Journal of Hydroinformatics

Journal of Hydroinformatics Vol 23 No 6, 1165 doi: 10.2166/hydro.2021.105

Evaluating the capability of hybrid data-driven approaches to forecast monthly streamflow using hydrometric and meteorological variables

Fariba Azarpira and Sajad Shahabi *

Department of Water Engineering, Faculty of Civil and Surveying Engineering, Graduate University of Advanced Technology, Kerman, Iran
*Corresponding author. E-mail: s.shahabi@kgut.ac.ir

SS, 0000-0003-1124-7685

ABSTRACT

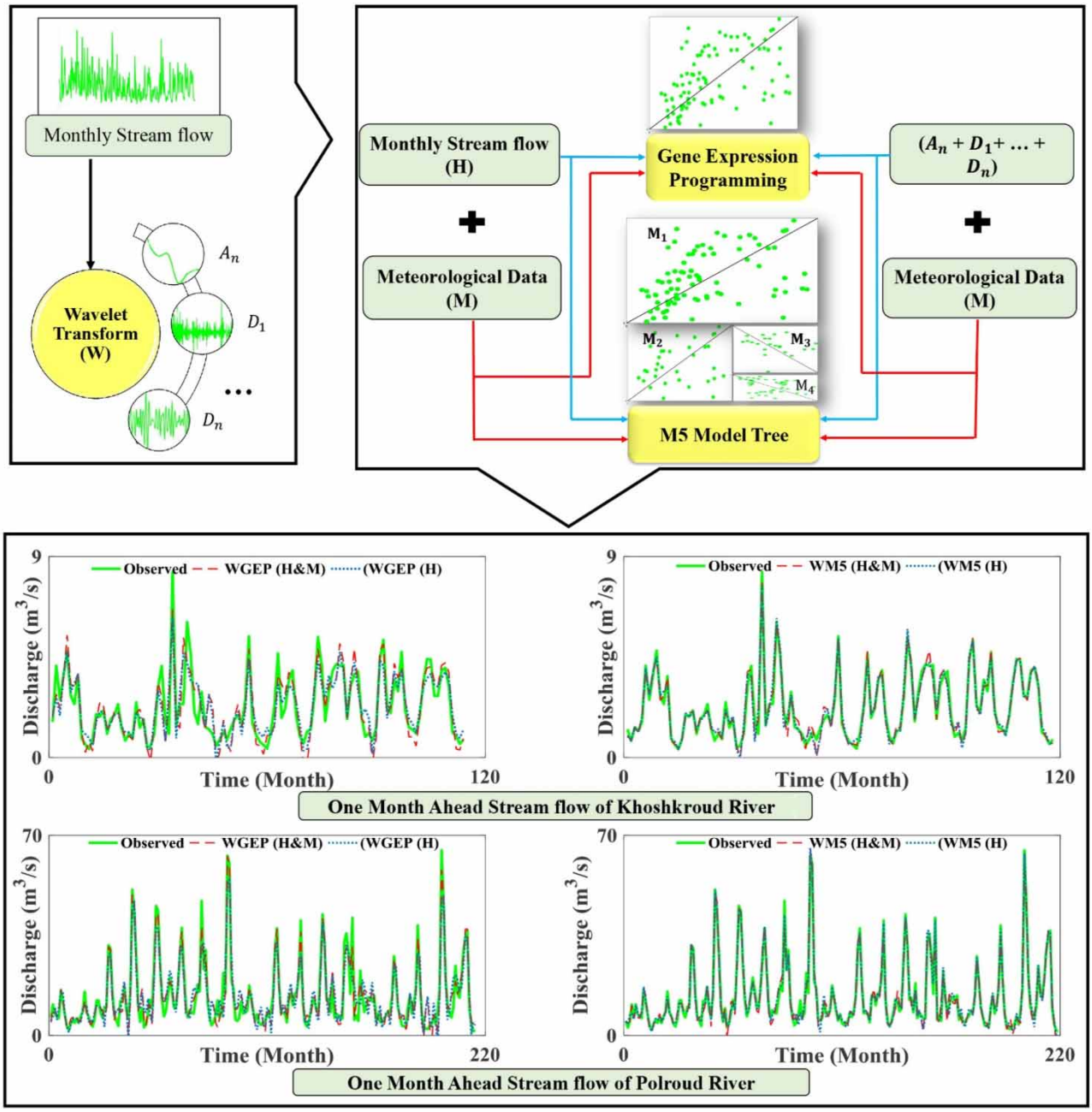
Streamflow forecasting, as one of the most important issues in hydrological studies, plays a vital role in several aspects of water resources management such as reservoir operation, water allocation, and flood forecasting. In this study, wavelet-gene expression programming (WGEP) and wavelet-M5 prime (WM5P) techniques, as two robust artificial intelligence (AI) models, were applied for forecasting the monthly streamflow in Khoshkroud and Polroud Rivers located in two basins with the same names. Results of hybrid AI techniques were compared with those achieved by two stand-alone models of GEP and M5P. Seven combinations of hydrological (H) and meteorological (M) variables were considered to investigate the effect of climatic variables on the performance of the proposed techniques. Moreover, the performance of both stand-alone and hybrid models were evaluated by statistical criteria of correlation of coefficient, root-mean-square error, index of agreement, the Nash–Sutcliffe model efficiency coefficient, and relative improvement. The statistical results revealed that there is a dependency between 'the M5P and GEP performance' and 'the geometric properties of basins (e.g., area, shape, slope, and river network density)'. It was found that a preprocessed technique could increase the performance of M5P and GEP models. Compared to the stand-alone techniques, the hybrid AI models resulted in higher performance. For both basins, the performance of the WM5P model was higher than the WGEP model, especially for extreme events. Overall, the results demonstrated that the proposed hybrid AI approaches are reliable tools for forecasting the monthly streamflow, while the meteorological and hydrometric variables are taken into account.

