

Cascade-based multi-scale AI approach for modeling rainfall-runoff process

Vahid Nourani, Gholamreza Andalib, Fahreddin Sadikoglu and Elnaz Sharghi

ABSTRACT

In this paper, runoff time series of the sub-basins in a cascade form were decomposed by Wavelet Transform (WT) to extract their dynamical and multi-scale features for modeling Multi-Station (MS) rainfall-runoff (R-R) process of the Little River Watershed (LRW) in USA. A Self-Organizing Map (SOM) clustering technique was also employed to find homogeneous extracted sub-series' clusters. As a complementary feature, extraction criterion of mutual information (MI) was utilized for proper cluster agent choice to impose to the artificial intelligence (AI) models (Feed Forward Neural Network, FFNN; Extreme Learning Machine, ELM; and Least Square Support Vector Machine, LSSVM) to predict the runoff of the LRW sub-basins. The performance of wavelet-based runoff prediction was compared to the Markovian-based MS model. The proposed method not only considers the prediction of the outlet runoff but also covers predictions of interior sub-basins behavior. The outcomes showed that the proposed AI-models combined with the SOM and MI tools enhanced the MS runoff prediction efficiency up to 23% in comparison with the Markovian-based models. Nevertheless, benefit of the seasonality of the process along with reduction of dimension of the inputs could help the AI-models to consume pure information of the recorded data.

Key words | artificial intelligence, Little River Watershed, mutual information, rainfall-runoff modeling, Self-Organizing Map, wavelet

Vahid Nourani (corresponding author)
Gholamreza Andalib
Elnaz Sharghi
 Department of Water Resources Engineering,
 Faculty of Civil Engineering,
 University of Tabriz,
 P.O. Box: 51666, Tabriz,
 Iran
 E-mail: vnourani@yahoo.com

Vahid Nourani
 Department of Civil Engineering,
 Near East University,
 P.O. Box: 99138, Nicosia,
 North Cyprus

Fahreddin Sadikoglu
 Department of Electrical and Electronic
 Engineering,
 Near East University,
 P.O. Box: 99138, Nicosia,
 North Cyprus

INTRODUCTION

Conversion of rainfall to runoff, according to the laws of gravity, vivifies earth, replenishes groundwater, keeps rivers and lakes full of water, and varies the landscape by the action of erosion. The benefit of rainfall-runoff (R-R) modeling as a role of science is providing information for engineers and decision makers to manage, protect, and enhance water resources. Large uncertainties and high non-linearity of the R-R barricade the process-based modeling and seek a black box relationship between driving and resultant variables. So, various black box methods such as artificial intelligence (AI) models have been already presented for R-R simulation (e.g. [Sun *et al.* 2014](#); [Yaseen *et al.* 2015](#); [Chadalawada *et al.* 2016](#); [Kwin *et al.* 2016](#); [Nourani 2017](#)). Among such AI models, Artificial Neural Network (ANN), Extreme

Learning Machine (ELM) and Support Vector Machine (SVM) as black box approaches have revealed their capabilities in modeling and forecasting non-linear hydrological processes in general and R-R modeling in particular (e.g. [Okkan & Serbes 2012](#); [Raghavendra & Deka 2014](#); [Taormina & Chau 2015](#); [Afan *et al.* 2016](#); [Gizaw & Gan 2016](#); [Hosseini & Mahjouri 2016](#); [Humphrey *et al.* 2016](#); [Lima *et al.* 2016](#); [Yaseen *et al.* 2016](#); [Yin *et al.* 2017](#)). On the other hand, available models often rely on empirical algorithms to represent processes that are not well defined, making multi-resolution wavelet analysis appealing, in which long term, high, or low frequency details can be made and used to deep understanding of hydro-environmental processes. The Wavelet Transform (WT) as a data pre-processor presents better

doi: 10.2166/nh.2017.045

interpretation of the process by decomposing non-stationary time series into sub-signals at different scales (levels) and can clarify spectral and temporal information of the signal by combination with ANN, ELM and SVM (WANN, e.g. Nourani et al. 2014; Gorgij et al. 2016); Wavelet-ELM (WELM, e.g. Deo et al. 2016) and Wavelet-SVM (WSVM, e.g. Nourani & Andalib 2015; Kim et al. 2017) models for hydro-environmental processes simulation. In this study, WANN, WELM and WSVM models are utilized for Multi-Station (MS) modeling of R-R. However in AI models, maybe there is not a paramount relation between inputs and outputs. So, dominant input selection is a main challenge of time series modeling, especially in WANN, WELM and WSVM. For this aim, in this study two pre-processing approaches of Self-Organizing Map (SOM) based clustering and mutual information (MI) concept are employed to extract main features and inputs of the WANN, WELM and WSVM methods. SOM as an effective clustering method recently has been used in various hydro-environmental studies (e.g. Kalteh et al. 2008; Adeloje & Rustum 2012; Nourani & Parhizkar 2013; Kar et al. 2015; Chang et al. 2016; Li et al. 2018). On the other hand, MI as a nonlinear measure of data substance can be a robust scheme in necessity of the choice of dominant inputs among numbers of wavelet-based sub-signals (Nourani et al. 2015). Altogether, due to the significance of the model's inputs to acquire the general pattern of R-R process, participation of watershed's effective data has the paramount importance. So, considering both spatial and temporal variations of runoff as a cascade-based MS can elevate prediction of runoff in watersheds. Although MS models were used in some hydrologic research (e.g. Turan & Yurdusev 2009; Nourani & Komasi 2013; Lee & Resdi 2016), to the best of our knowledge there was no study using a MS framework considering seasonality or periodic properties and its influence on the runoff transmission of the watershed for patterning the R-R process. Accordingly, this paper aims to not only predict the outlet runoff values of Little River Watershed (LRW) but also to simulate the runoff time series via a cascade manner at particular points inside the watershed. Therefore, R-R MS modeling is considered in cases where the inside runoff of the LRW is necessary to be predicted. In cascade modeling, the upper sub-basins runoff time series are applied to predict the interior sub-basins runoff, and consequently, central sub-basins are

attended in the LRW outlet runoff prediction. Therefore, the MS model can provide a promising platform regarding the runoff amount in LRW critical places. Hence, two different scenarios are considered for R-R MS modeling to identify an appropriate strategy in hydro-environmental studies. In the first scenario, R-R process Markovian property is suggested as the MS model base, where rainfall antecedent and sub-basins runoff time series are shared (Nourani & Komasi 2013). Furthermore, MI non-linear feature extraction criterion, which is a more appropriate measure compared to the Correlation Coefficient (CC) linear measure, is used for the suitable inputs selection of the Least Square SVM (LSSVM), Feed Forward Neural Network (FFNN), and ELM models to prevent from the laborious process of trial-error for input selection where FFNN is the most commonly used AI model in hydrology, LSSVM uses the concept of SVM classifier as a pre-processing tool which sometimes can lead to more accurate results and ELM as a newer generation of AI models were employed and compared in this study, of course other AI-based models (e.g. genetic programming, Ravansalar et al. 2017) can also be applied. In the second scenario, the seasonality-based or multi-scale property of the R-R process would be considered, where the sub-basins runoff time series using the WT are decomposed at an appropriate level to clarify temporal and spectral time series information. Consequently as a new feature extraction approach, both SOM and MI are used respectively for homogeneous sub-signals clustering and choosing the proper agents of clusters, to be fed into LSSVM, FFNN and ELM models for the LRW MS runoff modeling.

STUDY AREA AND DATA SET

The southeast USA is considered as an authoritative region from agricultural, social, and economical points of view, because of its rapidly increasing population which will increase environmental stress and water demand to currently degraded or stressed ecosystems. The LRW covers 334 km² and approximately 30% of the northern half and 40% of the southern half has been occupied by agricultural cropland including eight sub-basins of extent increasing from 2.62 to 334 km² (Bosch et al. 2004, Figure 1(a)). The observed rainfall and runoff time series from the sub-basins

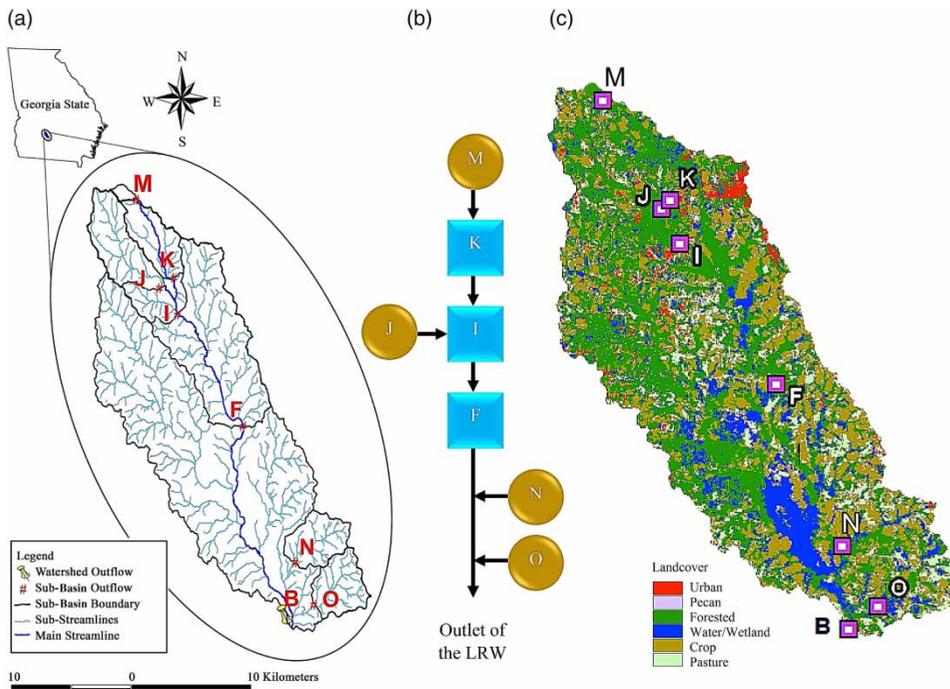


Figure 1 | (a) Study area. (b) Cascade representative of the LRU. (c) Land cover of the LRU.

