A two-stage multiple-point conceptual model to predict river stage-discharge process using machine learning approaches

Farhad Alizadeh, Alireza Faregh Gharamaleki and Rasoul Jalilzadeh

ABSTRACT

Due to the complex nature of river stage-discharge process, the present study tried to develop a unique strategy to predict it precisely. The proposed conceptual strategy has some advantages to cover the shortcomings. First, it uses one model instead of several models to predict multiple points instead of one point. On the one hand, the constructed model was inspired by physical-based model (to include time-space attributes of the catchment). On the other hand, ensemble empirical mode decomposition algorithm (EEMD), wavelet transform (WT), and mutual information (MI) were employed as a hybrid pre-processing approach conjugated to support vector machine. For this end, a conceptual strategy (multi-station model) was developed to forecast the Souris River discharge more accurately. The strategy used herein was capable of covering uncertainties and complexities of river discharge modeling. First, a classic model along with WT was performed to predict the 1-day-ahead river discharge for each single station. Therefore DWT-EEMD and feature selection were used for decomposed subseries using MI to be employed in conceptual models. In the proposed feature selection method, some useless subseries were deleted to achieve better performance. The results approved efficiency of the proposed WT-EEMD-MI approach to improve accuracy of different modeling strategies.

Key words | ensemble empirical decomposition mode (EEMD), machine learning, multi-station modeling, mutual information (MI), river stage–discharge process, wavelet transform (WT)

Farhad Alizadeh (corresponding author)
Alireza Faregh Gharamaleki
Rasoul Jalilzadeh
East Azerbaijan Regional Water Company,
Tabriz, East Azerbaijan,
Iran
E-mail: f.alizadeh.ce@gmail.com

INTRODUCTION

The collection of continuous discharge measurements is a very difficult mission and, therefore, a stage-discharge (Q-H) relationship is commonly used to estimate stream discharges as measured stage values. Rating curves are widely used to determine the Q-H relationship, although they are not able to provide sufficiently accurate results. A Q-H rating curve is a relationship between stream stage (water level) and discharge for a particular section of a stream. Such a relationship is usually established by a regression-based analysis, and the curves are generally expressed in the form of a power equation (Kisi & Cobenar 2009; Farhan *et al.* 2020). However, using the Q-H relationship for different river conditions might not be capable enough. In other words, to capture precise results, various researchers suggested applying AI-based models that have proved to provide more accurate outcome in comparison to the Q-H relationship (Potter *et al.* 2010; Ghorbani *et al.* 2016).

Artificial intelligence (AI) has been successfully applied in a number of diverse fields including water resources (e.g., Coulibaly *et al.* 1999; Anmala *et al.* 2000; Tokar & Markus 2000; Chou 2013; Singh *et al.* 2015), whereas there are many AI-based approaches that are driven by knowledge discovery (Babovic 2009; Meshgi *et al.* 2015). Likewise, machine learning approaches are remarkable AI methods

that have been implemented in the river modeling field in recent years (e.g., Yu *et al.* 2004; Azamathulla *et al.* 2009; Chen *et al.* 2010; Nourani *et al.* 2016). Also, some researchers performed reviews on good state-of-the-art application of AI and some pre-processing approaches to categorize different researches and their advantages (Labat 2005; Abrahart *et al.* 2012; Sang 2013).

Wavelet transform (WT) and ensemble empirical decomposition mode (EEMD) are pre-processing methods that give remarkable vision into the physical form of the data by representing information in both time and frequency domains (Roushangar & Alizadeh 2018). EEMD has been proven to be a quite effective method for extracting signals from data generated in noisy nonlinear and non-stationary processes (Huang *et al.* 1998). A time series in the WT breaks down into a series of linearly independent detail signals and one approximation signal by using a specific wavelet function. Mallat (1998) presented a complete theory for wavelet multi-resolution signal decomposition (also mentioned as a pyramid decomposition algorithm). Researches approved that proper data pre-processing by using the WT and EEMD can lead the models to adequately illustrate the real specifications of the basic system. WT along with EEMD decomposes a non-stationary signal into a given quantity of stationary sub-signals. Then, AI methods can be combined with WT and EEMD to improve preciseness of the prediction. In the present research, feature extraction and learning of historical data are the key to improve forecasting accuracy. For this reason, much of the actual effort is put into the processes of feature extraction, data pre-processing, and transformations. Maheswaran & Khosa (2013) also used wavelet technique to improve nonlinear model forecasting performances in Krishna River. Bai *et al.* (2015) developed a multiscale deep feature learning (MDFL) method with hybrid models to deal with the daily reservoir inflow forecasting. Precise hybrid models were employed in recent years to forecast hydrological and hydrogeological processes (e.g., Partal 2009; Tiwari & Chatterjee 2010; Wang *et al.* 2015).

In the modeling process based on AI approaches, some of the input variables might present correlation, noise, or have no significant relationship with target variables and generally are not equally informative. Therefore, one important step is to determine dominant input variables

that are independent, informative, and effective. Mutual information (MI) as a nonlinear measure of information content can be useful in obviating the selection of effective inputs among huge numbers of WT-based or WT-EEMD subseries (Sannasiraja *et al.* 2004; Sun *et al.* 2010; Roushangar *et al.* 2016). Shannon entropy-based measures are applied in this study to extract dominant inputs of the proposed models for discharge-stage modeling. Since the entropy measures arbitrary dependencies between random variables, it is suitable to be applied to complex classification tasks, where methods based on linear relations are prone to mistakes. Entropy can be regarded as a measure of information, disorder, chaos, and uncertainty (Shannon 1948; Roushangar & Alizadeh 2018). Meyer *et al.* (2008) present an original filter approach for effective feature selection in microarray data characterized by a large number of input variables and a few samples.

In most formerly held studies of AI-based hydrological modeling techniques, temporal variables (as historical time series) of the process were proposed as inputs of the model to predict runoff (Anmala *et al.* 2000). On the other hand, in some studies, the AI-based and conceptual models have been combined to estimate river flow (Tokar & Markus 2000; Ashu & Sanaga 2006; Chen & Adams 2006; Corzo *et al.* 2009; Roushangar *et al.* 2018). These models, which combine conceptual hydrological models and AI-based models, usually include two parts: one is a conceptual model for runoff of every sub-catchment and another is the AI-based model to ensemble the runoff from each catchment. Nourani *et al.* (2013) developed a geomorphological model to handle Eel River watershed runoff. The quality of the results supported the beneficial application of the approach in spatio-temporal modeling of the hydrological processes. Multi-station modeling is a conceptual approach whose capability was confirmed and, in some cases, imposing geomorphological features to the model caused an improvement in results (Zhang & Govindaraju 2003; Sarangi *et al.* 2005).

The present study aims to predict the daily river stage-discharge process and enhance the capability of modeling scenarios by using hybrid WT-EEMD-MI and machine learning models. To achieve this goal, the capability of WT-EEMD-MI based multi-station model was investigated and improvements were studied, and results were compared with classic rating curve (RC).

MATERIALS AND METHODS

Study area and used data

The Souris River is a river in central North America. It is about 700 km in length and drains about 61,100 km^2. It rises in the Yellow Grass Marshes north of Weyburn, Saskatchewan. It wanders south through North Dakota beyond Minot to its most southern point in the city of Velva, and then back north into Manitoba. The river passes through the communities of Melita, Hartney, Souris, and Wawanesa and on to its confluence with the Assiniboine River at Treesbank, about 40 km south east of Brandon. The main tributaries which flow into the Souris in Manitoba are the Antler River, the Gainsborough and Plum Creeks. Much of its drainage basin is fertile silt and clay deposited by former glacial Lake Souris. The channel capacity of the river in Manitoba varies from about 4.2 m^3/s near the border, to about 40 m^3/s through Melita, to about 31 m^3/s near Lauder and 48 m^3/s near Hartney. North of Hartney the capacity increases to more than 85 m^3/s. The drop between the border and Hartney is only about 9.5 cm/km. Two large dams in Saskatchewan, Rafferty Dam and Alameda Dam, were built, in part, to reduce flood peaks on the Souris River. Table 1 demonstrates the characteristics of the Souris River and used data.

Selected stations (Sherwood, Foxholm, Minot, Verendry, Bantry, and Westhope) as shown in Figure 1, are in continuous form which is suitable to be used in the proposed models. Discharge-stage data ranges from 1 January 2012 to 1 January 2013 in daily scale. Also, the dataset was separated into two parts: the first division including

70% of total data as calibration dataset and the rest was considered as verification dataset. The historical daily discharge data were available for six stations in the Souris River at the USGS website (http://co.water.usgs.gov/sediment).

Rating curves (RC)

A discharge-height rating curve (RC) technique consists of a graph or equation, relating discharge to stream stage (water level), which can be used to estimate the discharge from the water level record. The Q-H rating curve generally represents a functional relationship of the form given in Equation (1):

$$Q = aH^b \tag{1}$$

where Q is the stream discharge and H is the water level. The values of a and b for a particular stream are determined from data via a linear regression between log H and log Q. A major limitation of this approach is that it is not able to consider the hysteresis effect. In this study, the values of a and b are computed by using the least squares method.