Key words: gene expression programming, model tree, time series, wavelet

HIGHLIGHTS

- Wavelet-based approaches (WGEP and WM5P) are taken to forecast monthly river flow.
- Preprocessing of time series improves the capability of forecasting models.
- Meteorological parameters increase the model performance, especially for extremes.
- The proposed hybrid approaches successfully improve the performance of models.

GRAPHICAL ABSTRACT



INTRODUCTION

Streamflow forecasting has attracted wide attention among hydrologists owing to the fact that this issue plays a key role in decision-making cases related to water resources management and planning. The first attempts made to forecast the streamflow in various climatic conditions were based on time-series methods such as auto-regressive (AR), autoregressive moving average (ARMA), autoregressive moving integrated average (ARIMA), and *K*-nearest neighborhood (KNN) (Abrahart & See 1998; Cigizoglu 2003; Huang *et al.* 2004; Wu and Chau 2010). In these mathematical techniques, time-dependent variables should have complete time series so that results of forecasting stand at the reliable level. Accordingly, the performance of time-series models has a low potential of streamflow forecasting when time-series information is incomplete (Huang *et al.* 2004). With

the emergence of soft computing models, a broad range of artificial intelligence (AI) techniques was applied to predict the streamflow in rivers. AI methods have shown sufficient capacity to detect the hidden patterns in time series, and they have proven high reliability in comparison with traditional approaches (e.g., Solomatine & Ostfeld 2008; Yaseen *et al.* 2015; Karimi *et al.* 2017; Khairuddin *et al.* 2019; Najafzadeh & Saberi-Movahed 2019; Boucher *et al.* 2020; Hussain & Khan 2020). In the case of hydrological investigations, various meteorological variables (e.g., evapotranspiration, temperature, precipitation, and streamflow) are characterized by time series in which effects of many properties of time series (e.g., noise, sudden, and irrelevant patterns) on the performance of data-driven models (DDMs) may be inventible. There have been various investigations in which preprocessing of time series improved the efficiency level of DDMs. In the early 21st century, wavelet transform (WT) functions, as one of the most efficient preprocessors, have been widely applied to boost the intrinsic capabilities of DDMs in terms of precision level and generalization (e.g., Parasuraman & Elshorbagy 2005; Adamowski 2008; Jacquin & Shamseldin 2009; Yarar 2014; Shoaib *et al.* 2015; Fahimi *et al.* 2017; Shahabi *et al.* 2017; Nourani *et al.* 2019; Zakhrouf *et al.* 2020).

Additionally, some investigations indicated that the use of meteorological data increased the accuracy level of DDMs for the prediction of the peak flows (i.e., Abdollahi *et al.* 2017; Diop *et al.* 2018; Akcakoca & Apaydin 2020). Similarly, Hadi & Tombul (2018) used wavelet-multigene genetic programming (W-MGP) to take advantage of meteorological data to simulate rainfall–runoff. Furthermore, Kratzert *et al.* (2018) proposed a novel DDM utilizing the long short-term memory (LSTM) network to model the rainfall–runoff process of a large number of completely various catchments at the daily time scales. This study showed the LSTM capability as a regional hydrological model in which a model predicts the discharge for the different catchments. Adnan *et al.* (2019) selected streamflow, rainfall, and temperature via autocorrelation and partial autocorrelation. Through their research, this analysis not only forecasted monthly flow accurately but also showed rainfall and temperature alone was adequate inputs for 511 basins in the United States. Tyralis *et al.* (2019) and Chu *et al.* (2020) employed random forest and least absolute shrinkage and selection operator to optimally select climatic features for forecasting the streamflow.

Through the literature review, a broad range of investigations applied hydrological or meteorological data in order to forecast streamflow, while there are a few research in which both types of data were considered for modeling the streamflow. Simultaneous use of hydrological and meteorological variables increased the precision level of AI models in forecasting the streamflow for a long-time period (Hadi & Tombul 2018). Therefore, there is still an essential need for doing research in this respect in order to investigate the effects of both types of information on the accuracy level of AI techniques in the presence of wavelet models as preprocessors. In the present study, monthly streamflow is forecasted for two distinctive basins (Polroud and Khoshkroud) by using wavelet-M5 prime (WM5P) model tree and gene expression programming (GEP) techniques. Additionally, the efficiency of hydrological and meteorological data application as input variables for running the hybrid AI models and the performance of preprocessing process for both basins (with various length of time series or missing data) were investigated. These cases have not been frequently studied for the forecasting streamflow.

To the best of the authors' knowledge, WT is used to preprocess hydrometric and meteorological variables in forecasting monthly river flow. To be more specific, in order to identify the influential factors on the streamflow (discharge), hydrological (discharge) and meteorological (monthly maximum precipitation, average of monthly temperature, atmosphere pressure, and dew point) variables, accumulated from two Polroud and Khoshkroud basins, are considered to train and test M5P and GEP techniques. In addition, the performance of M5P and GEP models are statistically compared and then used to describe the beneficial impacts of various factors (i.e., geometry of basins, the number of recorded data, availability of meteorological data, and the preprocessing stage with various wavelet characteristics) on the proposed DDMs.