of LRU (Figure 1(a)) considered in this research ranging from January 1990 to December 2012 were compiled and recorded (<ftp://www.tifonars.org/databases/LREW>). Table 1 shows the used data statistics and Figure 2 indicates the runoff time series and recorded rainfall at the LRU outlet. The first 75% of total data (01/Jan/1990–02/Apr/2007, 6,301 days) were applied for the training and the remaining 25% data set (03/Apr/2007–31/Dec/2012, 2,100

days) used for verification purposes. In this way, higher values of maximum and standard deviation were considered in the training data set, due to the fact that the AI models, LSSVM, ELM and FFNN, can present accurate predictions for unseen data if their interpolator systems are familiar with the same patterns. To speed up training systems, the input and output data were normalized before entering into the training step.

Table 1 | Characteristics of the LRU sub-basins

Sub-basin (station)	USDA ^a ID	Drainage area (km ²)	Location	Mean					
				Rainfall (mm)	Runoff (m ³ /sec)	Channel slope %	Channel length (km)	Mean runoff × 100/area	
B	6840	334	83°03'W 31°54'N	3.12	3.085	0.10	39.10	0.92	
F	5760	115	83°53'W 31°17'N	3.16	1.183	0.14	24.02	1.03	
I	4950	50	83°26'W 31°28'N	3.18	0.592	0.22	12.71	1.18	
J	4850	22	83°09'W 31°32'N	3.18	0.258	0.25	10.30	1.16	
K	4860	17	83°51'W 31°47'N	3.18	0.19	0.29	8.73	1.14	
M	4360	2.62	83°27'W 31°20'N	3.07	0.02	0.33	2.45	0.76	
N	6550	16	83°11'W 31°05'N	3.09	0.15	0.32	10.44	0.96	
O	6860	16	83°03'W 31°37'N	3.09	0.17	0.37	6.11	1.07	

^aUnited States Department of Agriculture.

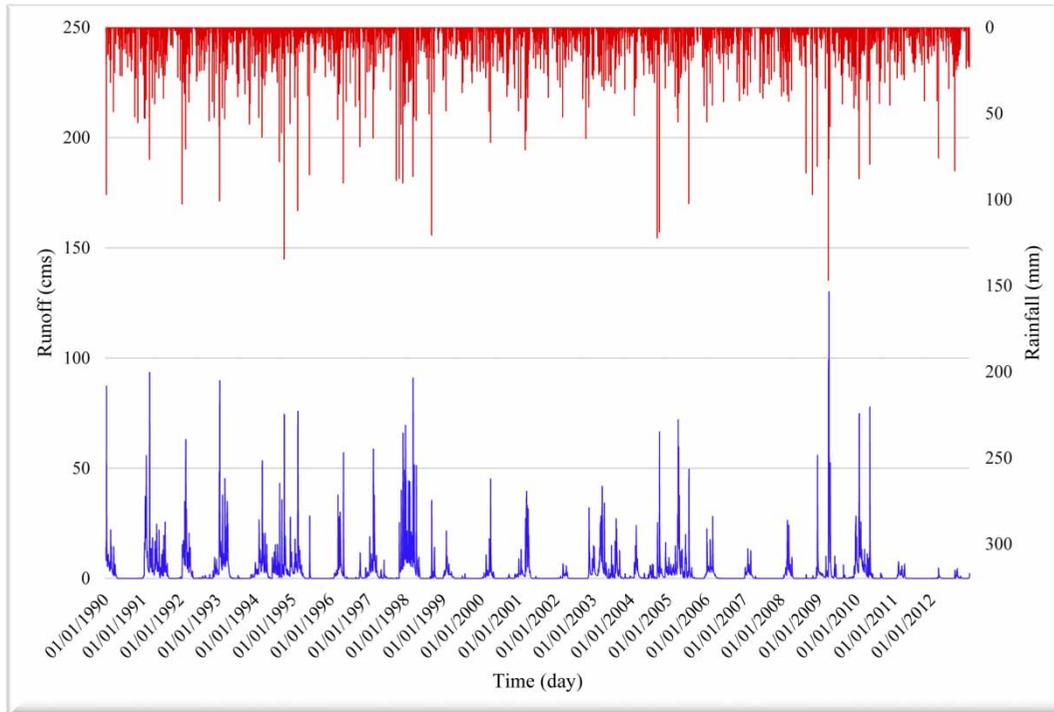


Figure 2 | Daily time series of rainfall and runoff at outlet of the LRW (station B).

MATERIALS AND METHODS

Proposed methodology

Due to the non-linear influential dynamical parameters in the R-R process, the LSSVM, FFNN, and ELM nonlinear AI-based models via two various scenarios were proposed in this research to predict runoff in the LRW outlet and some interior points. In a black-box modeling task of R-R, in a lumped routing manner, the flow is predicted at river specific locations (at sub-basin outlets). In this regard, as can be seen in [Figure 1\(b\)](#), a storage-routing strategy in a reservoir form (sub-basins) network (cascade) has been used in this research. This storage (reservoir)-routing approach is commonly applied in the watersheds for lumped flow routing (e.g. [Boyd et al. 1979](#)). The purpose of the proposed MS model was the runoff prediction of the watershed by relevant sub-basins data as AI models inputs to determine the runoff values at some points within the watershed. Therefore, it would be possible to determine runoff variations in the watershed by MS modeling. The proposed MS runoff modeling (a storage-routing approach) is shown in [Figure 1\(b\)](#) for the

LRW. In [Figure 1\(b\)](#), the circles illustrate sub-basins which transform the rainfall to runoff, and squares indicate the sub-basins not only transforming but also transmitting upper sub-basins runoff to the down sub-basins. In the suggested LRW MS modeling through a cascade manner, the sub-basin I runoff was predicted according to the sub-basins M, K and J runoff time series. In the next step, the sub-basin F runoff was predicted by the sub-basin I data and in the final step, the LRW outlet runoff at station B could be determined using the sub-basins F, N and O data. Therefore, the MS runoff modeling pattern of the LRW was completed using three AI modeling methods as a reservoir cascade. Two various scenarios of Markovian and seasonality-based (multi-scale) applied for LRW MS runoff modeling are explained in the following sub-sections (also see [Figure 3](#)).

Scenario 1

In this scenario, the LRW MS runoff modeling was set according to the Markovian property so that the runoff values at interior sub-basins I and F, as well as the LRW outlet, were simulated using relevant upstream sub-basins

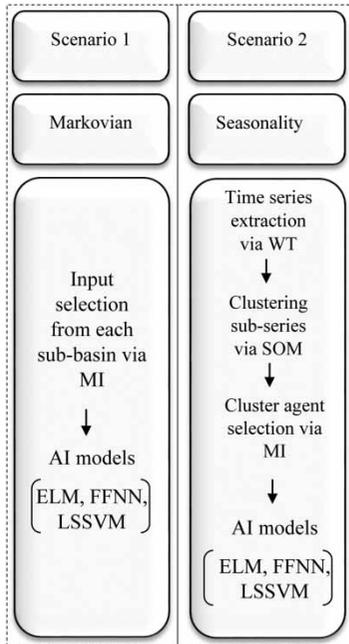


Figure 3 | Steps of the MS scenarios.

antecedent runoff values via Equations (1)–(3), respectively. The sub-basin I runoff was predicted through sub-basins M, K, J runoff values and sub-basin I rainfall as follows (see Figure 1(b)):

$$Q_t^I = f_n(I_t^I, Q_{t-p_M}^M, \dots, Q_{t-p_M}^M, Q_t^K, \dots, Q_{t-p_K}^K, Q_t^J, \dots, Q_{t-p_J}^J) \tag{1}$$

where interior sub-basin I runoff (Q) at time t is the function (f_n) of sub-basin I rainfall (I) at time t , and sub-basins M, J and K runoff values at time t up to lag time p_M , p_K and p_J , respectively. It is worth mentioning that since the influence of upstream sub-basins rainfall values (M, K and J) are implicitly considered at previous time steps of runoff values, just the current time rainfall value as secondary variable was considered in Equation (1) as potential input. However, suitable input selection among numerous potential inputs is a crucial step for the MS modeling. Therefore, the MI supervised feature extraction criterion was used here for proper input set identification instead of applying a trial-error method. Accordingly, from potential input variables, those with maximum MI values with the target (model output) were chosen and considered in the AI models for sub-basin I runoff prediction. Such a non-linear measure superiority in ANN-based input selection

over linear CC measure was previously investigated in several studies (e.g. Nourani et al. 2015). Similarly, the sub-basin F runoff was predicted using sub-basin I runoff via Equation (2), and the outlet runoff at station B was predicted via sub-basins F, N and O runoff values as Equation (3):

$$Q_t^F = f_n(I_t^F, Q_t^I, \dots, Q_{t-p_I}^I) \tag{2}$$

$$Q_t^B = f_n(I_t^B, Q_t^F, \dots, Q_{t-p_F}^F, Q_t^N, \dots, Q_{t-p_N}^N, Q_t^O, \dots, Q_{t-p_O}^O) \tag{3}$$

Scenario 2

In this scenario, the seasonality of R-R process of the LRW formed the MS model basis. The focus of the second scenario was on applying proper sub-basin dominant frequencies to remove redundant data.