Support vector machine (SVM)

Support vector machine (SVM) is a supervised learning model with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. A SVM model represents the samples as points in space, mapped so that the samples of the separate categories are divided by a clear gap that is as wide

Table 1 | Characteristics of six sub-basins in the Souris River

Hydrometric station	Station discharge (m³/s)	Stage (m) max.	Max. discharge (m³/s)	Min. discharge (m³/s)	Area (m²)	Distance from next hydrometric station (km)
Sherwood	4.1	1.02	10.8	0.36	23,154.4	–
Foxholm	6.74	2.15	11.44	0	24,527.2	145
Minot	7.28	1.65	12.32	0.07	27,453.9	42
Verendrye	13.74	1.96	14.78	0.31	29,266.9	126
Bantry	14.13	2.28	17.28	0.42	31,856.8	124
Westhope	15.66	3.05	14.8	0.12	43,770.8	121
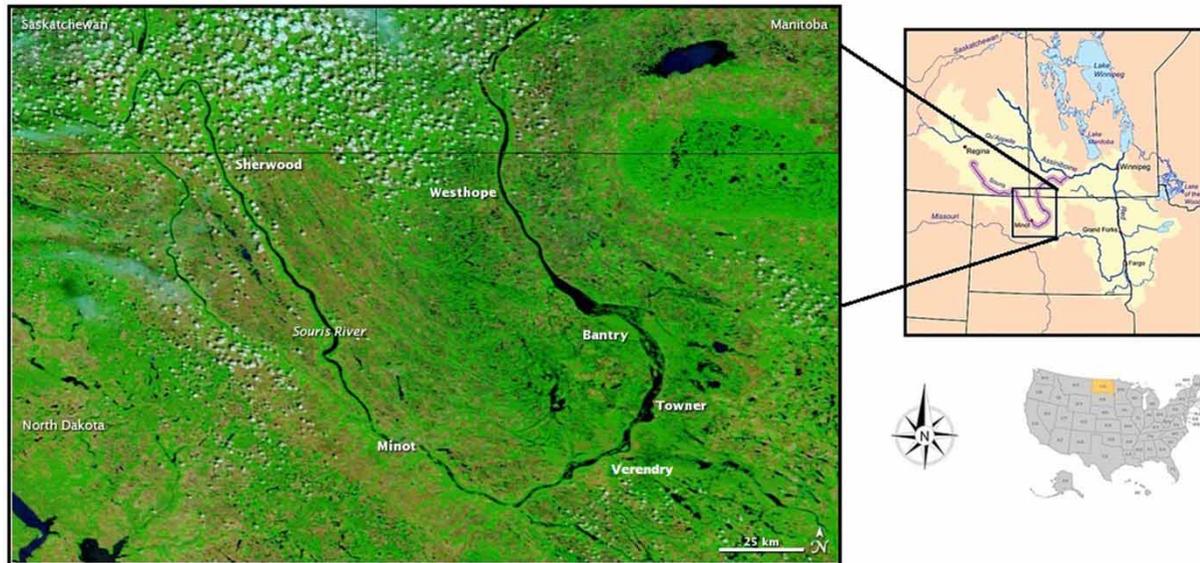
Figure 1 │ The geographic location of Souris River and location of hydrometric stations.

as possible. As an effective classification method, SVM has been proposed on the basis of statistical learning theory (SLT). SVM is a linear machine working in the high-dimensional feature space formed by the nonlinear mapping of the n-dimensional input vector x into a K-dimensional feature space (K > n) with a function. SVM neural network applies one hidden layer of nonlinear neurons, one output linear neuron and specialized learning procedure leading to the global minimum of the error function and excellent generalization ability of the trained network.

Alternative regression approaches

In this study, in order to measure the efficiency of SVMs' performance, five regression models were selected to compare the outcome as follows.

Auto regressive integrated moving average with exogenous input (ARIMAX)

Auto regressive integrated moving average with exogenous input (ARIMAX) is a classic time series forecasting model and therefore was used as a comparison model to evaluate the efficiency of SVM. Although a linear model like

ARIMAX is unlikely to model a complex nonlinear hydrological process, it is still a commonly used method in practice and, as such, is useful as a comparison model (Adamowski & Chan 2011). In this research, the ARIMAX (p, d, q) model was applied in which p, q, and d refer to orders of autoregressive and moving average components and the number of differencing operations, respectively. Temporal features were used as exogenous inputs to predict future discharge as the output.

Extreme learning machine (ELM)

Extreme learning machine (ELM), which is an algorithm developed for single hidden layer feed forward neural networks (SLFNs), can be thousands of times faster than traditional feed forward network learning algorithms like back-propagation algorithm, while obtaining a better generalization performance (Huang *et al.* 2004). Different from traditional learning algorithms, ELM tends to reach the smallest training error and the smallest norm of the weights. ELM, with higher scalability and less computational complexity, not only unifies different popular learning algorithms, but also provides a unified solution to different practical applications (e.g., regression, binary, and multi-class classifications).

Adaptive neuro-fuzzy inference system (ANFIS)

The ANFIS is a fuzzy Sugeno model put in the framework of adaptive systems to facilitate learning and adaptation (Jang 1993). Such framework makes the ANFIS modeling more systematic and less reliant on expert knowledge. To present the ANFIS architecture, two fuzzy if-then rules based on a first-order Sugeno model are considered:

> Rule 1: If x is A_1 and y is B_1; then $f_1 = p_1 x + q_1 y + r_1$
> Rule 2: If x is A_2 and y is B_2; then $f_2 = p_2 x + q_2 y + r_2$

where x and y are the inputs, A_i and B_i are the fuzzy sets, f_i are the outputs within the fuzzy region specified by the fuzzy rule, p_i; q_i; and r_i are the design parameters that are determined during the training process. The output of the ANFIS is calculated by employing the consequent parameters found in the forward pass. The detailed algorithm and mathematical background of the hybrid-learning algorithm can be found in Jang (1993).

Artificial neural network-radial basis function (ANN-RBF)

Radial basis function (RBF) neural network is based on supervised learning. RBF networks were independently proposed by many researchers and are a popular alternative to the multi-layer perceptron (MLP). RBF networks are also good at modeling nonlinear data and can be trained in one stage rather than using an iterative process as in MLP, and also learn the given application quickly (Venkatesan & Anitha 2006). The structure of RBF neural network is similar to that of MLP. It consists of layers of neurons. The main distinction is that RBF has a hidden layer which contains nodes called RBF units. Each RBF has two key parameters that describe the location of the function's center and its deviation or width. The hidden unit measures the distance between an input data vector and the center of its RBF. The RBF has its peak when the distance between its center and that of the input data vector is zero and declines gradually as this distance increases. There is only a single hidden layer in a RBF network, there are only two sets of weights, one connecting the hidden layer to the input layer and the other connecting the hidden layer to the output layer. Those weights connecting to the input layer contain the parameters of the basis functions. The weights connecting the hidden layer to the output layer are used to form linear combinations of the activations of the basis functions (hidden units) to generate the network outputs. Since the hidden units are nonlinear, the outputs of the hidden layer may be combined linearly and so processing is rapid. The output of the network is derived from Foody (2010).

Concept of modeling via AI-based approaches

With respect to the reviewed researches, most of the deficiencies about application of AI-based models are related to its input–output structure. A simple AI model might not be able to handle the natural uncertainty of specific hydrological processes completely. As an instance, an ad hoc AI model is not designed to predict the variable of interest in an arbitrary point across the river. Therefore, it is required to design models which have semi-distributed model characteristics. For this end, in this research, different scenarios in river discharge modeling context were proposed. Multi-station modeling via SVM as a conceptual model was used in this study, which is going to be discussed. Figure 2 shows the modeling flowchart in this study.

On the other hand, discharge data include a broad domain of values and fluctuations and have to be pre-processed. As is demonstrated in Figure 2, it is proposed to use the WT to decompose the original time series into two components: the approximation components and the detail components. Secondary decomposing was performed by applying the EEMD to the detail components obtained from WT into a new subseries of intrinsic mode functions (IMFs). The goal of this step is to additionally decrease the non-stationarity of the detail components captured from WT. Although decomposing twice might lead to the required outcome, reproduction of subseries might defect the SVM performance. For this reason, mutual information was used to select the dominant subseries. Such a feature selection could increase the performance of SVM.