MATERIALS AND METHODS

Study area and dataset

In the current study, Khoshkroud basin (50° 25′ E , 36° 55′ N) was utilized as the study area. The elevation of basin varies from 74 to 3,608 m altitude with an average slope of 15% whose the length of the main stream path and basin area is 20 km and 100 km^2, respectively. The time series of monthly streamflow was collected from Bajiguabar hydrometric station for a 32-year period beginning in 1985, provided by Guilan Regional Water Company (GRWC). Figure 1 displays the monthly discharge histogram of Bajiguabar station.

The question is: 'whether the proposed hybrid AI models can accurately forecast the monthly streamflow in any situation or its performance is good only in some ranges of input and output time series?' This is an important challenge of this research. To get rid of this challenge, the Polroud River basin whose characteristics are different from those of Khoshkroud is considered. The basin elevation varies between 52 and 3,931 m with an average slope of 5%. In the case of Polroud River, the

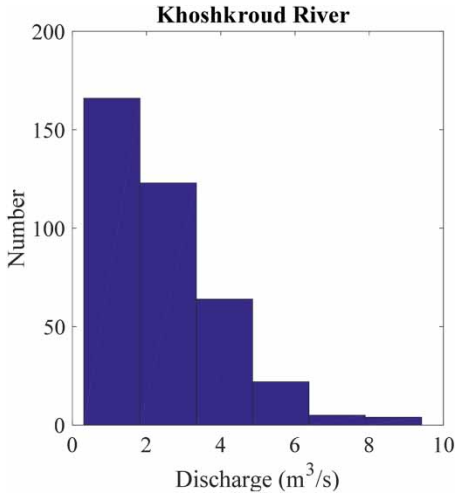


Figure 1 | Monthly discharge histogram of Khoshkroud River.

length of the main stream path and basin area are 67 km and 1,700 km^2, respectively. Additionally, the monthly streamflow time series of this river was collected from Tollat hydrometric station for a 60-year period beginning in 1956. Figure 2 depicts the monthly discharge histogram of Tollat station.

Both rivers flow through the city of Roudsar located at the east of Guilan province in the north of Iran. They rise in Chakroud Mountain, passing through coastal plains, and finally reaching the Caspian Sea. The geographical location of the study area is shown in Figure 3. Furthermore, the monthly meteorological data were acquired from the Ramsar synoptic station in the north of Iran, which was in the near vicinity of both hydrometric stations. In this paper, discharge (Q), pressure (P), average of monthly temperature (T_m), average of monthly dew point (T_d), and average of monthly precipitation (P_r) were considered as attributes. The pieces of statistical information about both basins are delineated in Tables 1 and 2.

In order to train the AI models, the first 70% data of streamflow time series for both stations were utilized. In addition, the 30% of the remaining dataset was dedicated to test the AI models.

Discrete WT

The wavelet function (t), called the mother wavelet, can be expressed as follows:

$$\int_{-\infty}^{+\infty} \Psi(t)dt = 0 \tag{1}$$

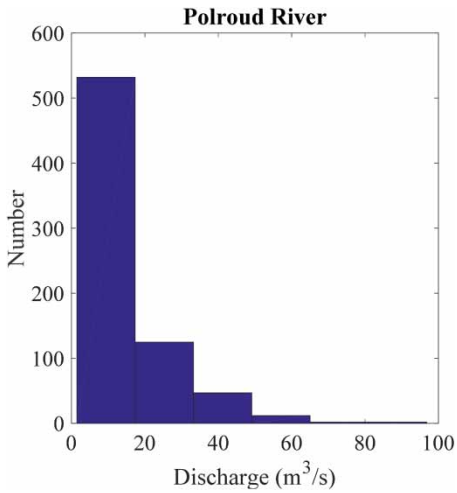


Figure 2 | Monthly discharge histogram of Polroud River.