For the interior sub-basin I runoff prediction, three steps were followed. First, the sub-basins M, J, K runoff time series were decomposed by WT at level q to control the process seasonal and non-stationarity influences, in a way that sub-basin I runoff was relevant to sub-signals of upstream sub-basins as:

$$Q_t^I = f_n(I_t^I, Q_a^M, Q_{dq}^M, \dots, Q_{dq}^M, Q_a^K, Q_{dq}^K, \dots, Q_{dq}^K, Q_a^J, Q_{dq}^J, \dots, Q_{dq}^J) \tag{4}$$

where Q_a^M and Q_{dq}^M are the approximation and detail runoff sub-signals of sub-basin M at level q , respectively. In the same way, other sub-signals are relevant to J and K sub-basins. In the next step, because of numerous potential inputs, the SOM clustering tool was applied for homogenous sub-signals spatio-temporal grouping. Finally, like scenario 1, MI was applied for dominant input sub-signals selection from each cluster to be imposed in the AI models for sub-basin I outlet (Figure 4). In the second scenario, it is worth mentioning that MI is not appropriate to directly choose dominant inputs as in the first scenario due to, first, when the numerous potential model inputs are ranked according to their MI with the output, the main problem is how to separate dominant inputs from the inputs of ranked potential model with methods such as maximum reduction rate. However, applying SOM before MI would solve the problem of dominant inputs number by clustering potential inputs into particular groups. Second, the criterion of MI can be considered to

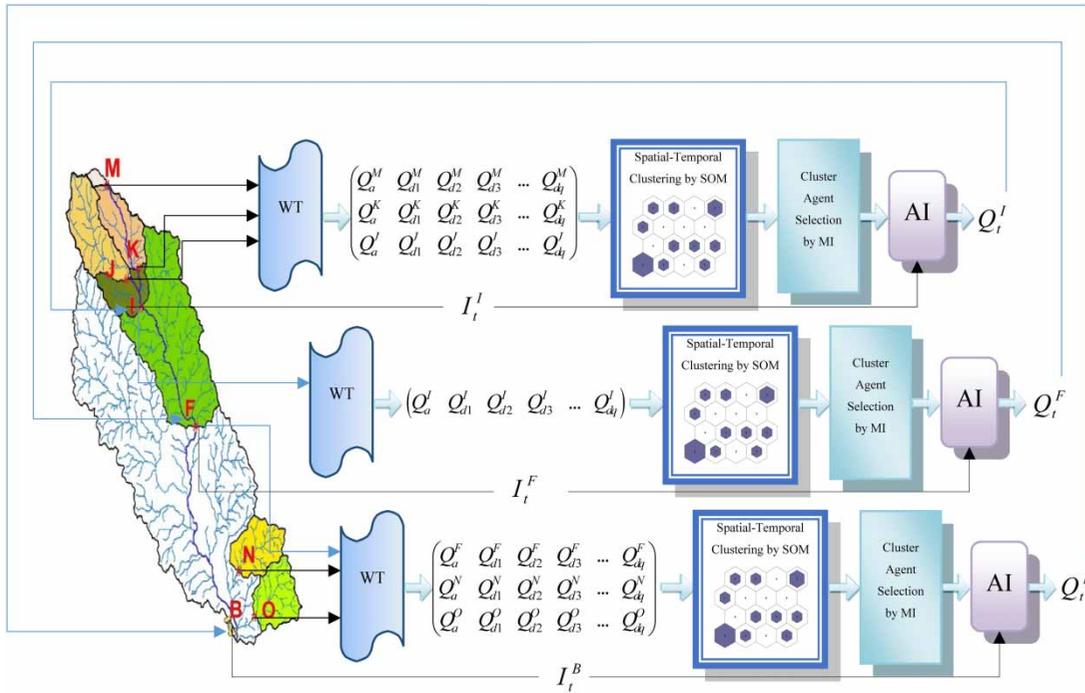


Figure 4 | The MS runoff prediction via second scenario for LRW.

determine suitable model inputs but it is not directly able to cover the problem of redundant inputs. Consequently, it is possible that MI chooses just one pattern and most likely the prediction accuracy in the verification step will collapse. However, the SOM is applied to cluster the input variables into similar inputs groups with particular patterns. After clustering using the SOM, MI chooses dominant inputs from each cluster with various patterns as inputs of the AI models to improve the prediction accuracy of unseen data in the verification step.

Similarly, for the runoff values prediction at stations F and B, Equations (5) and (6) were respectively applied (see Figure 4). After training the AI models using available historical rainfall and runoff data seen at different stations, it was possible to provide output of each sub-basin in the future just using the data from the relevant upstream sub-basins as:

$$Q_t^F = f_n(I_t^F, Q_a^F, Q_{d1}^F, \dots, Q_{dq}^F) \tag{5}$$

$$Q_t^B = f_n(I_t^B, Q_a^B, Q_{d1}^B, \dots, Q_{dq}^B, Q_a^N, Q_{d1}^N, \dots, Q_{dq}^N, Q_a^O, Q_{d1}^O, \dots, Q_{dq}^O) \tag{6}$$

The necessary tools for the two suggested scenarios are explained in the subsequent sections.

WT and Shannon entropy

For capturing the R-R process seasonality pattern in the second scenario, WT was applied for the time series decomposition into sub-signals at various time scales. The wavelet provides a time-scale localization of time series obtained from the compact support of its main function and relevance effects between the WT function and time series. In the hydro-environmental fields, the signals mainly have discrete forms, hence discrete WT was introduced by Mallat (1998) as Equation (7):

$$g_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} g^* \left(\frac{t - nb_0 a_0^m}{a_0^m} \right) \tag{7}$$

where $*$, $g(t)$, m and n are the complex conjugate, wavelet function or mother wavelet (MW), wavelet dilation and translation, respectively ($a_0 > 1$, $b_0 > 0$).

To select the suitable inputs of AI models with regard to the target in a non-linear process, it is necessary to use a robust

supervised tool. To this end, an entropy-based feature extraction tool of MI was utilized which is briefly described below.

Shannon entropy (H) or information content, for a discrete variable of X by sample size of N (bin number), that obtains values $x_1; x_2; \dots; x_N$ with probabilities of p_1, p_2, \dots, p_N , respectively, as (Shannon 1948):

$$H(X) = H(P) = - \sum_{i=1}^N p(x_i) \log [p(x_i)] \quad (8)$$

MI of X and Y is calculated by (Yang et al. 2000):

$$MI(A, B) = H(A) + H(B) - H(A, B) \quad (9)$$

where $H(A)$ and $H(B)$ are the entropy of A and B , and $H(A, B)$ their joint entropy by:

$$H(A, B) = - \sum_{a \in A} \sum_{b \in B} P_{AB}(a, b) \log p_{AB}(a, b) \quad (10)$$

Self-Organizing Map

Whereas redundant information can create problems in the training step of the AI models, it is important to identify the inputs which have similar information value by clustering them to utilize just one member of each cluster as the representative of all members, so this procedure can help to use just one signal instead of many similar signals. According to this, SOM, a powerful soft computing plan that performs a high-dimensional distribution regular mapping onto an orderly low-dimensional grid, was used in this study. Subsequently, SOM can change nonlinear, complicated relationships to simple geometric ones while retaining the data topology formation (Kohonen 1997).

FFNN, ELM and LSSVM models

In the simulation step, three non-linear FFNN, ELM and LSSVM data-driven models were utilized in various types of AI models, where the first is a common method in hydro-environmental modeling, the second is a new neural network approach and the third is currently gaining more attention. The aforementioned data-driven models are briefly explained below.

FFNN is known as the prominent kind of ANN, which is utilized in most research for examining different hydro-environmental processes. Three-layer FFNNs provide a general framework for representing nonlinear functional mapping between a set of input and output variables. The specific equation for one target variable of a three-layer FFNN was presented by Kim & Valdes (2003).

ELM is an elevated system of the single layer FFNN architecture, but is different from the typical neural network where tuning of parameters of the feed-forward networks (hidden layer biases and input weights) is not necessary. It should also be noted that ELM tends to remove some drawbacks of FFNN as over-fitting, local minima and unfit training.

