifiers: a study including novel predictive variables and observational data. *Atmospheric Research*, 139, 128–136.
- Ortiz-García, E., Salcedo-Sanz, S., Casanova-Mateo, C., Paniagua-Tineo, A., & Portilla-Figueras, J. (2012). Accurate local very short-term temperature prediction based on synoptic situation Support Vector Regression banks. *Atmospheric Research*, 107, 1–8.
- Ouyang, R., Liu, W., Fu, G., Liu, C., Hu, L., & Wang, H. (2014). Linkages between ENSO/PDO signals and precipitation, streamflow in China during the last 100 years. *Hydrology and Earth System Sciences*, 18, 3651–3661.
- Patterson DW (1998) Artificial neural networks: theory and applications. Prentice Hall PTR
- Paulescu, M., Tulcan-Paulescu, E., & Stefu, N. (2011). A temperature based model for global solar irradiance and its application to estimate daily irradiation values. *International Journal of Energy Research*, 35, 520–529.
- Piechota, T. C., Chiew, F. H., Dracup, J. A., & McMahon, T. A. (2001). Development of exceedance probability streamflow forecast. *Journal of Hydrologic Engineering*, 6, 20–28.
- Power, S., Casey, T., Folland, C., Colman, A., & Mehta, V. (1999). Inter-decadal modulation of the impact of ENSO on Australia. *Climate Dynamics*, 15, 319–324.
- Rajesh, R., & Prakash, J. S. (2011). Extreme learning machines—a review and state-of-the-art. *International Journal of Wisdom Based Computing*, 1, 35–49.
- Risbey, J. S., Pook, M. J., McIntosh, P. C., Wheeler, M. C., & Hendon, H. H. (2009). On the remote drivers of rainfall variability in Australia. *Monthly Weather Review*, 137, 3233–3253.
- Robertson D, Wang Q (2008) An investigation into the selection of predictors and skill assessment using the Bayesian joint probability (BJP) modelling approach to seasonal forecasting of streamflows Water for a Healthy Country flagship report, CSIRO Land and Water, Canberra
- Şahin, M. (2012). Modelling of air temperature using remote sensing and artificial neural network in Turkey. *Advances in Space Research*, 50, 973–985.
- Şahin, M., Kaya, Y., & Uyar, M. (2013). Comparison of ANN and MLR models for estimating solar radiation in turkey using NOAA/AVHRR data. *Advances in Space Research*, 51, 891–904.
- Şahin, M., Kaya, Y., Uyar, M., & Yıldırım, S. (2014). Application of extreme learning machine for estimating solar radiation from satellite data. *International Journal of Energy Research*, 38, 205–212.
- Saji, N., Goswami, B. N., Vinayachandran, P., & Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature*, 401, 360–363.
- Saji, N., & Yamagata, T. (2003). Possible impacts of Indian ocean dipole mode events on global climate. *Climate Research*, 25, 151–169.
- Salcedo-Sanz S, Deo RC, Carro-Calvo L, Saavedra-Moreno B (2015) Monthly prediction of air temperature in Australia and New Zealand with machine learning algorithms Theoretical and Applied Climatology: DOI: 10.1007/s00704-00015-01480-00704 doi:10.1007/s00704-015-1480-4
- Salcedo-Sanz, S., Pastor-Sánchez, A., Prieto, L., Blanco-Aguilera, A., & García-Herrera, R. (2014). Feature selection in wind speed prediction systems based on a hybrid coral reefs optimization—extreme learning machine approach. *Energy Conversion and Management*, 87, 10–18.
- Samui, P., & Dixon, B. (2012). Application of support vector machine and relevance vector machine to determine evaporative losses in reservoirs. *Hydrological Processes*, 26, 1361–1369.
- Sánchez-Monedero, J., Salcedo-Sanz, S., Gutiérrez, P., Casanova-Mateo, C., & Hervás-Martínez, C. (2014). Simultaneous modelling of rainfall occurrence and amount using a hierarchical nominal–ordinal support vector classifier. *Engineering Applications of Artificial Intelligence*, 34, 199–207.
- Shukla, R. P., Tripathi, K. C., Pandey, A. C., & Das, I. (2011). Prediction of Indian summer monsoon rainfall using Niño indices: a neural network approach. *Atmospheric Research*, 102, 99–109.
- Simpson, H., Cane, M., Herczeg, A., Zebiak, S., & Simpson, J. (1993). Annual river discharge in southeastern Australia related to El Niño-Southern Oscillation forecasts of sea surface temperatures. *Water Resources Research*, 29, 3671–3680.
- Taschetto, A. S., & England, M. H. (2009). El Niño Modoki impacts on Australian rainfall. *Journal of Climate*, 22, 3167–3174.
- Tiwari, M. K., & Adamowski, J. (2013). Urban water demand forecasting and uncertainty assessment using ensemble wavelet-bootstrap-neural network models. *Water Resources Research*, 49, 6486–6507.
- Tran H, Muttill N, Perera B Investigation of artificial neural network models for streamflow forecasting. In: 19th International Congress on Modelling and Simulation (MODSIM2011), 2011. Modelling and Simulation Society of Australia and New Zealand Inc.(MSSANZ), pp 1099–1105

- Trenberth, K. E. (1984). Signal versus noise in the Southern Oscillation. *Monthly Weather Review*, *112*, 326–332.
- Ulgun, K., & Hepbasli, A. (2002). Comparison of solar radiation correlations for Izmir. *Turkey International Journal of Energy Research*, *26*, 413–430.
- Verdon DC, Wyatt AM, Kiem AS, Franks SW (2004) Multidecadal variability of rainfall and streamflow: Eastern Australia Water Resources Research 40
- Vogl, T., Mangis, J., Rigler, A., Zink, W., & Alkon, D. (1988). Accelerating the convergence of the backpropagation method. *Biological Cybernetics*, *59*, 257–263.
- Wang E, Zhang Y, Luo J, Chiew F, Wang Q (2011) Monthly and seasonal streamflow forecasts using rainfall runoff modeling and historical weather data Water Resources Research 47
- Wang Q, Robertson D (2011) Multisite probabilistic forecasting of seasonal flows for streams with zero value occurrences Water Resources Research 47
- Wang Q, Robertson D, Chiew F (2009) A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites Water Resources Research 45
- Weng, H., Ashok, K., Behera, S. K., Rao, S. A., Yamagata, T. (2007). Impacts of recent El Niño Modoki on dry/wet conditions in the Pacific rim during boreal summer. *Climate Dynamics*, *29*, 113–129.
- Willmott, C. J. (1981). On the validation of models. *Physical Geography*, *2*, 184–194.
- Willmott, C. J. (1982). Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society*, *63*, 1309–1313.
- Willmott CJ (1984) On the evaluation of model performance in physical geography. In: Spatial statistics and models. Springer, pp 443–460
- Wu, C., & Chau, K. (2010). Data-driven models for monthly streamflow time series prediction. *Engineering Applications of Artificial Intelligence*, *23*, 1350–1367.
- Zhang Y, Wallace JM, Battisti DS (1997) ENSO-like interdecadal variability: 1900–93 *Journal of climate* 10:1004–1020
- Zubair, L., & Chandimala, J. (2006). Epochal changes in ENSO-streamflow relationships in Sri Lanka. *Journal of hydrometeorology*, *7*, 1237–1246.