Water Resources Management (2020) 34:4113–4131
https://doi.org/10.1007/s11269-020-02659-5

Employing Machine Learning Algorithms for Streamflow Prediction: A Case Study of Four River Basins with Different Climatic Zones in the United States

Peiman Parisouj[1] · Hamid Mohebzadeh[1] · Taesam Lee[1]

Received: 25 March 2020 / Accepted: 31 August 2020 /
Published online: 11 September 2020

Abstract
Streamflow estimation plays a significant role in water resources management, especially for flood mitigation, drought warning, and reservoir operation. Hence, the current study examines the prediction capability of three well-known machine learning algorithms (Support Vector Regression (SVR), Artificial Neural Network with backpropagation (ANN-BP), and Extreme Learning Machine (ELM)) for the monthly and daily streamflows of four rivers in the United States. For model development, three main predictor variables (P, T_{max}, and T_{min}) and their antecedent values were considered. The SVM-RFE feature selection method was used to select the most appropriate predictor variable.The performance of the developed models was tested using four evaluation statistics. The results indicate that (1) except some improvements, the accuracy of all models decreases at the daily scale compared to that at the monthly scale; (2) the SVR has the best performance among the three models at the monthly and daily scales, while the ANN-BP model has the worse performance; (3) the ELM has better generalization performance than the ANN-BP for streamflow simulation at the monthly and daily scales; and (4) all models fail to predict the streamflow for the Carson River as a snowmelt-dominated basin. Generally, findings of the current study indicate that the SVR model produces better results than the ELM and ANN-BP for streamflow simulation at the monthly and daily scales.

Keywords Streamflow prediction · Support vector regression · Artificial neural networks · Extreme learning machine

✉ Hamid Mohebzadeh
hamidmohebzadeh@gnu.ac.kr

✉ Taesam Lee
tae3lee@gnu.ac.kr

[1] Department of Civil Engineering, ERI, Gyeongsang National University, 501 Jinju-daero, Jinju, Gyeongnam 52828, South Korea

1 Introduction

The accurate prediction of a streamflow is challenging task due to the complex nonlinearity and stochastic characteristics of hydrological processes such as precipitation, temperature, evapotranspiration, and characteristics of the watershed (Adnan et al. 2019; Meng et al. 2019; Niu et al. 2019). However, due to their great importance for optimal management of water resources, the monthly and daily streamflow estimations have received considerable attention in the past decades (Wang et al. 2019; Adnan et al. 2019; Hadi and Tombul 2018). Streamflow prediction methods can be grouped into two categories (Wang 2006): physically based models and data-driven models. For example, conceptually hydrological models are considered a type of physically based models (Cheng et al. 2006); artificial neural network and support vector machine are categorized as data-driven models (Belayneh et al. 2014; Wu et al. 2010).

Physically based models require specific assumptions and large amount of hydrological data for calibration. However, the lack of accurate hydrological data such as the rainfall amount, intensity, and distribution may limit these models in practical applications (Noori and Kalin 2016; Peng et al. 2017; Ragettli et al. 2014). Data-driven models with minimum data requirement and strong universality can quickly model the linear and nonlinear relationships between a streamflow and its parameters and yield better performance than physically based methods (Karran et al. 2013; Liu et al. 2017). Therefore, data-driven models have been widely used in streamflow forecasting applications (Yang et al. 2017; Wang et al. 2019).

As Patel and Ramachandran (2015) indicated, since the mid-1990s, applying machine learning techniques to model the river water discharge has been far-reaching. Artificial neural networks (ANNs) have been developed for various hydrological applications with very successful results (Campolo et al. 2003; Hu et al. 2001; Patel and Ramachandran 2015; Wang et al. 2009; Wu et al. 2005; Zealand et al. 1999). Artificial neural networks (ANNs) are the most widely used nonlinear models for streamflow simulation (Yoon et al. 2007; Guo et al. 2011). However, the ANN results are not suitable for complex hydrological processes due to their drawbacks such as slow learning, finding the local minima instead of global minima, and overfitting (Meng et al. 2019).

Cortes and Vapnik (1995) proposed the support vector machine (SVM) based on Vapnik-Chervonenkis (Dibike et al.) dimension theory and the structural risk minimization principle. The SVM can overcome partial convergence and find the global optimal solutions (Meng et al. 2019). Several studies mentioned that the SVR could achieve much better simulation accuracy than ANNs in hydrological applications (Maity et al. 2010; Wang et al. 2009; Lafdani et al. 2013; Liu and Lu 2014). P.-S. Yu et al. (2006) applied SVR for real-time flood stage forecast; M. Liu and Lu (2014) showed that SVR could help the water quality forecast in an agricultural nonpoint source polluted river.

The extreme learning machine method proposed by Huang (2014) is a new learning method. ELM effectively enhances the performance of ANNs due to some capabilities. First, ELM does not require iteration for weight adjustment. Second, ELM can sharply decrease the learning speed in addition to the feasible generalization ability (Niu et al. 2019; Wang et al. 2019). Several studies have reported that ELM can outperform ANNs in different hydrological applications. Şahin et al. (2014) used ELM to estimate solar radiation in Turkey and found that ELM was approximately 26.5 times faster than ANN. Deo and Şahin (2015) predicted the monthly effective drought index using ELM and ANN and concluded that ELM was more efficient than the ANN model in drought prediction. Yaseen et al. (2015) investigated the

applicability of ELM for streamflow forecast and compared ELM results with ANN results. They concluded that ELM yielded better accuracy than the ANN model.

Based on the literature, although several studies reported the superiority of both SVR and ELM over the ANN in various hydrological applications, few studies compared SVR, ANN, and ELM in terms of simulation ability at the monthly and daily scales to select the best machine learning model. Therefore, the objective of the current study is to select the most suitable method among SVR, ANN, and ELM to simulate the streamflow at the monthly and daily scales.

2 Mathematical Background

2.1 Support Vector Regression (SVR)

Support vector regression (SVR) is used in support vector machine (SVM) for regression problems. As reported by Lin et al. (2006), this method combines structural risk reduction to reach the empirical risk reduction, which is generally used by statistical learning methods such as artificial neural networks (ANNs). The structural risk reduction method minimizes the empirical risk to obtain a good generalization capacity by minimizing the error of the generalization in contrast to the training error (Belayneh et al. 2014; Maity et al. 2010).

The fundamental hypothesis of SVR is the nonlinearly mapping of the primary data into a higher-dimensional feature space. The kernel is the function to perform linear regression in the feature space (Kalra et al. 2013; Maity et al. 2010; Yoon et al. 2011; Yu et al. 2017). Although several kernels such as polynomial, linear and sigmoid can be used in SVR, the radial basis function (RBF) performs better than other kernels (Barzegar et al. 2017; Lin et al. 2006; Liong and Sivapragasam 2002; Wu and Wang 2009; Yu and Liong 2007). Moreover, as illustrated by Dibike et al. (2001), the RBF kernel outperformed other kernels in the simulation of rainfall runoff. Thus, we used the RBF kernel in the current study.

2.2 Artificial Neural Network with Backpropagation (ANN-BP)

An ANN is built with nodes that are arranged in groups called layers (Sudheer et al. 2002). It does not require that the information about the complex nature of the underlying process under consideration is explicitly described in mathematical form. The structure of a typical three-layer feed-forward ANN is shown in Fig. 1. In Fig. 1, the structure includes one input, one hidden layer with n nodes, and one output layer, which is mathematically shown as follows:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{1}$$

where y is the output, f is the transfer function, x_i is the input variable, w_i is the weight, and b is the bias.

The core of neural network training is *back propagation* (Govindaraju 2000; Hydrology 2000)*, which calibrates the weights of a neural network dep*ending on the error rate in the previous iteration. Fixing the weights controls lower error rates and produces a valid model by increasing its generalization. Among many neural network architectures, the feed-forward neural network with back propagation has been widely used in hydrological modeling (Govindaraju 2000). Therefore, in the current study, a three-layer neural network with back propagation was applied for streamflow simulation.