The structure of ELM formed with a single hidden-layer FFNN in which the input weight matrix W is randomly chosen and the output weight matrix β is analytically determined. For a data set with N arbitrary distinct samples (x_i, t_i) that $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$ and $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in R^m$, the structure of ELM is formulated as (Huang et al. 2006):

$$\sum_{i=1}^{\bar{N}} \beta_i g_i(x_j) = \sum_{i=1}^{\bar{N}} \beta_i g_i(w_i x_j + b_i) = y_j, \quad j = 1, 2, \dots, N \quad (11)$$

where \bar{N} , $g(x)$, $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ and b_i are hidden nodes, activation function, the weight vector which connects the input nodes to the i th hidden node, the weight vector that connects the output nodes with the i th hidden node and the threshold of the i th hidden node, respectively.

LSSVM was emanated from SVM as a robust scheme for dealing with problems in nonlinear classification, function and density prediction (Suykens & Vandewalle 1999; Kumar & Kar 2002), where Suykens & Vandewalle (1999) presented the LSSVM non-linear regression. Between the various kernels of LSSVM, the Radial Basis Function (RBF) is usually employed in regression problems. The RBF presents the normal probability distribution and since most stochastic hydro-environmental phenomena follow the normal distribution probability or are capable of transforming to such normal distribution, the RBF kernel can be utilized as the LSSVM's basic kernel functions with

kernel width parameter σ as (Suykens & Vandewalle 1999):

$$K(X, X_i) = \exp\left(-\frac{\|X - X_i\|^2}{2\sigma^2}\right) \quad (12)$$

For AI modeling (e.g. FFNN, ELM and LSSVM), the codes were developed in the MATLAB[®] environment (MathWorks 2010).

Evaluation of models precision

The Determination Coefficient (DC) and Root Mean Square Error (RMSE) as two diverse criteria were used to assess the efficiency of the MS runoff values prediction. The DC and RMSE can be utilized to indicate differences between predictions and recorded values. Legates & McCabe (1999) revealed that hydro-environmental models could be effectively evaluated using Equations (13) and (14).

$$DC = 1 - \frac{\sum_{i=1}^T (O_{obs_i} - O_{com_i})^2}{\sum_{i=1}^T (O_{obs_i} - \bar{O}_{obs})^2} \quad (13)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^T (O_{obs_i} - O_{com_i})^2}{T}} \quad (14)$$

where T , O_{obs_i} , O_{com_i} and \bar{O}_{obs} are the number of recorded data, recorded data, computed values and mean of recorded data, respectively.

RESULTS AND DISCUSSION

The results of the proposed R-R MS modeling using LRW sub-basins information for two different scenarios are investigated in the following sub-sections. In this way obtained results of FFNN, ELM and LSSVM models are compared via the proposed scenarios.

Results of scenario 1

In data-driven FFNN, ELM and LSSVM models, the suitable detection of inputs has a crucial role in improving the

model's performance in both calibration and verification steps; also, it prevents over-training of the model. For this purpose, dominant inputs selection and their appropriate lag times for prediction of runoff values of the sub-basins was carried out by MI (Equation (9)). Table 2 shows the sensitivity analysis results via MI in selecting dominant inputs for MS predictions of runoff. Table 2 presents the outcomes of MS for interior sub-basins (at stations I and F) and outlet of the LRW (station B). Furthermore, by fitting different FFNN, ELM and LSSVM models to all sub-basins of the LRW (one AI model for each sub-basin, separately), it will be possible to update the runoff time series in the future, if necessary. On the other hand, for more understanding of the R-R process and evaluating the inputs of the black box MS models, conversion of rainfall to runoff was surveyed within the LRW. Since the LRW is a relatively small catchment, average rainfall of sub-basins is almost symmetric. So, outlet runoff would be affected by land use/cover. To this end, a reliable identification of hydrologic reaction to soil physical characteristics, such as saturated hydraulic conductivity, bulk density, and moisture retention, will elucidate the outlet runoff value of sub-basins. These characteristics can impact sub-basins hydrology by influencing pathways and transformation rates of rainfall to runoff networks (Price et al. 2010). In general, infiltration rate and rainfall absorption of soils from highest to lowest are forested, pasture, cropland and urban, where forested soils indicated significantly lower bulk densities and higher infiltration rates, and water holding capacities, than cropland and pasture soils. Via the LandSat scene, the LRW land cover was grouped into five categories of cropland, wetland/water, pasture, forested and urban (Figure 5). Considering Figures 1(c) and 6, high elevation lands have mostly forested cover whereas the lowlands are covered by pasture and cropland. After considering the general conditions and land cover/use arrangement of the LRW, upstream sub-basins M, K, J and I should have lower values of runoff per area compared to downstream sub-basins N and O. However, based on the statistics and land cover/use classification of the LRW sub-basins, soil type relation with outlet runoff per area is not observed in practice (Figure 5 and Table 1). So, it could be deduced that the inconformity between outlet runoff values per area with land cover/use classification has an anthropogenic source.

Table 2 | Results of scenario 1 for MS runoff predictions via FFNN, ELM and LSSVM models

Model	Station (sub-basin)	Output	Input variables selected via MI	Network structure	DC		RMSE (normalized)	
					Calibration	Verification	Calibration	Verification
FFNN	I	Q_t^I	I_t^I, Q_t^I Q_t^K, Q_{t-1}^M	(4.5.1) ^a	0.78	0.70	0.028	0.032
	F	Q_t^F	I_t^F, Q_{t-1}^I	(2.8.1)	0.77	0.68	0.029	0.033
	B	Q_t^B	I_t^B, Q_t^N Q_t^O, Q_{t-2}^F	(4.4.1)	0.80	0.73	0.026	0.031
ELM	I	Q_t^I	I_t^I, Q_t^I Q_t^K, Q_{t-1}^M	(4.5.1) ^a	0.75	0.74	0.029	0.030
	F	Q_t^F	I_t^F, Q_{t-1}^I	(2.5.1)	0.73	0.73	0.030	0.031
	B	Q_t^B	I_t^B, Q_t^N Q_t^O, Q_{t-2}^F	(4.9.1)	0.77	0.76	0.029	0.029
LSSVM	I	Q_t^I	I_t^I, Q_t^I Q_t^K, Q_{t-1}^M	(7.2) ^b	0.72	0.70	0.031	0.032
	F	Q_t^F	I_t^F, Q_{t-1}^I	(5.5)	0.70	0.70	0.033	0.032
	B	Q_t^B	I_t^B, Q_t^N Q_t^O, Q_{t-2}^F	(7.4)	0.79	0.74	0.027	0.030

^aThe result has been presented for the best structure. The first, second and third numbers represent input variable, hidden neurons and output variable, respectively.

^bRBF-kernel structure (γ, σ).

As in sub-basins with high pasture and cropland, water consumption is high which reduces outlet runoff from these sub-basins whereas regularly sub-basins with high pasture and cropland with lack of forested areas should allocate less runoff absorption value per area to themselves due to more bulk densities, less infiltration rates, and water holding capacities than forested land. In addition, the MS models input selection using MI was compatible with the LRW geomorphology. In the first model, regarding the sub-basin I runoff prediction, the sub-basins K and J runoff values at time t with no delay were chosen as the inputs due to the high slope and short distance of the river channels between sub-basins K and J outlet stations with the station of I. However, sub-basin M runoff with a 1 day delay was chosen by MI (Table 2). In the second model, sub-basin I runoff with 1 day lag was chosen to model the sub-basin F runoff (Table 2). In the third model, due to the long channel length and low channel slope between F and B stations,

runoff data with 2 days lag were selected for the model. Moreover, close to the LRW outlet, sub-basins O and N, shared their runoff with no lag for the runoff predictions at station B (Table 2). Comparison of the obtained results for runoff values predicted by MS models (see Table 2) revealed that although FFNN, ELM and LSSVM models could exhibit acceptable results, ELM could lead to more accuracy than FFNN and LSSVM models. An appropriate data pre-processing scheme such as the method proposed in scenario 2 could enhance the modeling performance. The following sub-section presents the result of MS modeling via scenario 2.