Multi-station modeling of river discharge

The multi-station modeling of river discharge used in this study was designed to catch the nonlinearity, uncertainty, and complexity of the river discharge. For this purpose, the multi-station model was established according to the
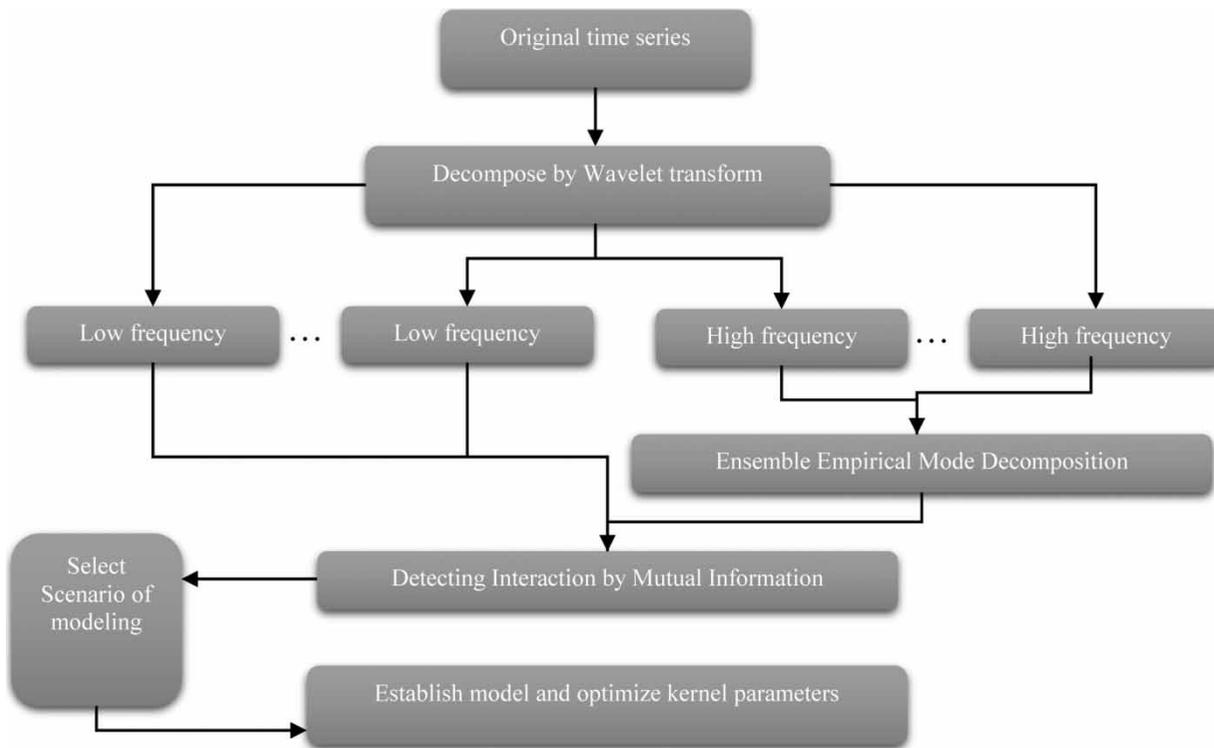
Figure 2 | Flowchart of pre-processing via WT-EEMD-MI in this study.

observed discharge time series. In other words, the model inputs consisted of the temporal features that include *Q-H* processes from the last days. A schematic of the proposed model is shown in Figure 3.

Since the model is in the multi-station form, the input matrix was determined in a way to cover the spatio-temporal variation of the river discharge uncertainties by using temporal features of the hydrometric stations. Therefore, multiple inputs were set in a way that all temporal features could establish a unit matrix. The input variables comprised different sets of antecedent and current gauge *Q-H* values of all the stations to forecast the discharge values ($Q_i(t)$, $i = 1$, 2, 3… *n*, where *n* is the number of hydrometric stations). The multi-station model presents a single model which is capable enough to be employed instead of several models within the watershed. The proposed model has unique temporal characteristics which could be applied to forecast the entire river discharge using only one model. Selection of an appropriate input combination for the model is the most important step in the modeling process (Nourani *et al.* 2013). In the multi-station model, the input variables of the

model are selected among the temporal features of the river to forecast the discharge values for each station ($Qi(t + 1)$) by assuming the Markovian property of the process. In addition, the proposed model is able to predict runoff in stations with lack of data or any desired point within the river due to considering the temporal features of the sub-basins as the model inputs. A MATLAB code was developed to calculate the model.

Proposed pre-processing approach

Discrete wavelet transform (DWT)

The wavelet transform (WT) is a popular method and a very precise method for time series processing (Aussem *et al.* 1998; Nourani *et al.* 2013; Farajzadeh & Alizadeh 2018). While the general theory behind WT is quite analogous to that of the short time Fourier transform (STFT), WT allows for a completely flexible window function (called the mother wavelet), which can be changed over time based on the shape and compactness of the signal. Given
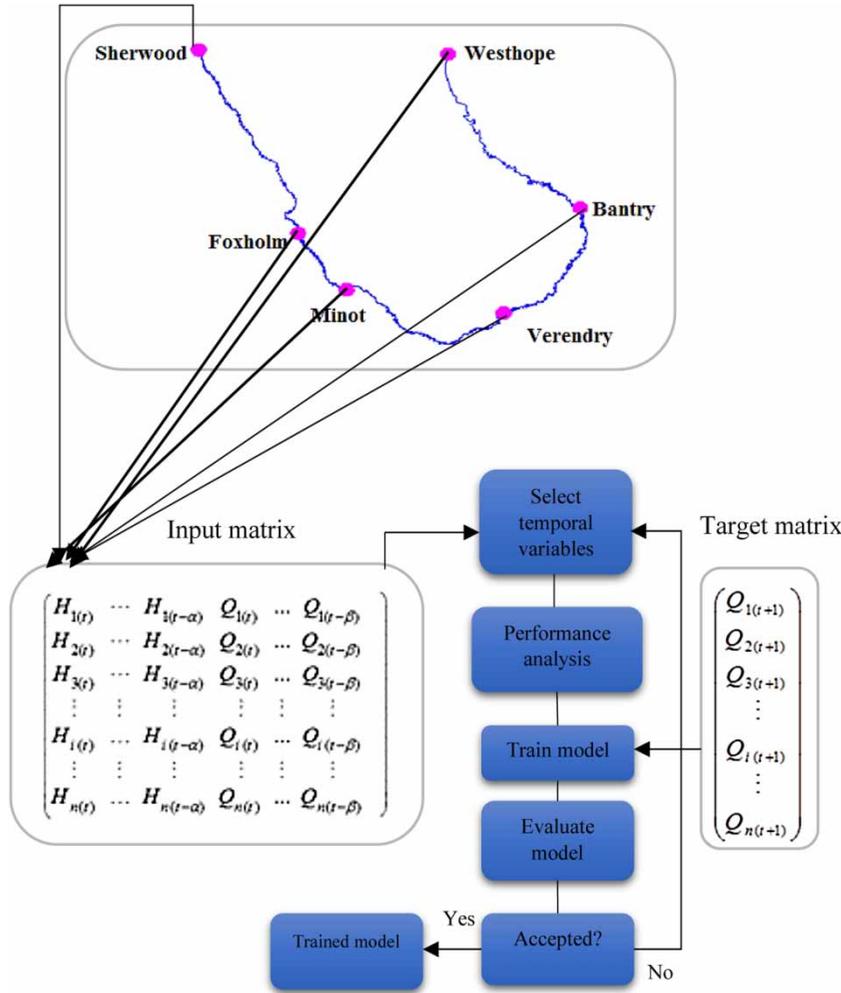
Figure 3 | Schematic of predicting via multi-station model. The figure displays the input data selection and process of prediction.

this property, WT can be used to analyze the time–frequency characteristics of any kind of time series. In recent years, WT has been widely used for the analysis of many hydro-meteorological time series (Nourani et al. 2013). As the mother wavelet moves across the time series during the WT process, it generates several coefficients that represent the similarity between the time series and the mother wavelet (at any specific scale). There are two main types of WT: continuous and discrete. Use of the continuous wavelet transform (CWT) can generate a large number of (often unnecessary) coefficients, making its use and interpretation more complicated. On the other hand, the discrete wavelet transform (DWT) method simplifies the transformation process while

still providing a very effective and precise analysis, since DWT is normally based on the dyadic calculation. DWT coefficients can be calculated by the following equation:

$$W_{\psi}(a, b) = \frac{1}{(2)^{a/2}} \sum_{t=0}^{N-1} x(t)\psi\left(\frac{t}{2^a} - b\right) \tag{2}$$

where 2^a represents the dyadic scale of the DWT. Applying DWT to a time series decomposes that time series into two ancillary time series shape components, called the approximation (A) and detail (D) components. Component A comprises the large-scale, low-frequency component of the

time series, while component D represents the small-scale, high-frequency component.

Ensemble empirical mode decomposition (EEMD)

EEMD was proposed to solve the mode mixing issue of empirical mode decomposition (EMD) which specifies the true intrinsic mode function (IMF) as the mean of an ensemble of trials (Wu & Huang 2009). Each trial consists of the decomposition results of the signal plus a white noise of finite amplitude. This new method is based on the insight from recent studies of the statistical properties of white noise (Wu & Huang 2004), which showed that the EMD method is an effective self-adaptive dyadic filter bank when applied to the white noise. On the other hand, studies demonstrated that noise could help data analysis in the EMD method. All these investigations promote the advent of the EEMD method.

The principle of the EEMD algorithm is the following: the added white noise would populate the whole time–frequency space uniformly with the constituting components of different scales. When a signal is added to this uniformly distributed white noise background, the components in different scales of the signal are automatically projected onto proper scales of reference established by the white noise in the background. Because each of the noise-added decompositions consists of the signal and the added white noise, each individual trial may certainly produce very noisy results, but the noise in each trial is different in separate trials (Wang *et al.* 2013; Roushangar & Alizadeh 2018). Thus, it can be decreased or even completely canceled out in the ensemble mean of enough trials. The ensemble mean is treated as the true answer because the only persistent part is the signal as more and more trials are added in the ensemble.