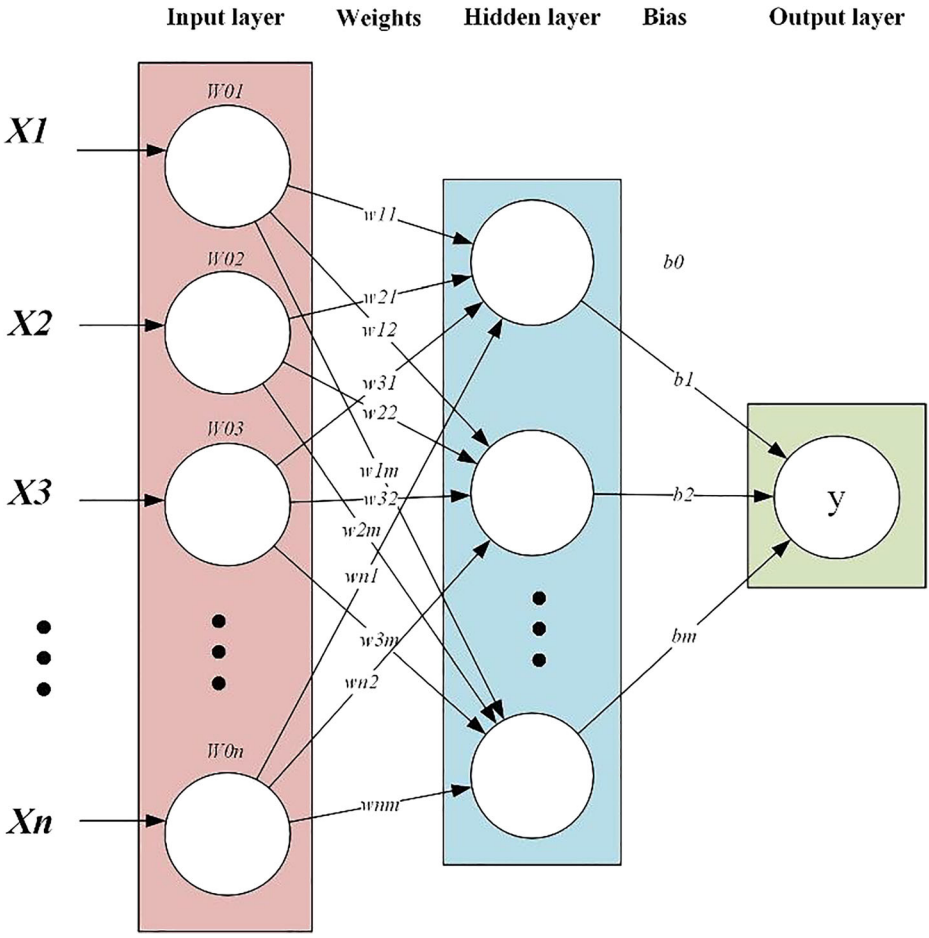

Fig. 1 Structure of the three-layer neural network with back propagation

2.3 Extreme Learning Machine (ELM)

Extreme Learning Machine (ELM) is a developed model of single-hidden-layer feed-forward neural networks (SLFNs) (Huang et al. 2004; Zhu et al. 2005). In ELM, hidden biases and input-hidden weights are randomly selected instead of being tuned; and the hidden-output weights can be directly determined by calculating Moore-Penrose generalized inverse of the hidden-output matrix. This computation procedure makes the ELM faster than the traditional ANN while satisfying the generalization ability.

The main procedure of ELM can be divided into four steps: (1) determining the hidden nodes values and activation function of each neuron; (2) randomly generating all input-hidden weights and hidden biases; (3) computing the hidden layer outputs based on all input variables; (4) directly determining the hidden-output weights by calculating the Moore-Penrose generalized inverse of the hidden-output matrix and predicting the data corresponding to the specified input variables. More information about the ELM algorithm can be found in Huang et al. (2006).

3 Study Area and Data Description

Four case studies were selected in this study: North Fork River, Chehalis River, Carson River, and Sacramento River basins based on the different sizes, distinct topographies, and hydroclimatic regimes in western U.S. The locations of all basins are shown in Fig. 2.

The North Fork American River watershed with a drainage area of 886 km^2 is the longest tributary of the American River in northern California (Kang and Lee 2014). Two studies (Dettinger et al. 2004; Jeton et al. 1996) showed that the effects of topographic relief of the mountains are the main cause of spatial variations of precipitation in North Fork watershed. For example, this effect causes the mean annual precipitations of approximately 1651 mm and 813 mm in Blue Canyon and Auburn at 1676-m and 393-m altitude, respectively. Since the North Fork River flows in a warm basin with low elevation, two-thirds of its streamflow originate from the wintertime rainfall and snowmelt runoff, and less than one-third originate

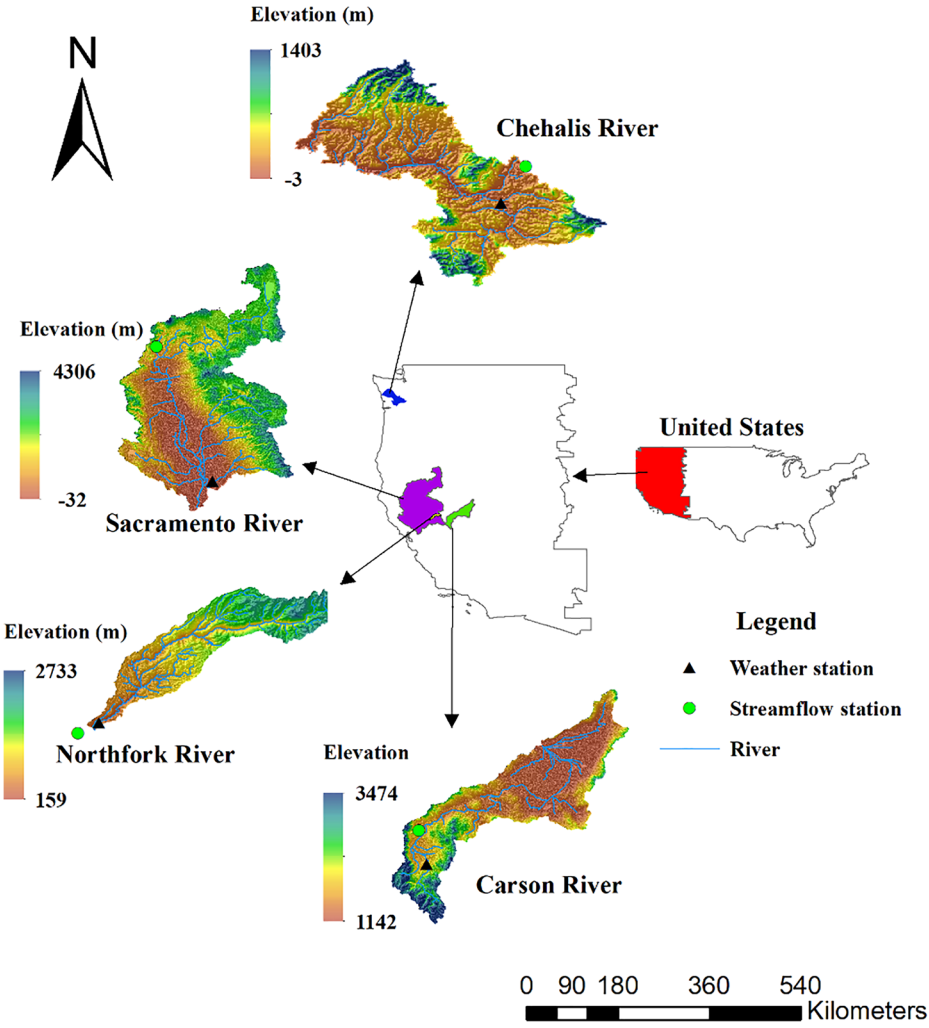

Fig. 2 Locations of the selected river basins; streamflow and weather stations

from the springtime snowmelt runoff (Dettinger et al. 2004). According to the time series of 1950–2019, the maximum daily streamflow was 1418 m^3/s on 2 February 1997, and the average daily streamflow was 23.7 m^3/s.

The Chehalis River basin with a drainage area of 6900 km^2 is the largest river basin in Washington state. The basin can be classified as rain-fed with very little winter precipitation as snow (Mauger et al. 2016). Therefore, seasonal cycles in streamflow are primarily supported by seasonal variations in precipitation; the mean annual precipitation levels were 1520–2030 mm (Henning et al. 2007). According to the time series of 1945–2019, the peak flood discharge was 1817 m^3/s on 2 September 1996. High flows in the Chehalis River basin are directly affected by intense rainfall from November to March, and low flows occur in late spring and summer months (Kimbrough et al. 2006).

The East Fork of the Carson River basin is located in Nevada and California in western United States. It originates in the California Sierra Nevada and passes through the semi-arid area of western Nevada in the northeastward direction (Grantz et al. 2007). The basin has a drainage area of 920 km^2 and an elevation range of approximately 1600–3000 m (Hay et al. 2000). The Carson river is snowmelt-dominated, and the accumulated snow in the mountains is melted by warm temperatures in the spring and supply the river flow. The ratio between potential evaporation and precipitation is approximately 12:1, and the bulk of the water supply to the Carson River accumulates over only 4 months (April–July) (Grantz et al. 2007). According to the time series of 1980–2019, the maximum daily streamflow was 481 m^3/s on 2 January 1997, and the average daily streamflow was 10.73 m^3/s.