Results of scenario 2

For a better R-R process understanding, the time series clarifying in both spectral and temporal terms would be helpful. Therefore, to provide such clarifications and to extract

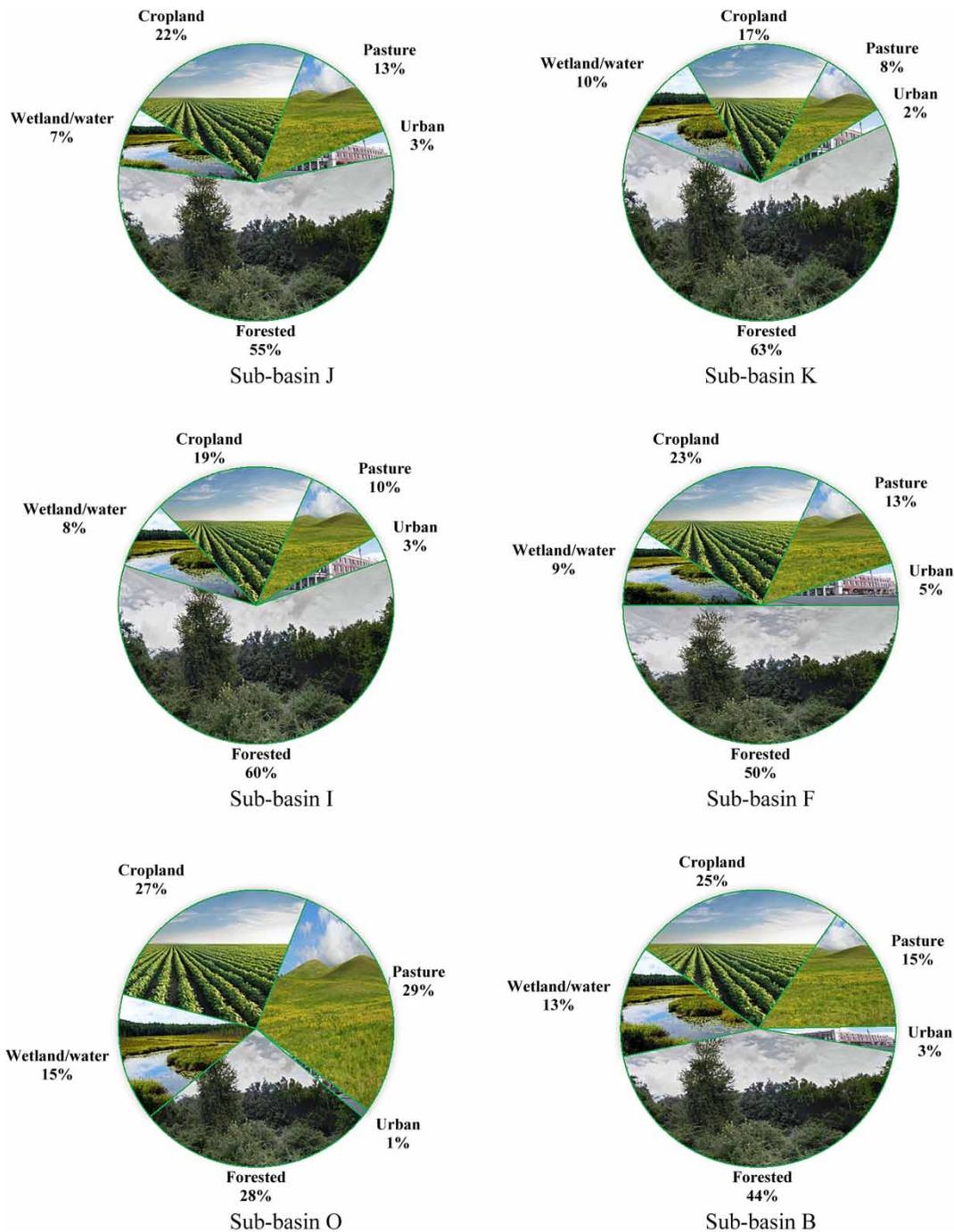


Figure 5 | The land-use classification of LRW.

seasonal multi-scale properties, WT was attached to the MS model to enhance the accuracy of R-R modeling. Appropriate MW selection would provide a challenge in hybrid wavelet-AI modeling as the modeling outcomes may be influenced by the applied type of MW. The main goal of WT is to find the

similarity between the used wavelet prototype and analyzed time series and because of the *coif2* MW structure, which is mainly the same as the runoff signal compared to other MWs like Daubechies-2, 5 (*db2*, *db5*) (Figure 7), it could well capture the features of signal and was chosen as MW for the

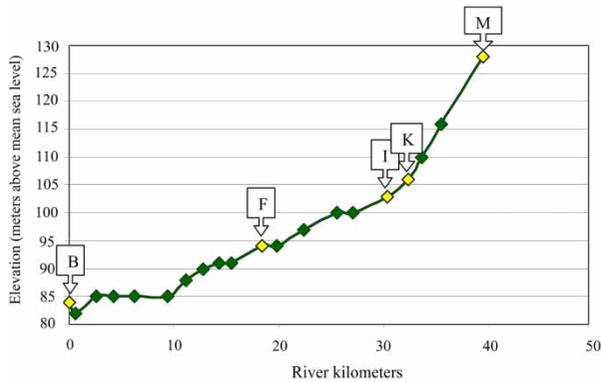


Figure 6 | The longitudinal profile of main river channel of LRW.

runoff time series decomposition in this research. Moreover, the suitable decomposition level in the wavelet-AI models is mainly achieved by sensitivity analysis. Whereas the LRW R-R process could have various seasonality patterns, the decomposition level 8 contains eight details such as 2^1 -day mode, 2^2 -day mode, 2^3 -day mode (nearly weekly), 2^4 -day mode, 2^5 -day mode, 2^6 -day mode, 2^7 -day mode and 2^8 -day mode (approximately annually) and one approximation signal as suitable level for the observed LRW runoff time series decomposition. The following step in the second scenario would be feeding the AI models using the obtained sub-signals (details and approximation) after the time series decomposition via the wavelet. However, such numerous

sub-signals consideration (decomposition at level 8 would give eight details and one approximation sub-signal, and in total nine runoff sub-signals for each station) as the AI models inputs and redundant data may prompt network over fitting, divergence, poor precision and obscurity. Consequently, for the optimization of input layer of the models and improving the rate of model training and performance, it was attempted to select the main sub-signals of the decomposed runoff time series of the sub-basins to be imposed to AI models. Dominant input variables selection could be an efficient alternative for the conventional method of trial and error in complicated hydro-environmental systems. In the second proposed scenario, SOM and MI hybrid was applied to pick dominant seasonalities. After the sub-basins runoff multi-scale sub-signals extraction, spatio-temporal clustering was performed using SOM (e.g. Nourani & Parhizkar 2013). The clustering results for the decomposed time series of LRW can be seen in Table 3. The decomposed runoff sub-signals were classified into six groups for sub-basin I and B modelling, and four groups for sub-basin F. The clustering outcomes by unsupervised SOM indicated that mainly the sub-signals were grouped according to their frequency scale, whereas low frequency sub-signals and approximation were grouped in the same groups and high and low frequencies were separated. In addition to the clustering (by SOM), MI was employed to pick dominant sub-signals from each cluster.

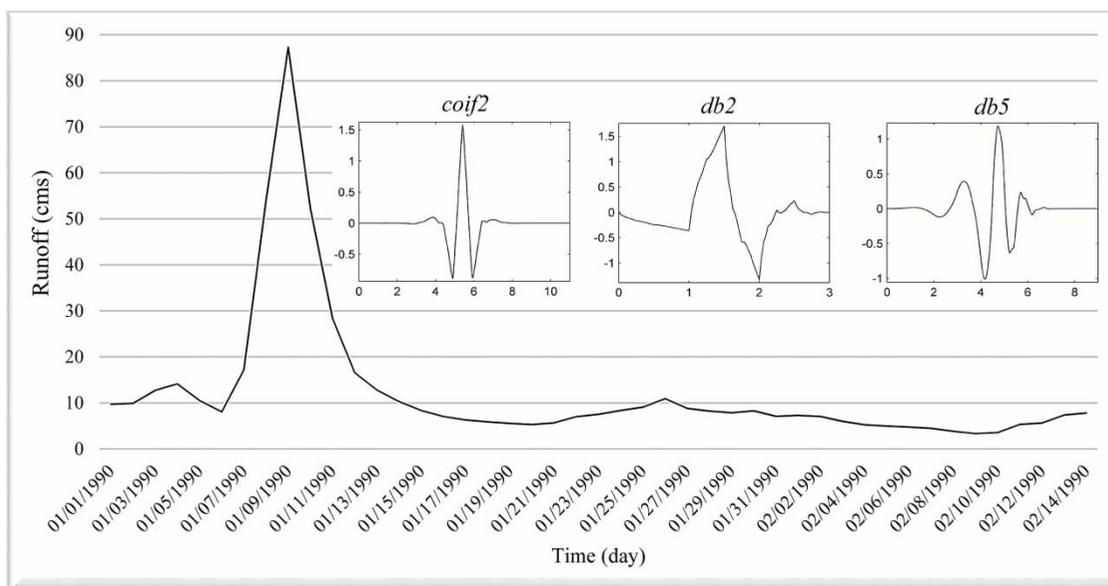


Figure 7 | Comparison of similarity between outlet runoff time series of LRW with *coif2*, *db2* and *db5* MWs.