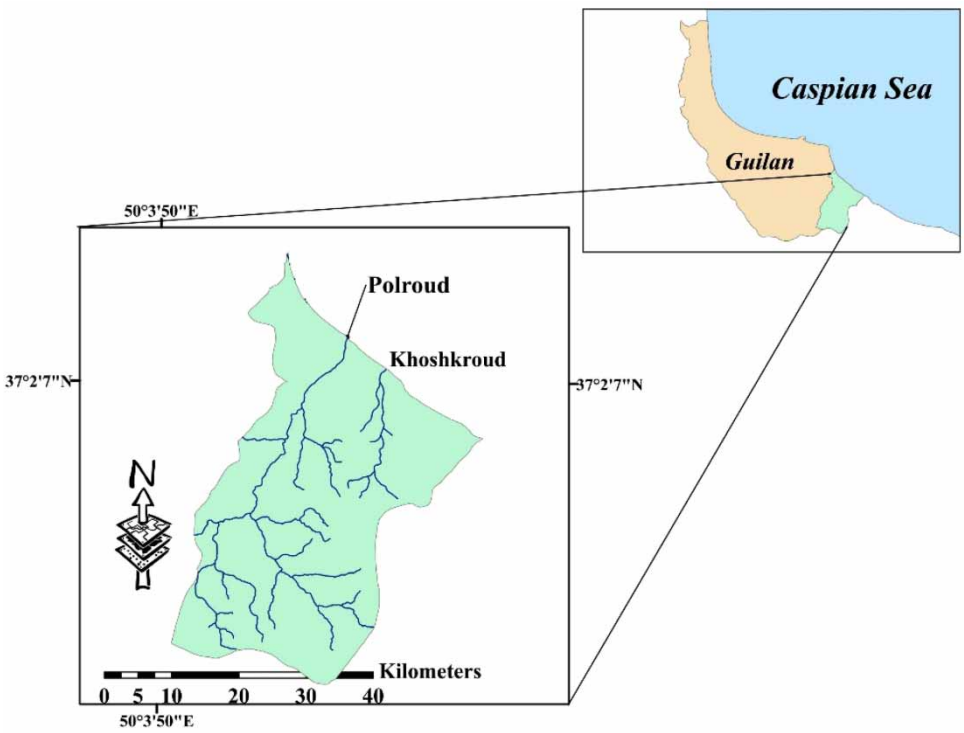


Figure 3 | Location of Polroud and Khoshkroud Rivers in the north of Iran.

Table 1 | Statistical parameters of the Khoshkroud River monthly dataset

Variable	Minimum	Maximum	Average	SD	Skewness
Q (m^3/s)	0.31	9.41	2.46	1.61	1.30
P_r (mm)	2.20	784.01	146.09	126.30	2.16
T_m (°C)	4.47	28.68	16.52	6.86	0.08
P (mbar)	1,007.78	1,054.29	1,017.74	5.14	0.88
T_d (°C)	0.91	23.68	13.06	6.38	−0.02

Table 2 | Statistical parameters of the Polroud monthly dataset

Variable	Minimum	Maximum	Average	SD	Skewness
Q (m^3/s)	1.47	96.81	14.73	12.23	2.34
P_r (mm)	0.00	784.01	118.04	114.07	2.45
T_m (°C)	0.98	28.44	16.16	40.40	2.30
P (mbar)	920.92	1,054.29	1,017.67	6.15	−5.12
T_d (°C)	−0.96	23.68	12.84	6.33	−0.05

Then the function might be identified through compressing and expanding (t),

$$\Psi_{a,b}(t) = |a|^{-\frac{1}{2}}\Psi\left(\frac{t-b}{a}\right) \quad a, b \in R\ ,\ a \neq 0 \tag{2}$$

where $\Psi_{a,b}(t)$ is the successive wavelet, and a and b represent the scale (frequency) and time factors, respectively. Also, R denotes the domain of real numbers.

If $\Psi_{a,b}(t)$ satisfies Equation (2), for the time series $f(t) \in L^2(R)$, the successive WT of $f(t)$ is expressed as follows:

$$W_\Psi f(a, b) = |a|^{-\frac{1}{2}} \int_R f(t)\bar{\Psi}\left(\frac{t-b}{a}\right)dt \tag{3}$$

where $\bar{\Psi}(t)$ denotes the complex conjugate functions of (t). It is clear from Equation (3) that the WT is the decomposition of $f(t)$ under different resolution levels (scales). In other words, to filter the wave for $f(t)$ with various filters is the essence of the WT. The successive is frequently discrete in real applications. Assume $a = a_0^j$, $b = kb_0a_0^j$, $a_0 > 1$, $b_0 \in R$, k, and j are integer numbers. DWT of $f(t)$ is written as follows:

$$W_\Psi f(j, k) = a_0^{-\frac{j}{2}} \int_R f(t)\bar{\Psi}(a_0^{-j}t - kb_0)dt \tag{4}$$

The most common (and simplest) choices for the parameters a_0 and b_0 are 2 and 1-time steps, respectively (Mallat 1989). Thus, Equation (4) is transformed to the binary form as follows:

$$W_\Psi f(j, k) = 2^{-\frac{j}{2}} \int_R f(t)\bar{\Psi}(2^{-j}t - k)dt \tag{5}$$

The characteristics of the original time series in frequency (a or j) and time-domain (b or k) at the same time are presented by $W_\Psi f(a, b)$ or $W_\Psi f(j, k)$.

For a discrete-time series $f(t)$, where it occurs at different time t, the DWT can be considered as follows:

$$W_\Psi f(j, k) = 2^{-\frac{j}{2}} \sum_{t=0}^{N-1} f(t)\bar{\Psi}(2^{-j}t - k) \tag{6}$$

where $W_\Psi f(j, k)$ shows the wavelet coefficients for the DWT of scale $a = 2j$ and $b = 2j \times k$.

The DWT operates two sets of functions, e.g., high-pass and low-pass filters, as seen in Figure 2. The original time series are passed through these filters and then separated at different levels. The time series is decomposed into one comprising its trend (the approximation) and one comprising the high frequencies and the fast events (the detail) (Karimi *et al.* 2015). In the present research, integer time steps are utilized, and additionally, the detail (D) coefficients and approximation (A) sub-time series are extracted using Equation (5).

Gene expression programming

GEP was introduced by Ferreira (2001). The GEP is known as an evolutionary algorithm with the potential to produce explicit formulations of the relationship that describes the physical phenomena. It is one of the supremacies of the GEP model over other DDTs (Martí *et al.* 2013). In general, similar to other evolutionary algorithms, the GEP utilized population of individuals, chooses them according to fitness, and conducts genetic variations by exploiting one or more genetic operators (Ferreira 2006). The GEP procedure is as follows. The first step is defining the fitness function. Here, the root-mean-square error (RMSE) fitness function was utilized. The second step consists of choosing the terminal and the set of functions. In the current research, the terminal sets include the river flow records (with various lag times) and meteorological variables. The selection of the appropriate function was done based on the viewpoints of the user. Here, various mathematical functions (i.e., $+$, $-$, $\times$, $\div$, $\sqrt[2]{\ }$, $\sqrt[3]{\ }$, ln, e^x, x^2, x^3, $\sin x$, $\cos x$, Arctg$x$) were used based on the previous pieces of experience (Karimi *et al.* 2018). The next step is to select the linking function. In the present investigation, an 'addition' function was used to link genes as applied by Karimi *et al.* (2018). The final step is to select the genetic operators. Table 3 illustrates the GEP genetic operators.