Sacramento River with length of 719 km is the largest river of Northern California and the second largest river in the United States, which drains into the Pacific Ocean. Shasta lake located 32 km downstream of the Sacramento River is the largest storage reservoir in California. The reservoir supplies water for agricultural, municipal, and industrial use and fishery of both cold- and warm-water species (Sapin et al. 2017). The watershed has a Mediterranean-climate region with hot, dry summer and wet, cool winters (Bonada and Resh 2013). The monthly temperature average is 4.5–22.4 °C, and the mean annual precipitation is 100–2000 mm (Benke and Cushing 2011). According to the time series of 1944–2019, the maximum daily streamflow was 1525 m^3/s on 16 January 1974, and the average daily streamflow was 33 m^3/s.

In the current study, the daily streamflow data set from 1 January 1990 to 31 July 2019 was acquired from the U.S Geological Survey for four measurement stations at each selected basin. Three weather records of the daily precipitation (P), daily maximum (T_{max}), and minimum temperature (T_{min}) were acquired from four nearby weather stations for the same period of streamflow data. Weather records of all weather stations are available at the NCDC (National Climatic Data Center) website, which was created by the National Oceanic and Atmospheric Administration (NOAA). The location and information of the streamflow and weather stations are shown in Fig. 2 and Table 1, respectively.

4 Application Methodology

4.1 Model Development and Inputs

This study applied three machine learning methods (SVR, ANN-BP, and ELM) to simulate the streamflow at the monthly and daily scales in the four selected basins. To present the

Table 1 Detailed information of the streamflow and weather data stations

Basin	Streamflow Station			Weather Station		
	Code	Latitude	Longitude	Code	Latitude	Longitude
North Fork River	11,427,000	38.93	−121.02	USC00040383	38.90	−121.08
Chehalis River	12,027,500	46.77	−123.03	USW00024227	46.97	−122.90
Carson River	10,309,000	38.84	−119.70	USC00261485	39.12	−119.76
Sacramento River	11,342,000	40.93	−122.41	USW00023271	38.55	−121.41

catchment characteristics, antecedent values of input variables (P, T_{max}, and T_{min}) were used. Therefore, in addition to the input variables, a 1- to 10-day lag of all input variables was used, and the most significant combination of the variables was selected using the Support Vector Machine-Recursive Feature Elimination (SVM-RFE) method (Lin et al. 2006; Luo et al. 2019). The SVM-RFE method was applied for each basin to seperately find the best input variables in each basin. Eqs. (2)–(5) show the selected input structure for the North Fork, Chehalis, Carson, and Sacramento river basins. The number of input variables for each basin is shown in Table 2.

$$Q_t = f(P, \mathrm{Tmax}, \mathrm{Tmin}, \mathrm{P}_{t-1}, \mathrm{P}_{t-2}, \mathrm{P}_{t-7}, \mathrm{P}_{t-8}, \mathrm{P}_{t-9}, \mathrm{P}_{t-10}, \mathrm{Tmax}_{t-1}, \mathrm{Tmax}_{t-2}, \mathrm{Tmax}_{t-3}, \mathrm{Tmax}_{t-9}, \mathrm{Tmax}_{t-10}, \mathrm{Tmin}_{t-9}, \mathrm{Tmin}_{t-10}) \tag{2}$$

$$Q_t = f(P, \mathrm{Tmax}, \mathrm{P}_{t-1}, \ldots, \mathrm{P}_{t-10}, \mathrm{Tmax}_{t-1}, \mathrm{Tmax}_{t-9}, \mathrm{Tmax}_{t-10}, \mathrm{Tmin}_{t-9}, \mathrm{Tmin}_{t-10}) \tag{3}$$

$$Q_t = f(\mathrm{Tmax}, \mathrm{Tmin}, \mathrm{P}_{t-1}, \mathrm{P}_{t-3}, \mathrm{P}_{t-8}, \mathrm{Tmax}_{t-1}, \mathrm{Tmax}_{t-4}, \mathrm{Tmax}_{t-5}, \mathrm{Tmax}_{t-7}, \mathrm{Tmax}_{t-8}, \mathrm{Tmax}_{t-10}\ \mathrm{Tmin}_{t-1}, \mathrm{Tmin}_{t-2}, \mathrm{Tmin}_{t-5}, \mathrm{Tmin}_{t-10}) \tag{4}$$

$$Q_t = f(\mathrm{Tmax}, \mathrm{Tmin}, \mathrm{P}_{t-2}, \mathrm{P}_{t-4}, \mathrm{P}_{t-5}, \mathrm{P}_{t-8}, \mathrm{P}_{t-9}, \mathrm{P}_{t-10}, \mathrm{Tmax}_{t-1}, \mathrm{Tmax}_{t-8}, \mathrm{Tmax}_{t-9}, \mathrm{Tmax}_{t-10}, \mathrm{Tmin}_{t-1}, \mathrm{Tmin}_{t-2}, \mathrm{Tmin}_{t-9}, \mathrm{Tmin}_{t-10}) \tag{5}$$

The mean monthly and daily values of the main input variables (P, T_{max}, and T_{min}) and their selected antecedent values were imported as the model inputs (predictor variable), and the streamflow data set was the output (target variable). The data of 1990–2009 were employed to train the models, and those of 2010–2019 were employed for validation. To match the consistency of the models, normalization was applied to all input variables using the mean (μ) and standard deviation (σ) of the train data set as described in Eq. (6). The mean and standard deviation of the train data set was used to normalize the validation data to prevent using any values computed on the validation data in the model development.

$$\text{normalized } x = \frac{x-\mu}{\sigma} \tag{6}$$

Table 2 Optimized parameters of each ML model for four catchments;the tolerance threshold (ε), structural parameter (γ), and penalty coefficients (C) are SVR-RBF kernel coefficients

	Station	Input variables	SVR			ANN-BP	ELM
			ε	γ	C	Nodes	Nodes
Monthly scale	North Fork River	16	0.556	0.0011	3000	7	11
	Chehalis River	16	0.223	0.0564	334.22	1	30
	Carson River	15	0.889	0.0564	334.22	6	8
	Sacramento River	16	1	0.0564	334.22	1	15
Daily scale	North Fork River	16	0.001	0.0010	3000	7	89
	Chehalis River	16	0.223	0.0564	334.22	1	85
	Carson River	15	0.001	0.0564	334.22	5	22
	Sacramento River	16	1	0.0564	334.22	1	94

4.2 Model Parameterization

In the current study, the optimal parameters of SVR, ANN-BP, and ELM were determined using 10-fold cross validation by minimizing the Root Mean Square Error (RMSE) function for the training period. Then, the selected optimal parameters of each model were used to predict the streamflow for the validation period.

In SVR, the RBF kernel function, which is the most widely used kernel, was selected and its three main parameters were optimized: structural parameter (γ), penalty coefficient (C), and tolerance threshold or ε -precision. The best choice of set values for parameters C and ε is unknown in most cases and should be experimentally specified (Patel and Ramachandran 2015). For this purpose, different values of γ), (10 values were selected from 0.0001 to 0.5), C (10 values were selected from 1 to 3000), and ε -precision (10 values were selected from 0.001 to 1) in the range of the previous studies (Zhang et al. 2018) were evaluated, and best values were selected based on the lowest RMSE.

With respect to the ANN-BP, a three-layer feed-forward ANN model with the Limited-memory Broyden-Fletcher-Goldfarb-Shanno backpropagation algorithm (LBFGS) was used. The logistic sigmoid function was used as the activation function between the input layer and the hidden layer, and a linear transfer function was adopted to connect the hidden layer to the output layer. The number of neurons in the hidden layer is an important parameter of the ANN-BP, which should be determined. Thus, several numbers of hidden neurons (1–10) were tested, and the best number of hidden neurons was selected using a 10-fold cross validation on the training data set (Genç and Dağ 2016; Rezaie-Balf et al. 2017). The random initialization of the weights in ANNs can result in different outputs of the networks for identical numbers of neurons. To find the best weights, 10,000 ANNs were trained in the training period with the selected number of hidden neurons, and the best weights that minimized the objective function (RMSE) of the validation period were maintained.

In contrast with traditional learning algorithms in single-hidden-layer feedforward networks (SLFNs), ELM aims to generate the smallest training error and smallest norm of output weights (Huang et al. 2011). In this algorithm, random computation is used to generate the weights and biases of the hidden layer and obtain the highest accuracy of the model; the number of hidden neurons must be tuned by the user. Therefore, a three-layer ELM model with a sigmoid activation function was developed; similar to SVR and ANN-BP, 10-fold cross validation was adopted to find the optimal number of hidden neurons (Atiquzzaman and Kandasamy 2016).