Shannon entropy and mutual information (MI) and selecting dominant subseries

Shannon formulated entropy (H), now called the Shannon entropy or information content, for a discrete random variable of X with sample size of N (bin number), which takes values $x_1; x_2; . . .; x_N$ with probabilities of $p_1, p_2, . . ., p_N$,

respectively, as (Shannon 1948; Roushangar *et al.* 2020):

$$H(x) = H(p) = -\sum_{i=1}^{N} p(x_i)[p(x_i)] \tag{3}$$

Feature selection is more appropriate by MI than other linear correlation measuring tools since sometimes there may be a weak linear but strong nonlinear relationship between input(s) and output(s). Such a usage of MI along with SVM is favorable since both tools are beneficial in finding nonlinear properties of features. MI between two random variables X and Y is defined as (Yang *et al.* 2000):

$$MI(A, B) = H(A) + H(B) - H(A, B) \tag{4}$$

with $H(A)$ and $H(B)$ being the entropy of A and B, respectively, and $H(A,B)$ their joint entropy as:

$$H(A, B) = -\sum_{a \in A} \sum_{b \in B} P_{AB}(a, b)\log_{AB}(a, b) \tag{5}$$

Efficiency criteria

In this study, the total data were separated into calibration and verification sets. Two different criteria were selected to evaluate the revenue of the proposed forecasting methods: the root mean square error (RMSE) and the determination coefficient (DC). The RMSE and DC were applied to exhibit discrepancies between predicted and observed values (Adamowski & Chan 2011). The RMSE were used to quantify the modeling accuracy by generating a positive value. The RMSE grows from zero for perfect forecasts through large positive values as the differences between forecasts and observations become increasingly large. Obviously, a high value for DC (up to 1) and small value for RMSE indicate high efficiency of the model. Legates & McCabe (1999) indicated that a hydrological model can be sufficiently evaluated by these statistics. These measures are not oversensitive to extreme values (outliers) and are sensitive to additive and proportional differences between model predictions and observations. Therefore, correlation-based measures (e.g., DC statistic) can indicate that a model is a good predictor (Legates & McCabe 1999).

RMSE is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points are; RMSE is a measure of how spread out these residuals are. In other words, it tells you how concentrated the data are around the line of best fit. RMSE is commonly used in climatology, forecasting, and regression analysis to verify experimental results. On the other hand, DC is a statistic used in the context of statistical models whose main purpose is either the prediction of future outcomes or the testing of hypotheses, on the basis of other related information. It provides a measure of how well observed outcomes are replicated by the model, based on the proportion of total variation of outcomes explained by the model. Most researchers use these statistics since they provide thorough statistical information about models (Legates & McCabe 1999; Roushangar & Alizadeh 2018).

Input and output variables are normalized by scaling between 0 and 1 to eliminate their dimensions and to ensure that all variables receive equal attention during calibration of the model. Therefore, RMSE values are to be transformed between 0 and 1 (Nourani *et al.* 2016).

RESULTS AND DISCUSSION

The RC is an empirical model which extracts information from recorded stage values. The RC was applied for all six hydrometric stations. As an instance, Figure 4(a) and 4(b) demonstrate the RC for Sherwood and Minot stations. Results of RC using least square method led to the following results:

$$Q_{(Sherwood)} = 2.7422H^{2.8127}$$

$$Q_{(Minot)} = 0.0281H^{11.313}$$

Results of modeling are demonstrated in Figure 4(c). Based on the results it was observed that results need to be strengthened in terms of DC, RMSE, and MAE. We are going to discuss the results of modeling using SVM in the next section.

Optimizing of SVM parameters

In this study, SVM was used to predict the discharge values for hydrometric stations of the Souris River. A correct use of SVM is to select an appropriate kernel function and tune the related hyper-parameters. A number of kernels were discussed by researchers, however, studies suggested the effectiveness of radial basis function (RBF) kernel in the case of SVM in the majority of civil engineering applications (Gill *et al.* 2006; Goel & Pal 2009). For a fair comparison of results, a RBF kernel was utilized with SVM in this study, wherein, σ is a kernel specific and γ is margin parameter. The tuned parameters (σ, γ) were optimized via grid search procedure. Therefore, we will call the model optimized SVM (OSVM).

The RBF kernel function was selected as the core tool of OSVM used for all the rest of the models. The OSVM needs the choice of three crucial variables, that are ε, and (γ), in which γ is a variable of the RBF kernel function. In order to implement the above-mentioned variables, a systematic
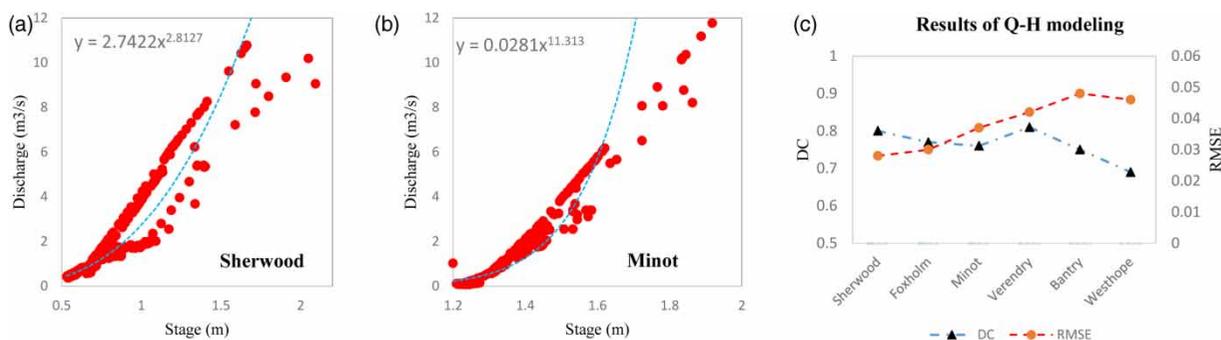

Figure 4 │ (a) Results of modeling via RC for Sherwood station, (b) results of modeling via RC for Minot station, and (c) overall results of RC for all stations.

grid search of the variables employing cross-validation on the training stage has been performed (OSVM procedure). To consider this grid search, a normal range of variables settings are considered. Initially, modified values of ε and C for a determined γ were obtained and then γ was altered. Performance optimization procedures were applied to find the best values. Figure 5 illustrates the statistics parameters' values of various Gamma values of the SVM model (fed with the input configuration). Accordingly, optimal parameters were found out for the rest of the models.

A data-driven model's performance mostly depends on its structure. Hence, in this study, a conceptual model was developed to predict the Q-H process of the Souris River using only one model. First, different types of classic and wavelet-based modeling strategies based on temporal features (i.e., stage-discharge time series) were developed to assess the connection between discharge and river stage over the Souris River. Next, a unique multi-station model was constructed for the hydrometric stations.

Results of multi-station model

The river flow process in any cross section of a river system can be characterized as the function of various variables, such as discharge, catchment and river physical characteristics. The relationship between river discharge and influential variables can be expressed by:

$$Q(t+1) = f(x_{t+1}) + \varepsilon_{t+1} \tag{6}$$

where $Q(t+1)$ denotes the river discharge in any cross section of the river system, ε_{t+1} is the random error, $x(t+1)$ is the input vector, which may include many variables such as
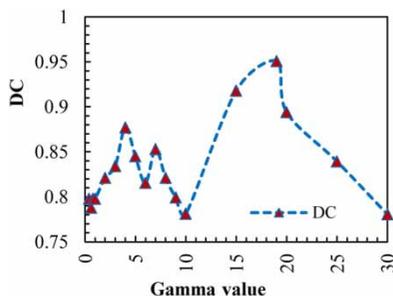
antecedent flow data and other relevant variables at various time lags, but in the present application, it consists of combinations of antecedent daily Q-H records. It is evident that the training datasets should cover all the characters of the problem in order to capture effective estimation. One of the most important steps in developing a satisfactory forecasting model is the selection of the input variables. Hence, different combinations of the antecedent Q-H are used to construct the appropriate input structure for all scenarios.

In order to determine input variables for AI-based models, the input data were extracted based on the principle underlying the RC, in which different model structures were selected by a combination of stage and/or discharge variables including: H_t, H_{t-1}, $H_{t-2, \ldots}$ and H_{t-7} representing stage values at times (t), $(t-1)$, $(t-2)$, … and $(t-7)$ and Q_{t-1}, $Q_{t-2\ldots}$ and Q_{t-7} representing discharge values at times $(t-1)$, $(t-2)$, … and $(t-7)$. Models 1–3 extract information from stage values alone; Models 4–6 extract information from discharge values alone with a structure similar to autocorrelation; and Models 7–10 extract information from both stage and discharge values. While it is clear that the RC is not the only technique, Table 2 shows the possible different choices of dependent variables for each modeling technique, but these are required to have optimized structures by avoiding over-fitting. In the present study, only the best structures are used to report the performance of the proposed models. The performance measures outlined in 'Performance criteria' were used to select the particular model structure but other techniques will be used to assist the choice, e.g., scatter diagram, or plotting difference between individual model performances. Based on the correlation analysis and modeling outcome Comb. 1, 2, 8, and 10 had better performance and this is going to be explained.

Figure 5 │ Statistics variables via γ values to obtain SVM optimization (OSVM).