Table 3 | Results of SOM-based clustering of decomposed runoff sub-series by *coif2* MW

Sub-basin	Cluster 1	Cluster 2	Cluster 3
I	Q_a^K, Q_a^J, Q_a^M	$Q_{d1}^K, Q_{d1}^J, Q_{d1}^M$	$Q_{d8}^K, Q_{d8}^J, Q_{d8}^M$
	Cluster 4 $Q_{d7}^K, Q_{d7}^J, Q_{d7}^M$	Cluster 5 $Q_{d5}^K, Q_{d5}^J, Q_{d5}^M$ $Q_{d5}^K, Q_{d5}^J, Q_{d5}^M$	Cluster 6 $Q_{d2}^K, Q_{d2}^J, Q_{d2}^M$ $Q_{d6}^K, Q_{d6}^J, Q_{d6}^M$ $Q_{d4}^K, Q_{d4}^J, Q_{d4}^M$
	Cluster 1 Q_a^J	Cluster 2 Q_{d8}^J	Cluster 3 $Q_{d1}^J, Q_{d2}^J, Q_{d3}^J, Q_{d5}^J$
F	Cluster 4 $Q_{d4}^J, Q_{d6}^J, Q_{d7}^J$		
	Cluster 1 Q_a^O, Q_a^N, Q_a^F	Cluster 2 $Q_{d8}^O, Q_{d8}^N, Q_{d8}^F$	Cluster 3 $Q_{d7}^O, Q_{d7}^N, Q_{d7}^F$
B	Cluster 4 $Q_{d6}^O, Q_{d6}^N, Q_{d6}^F$	Cluster 5 $Q_{d4}^O, Q_{d5}^F, Q_{d4}^F$ $Q_{d5}^O, Q_{d5}^N, Q_{d5}^F$	Cluster 6 $Q_{d1}^O, Q_{d1}^N, Q_{d1}^F$ $Q_{d2}^F, Q_{d4}^N, Q_{d2}^O$ $Q_{d5}^F, Q_{d5}^N, Q_{d5}^O$

Similar to the first scenario, the sub-signals from each cluster which had higher non-linear correlation (high MI) with the main time series (target) were consequently entered into the AI models. Table 4 presents the dominant inputs of MS model in the second scenario picked by MI. In the first model, for sub-basin I runoff prediction, sub-basin J runoff time series participated with four sub-signals, i.e. 2^8 -day, 2^3 -day, 2^2 -day and 2^1 -day modes, and station I rainfall time series at time t . It is worth mentioning that sub-basin J includes more runoff and a larger area in comparison with sub-basins M and K. As the second effective area, sub-basin K in producing sub-basin I runoff, shared its approximation sub-signal and 2^7 -day mode. It can be noted that the furthest sub-basin M with runoff poverty due to huge forested lands was not desired in the runoff prediction of sub-basin I. In the second model, for the sub-basin F runoff prediction, the approximation, 2^8 -day, 2^7 -day and 2^5 -day, modes of sub-basin I as well as station F rainfall time series at time t were chosen as inputs. In the third model, for sub-basin B runoff prediction, sub-basin F as a sub-basin including higher runoff and far from the LRW outlet, and sub-basins O and N with low runoff values per area and close to station B, shared their suitable sub-signals

as inputs. It can be observed that according to the sub-basins properties, high and low frequencies from the near sub-basins and the far sub-basin respectively were chosen as the inputs using the suggested SOM-MI method. The MS models results through LSSVM, FFNN, and ELM methods are shown in Table 4 for scenario 2. The comparison of results showed that LSSVM has slightly more accuracy than FFNN, this is perhaps because firstly FFNNs often converge on local minima rather than global minima, and secondly FFNNs often overfit if training goes on too long, meaning that for any given pattern, an FFNN might start to consider the noise as part of the pattern. Also, superiority of ELM was proved to LSSVM and FFNN models due to presenting good generalization performance. For the second scenario, the observed and predicted runoff value in the LRW outlet obtained via ELM model can be seen in Figure 8.

Comparison of models

The MS runoff predictions for the LRW indicated differences between two proposed scenarios of Markovian and seasonality-based modeling (Figure 9, Table 5). The AI models results indicated that scenario 2 can provide more precision outcomes as far as the R-R process seasonal (multi-scale) pattern was concerned. The justification would be that in scenario 2 the models were fed by dominant pre-processed data, whereas in scenario 1, the worthiness of sub-basins was not compared. Meanwhile, the accuracy of scenario 1 was lost because of the absence of temporal pre-processing and data robust purge. The prediction results of the LRW runoff via ELM can be seen in Figure 9. The second scenario results, employing both spatial and temporal pre-processing methods using feature extraction criterion, were more accurate compared to the first scenario results. Consequently, the wavelet-AI models, using SOM and MI to capture the space and/or time variations in the process, are a suitable and reliable MS runoff modeling method. After decomposition of time series by wavelet, the sub-basins multi-scale sub-signals entered the SOM and the sub-signals with similar properties were grouped in one cluster. Subsequently, using a suitable feature extraction criterion (e.g. MI), effective sub-signals from each cluster were picked, and those reflecting the runoff process dominant seasonality were

Table 4 | Results of scenario 2 for MS runoff predictions via FFNN, ELM and LSSVM models

Model	Station (sub-basin)	Output	Input variables selected via MI	Network structure	DC		RMSE (normalized)	
					Calibration	Verification	Calibration	Verification
FFNN	I	Q_t^I	$Q_a^K, Q_{d1}^J,$ $Q_{d8}^J, Q_{d3}^J,$ $Q_{d2}^J, Q_{d7}^K,$ I_t^I	(7.9.1)	0.88	0.83	0.021	0.024
	F	Q_t^F	$Q_a^I, Q_{d8}^I,$ $Q_{d3}^I, Q_{d7}^I,$ I_t^F	(5.7.1)	0.90	0.84	0.019	0.024
	B	Q_t^B	$Q_a^F, Q_{d4}^N,$ $Q_{d8}^F, Q_{d3}^O,$ $Q_{d7}^F, Q_{d6}^F,$ I_t^B	(7.8.1)	0.89	0.84	0.020	0.024
ELM	I	Q_t^I	$Q_a^K, Q_{d1}^J,$ $Q_{d8}^J, Q_{d3}^J,$ $Q_{d2}^J, Q_{d7}^K,$ I_t^I	(7.7.1)	0.88	0.88	0.021	0.020
	F	Q_t^F	$Q_a^I, Q_{d8}^I,$ $Q_{d3}^I, Q_{d7}^I,$ I_t^F	(5.7.1)	0.88	0.88	0.021	0.020
	B	Q_t^B	$Q_a^F, Q_{d4}^N,$ $Q_{d8}^F, Q_{d3}^O,$ $Q_{d7}^F, Q_{d6}^F,$ I_t^B	(7.5.1)	0.91	0.90	0.018	0.018
LSSVM	I	Q_t^I	$Q_a^K, Q_{d1}^J,$ $Q_{d8}^J, Q_{d3}^J,$ $Q_{d2}^J, Q_{d7}^K,$ I_t^I	(4.3)	0.86	0.85	0.022	0.023
	F	Q_t^F	$Q_a^I, Q_{d8}^I,$ $Q_{d3}^I, Q_{d7}^I,$ I_t^F	(9.3)	0.85	0.84	0.023	0.024
	B	Q_t^B	$Q_a^F, Q_{d4}^N,$ $Q_{d8}^F, Q_{d3}^O,$ $Q_{d7}^F, Q_{d6}^F,$ I_t^B	(4.2)	0.86	0.86	0.022	0.022

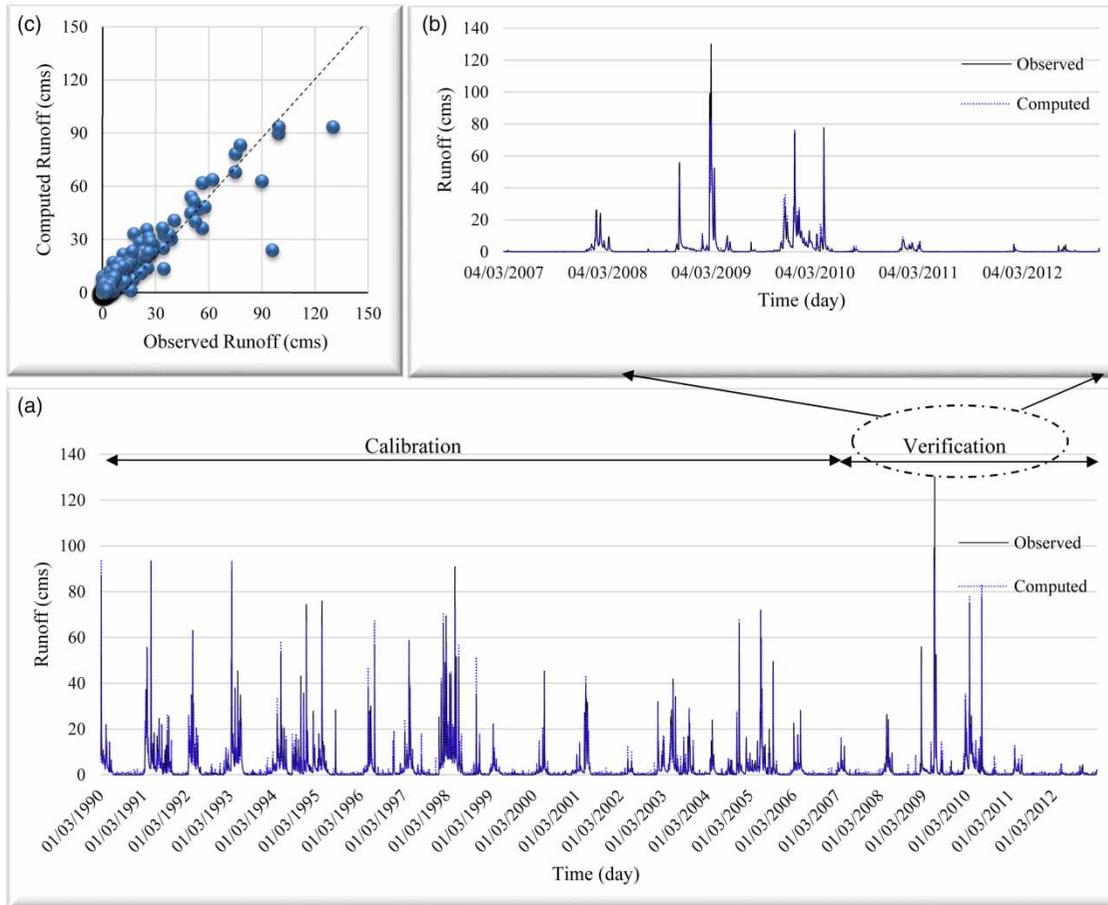


Figure 8 | (a) Computed (by ELM) versus observed runoff time series of LRW outlet via scenario 2. (b) Verification step. (c) Scatter plot.