Table 3 Performance indices computed from the models for training and validation set at the monthly scale

Station	ML Method	Training				Validation			
		RMSE	R	NSE	MARE	RMSE	R	NSE	MARE
North Fork River	SVR	20.00	0.80	0.60	159.24	**31.96**	**0.59**	**0.32**	219.08
	ANN-BP	**8.23**	**0.96**	**0.93**	**84.89**	37.93	0.45	0.03	241.81
	ELM	15.75	0.87	0.75	185.50	36.41	0.50	0.11	**206.22**
Chehalis River	SVR	**17.16**	**0.98**	**0.96**	**20.26**	29.21	**0.94**	0.87	35.14
	ANN-BP	20.35	0.97	0.94	30.68	**27.77**	**0.94**	**0.89**	**33.97**
	ELM	17.66	**0.98**	0.95	32.55	32.30	0.92	0.85	45.64
Carson River	SVR	7.29	0.84	0.68	**44.17**	**9.15**	**0.66**	**0.16**	**111.69**
	ANN-BP	**4.17**	**0.95**	**0.90**	56.16	17.08	0.46	−1.92	213.11
	ELM	6.30	0.87	0.76	94.24	12.28	0.57	−0.50	181.15
Sacramento River	SVR	**19.98**	**0.86**	**0.74**	**42.56**	22.43	0.81	0.63	**75.21**
	ANN-BP	24.20	0.78	0.61	83.15	24.13	0.76	0.57	93.23
	ELM	21.47	0.83	0.70	80.88	**20.66**	**0.84**	**0.69**	100.64

The bold values indicate the best statistics among all models

4.3 Model Evaluation Statistics

To quantify the accuracy of the ML techniques during the training and validation periods, four commonly used statistical indices were calculated: root mean square error (RMSE), coefficient of correlation (R), Nash-Sutcliffe efficiency (NSE), and mean absolute relative error (MARE).

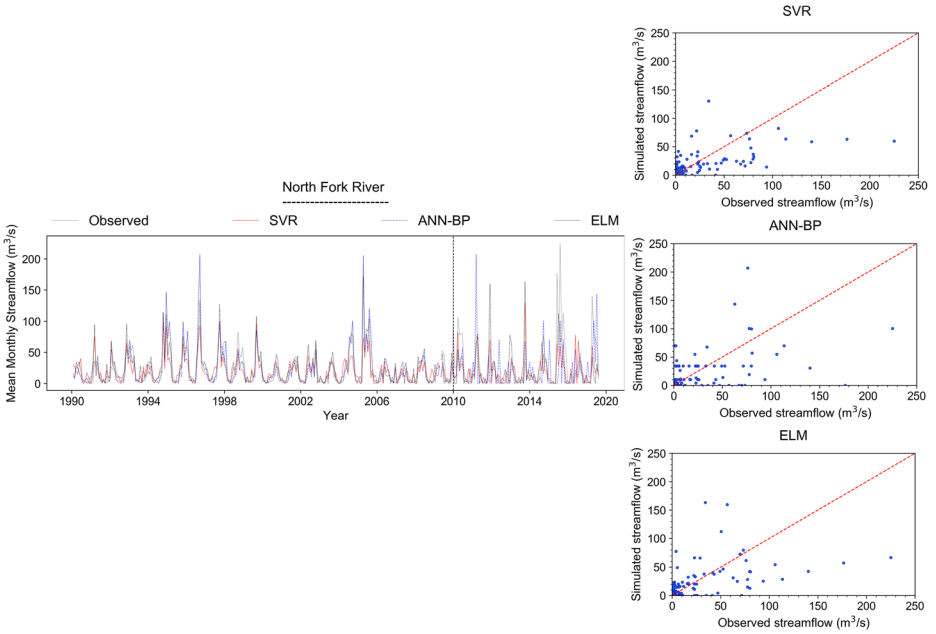

Fig. 3 Comparison of the observed and simulated streamflow of the North Fork River at the monthly scale. The black dash line separates the training and validation period in hydrographs; the scatter plots show the accuracy assessment for the validation period

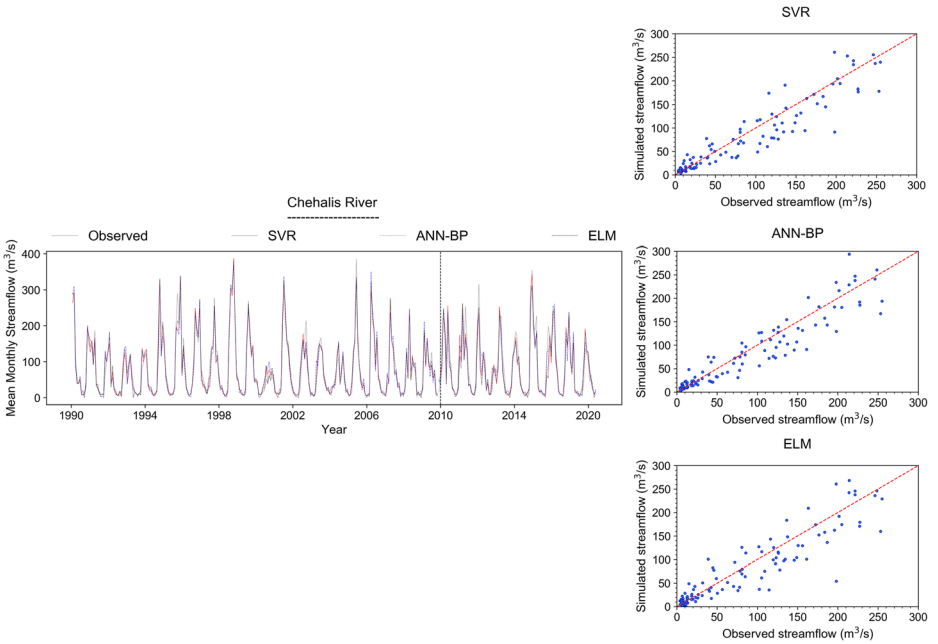

Fig. 4 Comparison of the observed and simulated streamflow of the Chehalis River at the monthly scale. The black dash line separates the training and validation period in hydrographs; the scatter plots show the accuracy assessment for the validation period

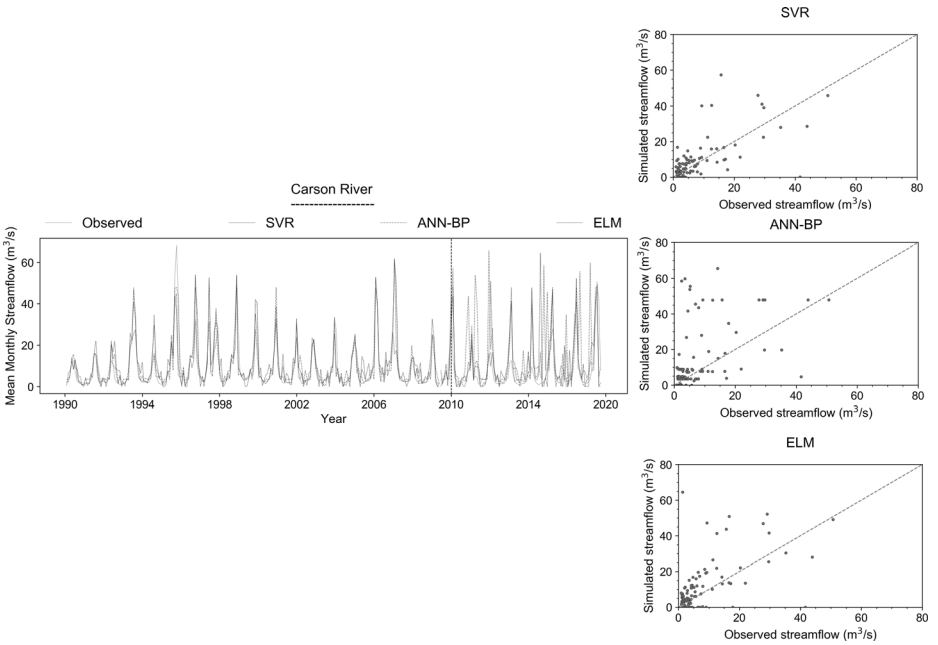

Fig. 5 Comparison of the observed and simulated streamflow of the Carson River at the monthly scale. The black dash line separates the training and validation period in hydrographs; the scatter plots show the accuracy assessment for the validation period