Table 3 | Genetic operators used in GEP modeling

Variable	Setting
Chromosome number	30
Head size	10
Gene number	4
Generation number	1,000
Fitness function	RMSE
Mutation	0.00138
One point recombination	0.00277
Two point recombination	0.00277
Inversion	0.00546

Model tree

The original M5 model tree was first introduced by Quinlan (1992). That is based on binary decision tree having linear regression functions at the leaves (terminals) nodes. This type of model tree evaluates the relationship between input and output parameters. The M5P is an expanded version of M5 and requires four main phases. In the first phase, a splitting criterion is utilized to build a decision tree. The standard deviation of the desired parameter values (the measured river flow values) that reach that node is utilized to measure the variability. The desired reduction in error value is calculated by utilizing the standard deviation reduction (SDR) factor which maximizes the expected error reduction at the node (Behnood *et al.* 2017):

$$\mathrm{SDR} = \mathrm{sd}(S) - \sum_i \frac{S_i}{S} \cdot \mathrm{sd}(S_i) \tag{7}$$

in which S is the set of the desired data that reach the node, S_i shows the resulted from splitting the node according to a given attribute, and sd denotes the standard deviation. Then, a linear regression model is developed in each of the sub-spaces utilizing the data associated with that sub-space. In the third phase, pruning is conducted where the initially built tree is simplified (pruned) by omitting different branches. The pruning phase can cause sharp discontinuities between the adjacent linear models. Finally, to compensate for this issue, a smoothing process is performed. The smoothing process combines overall the model from the leaf to the root to build the final model of the leaf. So, the forecasted leaf value is filtered, as it paths back to the root (Behnood *et al.* 2017). For additional information regarding M5 and M5P, their applications and the model-building process are found in the literature (Quinlan 1992; Wang & Witten 1996; Solomatine & Dulal 2003; Solomatine & Xue 2004).

Development of hybrid AI models

In this study, two hybrid AI techniques were developed by combining DWT with two AI models, i.e., GEP and M5P. The WGEP and WM5P models, which use the sub-time-series components, compute through DWT on original data. For processing the input variables (i.e., hydrological data with different lags), the original time series were decomposed into a certain number of sub-time-series components (Ds) using the algorithm proposed by Mallat (1989). The hybrid models were built in which the Ds of the original time series are considered as input variables of the AI techniques, and the original output time series are the output of the WGEP and WM5P. Figure 4 provides a schematic representation of both hybrid AI models that are used in the current study.

Model assessment criteria

Several statistical indices have been drawn to measure the efficiency of models in terms of various aspects. Mean absolute error (MAE) and RMSE compute errors with distinctive weights. Both statistical measures have the same unit as a measurable factor within the range of [0, ∞], although the best value for both statistical indices is zero. Correlation coefficient (R) is indicative of the linear relationship between the observed and forecasted values which lies in [−1, 1], and a close positive

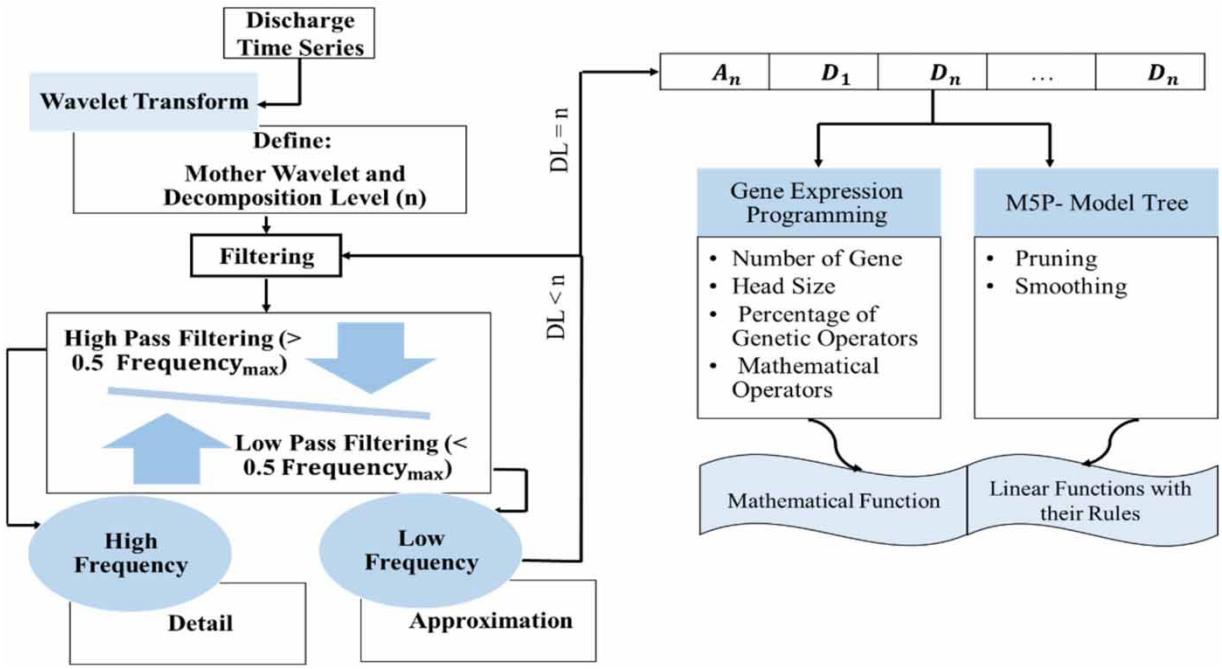


Figure 4 | Schematic representation of hybrid AI models.