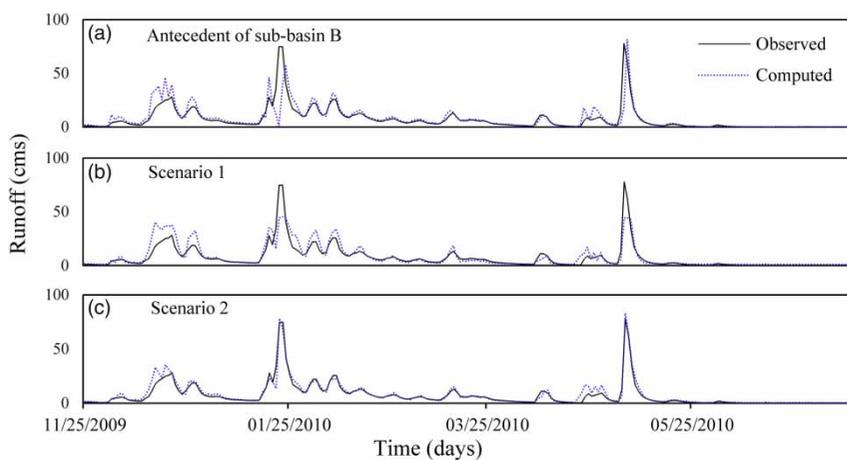


Figure 9 | Computed (by ELM) versus observed runoff time series of LRW outlet in verification step. (a) Antecedent of B. (b) Scenario 1. (c) Scenario 2.

imposed into AI models. In this way, just dominant information was fed to the models. To assess the suggested MS method, the station B runoff was also predicted using

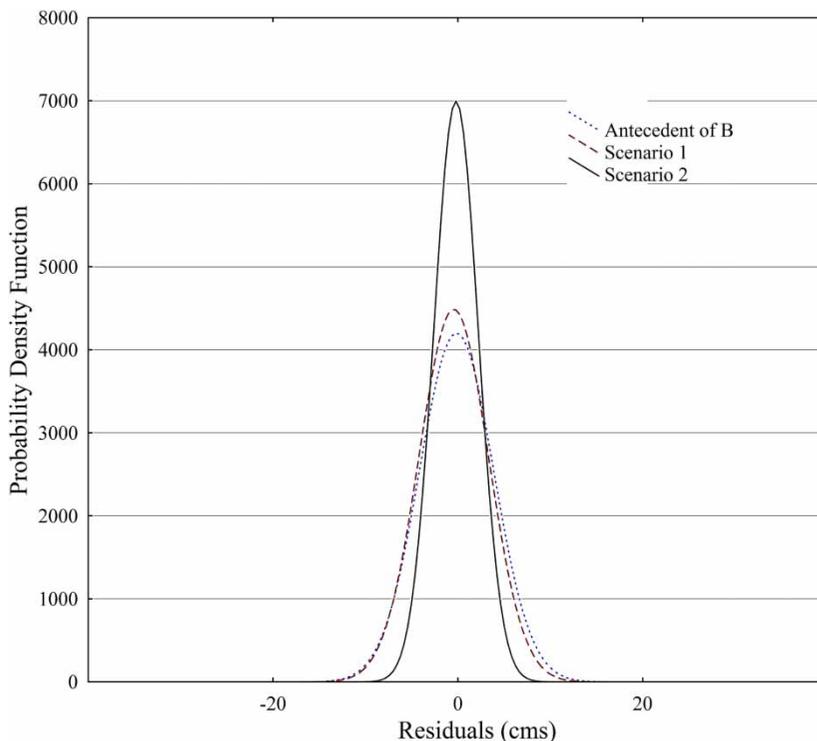
just rainfall and runoff time series antecedent of this station, where up to 3 days' lag time was considered through the trial-error process (i.e. $I_t^B, Q_{t-1}^B, Q_{t-2}^B, Q_{t-3}^B$) as

Table 5 | Comparison of runoff predictions at outlet of the LRW

Model	Method	Network structure	DC		RMSE (normalized)	
			Calibration	Verification	Calibration	Verification
FFNN	Antecedent of sub-basin B	(4.5.1)	0.80	0.73	0.026	0.031
	Scenario 1	(4.4.1)	0.80	0.73	0.026	0.031
	Scenario 2	(7.8.1)	0.89	0.84	0.020	0.024
ELM	Antecedent of sub-basin B	(4.6.1)	0.79	0.74	0.027	0.030
	Scenario 1	(4.9.1)	0.77	0.76	0.029	0.029
	Scenario 2	(7.5.1)	0.91	0.90	0.018	0.018
LSSVM	Antecedent of sub-basin B	(2.7)	0.80	0.73	0.026	0.031
	Scenario 1	(7.4)	0.79	0.74	0.027	0.030
	Scenario 2	(4.2)	0.86	0.86	0.022	0.022

inputs. Such antecedent modeling and MS modeling results via both scenarios can be observed in Table 5. According to the obtained assessment criteria shown in Table 5, it is evident that the cascade-based MS method has more accuracy in comparison with the lonely station B modeling. Moreover, Figure 9 shows a comparison between the computed and the observed runoff time series at the LRW outlet. As can be seen, the seasonality-based MS modeling

could provide more accurate runoff predictions (especially peak values), as it considers runoff spatial variations within the watershed and various upper sub-basins dominant frequencies contributions. The residuals (error) Probability Density Function (PDF) of the proposed scenarios via ELM model in the outlet runoff prediction can be observed in Figure 10 where the scenario 2 error PDF with lower values of mean and standard deviation is well-proportioned

**Figure 10** | Probability density function of residuals runoff values at outlet of the LRW outlet obtained using different scenarios via ELM model.

having high errors accumulation in zero values and, regarding the B antecedent, indicates the highest errors standard deviation.

CONCLUDING REMARKS

In this study, three AI-models (FFNN, ELM and LSSVM) were applied for LRW cascade-based MS runoff prediction (located in Georgia, USA) via two various scenarios (Markovian and seasonality-based scenarios). Such MS modeling could reveal information about the watershed internal hydrometric stations and LRW outlet. The sub-basins daily rainfall and runoff time series were considered over the study area as the potential inputs and the LRW interior and outlet runoff values were considered as the AI models outputs. In such a cascade manner the output runoff was predicted using data from upper sub-basins. Before modeling, the conversion of rainfall to runoff in LRW was investigated by land cover classification and data statistics. It was deduced that in the sub-basins with high cropland and pasture, water consumption is high which reduces outlet runoff from such sub-basins. In the first scenario, the R-R time series antecedent were used as inputs, but for appropriate input selection, supervised feature extraction criterion of MI was applied to prevent the trial-error process. The MI input selection results indicated conformity with the LRW geomorphology (i.e. the input sub-basins far from the output sub-basin and low slope would have more lagged time in comparison with the high slope or near sub-basins). Considering the second scenario and to improve the first scenario results, data pre-processing using feature extraction methods of WT and SOM-MI led to important hydrological parameters detection which proved helpful in enhancing AI-based MS runoff predictions. The second scenario success was because of the fact that the LRW R-R process obeys from a multi-scale seasonal pattern which can be covered using WT. The scenarios comparison indicated that using WT to capture sub-basins multi-scale features could enhance the model accuracy if it is combined with a promising feature extraction method such as SOM-MI, compared to ad hoc AI-models. Runoff time series of sub-basins were decomposed at level 8, which not only considers the dominant seasonality but also does not mar the clustering. Afterwards,

decomposed sub-signals were clustered via SOM, and then the MI feature extraction criterion picked the effective sub-signals of clusters as AI models inputs. Moreover, it was concluded that for the near upstream sub-basins, higher frequency sub-signals participate in LRW runoff predictions, while for far sub-basins it is the low frequency sub-signals. Finally, the comparison of the results achieved by LSSVM, FFNN, and ELM models indicated that ELM had superior efficiency than FFNN and LSSVM as ELM could reveal well generalization efficiency and learning process speed up by simplifying the training step avoiding over-fitting and improving the accuracy of runoff prediction (Deo et al. 2016). For future research, it can be suggested to use and compare the results of the theory-driven (conceptual and physical based) models with the outcomes of the proposed methodology for other watersheds. Moreover, other feature extraction methods and alternative AI models may be examined via the proposed scenarios.