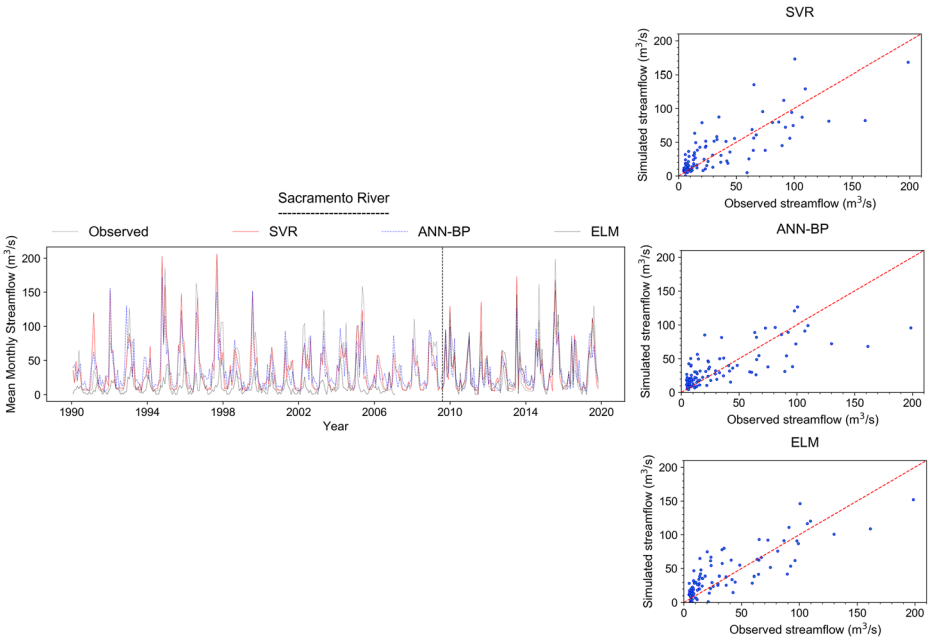

Fig. 6 Comparison of the observed and simulated streamflow of the Sacramento River at the monthly scale. The black dash line separates the training and validation period in hydrographs; the scatter plots show the accuracy assessment for the validation period

5 Results and Discussion

5.1 Comparison of the Simulation Results at the Monthly Time Scale

The optimal parameters of each model for the two time scales are shown in Table 2. Identical numbers of input variables were used for each catchment at the monthly and daily scales. The statistical results of all models for the monthly scale are summarized in Table 3. In Table 3, in the training period, the ANN-BP shows higher accuracy for the North Fork and Carson rivers, while the SVR shows higher simulation accuracy for the Chehalis and Sacramento rivers. However, in the validation period, SVR shows higher accuracy than ANN-BP for the North Fork and Carson rivers, and ANN-BP outperforms SVR for the Chehalis River. ELM also shows better performance in the validation period for the Sacramento River.

Overall, although ANN-BP outperforms other models for the Chehalis River, SVR and ELM outperform ANN-BP with the average values of the model performance indices for all catchments in the validation period. Therefore, according to the comparison of the models, the accuracy ranking is SVR > ELM> ANN-BP for all catchments at the monthly scale.

This finding shows that SVR has better generalization performance than the ELM and ANN-BP models, possibly since the structural risk minimization procedure can lead to an optimal global solution (Gizaw and Gan 2016; Kumar et al. 2016; Shrestha and Shukla 2015). In addition, ELM shows better accuracy than ANN-BP, which indicates the better generalization performance of ELM, as reported in other studies (Huang et al. 2011; Huang et al. 2006; Ding et al. 2015).

Table 4 Performance indices computed from the models for the training and validation sets at the daily scale

Station	ML Method	Training				Validation			
		RMSE	R	NSE	MARE	RMSE	R	NSE	MARE
North Fork River	SVR	39.16	0.70	0.40	**174.50**	**32.26**	**0.56**	**0.23**	**474.31**
	ANN-BP	**31.91**	**0.78**	**0.60**	486.86	33.68	0.42	0.16	1208.19
	ELM	32.90	0.76	0.58	333.41	34.41	0.46	0.13	4001.56
Chehalis River	SVR	**34.00**	**0.96**	**0.91**	**40.84**	54.84	0.88	0.78	491.69
	ANN-BP	49.28	0.91	0.82	69.46	**53.37**	**0.89**	**0.79**	564.35
	ELM	47.77	0.91	0.83	72.94	54.86	0.88	0.78	**252.47**
Carson River	SVR	11.66	0.71	0.48	**73.23**	**9.66**	**0.59**	**−0.04**	**173.74**
	ANN-BP	**10.48**	**0.76**	**0.58**	150.40	21.41	0.31	−4.13	249.05
	ELM	12.94	0.60	0.36	155.13	10.42	0.56	−0.22	258.00
Sacramento River	SVR	**41.82**	**0.71**	**0.47**	**61.24**	**41.86**	**0.58**	**0.33**	**87.34**
	ANN-BP	46.00	0.60	0.36	161.38	44.26	0.50	0.25	181.50
	ELM	43.60	0.65	0.42	126.05	41.87	0.58	0.32	143.07

The bold values indicate the best statistics among all models

The comparison between observed and simulated streamflows for the optimal models at the monthly scale is presented in Figs. 3, 4, 5, and 6. In Figs. 3, 4, 5, and 6, all hydrographs show the performance of all models during the training and validation periods, while the scatter plots show the results of the validation period for all models. From Figs. 3, 4, 5, and 6 and Table 3, all three models with R values of 0.92–0.94, NSE of 0.85–0.89, and MARE of 33.97–45.64 present their best performance for the Chehalis River.

All models show the lowest RMSE values (9.15–17.08) for the Carson River, but their NSE values (−1.92–0.16) lead to the worst performance of all models for this river in the validation period. This worst performance of the models for the Carson River was caused by the catchment dynamics of the Carson River. The Carson River is a snowmelt basin, and snowmelt runoff is the main source of the annual streamflow (Hay et al. 2000). Therefore, the input variables (P, T_{max}, and T_{min}) and their lags are not sufficient to capture the relation between snowmelt and runoff (Tongal and Booij 2018).

The hydrographs in Figs. 3, 4, 5, and 6 show that in the validation period of all rivers except the Carson River, the models can capture the temporal variation of the streamflow. However, for the Carson River, ANN-BP and ELM fail to predict the temporal variation of the streamflow, in particular, both models tend to overestimate the peak flow values. Considering the results in Figs. 3, 4, 5, and 6 and Table 3, the SVR model is more accurate than both ANN-BP and ELM models. The SVR model can estimate the temporal variations of streamflow and has a strong learning ability to capture peak flow values. To compare the efficiency of the developed models for a large dataset, the accuracies of the models are compared at the daily scale in the following section.

5.2 Comparison of Simulation Results at the Daily Time Scale

The statistical results of all models for the daily scale are shown in Table 4. Comparison of the performance indices at the monthly scale (Table 3) and the daily scale (Table 4) indicates that the accuracy of all models at the daily scale decreases in both training and validation period,

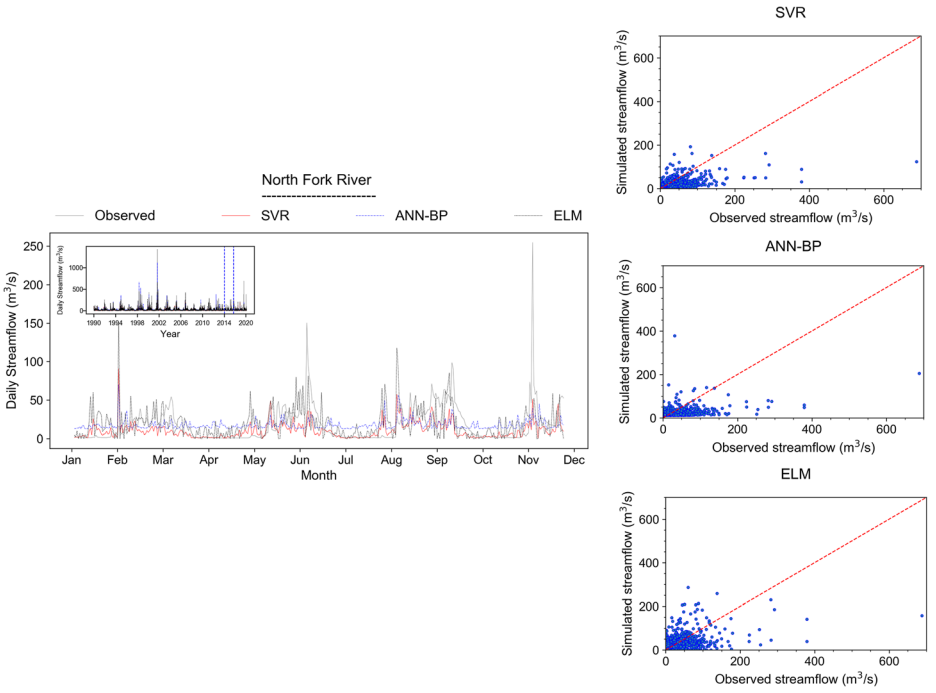

Fig. 7 Comparison of the observed and simulated streamflow of the North Fork River at the daily scale. The two blue dash lines separate 1 year of the validation period in hydrographs for larger presentation; the scatter plots show the accuracy assessment for the validation period

compared with the monthly scale. Some exceptions can be seen in RMSE and NSE values of the ANN-BP and ELM for the North Fork River, and RMSE and NSE values of the ELM for the Carson River at the daily scale in the validation period that show some improvements in accuracy compared with the monthly scale.