Table 2 │ Input dataset applied in various models

Input number	Input definition	Input number	Input definition
1	$H_{(t)}$	6	$Q_{(t)}$, $Q_{(t-1)}$, $Q_{(t-2)}$
2	$H_{(t)}$, $H_{(t-1)}$,	7	$Q_{(t)}$, $H_{(t)}$
3	$H_{(t)}$, $H_{(t-1)}$, $H_{(t-2)}$	8	$Q_{(t)}$, $Q_{(t-1)}$, $H_{(t)}$
4	$Q_{(t)}$	9	$Q_{(t)}$, $Q_{(t-1)}$, $H_{(t)}$, $H_{(t-1)}$
5	$Q_{(t)}$, $Q_{(t-1)}$	10	$Q_{(t)}$, $Q_{(t-1)}$, $Q_{(t-2)}$, $H_{(t)}$, $H_{(t-1)}$

For all determined combinations, t demonstrates the time step (daily step), Q represents discharge features and H represents river stage features. The output consists of sole variable, i.e., Q at future time steps $(Q(t+1))$. A SVM through RBF kernel was trained for all input combinations to predict 1-day-ahead discharge employing normalized discharge and river stage features for hydrometric stations. In order to determine the SVM results, the outcomes based on DC values as overall results have been demonstrated for different input combinations in Figure 6. According to Figure 6, the Combs. 1 and 2 did not perform adequately in the calibration and verification steps. It can be inferred

that target discharge could not correlate with river stage-discharge values of 1 or 2 days ago. Among the Combs. 3 and 4, Comb. 4 could predict the discharge relatively accurately. It could be concluded that $Q(t+1)$ established a good relation with Q and H in previous time steps. Comb. 3 also employed 1e and 2 days ago values of river stage-discharge features. Although results showed improvement using Comb. 4 in comparison to other antecedent-based combinations, nevertheless, generally results are not assured. Towards recovering the efficiency of classic modeling strategy, WT was conjugated to the modeling process in this step. The WT was employed to decompose discharge and river stage
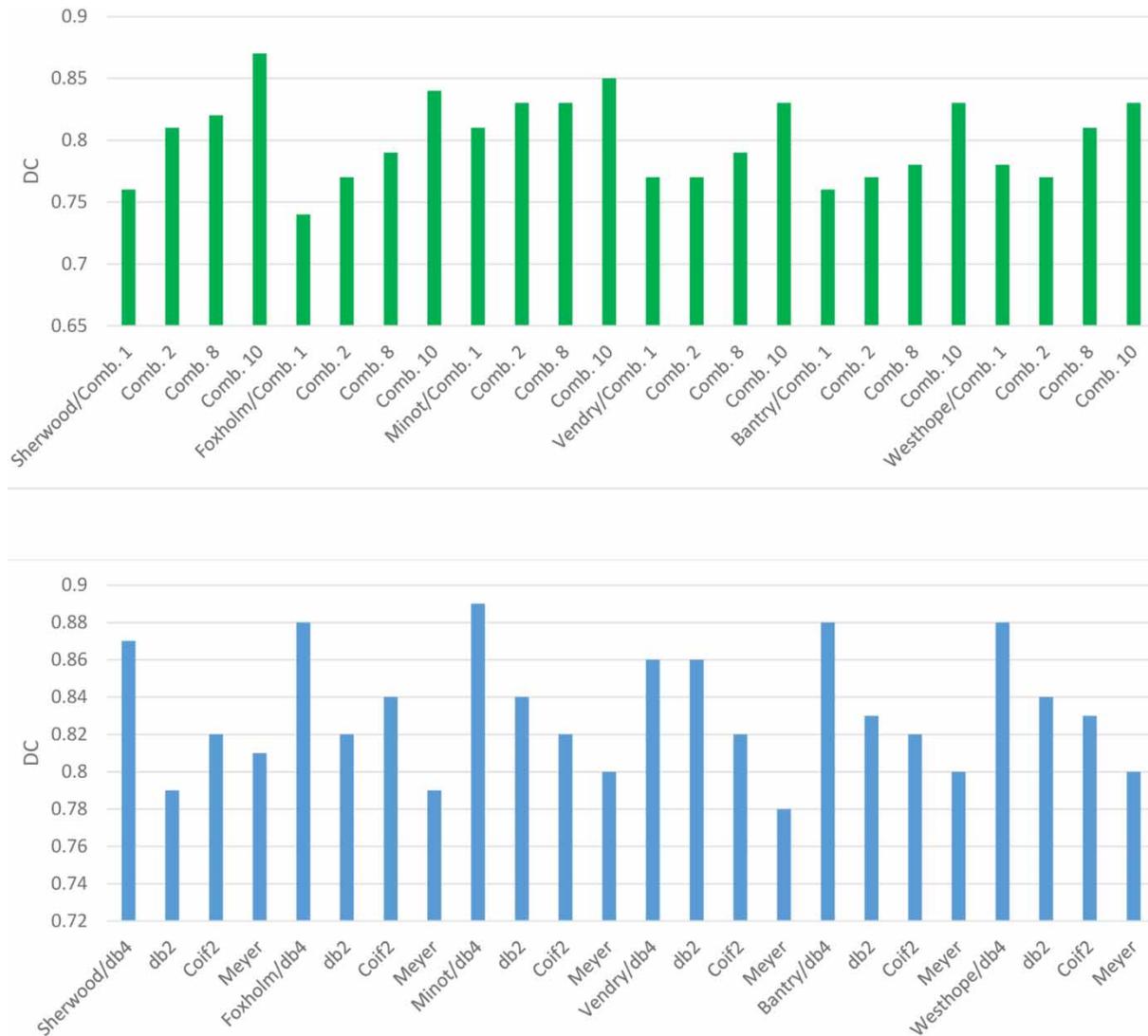

Figure 6 │ Sensitivity analysis of classic and WT-based LSSVM model based on DC values.

time series. Using a trial-and-error process to obtain the optimum decomposition level eventuated to select level 5 (six subseries, one approximation and five details). In order to have a comprehensive overview on decomposition level, initially the following formula (which offers a minimum level of decomposition) was employed (Aussem *et al.* 1998). As studied by Aussem *et al.* (1998), on finding the correct decomposition level, initially the following formula which offers a minimum level of decomposition was introduced:

$$L = \text{int}[\log N] \tag{7}$$

where L is decomposition level and N is the number of time series features. This experimental equation was derived for fully autoregressive signals, and only considering time series length without considering seasonal efficacy of the process. In order to handle the seasonality of the process, decomposition level up to 5 (25; approximately monthly mode) was selected as the optimum level to decompose. Therefore, the seasonality of the process up to one month

could be handled by the model. Due to the proportional relationship between amount of discharge and river stage, these signals were supposed to have the same seasonality level and both time series were decomposed at the same level (i.e., level 5). By selecting Daubechies-2 (db2), Daubechies-4 (db4), Meyer, and coif2 mother wavelets, both discharge and river stage time series were decomposed. In order to prevent probable issues from overtraining or other problems, it was attempted to select the dominant subseries from the decomposed time series. In this study, feature selection was performed using MI approach. MI as a supervised measuring tool was computed with respect to input and target values. The input matrix was established by considering general ranking of subseries according to MI values. Results are presented in Figure 6. According to Figure 6, the db(4,5) outperformed the classic modeling (antecedent input) and other wavelet-based inputs. Also, Figure 7 represents the scatterplots of observed against predicted models (best structures in verification period) for Foxholm station.

The proposed multi-station model, in addition to Q-H time series, took advantage of spatially varying parameters
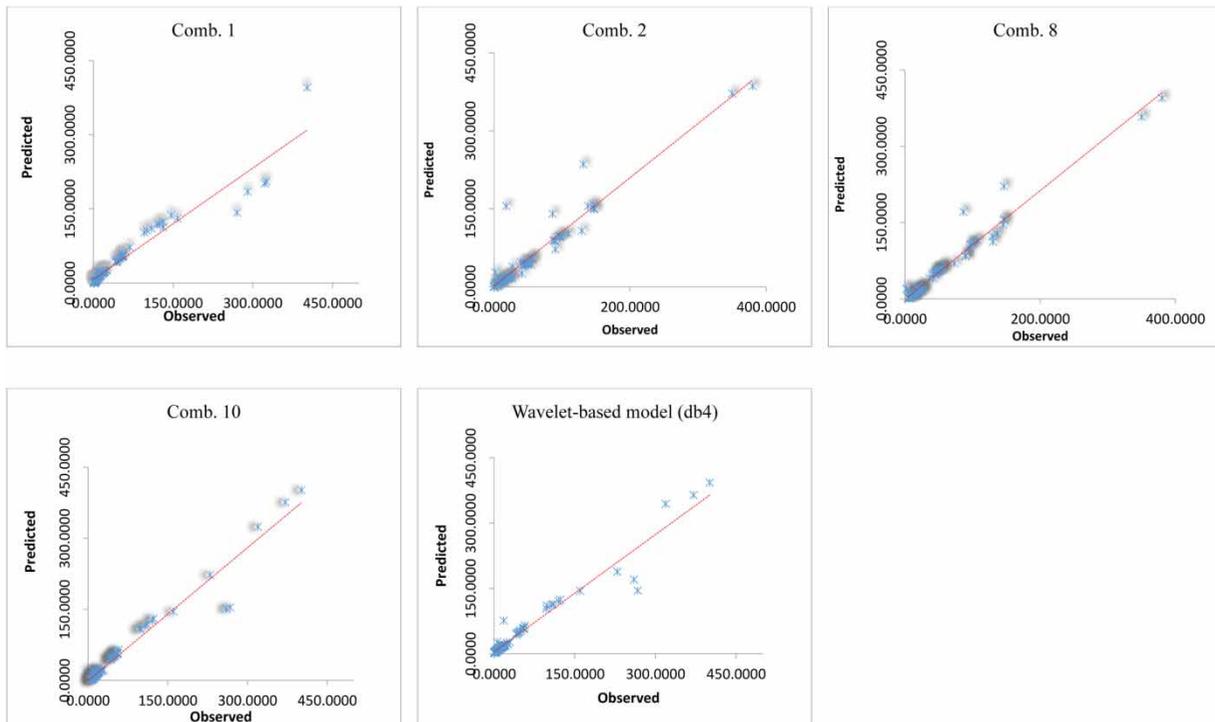
Figure 7 │ Scatterplots of modeling of classic and WT-based LSSVM model (best combination) for Foxholm station in Souris River (verification period).