correlation exists between them at 1 (Najafzadeh *et al.* 2018; Saberi-Movahed *et al.* 2020). Two other dimensionless criteria, Nash–Sutcliffe model efficiency coefficient (NSE) and index of agreement (I_a), assisted to measure the error degrees of freedom and the rate of accuracy level, respectively. I_a and NSE values vary between 0 and 1, and ∞ and 1, respectively (Willmott 1981). With the aid of the relative improvement (RI) index, the performance of each hybrid AI model is compared with those obtained by the AI model without processing input variables (Amini *et al.* 2005). All the statistical measures are expressed as follows:

$$\mathrm{MAE} = \frac{\sum_{1}^{n} |Q_i^o - Q_i^f|}{n} \tag{8}$$

$$\mathrm{RMSE} = \sqrt{\frac{(Q_i^o - Q_i^f)^2}{n}} \tag{9}$$

$$R = \frac{\sum_{1}^{n} (Q_i^o - Q_m^o)(Q_i^f - Q_m^f)}{\sqrt{\sum_{1}^{n} (Q_i^o - Q_m^o)^2 \sum_{1}^{n} (Q_i^f - Q_m^f)^2}} \tag{10}$$

$$I_a = 1 - \frac{\sum_{1}^{n} (Q_i^o - Q_i^f)^2}{\sum_{1}^{n} (|Q_i^f - Q_m^o| + |Q_i^o - Q_m^o|)^2} \tag{11}$$

$$\mathrm{NSE} = 1 - \frac{\sum_{1}^{n} (Q_i^o - Q_i^f)^2}{\sum_{1}^{n} (Q_i^o - Q_m^o)^2} \tag{12}$$

$$\mathrm{RI} = \frac{\mathrm{RMSE_{M1}} - \mathrm{RMSE_{M2}}}{\mathrm{RMSE_{M1}}} \times 100 \tag{13}$$

where Q_i^f, Q_i^o, Q_m^f, Q_m^o, and n represent the forecasted discharge, observed discharge, average of forecasted discharge, average value of observed discharge, and number of observed discharge, respectively. In Equation (13), subscripts of M1 and M2 are associated with the hybrid AI model (WGEP or WM5P) and the AI technique (GEP or M5P), respectively.

Modeling process

In this paper, Khoshkroud and Polroud Rivers have been selected to determine whether the performance of the AI models (GEP, WGEP, M5P, and WM5P) can satisfy the sufficiency of statistical criteria. GEP and M5P have been considered due to the representation of mathematical functions which simplify the analysis of their performance. Furthermore, their mathematical expressions determine the relationships among input variables and their relative importance in modeling. As indicated in Table 4, seven combinations (I_1 to I_7) were composed of hydrological (H) and climatic (M) attributes to demonstrate the benefits of attributes on the precision level of hybrid AI models.

In the hybrid AI models (WGEP and WM5P), in order to explore the effectiveness of two wavelet features on the wavelet modeling ability, three mother wavelets (haar, db7, and dmey) with four decomposition levels (3–6) have been employed to preprocess streamflow time series. M5P is a linear technique that has required more delay times to satisfy classification bands, while GEP is a nonlinear method that could have been run in a global space with every shape. According to this, all seven combinations (I_1–I_7) have been prepared in five frame times as seen in Table 4.

According to the linear structure of M5, the use of all input variables needs a large number of lags for the time series of streamflow. Additionally, the high capability of wavelet functions in the detection of time series caused a low number of lags. Accordingly, M5P and WM5P have been modeled for 'T_1–T_5' (35 scenarios), GEP for 'T_1–T_3' (21 scenarios), and WGEP for 'T_2–T_3' (14 scenarios).

All mentioned scenarios have been investigated for three lead times, including 1-, 2-, and 3-month-ahead (L1, L2, and L3). Additionally, 70% of samples were considered for the training phase and the remaining samples (30%) were employed for the testing phase. The methods have utilized the Polroud and Khoshkroud data based on 504 and 269 months for the training phase, and 216 and 115 months for the testing phase, respectively. Finally, the performance of AI techniques was assessed by statistical indices and diagrams. Table 4 shows all scenarios' properties used in this study.