From Table 4, in the training period, the better performance of ANN-BP than that of SVR and ELM at the daily scale is similar to the monthly scale for the North Fork and Carson rivers. The identical is obtained for the SVR model that outperforms the other models for the Chehalis and Sacramento rivers in the training period at the daily scale. In the validation period, SVR for the North Fork and Carson rivers, and ANN-BP for the Chehalis River outperforms the other models at the daily scale similar to the monthly scale. However, for the Sacramento River, in the validation period, SVR shows better performance than that of ELM, which is not consistent with the better performance of ELM at the monthly scale in the validation period.

Table 4 shows that all models have the best performance indices for the Chehalis River in the training period at the daily scale, except RMSE for the Carson River which has the lowest values for the all methods in this river (from 10.48 to 12.94). In the validation period, the obtained results show stronger deviations from the observed streamflow. However, the best performance indices are obtained for the Chehalis River, except RMSE for the Carson river (from 9.66 to 21.41) and MARE for the Sacramento River (from 87.34 to 181.50) that has the lowest values for the all models in these rivers. Additionally, all models show their worst performance for the Carson River with NSE values ranged from −4.13 (for the ANN-BP) to −0.04 (for the SVR).

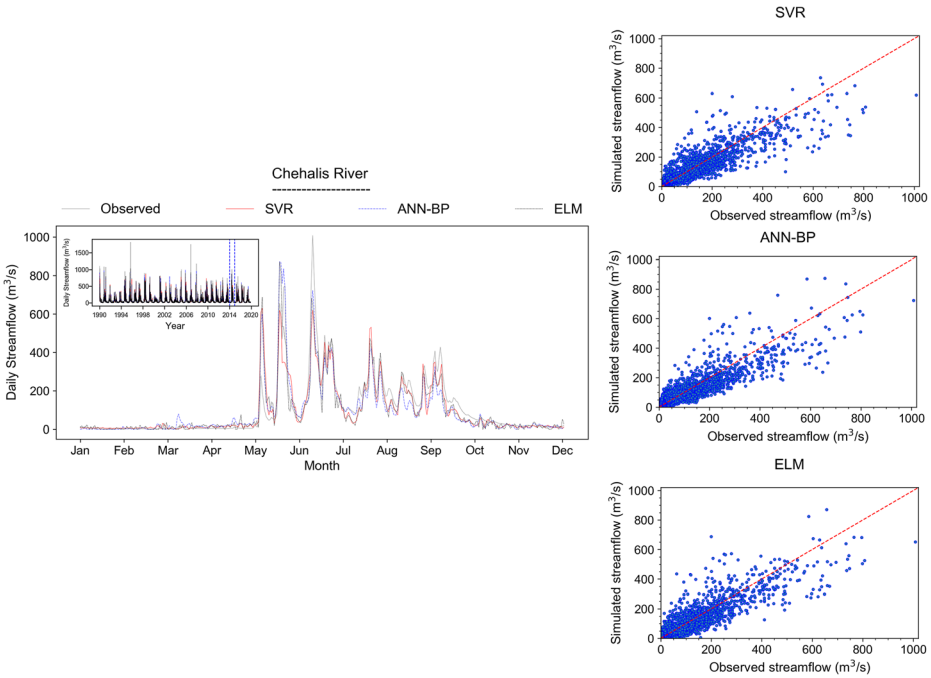

Fig. 8 Comparison of the observed and simulated streamflow of the Chehalis River at the daily scale. The two blue dash lines separate 1 year of the validation period in hydrographs for a larger presentation; the scatter plots show the accuracy assessment for the validation period

Figures 7, 8, 9, and 10 display the observed and simulated streamflow from the SVR, ANN-BP, and ELM models. The scatter plots in Fig. 7 indicate that all three models underestimate streamflow values for the North Fork River, especially at the peak flow values. The similar phenomenon can be seen in Fig. 10 that all models tend to underestimate peak values for the Sacramento River. From Fig. 9, the scatter plot of the ANN-BP model shows that this model tends to overestimate the low streamflow values for the Carson River and this is the main reason for the worst performance of this model for the Carson River at the daily scale, compared with the SVR and ELM. This finding indicates the poor generalization of the ANN-BP model that is less than that of SVR and ELM and is consistent with other related studies (Shortridge et al. 2016; Goyal et al. 2014; Yoon et al. 2011). While the ANN-BP model shows the best accuracy in the training period of the Carson River, has the worst performance in the validation period of this river.

The scatter plots in Fig. 8 indicate that all models present their best performance for the Chehalis River at the daily scale, and the ANN-BP slightly outperforms SVR and ELM for this river in the validation period (Table 4).

In conclusion, although the ANN-BP has a strong ability to simulate the streamflow in the training period, it cannot predict the streamflow as accurately as the SVR and ELM in the validation period. This issue further confirms the superiority of the SVR and ELM over the ANN-BP. Therefore, according to the comparison of the models, the accuracy ranking is SVR > ELM> ANN-BP for all catchments at the daily scale.

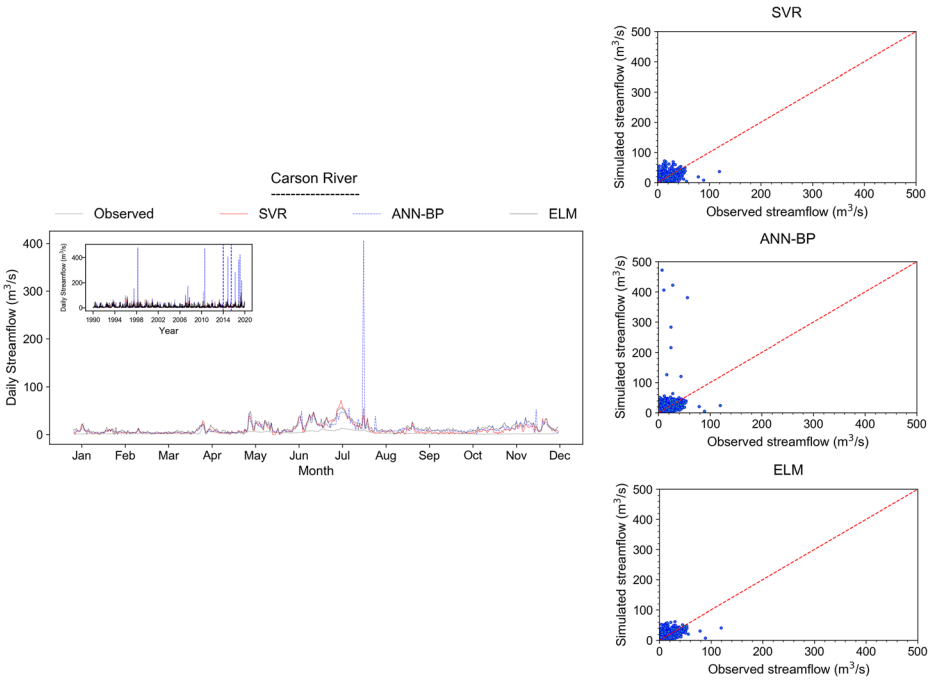

Fig. 9 Comparison of the observed and simulated streamflow of the Carson River at the daily scale. The two blue dash lines separate 1 year of the validation period in hydrographs for a larger presentation; the scatter plots show the accuracy assessment for the validation period

6 Summary and Conclusions

The current study investigated the efficiency of three well-known machine learning algorithms (support vector regression, artificial neural network with backpropagation, and extreme learning machine) in streamflow estimation at the monthly and daily scales. For this purpose, four river catchments with different areas and hydrological characteristics were selected in the United States. Various conclusions can be drawn as follows:

Overall, the accuracy of all models decreases at the daily scale compared to the monthly scale. However, some exceptions are found in RMSE and NSE values of the ANN-BP and ELM models for the North Fork River in the validation period, which indicate some improvements in accuracy of the ANN-BP and ELM at the daily scale compared to the monthly scale. These improvements in RMSE and NSE values of the ELM are also found for the Carson River in the validation period at the daily scale.