as model inputs. The geomorphological variables used in the multi-station model had an issue of incompatibility in dimension. In order to avoid possible conflicts with incorrect dimensionality of obtained formulations, it was suggested to use the dimensionless values. This is a prevalent scientific application where units of measurements are effectively omitted through the introduction of dimensionless ratios (Nourani *et al.* 2013). In this way, the dimensionless geomorphological parameters were applied in the model. Therefore, the model construction for each sub-basin could be represented as:

$$Qi(t + 1) = f\left[H_i(t), H_i(t - 1), Q_i(t), Q_i(t - 1), \frac{A_i}{A_t}\right] \quad (8)$$

In Equation (8), *i* refers to the sub-basin number ($i = 1$, 2, 3. . . 6). $H_i(t - \alpha)$ is river stage values with 1-day lag time at *ith* sub-basin. A_i is the drainage area for each sub-basin and A_t is the total watershed area. In order to develop the multi-station model for the entire watershed, the data of all six stations were imposed to the SVM framework. Since the *Q-H* values are sorely affected by recent conditions of watershed at daily time scale, only values with 2-days' lag times (i.e., $\alpha = 2$) were used in the multi-station model. By

considering the aforementioned issues for the SVM modeling, the multi-station model was calibrated and verified.

In order to find the efficient structure of the proposed multi-station model, sensitivity analysis was performed. To this end, six best input combinations were considered to create the input matrix (Figure 8). Figure 8 shows the multi-station model's performance for different input combinations. In order to explore the efficiency of physical parameters, Comb. (1) was designed using only temporal variables (i.e., discharge-gauge stage data).

According to the obtained results, imposing dimensionless physical parameters along with temporal dataset caused an increase in modeling accuracy (Combs. (2)–(6)). Generally, Combs. (2)–(6) provided higher efficiency in comparison to Comb. (1). Conjugating WT to temporal features in Comb. 5 gave an incredible insight into temporal features which caused Comb. 5 to have a better outcome. By considering improvement in Comb. 5, WT-EEMD-MI pre-processing approach was applied to find more dominant subseries. For this reason, *detail 1* and *2* of *db(4,5)* was further decomposed using EEMD. MI was used to select the appropriate IMFs from decomposed subseries. According to Comb. 6, *imf 8* and *9* from *detail subseries 1* and *imf 6* and *7* from *detail subseries 2* were selected via the proposed pre-processing approach. Figure 9 presents
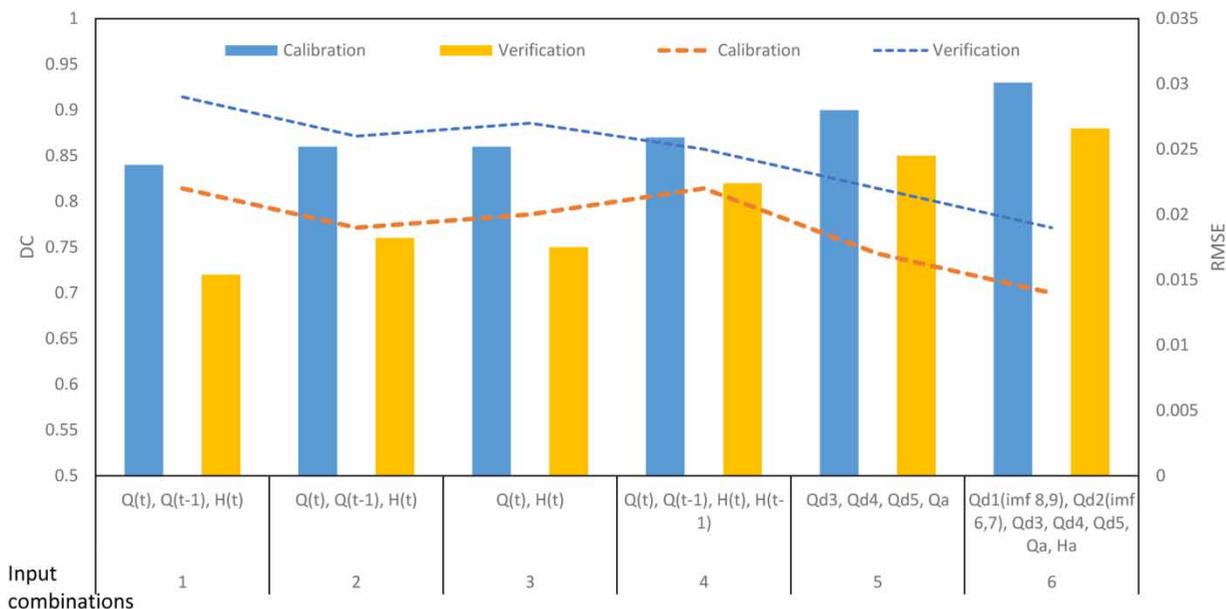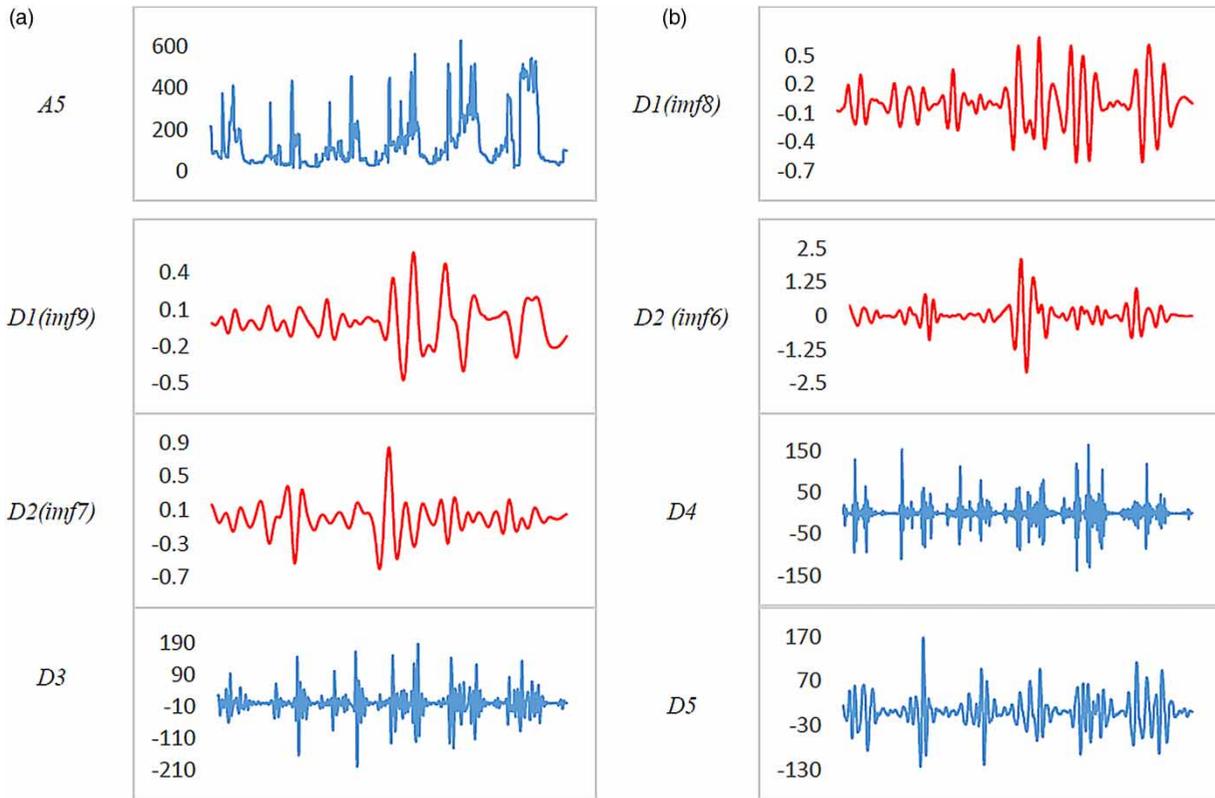

Figure 8 | Results of multi-station modeling in terms of DC and RMSE for various input combinations.

Figure 9 │ Decomposed time series using *db4-EEMD-MI* (eight subseries in multiple-point form).

decomposed time series of the discharge values by db(4,5)-EEMD-MI for the multiple station form and Figure 10 the predicted time series using the proposed model (Comb. 6).