RESULTS AND DISCUSSION

The scenarios presented in this study were operated with their optimized results for both rivers. in Tables 5 and 6, the best results are highlighted in each lead time. From Tables 5 and 6, GEP and M5P had a poor performance with regard to hydrological factors in the multi-month-ahead, notably when the Khoshkroud River datasets were examined. In addition, as meteorological data is merged to the hydrologic data, the capabilities of both AI methods in streamflow forecasting increase. Adding the meteorological elements (i.e., T_m, T_d, P, and P_r) to hydrological input (Q) has no tangible variations in the accuracy level of hybrid AI models (WM5P and WGEP), whereas this is remarkable for GEP and M5P especially in high values of

Table 4 | Definition of scenarios

Lead time (L) (month)	Model		Mother wavelet (MW)	Decomposition level (DL)	Effective times (T)	Input type (I)
$t+1$	Stand-alone	GEP	–	–	$T_1 = t$	$I_1 = Q$
$t+2$					$T_2 = t, t-1$	$I_2 = Q, P$
$t+3$					$T_3 = t, t-1, t-2$	$I_3 = Q, P, T_m$
		M5	–	–	$T_1 = t$	$I_4 = Q, P, T_d$
					$T_2 = t, t-1$	$I_5 = Q, P, P_r$
					$T_3 = t, t-1, t-2$	$I_6 = Q, P, T_m, P_r$
					$T_4 = t, t-1, t-2, t-3$	$I_7 = Q, P, T_m, T_d, P_r$
					$T_5 = t, t-1, t-2, t-3, t-4$	
	Hybrid	WGEP	Haar	3	$T_2 = t, t-1$	
			Db7	4	$T_3 = t, t-1, t-2$	
		WM5	dmey	5	$T_1 = t$	
				6	$T_2 = t, t-1$	
					$T_3 = t, t-1, t-2$	
					$T_4 = t, t-1, t-2, t-3$	
					$T_5 = t, t-1, t-2, t-3, t-4$	

Table 5 | Models evaluation for river flow forecasting at Bajiguabar station

				Train						Test					
L	Model	IT	LT	MAE (m^3/s)	RMSE (m^3/s)	*R*	NSE	I_a	RI (%)	MAE (m^3/s)	RMSE (m^3/s)	*R*	NSE	I_a	RI (%)
$t+1$	GEP	I_2	T_2	1.18	1.64	0.31	0.08	0.47	–	1	1.27	0.50	0.23	0.61	–
	M5P	I_1	T_1	1.18	1.55	0.30	0.09	0.39	5	1.09	1.37	0.40	0.13	0.42	−8
	Haar (5)-GEP	I_3	T_3	0.94	1.28	0.64	0.41	0.76	22	0.82	1.10	0.67	0.44	0.78	13
	Haar (3)-M5P	I_1	T_4	0.83	1.15	0.73	0.53	0.83	30	0.73	0.98	0.74	0.55	0.84	23
	Db7 (3)-GEP	I_4	T_2	0.62	0.82	0.88	0.76	0.94	50	0.61	0.80	0.86	0.71	0.92	37
	Db7 (3)-M5P	I_1	T_5	0.36	0.47	0.96	0.92	0.98	71	0.34	0.46	0.95	0.90	0.98	64
	Dmey (3)-GEP	I_1	T_3	0.55	0.70	0.91	0.82	0.95	57	0.56	0.74	0.87	0.75	0.92	42
	Dmey (3)-M5P	***I*₂**	***T*₅**	**0.23**	**0.31**	**0.98**	**0.97**	**0.99**	**81**	**0.20**	**0.28**	**0.98**	**0.96**	**0.99**	**78**
$t+2$	GEP	I_3	T_1	1.18	1.64	0.26	0.06	0.39	–	1.18	1.46	0.23	0.02	0.40	–
	M5P	I_6	T_3	1.09	1.46	0.51	0.23	0.53	11	1.18	1.46	0.23	0.00	0.43	0
	Haar (3)-GEP	I_5	T_3	1.11	1.47	0.48	0.23	0.63	10	0.96	1.25	0.53	0.28	0.65	14
	Haar (3)-M5P	I_3	T_5	1.01	1.34	0.60	0.37	0.72	18	0.92	1.20	0.59	0.33	0.72	18
	Db7 (5)-GEP	I_1	T_3	0.86	1.15	0.72	0.52	0.82	30	0.71	0.96	0.77	0.58	0.86	34
	Db7 (3)-M5P	I_2	T_4	0.76	0.98	0.81	0.66	0.89	40	0.66	0.92	0.79	0.61	0.88	37
	Dmey (5)-GEP	I_6	T_3	0.81	1.05	0.78	0.60	0.87	36	0.73	1.01	0.73	0.53	0.83	31
	Dmey (5)-M5P	***I*₁**	***T*₅**	**0.55**	**0.73**	**0.90**	**0.81**	**0.95**	**55**	**0.45**	**0.65**	**0.90**	**0.80**	**0.94**	**55**
$t+3$	GEP	I_6	T_1	1.27	1.64	0.26	0.07	0.38	–	1.18	1.46	0.31	0.07	0.46	–
	M5P	I_5	T_2	1.18	1.55	0.33	0.11	0.42	5	1.18	1.46	0.30	0.06	0.44	0
	Haar (5)-GEP	I_3	T_2	1.09	1.48	0.47	0.22	0.61	10	0.98	1.25	0.53	0.28	0.65	14
	Haar (3)-M5P	I_5	T_2	1.11	1.50	0.44	0.19	0.55	9	0.99	1.26	0.52	0.27	0.62	14
	Db7 (4)-GEP	I_7	T_2	0.91	1.21	0.70	0.48	0.82	26	0.79	1.05	0.70	0.49	0.81	28
	Db7 (3)-M5P	I_1	T_5	0.84	1.10	0.76	0.57	0.85	33	0.71	0.93	0.77	0.60	0.86	36
	Dmey (3)-GEP	I_1	T_3	0.84	1.11	0.75	0.56	0.86	32	0.78	1.06	0.72	0.49	0.84	27
	Dmey (5)-M5P	***I*₁**	***T*₅**	**0.66**	**0.87**	**0.86**	**0.73**	**0.92**	**47**	**0.51**	**0.76**	**0.86**	**0.74**	**0.92**	**48**