All models show their best performance for Chehalis River in the training and validation periods at both monthly and daily scales. However, the simulation results of all models worsen for the Carson River, possibly due to the lag between the streamflow and the snowmelt process in this river.

Considering the simulation results at the daily and monthly scales, one can conclude that the ELM model performsbetter than ANN-BP model in simulating the streamflow, which has been reported in various studies.

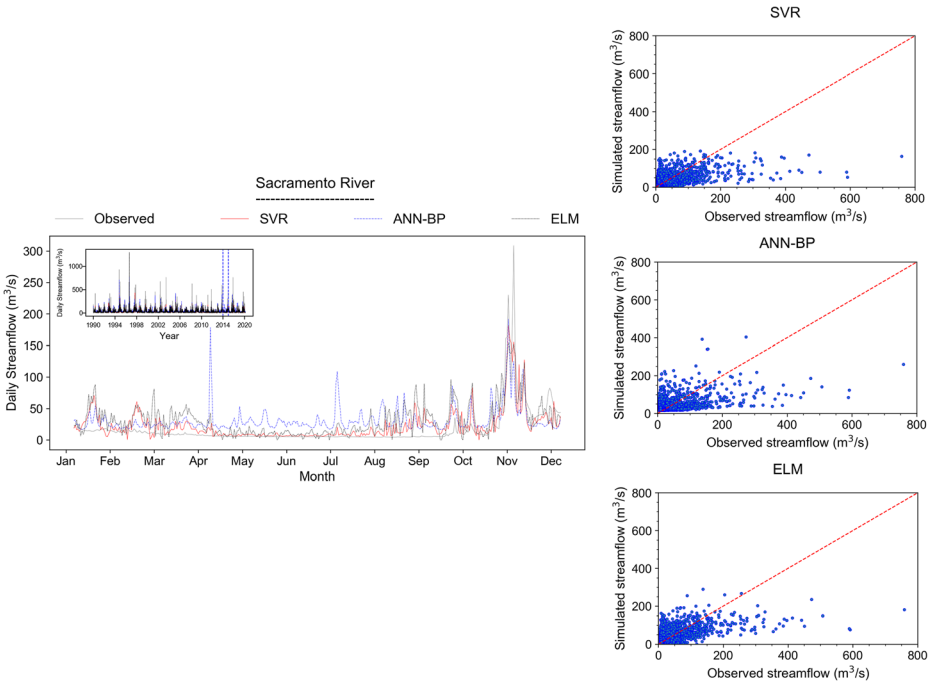

Fig. 10 Comparison of the observed and simulated streamflow of the Sacramento River at the daily scale. The two blue dash lines separate 1 year of the validation period in hydrographs for larger presentation; the scatter plots show the accuracy assessment for the validation period

Athough the ANN-BP shows fair accuracy in the training period for Carson River at the monthly and daily scales, it performs worse than the SVR and ELM with overestimation of the low streamflow values in this river. This finding indicates the poor generalization of the ANN-BP model compared to the SVR and ELM.

Hadi and Tombul (2018) forecased the daily streamflow for basins with different physical characteristics using data-driven methods and indicated that ANN showed better results in the daily scale than SVM, AR, and Adaptive Network-based Fuzzy Inference System (ANFIS). Yaseen et al. (2015) concluded that ELM was a useful alternative to the SVR and Generalized Regression Neural Network (GRNN) in a semi-arid area. Ghumman et al. (2018) found that the SVR-RBF had the best performance among three ANN and four SVR models. However, based on our study on both daily and monthly scales, the results reveal that the SVR performs better than both ELM and ANN-BP. Thus, the generalization performance of the SVR can lead to more accurate and stable simulation results for both monthly and daily scales. Furthermore, the quality of the data or architecture of the models affect the selection of the best model in an area. The acceptable accuracy of the models for the monthly and daily scales proves the suitability of the SVR-RFE method to select the most appropriate predictor variables.

Although all three machine learning models, especially SVR, present acceptable performance for the streamflow simulation, they yield worse simulation results for the Carson River as a snowmelt-dominated basin. In future studies, the suitability of the satellite snow cover products such as MODIS snow cover should be investigated as an additional predictor variable to improve the simulation capabilities of machine learning methods for snowmelt-dominated basins.

The overall results of the current study conclude that the SVR performs better than both ELM and ANN-BP at the monthly and daily scales. This finding proves that the generalization performance of the SVR can have more accurate and stable simulation results for both monthly and daily scales.

Most hydrology methods require easily accessible variables such as the precipitation and temperature as the input to simulate a streamflow process. Additionally, the results of the current study show that data-driven methods using precipitation and temperature can produce reliable results for streamflow simulation and provide an interesting tool for decision makers, especially in data-scarce basins.

Acknowledgments We appreciate the U.S Geological Survey for providing the daily streamflow data of the four stations in this study. We would also like to express our appreciation to the National Oceanic and Atmospheric Administration (NOAA) for providing the daily historical weather information. This work was supported by the National Research Foundation of Korea (NRF) grant (2018R1A2B6001799) funded by the Korean Government (MEST).

Compliance with Ethical Standards

Conflict of Interest None.

References

Adnan RM, Liang Z, Trajkovic S, Zounemat-Kermani M, Li B, Kisi O (2019) Daily streamflow prediction using optimally pruned extreme learning machine. J Hydrol 577:123981

Atiquzzaman M, Kandasamy J (2016) Prediction of hydrological time-series using extreme learning machine. J Hydroinf 18(2):345–353

Barzegar R, Moghaddam AA, Adamowski J, Fijani E (2017) Comparison of machine learning models for predicting fluoride contamination in groundwater. Stoch Env Res Risk A 31(10):2705–2718

Belayneh A, Adamowski J, Khalil B, Ozga-Zielinski B (2014) Long-term SPI drought forecasting in the Awash River basin in Ethiopia using wavelet neural network and wavelet support vector regression models. J Hydrol 508:418–429

Benke AC, Cushing CE (2011) Rivers of North America. Elsevier, Amsterdam

Bonada N, Resh VH (2013) Mediterranean-climate streams and rivers: geographically separated but ecologically comparable freshwater systems. Hydrobiologia 719(1):1–29

Campolo M, Soldati A, Andreussi P (2003) Artificial neural network approach to flood forecasting in the river Arno. Hydrol Sci J 48(3):381–398

Cheng C-T, Zhao M-Y, Chau K, Wu X-Y (2006) Using genetic algorithm and TOPSIS for Xinanjiang model calibration with a single procedure. J Hydrol 316(1–4):129–140

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297

Deo RC, Şahin M (2015) Application of the extreme learning machine algorithm for the prediction of monthly effective drought index in eastern Australia. Atmos Res 153:512–525

Dettinger MD, Cayan DR, Meyer MK, Jeton AE (2004) Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American River basins, Sierra Nevada, California, 1900–2099. Clim Chang 62(1–3):283–317

Dibike YB, Velickov S, Solomatine D, Abbott MB (2001) Model induction with support vector machines: introduction and applications. J Comput Civ Eng 15(3):208–216

Ding S, Zhao H, Zhang Y, Xu X, Nie R (2015) Extreme learning machine: algorithm, theory and applications. Artif Intell Rev 44(1):103–115

Genç O, Dağ A (2016) A machine learning-based approach to predict the velocity profiles in small streams. Water Resour Manag 30(1):43–61

Ghumman AR, Ahmad S, Hashmi HN (2018) Performance assessment of artificial neural networks and support vector regression models for stream flow predictions. Environ Monit Assess 190(12):704

Gizaw MS, Gan TY (2016) Regional flood frequency analysis using support vector regression under historical and future climate. J Hydrol 538:387–398