By comparing the obtained results for Combs. (1)–(6) (Figure 7), it could be inferred that Comb. (6) was more efficient than others. The input structure of this combination
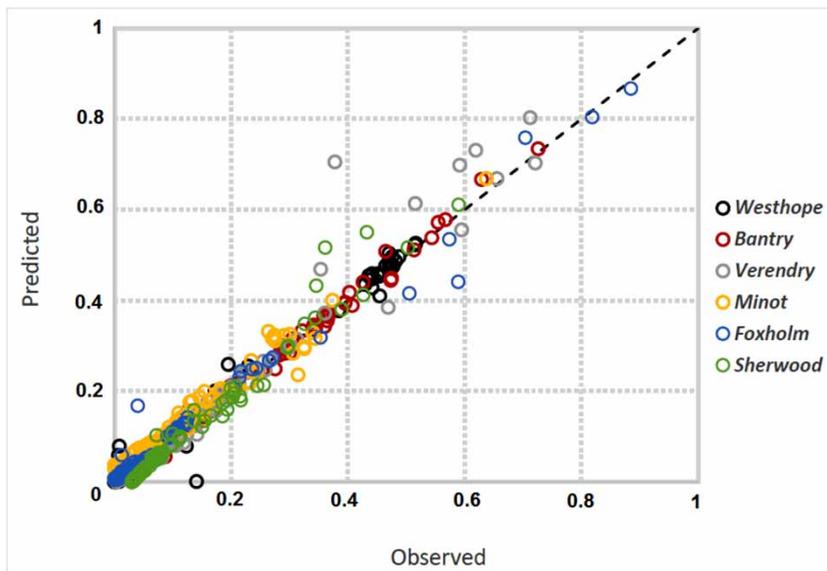
Figure 10 │ Scatterplot of multi-station forecasting of Souris River discharge (separated by colors). Please refer to the online version of this paper to see this figure in color: http://dx.doi.10.2166/wcc.2020.006.

consisted of two temporal variables. In Figures 8 and 9, *Qd1* means *detail subseries 1* and *Qd2(imf 6, 7)* means decomposed *detail subseries 2 via* EEMD and selected *imf 6* and *7 via* MI and the rest is the same.

The proposed pre-processing approach comprised WT, EEMD, and MI methods. WT decomposed the time series into subseries which were grown in a dyadic form. Further breakdown was performed via EMMD and caused subseries with more stationary properties and less noise to be obtained. Since decomposing in two levels led to a number of subseries (about 40–50 subseries), feeding them into SVM might have caused over-training and other issues. Therefore, MI was used to decrease the number of subseries and omit noise and select dominant subseries. Using WT-EEMD-WT caused improvement in results, about 3–19% on average in comparison with other models.

Other machine learning methods

ARIMAX, ELM, ANFIS, and ANN-RBF are other remarkable forecasting approaches which were employed as benchmark models in order to find out the revenue of the proposed hybrid models via OSVM. Such acknowledged

models' performance could be good criteria to find OSVMs' capability as a nonlinear prediction approach. The structure of ARIMAX and ANN-RBF was optimized by trial-and-error process; on the other hand, ELM is a self-tuning regression tool. In a similar manner and using the selected input combinations, the ANFIS model was also used for prediction of the river discharge. In the ANFIS modeling, two substantial points must be considered: first, the ANFIS architecture (e.g., number and type of MFs) and second, the calibration procedure (e.g., selection of iteration number, efficiency criteria). Based on the models, the Gaussian membership function (gaussmf) was selected because of its accurate results. Figure 11 demonstrates the overall results by OSVMs via different scenarios compared with ARIMAX, SVM, ELM, ANFIS, and ANN-RBF. The best input in classic modeling for OSVM (Comb. 10) was used for the selected tools to have a fair comparison. According to the results from Figure 11, it is concluded that OSVMs' outcome led to a superior performance in comparison to other regression tools. Also, it was concluded that ELM and ANFIS performed well, respectively, and ANN-RBF performed relatively weakly and ARIMAX was the weakest among all approaches.
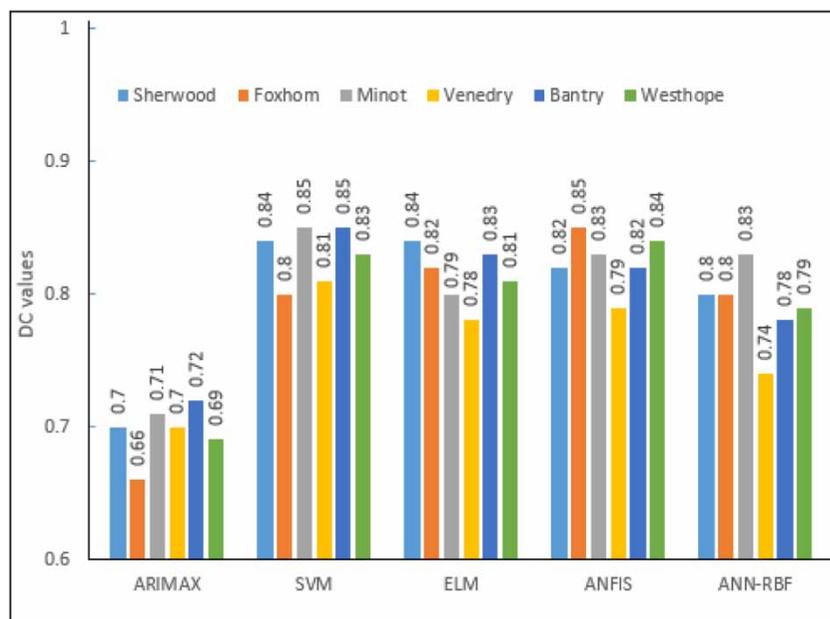

Figure 11 │ Comparison of results of various machine learning models with OSVM (SVM).

CONCLUSIONS

In this research, capability of a two-stage conceptual model was proposed and verified to discover its quantitative and qualitative aspects in prediction of river Q-H process. At the first stage, a strong novel DWT-EEMD based pre-processing approach was used to capture signals with more stationary properties. Application of the model led to interesting results which can be presented in two different categories: proposed hydrological models and proposed pre-processing approach. A conceptual strategy was proposed to study the Q-H relation of the river, namely, a multi-station model. All proposed models have unique capabilities and some deficiencies to be discussed.

- Rating curve: RC is a physical-based model which is used to extract information from river stage datasets. The best outcome was captured for Verendry station ($DC = 0.81$, $RMSE = 0.042$, and $MAE = 0.034$), and the worst result was captured for Westhope station ($DC = 0.69$, $RMSE = 0.046$, and $MAE = 0.0.36$).
- Multi-station model: The proposed model is in continuous and multi-station form and since this model took advantage of watershed morphology data (upstream drainage area of hydrometric stations) it behaved like a spatio-temporal model that could identify and predict the discharge values more accurately when employed with the proposed WT-EEMD-MI pre-processing approach. The multi-station model could handle the uncertainty and nonlinearity of the Souris River by exploring the interaction of stations by using the spatial and temporal variables and it is capable of being applied to predict discharge values in an arbitrary point of the watershed because of its structure (input data, form of prediction) ($DC = 0.89$ and $RMSE = 0.18$, in the verification period).
- WT-EEMD-MI: The proposed pre-processing approach consisted of wavelet transform, ensemble empirical mode decomposition, and mutual information approaches. WT decomposes the time series into subseries which grow in a dyadic form. The subseries carry the seasonal properties which strengthen any AI-based model to predict the time series more accurately. Further breakdown via EMMD led to obtaining subseries with more stationary properties

and less noise in data. Since decomposing in two levels led to a number of subseries (about 40–50 subseries), feeding them into SVM might lead to over-training or other issues. Therefore, MI was used to select the dominant subseries and decrease the number of subseries and omit noise. Using the WT-EEMD-MI caused an improvement in results (about 3–19%) in comparison with other models.

This study also proved the advantages of a new conceptual model. The multi-station model which used multiple-point measures, was designed to estimate the river discharge in multiple-station form by training only one SVM model instead of several SVMs. This strategy was capable of forecasting at the point of interest in a river or watershed by including temporal properties of the case study.

Developing such a model could be very useful in terms of the accuracy of least time required for model development. For future studies, it is suggested to apply hydraulic properties of the Q-H model such as Froude number, Reynolds number, Manning coefficient, etc., to verify these variables' capability in increasing the accuracy of multi-stations' model.

ACKNOWLEDGEMENTS

REFERENCES

Abrahart, R. J., Anctil, F., Coulibaly, P., Dawson, C. W., Mount, N. J., See, L. M., Shamseldin, A. Y., Solomatine, D. P., Toth, E. & Wilby, R. L. 2012 Two decades of anarchy? emerging themes and outstanding challenges for neural network river forecasting. *Progress in Physical Geography* **36** (4), 480–513.

Adamowski, J. & Chan, H. F. 2011 A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology* **407**, 28–40.

Anmala, J., Zhang, B. & Govindaraju, R. 2000 Comparison of ANNs and empirical approaches for predicting watershed

runoff. *Journal of Water Resources Planning and Management* **126**, 156–166.

Ashu, J. & Sanaga, S. 2006 Integrated approach to model decomposed flow hydrograph using artificial neural network and conceptual techniques. *Journal of Hydrology* **317**, 291–306.

Aussem, A., Campbell, J. & Murtagh, F. 1998 Wavelet-based feature extraction and decomposition strategies for financial forecasting. *Journal of Computational Finance* **6** (2), 5–12.

Azamathulla, H. M., Chang, C. K., Ab Ghani, A., Ariffin, J., Zakaria, N. A. & Abu-Hasan, Z. 2009 An ANFIS-based approach for predicting the bed load for moderately sized rivers. *Journal of Hydro-Environmental Research* **3** (1), 35–44.

Babovic, V. 2009 Introducing knowledge into learning based on genetic programming. *Journal of Hydroinformatics* **11** (3–4), 181–193.

Bai, Y., Wang, P., Xie, J. J., Li, J. T. & Li, C. 2015 An additive model for monthly reservoir inflow forecast. *Journal of Hydrologic Engineering* **20** (7). HE.1943-5584.0001101.