REFERENCES

- Adeloye, A. J. & Rustum, R. 2012 Self-organising map rainfall-runoff multivariate modeling for runoff reconstruction in inadequately gauged basins. *Hydrol. Res.* **43**, 603–617.
- Afan, H. A., El-Shafie, A., Wan Mohtar, W. H. M. & Yaseen, Z. M. 2016 Past, present and prospect of an artificial intelligence (AI) based model for sediment transport prediction. *J. Hydrol.* **541**, 902–913.
- Bosch, D. D., Sheridan, J. M., Batten, H. L. & Arnold, J. G. 2004 Evaluation of the SWAT model on a coastal plain agricultural watershed. *Trans. ASABE* **47** (5), 1493–1506.
- Boyd, M. J., Pilgrim, D. H. & Cordery, I. 1979 A storage routing model based on catchment geomorphology. *J. Hydrol.* **42**, 209–230.
- Chadalawada, J., Havlicek, V. & Babovic, V. 2016 Genetic programming based approach towards understanding the dynamics of urban rainfall-runoff process. *Procedia Eng.* **154**, 1093–1102.
- Chang, F. J., Chang, L. C., Huang, C. W. & Kao, I. F. 2016 Prediction of monthly regional groundwater levels through hybrid soft-computing techniques. *J. Hydrol.* **541**, 965–976.
- Deo, R. C., Tiwari, M. K., Adamowski, J. F. & Quilty, J. M. 2016 Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. *Stoch. Environ. Res. Risk Assess.* **31**, 1211–1240.
- Gizaw, M. S. & Gan, T. Y. 2016 Regional flood frequency analysis using support vector regression under historical and future climate. *J. Hydrol.* **538**, 387–398.
- Gorgij, A. D., Kisi, O. & Moghaddam, A. A. 2016 Groundwater budget forecasting, using hybrid wavelet-ANN-GP modelling:

- a case study of Azarshahr Plain, East Azerbaijan, Iran. *Hydrol. Res.* **48** (2), 455–467.
- Hosseini, S. M. & Mahjouri, N. 2016 Integrating support vector regression and a geomorphologic artificial neural network for daily rainfall-runoff modeling. *J. Hydrol.* **38**, 329–345.
- Huang, G. B., Zhu, Q. Y. & Siew, C. K. 2006 Extreme learning machine: theory and applications. *Neurocomputing* **70**, 489–501.
- Humphrey, G. B., Gibbs, M. S., Dandy, G. C. & Maier, H. R. 2016 A hybrid approach to monthly streamflow forecasting: integrating hydrological model outputs into a Bayesian artificial neural network. *J. Hydrol.* **540**, 623–640.
- Kalteh, A. M., Hjorth, P. & Berndtsson, R. 2008 Review of self-organizing map in water resources: analysis, modeling, and application. *Environ. Model. Softw.* **23**, 835–845.
- Kar, A. K., Lohani, A. K., Goel, N. K. & Roy, G. P. 2015 Rain gauge network design for flood forecasting using multi-criteria decision analysis and clustering techniques in lower Mahanadi river basin, India. *J. Hydrol.* **4**, 313–332.
- Kim, T. & Valdes, J. B. 2003 Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *J. Hydrol. Eng.* **6**, 319–328.
- Kim, S. W., Kisi, O., Seo, Y. M., Singh, V. P. & Lee, C.-J. 2017 Assessment of rainfall aggregation and disaggregation using data-driven models and wavelet decomposition. *Hydrol. Res.* **48** (1), 99–116.
- Kohonen, T. 1997 *Self-Organizing Maps*. Springer-Verlag, Berlin, Heidelberg.
- Kumar, M. & Kar, I. N. 2002 Non-linear HVAC computations using least square support vector machines. *Energy Convers. Manage.* **50**, 1411–1418.
- Kwin, C. T., Talei, A., Alaghamand, S. & Chua, L. H. C. 2016 Rainfall-runoff modeling using dynamic evolving neural fuzzy inference system with online learning. *Procedia Eng.* **154**, 1103–1109.
- Lee, W. K. & Resdi, T. A. T. 2016 Simultaneous hydrological prediction at multiple gauging stations using the NARX network for Kemaman catchment, Terengganu, Malaysia. *Hydrolog. Sci. J.* **61**, 2930–2945.
- Legates, D. R. & McCabe Jr, G. J. 1999 Evaluating the use of goodness-of-fit measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **35** (1), 233–241.
- Li, J., Sharma, A., Evans, J. & Johnson, F. 2018 Addressing the mischaracterization of extreme rainfall in regional climate model simulations a synoptic pattern based bias correction approach. *J. Hydrol.* **556**, 901–912.
- Lima, A. R., Cannon, A. J. & Hsieh, W. W. 2016 Forecasting daily streamflow using online sequential extreme learning machines. *J. Hydrol.* **537**, 431–443.
- Mallat, S. G. 1998 *A Wavelet Tour of Signal Processing*, 2nd edn. Academic Press, San Diego.
- MathWorks, Inc. 2010 *MATLAB: User's Guide, Version 7*. The Math Works, Inc., Natick, MA.
- Nourani, V. 2017 An Emotional ANN (EANN) approach to modeling rainfall-runoff process. *J. Hydrol.* **544**, 267–277.
- Nourani, V. & Andalib, G. 2015 Daily and monthly suspended sediment load predictions using wavelet based artificial intelligence approaches. *J. Mt. Sci.* **12** (1), 85–100.
- Nourani, V. & Komasi, M. 2013 A geomorphology-based ANFIS model for multi-station modeling of rainfall-runoff process. *J. Hydrol.* **490**, 41–55.
- Nourani, V. & Parhizkar, M. 2013 Conjunction of SOM-based feature extraction method and hybrid wavelet-ANN approach for rainfall-runoff modeling. *J. Hydroinform.* **15**, 829–848.
- Nourani, V., Baghanam, A. H., Adamowski, J. & Kisi, O. 2014 Applications of hybrid wavelet-artificial intelligence models in hydrology: a review. *J. Hydrol.* **514**, 358–377.
- Nourani, V., Khanghah, T. R. & Baghanam, A. H. 2015 Application of entropy concept for input selection of wavelet-ANN based rainfall-runoff modeling. *J. Environ. Inform.* **26**, 52–70.
- Okkan, U. & Serbes, Z. A. 2012 Rainfall-runoff modeling using least squares support vector machines. *Environmetrics* **23**, 549–564.
- Price, K., Jackson, C. R. & Parker, A. J. 2010 Variation of surficial soil hydraulic properties across land uses in the southern Blue Ridge Mountains, North Carolina, USA. *J. Hydrol.* **383**, 256–268.
- Raghavendra, S. N. & Deka, P. C. 2014 Support vector machine applications in the field of hydrology: a review. *Appl. Soft Comput.* **19**, 372–386.
- Ravansalar, M., Rajaei, T. & Kisi, O. 2017 Wavelet-linear genetic programming: a new approach for modeling monthly streamflow. *J. Hydrol.* **549**, 461–475.
- Shannon, C. E. 1948 A mathematical theory of communications I and II. *Bell Syst. Tech. J.* **27**, 379–443.
- Sun, A. Y., Wang, D. & Xu, X. 2014 Monthly streamflow forecasting using Gaussian process regression. *J. Hydrol.* **511**, 72–81.
- Suykens, J. A. K. & Vandewalle, J. 1999 Least square support vector machine classifiers. *Neural Process. Lett.* **9** (3), 293–300.
- Taormina, R. & Chau, K. W. 2015 Data-driven input variable selection for rainfall-runoff modeling using binary-coded particle swarm optimization and extreme learning machines. *J. Hydrol.* **529**, 1617–1632.
- Turan, M. E. & Yurdusev, M. A. 2009 River flow estimation from upstream flow records by artificial intelligence methods. *J. Hydrol.* **369**, 71–77.
- Yang, H. H., Vuuren, S. V., Sharma, S. & Hermansky, H. 2000 Relevance of time-frequency features for phonetic and speaker-channel classification. *Speech Commun.* **31**, 35–50.
- Yaseen, Z. M., El-Shafie, A., Jaafar, O., Afan, H. A. & Sayl, K. N. 2015 Artificial intelligence based models for stream-flow forecasting: 2000–2015. *J. Hydrol.* **530**, 829–844.
- Yaseen, Z. M., Jaafar, O., Deo, R. C., Kisi, O., Adamowski, J., Quilty, J. & El-Shafie, A. 2016 Streamflow forecasting model with extreme learning machine data-driven: a case study in a semi-arid region in Iraq. *J. Hydrol.* **542**, 603–614.
- Yin, Z. L., Wen, X. L., Feng, Q., He, Z. B., Zou, S. B. & Yang, L. S. 2017 Integrating genetic algorithm and support vector machine for modeling daily reference evapotranspiration in a semi-arid mountain area. *Hydrol. Res.* **48** (5), 1177–1191.