Govindaraju RS (2000) Artificial neural networks in hydrology. II: hydrologic applications. J Hydrol Eng 5(2): 124–137
Goyal MK, Bharti B, Quilty J, Adamowski J, Pandey A (2014) Modeling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, fuzzy logic, and ANFIS. Expert Syst Appl 41(11):5267–5276
Grantz K, Rajagopalan B, Zagona E, Clark M (2007) Water management applications of climate-based hydrologic forecasts: case study of the Truckee-Carson River basin. J Water Resour Plan Manag 133(4): 339–350
Guo J, Zhou J, Qin H, Zou Q, Li Q (2011) Monthly streamflow forecasting based on improved support vector machine model. Expert Syst Appl 38(10):13073–13081
Hadi SJ, Tombul M (2018) Forecasting daily streamflow for basins with different physical characteristics through data-driven methods. Water Resour Manag 32(10):3405–3422
Hay LE, Wilby RL, Leavesley GH (2000) A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States 1. JAWRA J Am Water Res Assoc 36(2):387–397
Henning JA, Gresswell RE, Fleming IA (2007) Use of seasonal freshwater wetlands by fishes in a temperate river floodplain. J Fish Biol 71(2):476–492
Hu T, Lam K, Ng S (2001) River flow time series prediction with a range-dependent neural network. Hydrol Sci J 46(5):729–745
Huang G-B (2014) An insight into extreme learning machines: random neurons, random features and kernels. Cogn Comput 6(3):376–390
Huang G-B, Zhou H, Ding X, Zhang R (2011) Extreme learning machine for regression and multiclass classification. IEEE Trans Syst Man Cyber Part B (Cybernetics) 42(2):513–529
Huang G-B, Zhu Q-Y, Siew C-K (2004) Extreme learning machine: a new learning scheme of feedforward neural networks. Neural Netw 2:985–990
Huang G-B, Zhu Q-Y, Siew C-K (2006) Extreme learning machine: theory and applications. Neurocomputing 70(1–3):489–501
ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000) Artificial neural networks in hydrology. I: preliminary concepts. J Hydrol Eng 5(2):115–123
Jeton AE, Dettinger MD, Smith J (1996) Potential effects of climate change on streamflow, eastern and western slopes of the Sierra Nevada, California and Nevada. Water Resourc Invest Rep 95:4260
Kalra A, Ahmad S, Nayak A (2013) Increasing streamflow forecast lead time for snowmelt-driven catchment based on large-scale climate patterns. Adv Water Resour 53:150–162
Kang K, Lee JH (2014) Hydrologic modelling of the effect of snowmelt and temperature on a mountainous watershed. J Earth Syst Sci 123(4):705–713
Karran DJ, Morin E, Adamowski J (2013) Multi-step streamflow forecasting using data-driven non-linear methods in contrasting climate regimes. J Hydroinf 16(3):671–689
Kimbrough R, Ruppert G, Wiggins W, Smith R, Kresch D (2006) Water resources data-Washington water year 2005. U. S. Geological Survey
Kumar D, Pandey A, Sharma N, Flügel W-A (2016) Daily suspended sediment simulation using machine learning approach. Catena 138:77–90
Lafdani EK, Nia AM, Ahmadi A (2013) Daily suspended sediment load prediction using artificial neural networks and support vector machines. J Hydrol 478:50–62
Lin J-Y, Cheng C-T, Chau K-W (2006) Using support vector machines for long-term discharge prediction. Hydrol Sci J 51(4):599–612
Liong SY, Sivapragasam C (2002) Flood stage forecasting with support vector machines 1. JAWRA J Am Water Res Assoc 38(1):173–186
Liu M, Lu J (2014) Support vector machine—an alternative to artificial neuron network for water quality forecasting in an agricultural nonpoint source polluted river? Environ Sci Pollut Res 21(18):11036–11053
Liu Y, Sang Y-F, Li X, Hu J, Liang K (2017) Long-term streamflow forecasting based on relevance vector machine model. Water 9(1):9
Luo X, Yuan X, Zhu S, Xu Z, Meng L, Peng J (2019) A hybrid support vector regression framework for streamflow forecast. J Hydrol 568:184–193
Maity R, Bhagwat PP, Bhatnagar A (2010) Potential of support vector regression for prediction of monthly streamflow using endogenous property. Hydrol Proc: Int J 24(7):917–923
Mauger G, Lee S-Y, Bandaragoda C, Serra Y, Won J (2016) Effect of climate change on the Hydrology of the Chehalis Basin. Prepared for anchor QEA. Climate Impacts Group, University of Washington, Seattle
Meng E, Huang S, Huang Q, Fang W, Wu L, Wang L (2019) A robust method for non-stationary streamflow prediction based on improved EMD-SVM model. J Hydrol 568:462–478
Niu W-j, Feng Z-k, Zeng M, Feng B-f, Min Y-w, Cheng C-t et al (2019) Forecasting reservoir monthly runoff via ensemble empirical mode decomposition and extreme learning machine optimized by an improved gravitational search algorithm. Appl Soft Comput 82:105589

Noori N, Kalin L (2016) Coupling SWAT and ANN models for enhanced daily streamflow prediction. J Hydrol 533:141–151
Patel SS, Ramachandran P (2015) A comparison of machine learning techniques for modeling river flow time series: the case of upper Cauvery river basin. Water Resour Manag 29(2):589–602
Peng T, Zhou J, Zhang C, Fu W (2017) Streamflow forecasting using empirical wavelet transform and artificial neural networks. Water 9(6):406
Ragettli S, Cortés G, McPhee J, Pellicciotti F (2014) An evaluation of approaches for modelling hydrological processes in high-elevation, glacierized Andean watersheds. Hydrol Process 28(23):5674–5695
Rezaie-Balf M, Zahmatkesh Z, Kim S (2017) Soft computing techniques for rainfall-runoff simulation: local non–parametric paradigm vs. model classification methods. Water Resour Manag 31(12):3843–3865
Şahin M, Kaya Y, Uyar M, Yıldırım S (2014) Application of extreme learning machine for estimating solar radiation from satellite data. Int J Energy Res 38(2):205–212
Sapin J, Rajagopalan B, Saito L, Caldwell RJ (2017) A K-nearest neighbor based stochastic multisite flow and stream temperature generation technique. Environ Model Softw 91:87–94
Shortridge JE, Guikema SD, Zaitchik BF (2016) Machine learning methods for empirical streamflow simulation: a comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds. Hydrol Earth Syst Sci 20(7):2611–2628
Shrestha N, Shukla S (2015) Support vector machine based modeling of evapotranspiration using hydro-climatic variables in a sub-tropical environment. Agric For Meteorol 200:172–184
Sudheer K, Gosain A, Ramasastri K (2002) A data-driven algorithm for constructing artificial neural network rainfall-runoff models. Hydrol Process 16(6):1325–1330
Tongal H, Booij MJ (2018) Simulation and forecasting of streamflows using machine learning models coupled with base flow separation. J Hydrol 564:266–282
Wang L, Li X, Ma C, Bai Y (2019) Improving the prediction accuracy of monthly streamflow using a data-driven model based on a double-processing strategy. J Hydrol 573:733–745
Wang W-C, Chau K-W, Cheng C-T, Qiu L (2009) A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. J Hydrol 374(3–4):294–306
Wang W (2006) Stochasticity, nonlinearity and forecasting of streamflow processes. IOS Press, Amsterdam
Wu C, Chau K, Fan C (2010) Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques. J Hydrol 389(1–2):146–167
Wu JS, Han J, Annambhotla S, Bryant S (2005) Artificial neural networks for forecasting watershed runoff and stream flows. J Hydrol Eng 10(3):216–222
Wu K-P, Wang S-D (2009) Choosing the kernel parameters for support vector machines by the inter-cluster distance in the feature space. Pattern Recogn 42(5):710–717
Yang T, Asanjan AA, Welles E, Gao X, Sorooshian S, Liu X (2017) Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information. Water Resour Res 53(4):2786–2812
Yaseen ZM, El-Shafie A, Jaafar O, Afan HA, Sayl KN (2015) Artificial intelligence based models for stream-flow forecasting: 2000–2015. J Hydrol 530:829–844
Yoon H, Hyun Y, Lee K-K (2007) Forecasting solute breakthrough curves through the unsaturated zone using artificial neural networks. J Hydrol 335(1–2):68–77
Yoon H, Jun S-C, Hyun Y, Bae G-O, Lee K-K (2011) A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. J Hydrol 396(1–2):128–138
Yu P-S, Chen S-T, Chang I-F (2006) Support vector regression for real-time flood stage forecasting. J Hydrol 328(3–4):704–716
Yu P-S, Yang T-C, Chen S-Y, Kuo C-M, Tseng H-W (2017) Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting. J Hydrol 552:92–104
Yu X, Liong S-Y (2007) Forecasting of hydrologic time series with ridge regression in feature space. J Hydrol 332(3–4):290–302
Zealand CM, Burn DH, Simonovic SP (1999) Short term streamflow forecasting using artificial neural networks. J Hydrol 214(1–4):32–48
Zhang D, Lin J, Peng Q, Wang D, Yang T, Sorooshian S, Liu X, Zhuang J (2018) Modeling and simulating of reservoir operation using the artificial neural network, support vector regression, deep learning algorithm. J Hydrol 565:720–736
Zhu Q-Y, Qin AK, Suganthan PN, Huang G-B (2005) Evolutionary extreme learning machine. Pattern Recogn 38(10):1759–1763