Chen, J. & Adams, B. J. 2006 Integration of artificial neural networks with conceptual models in rainfall–runoff modeling. *Journal of Hydrology* **318**, 232–249.

Chen, S. T., Yu, P. S. & Tang, Y. H. 2010 Statistical downscaling of daily precipitation using support vector machines and multivariate analysis. *Journal of Hydrology* **385**, 13–22.

Chou, C.-M. 2013 Enhanced accuracy of rainfall–runoff modeling with wavelet transform. *Journal of Hydroinformatics* **15** (2), 392–404.

Corzo, G. A., Solomatine, D. P., Hidayat de Wit, M., Werner, M., Uhlenbrook, S. & Price, R. K. 2009 Combining semi-distributed process-based and data-driven models in flow simulation: a case study of the Meuse river basin. *Hydrology and Earth System Sciences* **13**, 1619–1634.

Coulibaly, P., Anctil, F. & Bobee, B. 1999 Hydrological forecasting with artificial neural networks: the state of the art. *Canadian Journal of Civil Engineering* **26**, 293–304.

Farajzadeh, J. & Alizadeh, F. 2018 A hybrid linear–nonlinear approach to predict the monthly rainfall over the Urmia Lake watershed using wavelet-SARIMAX-LSSVM conjugated model. *Journal of Hydroinformatics* **20** (1), 246–262.

Farhan, S. B., Zhang, Y., Aziz, A., Gao, H., Ma, Y., Kazmi, J., Shahzad, A., Hussain, I., Mansha, M., Umar, M., Nasir, J., Shafiq, M., Farhan, Y., Shaik, S., Bin Zamir, U., Asad, F. & Ahmed, R. 2020 Assessing the impacts of climate change on the high altitude snow- and glacier-fed hydrological regimes of Astore and Hunza, the sub-catchments of Upper Indus Basin. *Journal of Water and Climate Change* **11** (2), 479–490. https://doi.org/10.2166/wcc.2018.107.

Foody, G. M. 2010 Supervised image classification by MLP and RBF neural networks with and without an exhaustively defined set of classes. *International Journal of Remote Sensing* **25**, 3091–3104.

Ghorbani, M. A., Khatibi, R., Goel, A., FazeliFard, M. H. & Azani, A. 2016 Modeling river discharge time series using support vector machine and artificial neural networks. *Environmental Earth Sciences* **75**, 685.

Gill, M. K., Asefa, T., Kemblowski, M. W. & Makee, M. 2006 Soil moisture prediction using Support Vector Machines. *Journal of the American Water Resources Association* **42** (4), 1033–1046.

Goel, A. & Pal, M. 2009 Application of support vector machines in scour prediction on grade-control structures. *Engineering Application of Artificial Intelligence* **22** (2), 216–223.

Huang, G.-B., Zhu, Q.-Y. & Siew, C. K. 2004 Extreme learning machine: A new learning scheme of feedforward neural networks. IEEE International Proceedings Joint Conference on Neural Networks, Vol. 2.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. L. C., Shih, H. H., Zheng, Q. N., Yen, N. C., Tung, C. C. & Liu, H. H. 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A Mathematical, Physical, and Engineering Sciences* **454**, 903–995.

Jang, S. H. R. 1993 ANFIS Adaptive-Network-based Fuzzy Inference System. *IEEE Transactions on Systems Man and Cybernetics*. 23 (3), 665–685. doi:10.1109/21.256541

Kisi, O. & Cobenar, M. 2009 Modeling river stage-discharge relationships using different neural network computing techniques. *Clean* **37** (2), 160–169.

Labat, D. 2005 Recent advances in wavelet analyses: part 1. A review of concepts. *Journal of Hydrology* **314**, 275–288.

Legates, D. R. & McCabe, G. J. 1999 Evaluating the use of goodness of fit measures in hydrologic and hydroclimatic model validation. *Water Resources Research* **35**, 233–241.

Maheswaran, R. & Khosa, R. 2013 Wavelets-based non-linear model for real-time daily flow forecasting in Krishna River. *Journal of Hydroinformatics* **15** (3), 1022–1041.

Mallat, S. G. 1998 *A Wavelet Tour of Signal Processing*, 2nd edn. Academic Press, San Diego, CA, USA.

Meshgi, A., Schmitter, P., Chui, T. F. M. & Babovic, V. 2015 Development of a modular streamflow model to quantify runoff contributions from different land uses in tropical urban environments using genetic programming. *Journal of Hydrology* **525**, 711–723.

Meyer, P. E., Schretter, C. & Bontempi, G. 2008 Information-theoretic feature selection in microarray data using variable complementarity. *IEEE Journal of Selected Topics in Signal Processing* **2**, 261–274.

Nourani, V., Komasi, M. & Alami, M. T. 2013 Geomorphology-based genetic programming approach for rainfall–runoff modeling. *Journal of Hydroinformatics* **15** (2), 427–445.

Nourani, V., Alizadeh, F. & Roushangar, K. 2016 Evaluation of a two-stage SVM and spatial statistics methods for modeling monthly river suspended sediment load. *Water Resources Management* **30**, 393–407.

Partal, T. 2009 Modelling evapotranspiration using discrete wavelet transform and neural networks. *Hydrological Processes* **23**, 3545–3555.

Potter, C., Shupe, P., Gross, P., Genovese, V. & Klooster, S. 2010 Modeling river discharge rates in California watersheds. *Journal of Water and Climate Change* **1** (1), 36–54.

Roushangar, K. & Alizadeh, F. 2018 Entropy-based analysis and regionalization of annual precipitation variation in Iran during 1960–2010 using ensemble empirical mode decomposition. *Journal of Hydroinformatics* **20** (2), 468–485.

Roushangar, K., Garekhani, S. & Alizadeh, F. 2016 Forecasting daily seepage discharge of an earth dam using wavelet–mutual information–Gaussian process regression approaches. *Geotechnical and Geological Engineering* **34** (5), 1313–1326.

Roushangar, K., Alizadeh, F. & Nourani, V. 2018 Improving capability of conceptual modeling of watershed rainfall–runoff using hybrid wavelet-extreme learning machine approach. *Journal of Hydroinformatics* **20** (1), 69–87.

Roushangar, K., Alizadeh, F., Adamowski, J. & Saghebian, M. 2020 Exploring the multiscale changeability of precipitation using the entropy concept and self-organizing maps. *Journal of Water and Climate Change* **11** (3), 655–676.

Sang, Y. F. 2013 A review on the applications of wavelet transform in hydrology time series analysis. *Atmospheric Research* **122**, 8–15.

Sannasiraja, S. A., Zhang, H., Babovic, V. & Chand, E. S. 2004 Enhancing tidal prediction accuracy in a deterministic model using chaos theory. *Advances in Water Resources* **27** (7), 761–772.

Sarangi, A., Madramootoo, C. A., Enright, P., Prasher, S. O. & Patel, R. M. 2005 Performance evaluation of ANN and geomorphology–based models for runoff and sediment yield prediction for a Canadian watershed. *Current Science* **89** (12), 2022–2033.

Shannon, C. E. 1948 The mathematical theory of communications, I and II. *Bell System Technical Journal* **27**, 379–423.

Singh, G., Panda, R. K. & Lamers, M. 2015 Modeling of daily runoff from a small agricultural watershed using artificial neural network with resampling techniques. *Journal of Hydroinformatics* **17** (1), 56–74.

Sun, Y., Babovic, V. & Chan, E. S. 2010 Multi-step-ahead model error prediction using time-delay neural networks combined with chaos theory. *Journal of Hydrology* **395** (1–2), 109–116.

Tiwari, M. K. & Chatterjee, C. 2010 Development of an accurate and reliable hourly flood forecasting model using waveletbootstrap-ANN (WBANN) hybrid approach. *Journal of Hydrology* **394**, 458–470.

Tokar, A. S. & Markus, M. 2000 Precipitation–runoff modeling using artificial neural networks and conceptual models. *Journal of Hydrologic Engineering* **5** (2), 156–161.

Venkatesan, P. & Anitha, S. 2006 Application of a radial basis function neural network for diagnosis of diabetes mellitus. *Current Science* **91** (9), 1195–1199.

Wang, W.-C., Xu, D.-M., Chau, K. W. & Chen, S. 2013 Improved annual rainfall-runoff forecasting using PSO–SVM model based on EEMD. *Journal of Hydroinformatics* **15** (4), 1377–1390.

Wang, W. C., Chau, K. W., Xu, D. M. & Chen, X. Y. 2015 Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition. *Water Resources Management* **29** (8), 2655–2675.

Wu, Z. H. & Huang, N. E. 2004 A study of the characteristics of white noise using the empirical mode decomposition method. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **460A**, 1597–1611.

Wu, Z. H. & Huang, N. E. 2009 Ensemble empirical mode decomposition: a noise assisted data analysis method. *Advances in Adaptive Data Analysis* **1**, 1–41.

Yang, H. H., Vuuren, S. V., Sharma, S. & Hermansky, H. 2000 Relevance of time frequency features for phonetic and speaker-channel classification. *Speech Communication* **31**, 35–50.

Yu, X., Liong, S. Y. & Babovic, V. 2004 EC-SVM approach for real-time hydrologic forecasting. *Journal of Hydroinformatics* **6** (3), 209–223.

Zhang, B. & Govindaraju, R. S. 2003 Geomorphology–based artificial neural networks (GANNs) for estimation of direct runoff over watersheds. *Journal of Hydrology* **273**, 18–34.