The bold values show the best results at each lead time.

lead time. Furthermore, Tables 5 and 6 indicated that the precision level of AI models has an inverse direction with an increase in the number of lead times. Furthermore, it can be seen from Figure 5 that both graphical performances of $M5P_{Polroud}$ and $GEP_{Polroud}$ illustrated convergence toward the 45° line, while $M5P_{Khoshkroud}$ and $GEP_{Khoshkroud}$ were scattered in a horizontal area. Careful consideration of the time-series length for both rivers demonstrated that the effects of time-series length on the performance of the M5P technique are higher than the GEP model. In the case of M5P efficiency, the stream-flow forecasting for Polroud River ($I_{a(M5P,\ Polroud,\ L1)} = 0.85$) with a 720-month time series is more accurate than that predicted for Khoshkroud River ($I_{a(M5P,\ Khoshkroud,\ L1)} = 0.42$) with a 360-month time series.

It can be inferred from the results of M5P and GEP techniques that wavelets had a significant effect on the performance of these AI methods, which were changeable in their WT properties. In the case of the case study basins, discrete and step Haar functions have made a marginal impact compared to other mother wavelet functions (i.e., dmey and db) and could have not dominated frequencies correctly, out of discontinues shape and dissimilar style to oscillation flow series and short band length ($NSE_{Khoshkroud,\ haar\text{-}GEP,\ M5P,\ L1\text{-}3} > 0.27$ and $NSE_{Polroud,\ haar\text{-}GEP,\ M5P,\ L1\text{-}3} > 0.35$). In the preference, db7 and dmey have had a notable role in improving the performance of GEP and M5P techniques, in order for the NSE greater than 50% of all the lead times ($NSE_{Khoshkroud,\ GEP,\ M5P,\ L1\text{-}3} > 0.49$ and $NSE_{Polroud,\ GEP,M5P,\ L1\text{-}3} > 0.68$). Dmey and db7 functions had almost the same impacts on GEP and M5P approaches; nonetheless, a decrease in differences between RMSE and MAE values has attracted attention and has observed less contrast with Koshkroud River, in particular GEP technique. Dmey wavelet function includes a lengthier band than db7 (7.7 times) and focused on extremes as well as sudden jumps primarily; however, db7 has a smooth and expandable curve shape and behaves closely toward the entire time series. These differences discovered how the db7-GEP could have forecasted the 10 m^3/s maximum discharges (low flow) of Khoshkroud River as well as dmey-GEP; however, dmey-M5P has grown results toward db7-M5P which M5P has recognized patterns of extreme flow simulations. Figure 6 illustrates the better graphical performance of M5P for the 1-month-ahead, where WM5P has decreased the RMSE index 78% toward the GEP technique and could simulate peak flows accurately as same as the GEP model.

Table 6 | Models evaluation for river flow forecasting at Tollat station

				Train						Test					
L	Model	IT	LT	MAE (m^3/s)	RMSE (m^3/s)	*R*	NSE	I_a	RI (%)	MAE (m^3/s)	RMSE (m^3/s)	*R*	NSE	I_a	RI (%)
$t+1$	GEP	I_7	T_2	5.72	9.53	0.64	0.41	0.75	–	5.72	8.58	0.68	0.46	0.79	–
	M5P	I_4	T_3	5.72	8.58	0.73	0.53	0.82	10	5.72	7.63	0.75	0.55	0.85	11
	Haar (3)-GEP	I_7	T_2	4.12	6.02	0.87	0.76	0.93	37	3.90	5.74	0.88	0.76	0.93	33
	Haar (3)-M5P	I_1	T_4	3.16	4.88	0.92	0.85	0.96	49	3.01	4.01	0.93	0.86	0.96	53
	Db7 (3)-GEP	I_4	T_2	3.00	4.35	0.94	0.88	0.97	54	2.90	4.02	0.94	0.88	0.97	53
	Db7 (4)-M5P	I_1	T_5	2.02	2.85	0.97	0.95	0.99	70	2.06	2.75	0.97	0.95	0.99	68
	Dmey (3)-GEP	I_6	T_3	2.83	3.78	0.95	0.91	0.97	60	2.76	3.59	0.95	0.91	0.98	58
	Dmey (3)-M5P	$\boldsymbol{I_1}$	$\boldsymbol{T_5}$	**1.19**	**1.60**	**0.99**	**0.98**	**1.00**	**83**	**1.16**	**1.52**	**0.99**	**0.98**	**1.00**	**82**
$t+2$	GEP	I_4	T_2	6.67	10.49	0.51	0.25	0.57	–	7.63	9.53	0.56	0.29	0.60	–
	M5P	I_7	T_5	5.72	8.58	0.72	0.52	0.81	18	5.72	8.58	0.71	0.50	0.80	10
	Haar (3)-GEP	I_7	T_2	4.69	6.96	0.83	0.68	0.90	34	4.34	6.55	0.83	0.69	0.90	31
	Haar (3)-M5P	I_7	T_5	4.27	6.53	0.85	0.72	0.91	38	4.19	5.92	0.87	0.75	0.92	38
	Db7 (5)-GEP	I_7	T_3	4.38	6.22	0.87	0.75	0.93	41	4.21	5.77	0.87	0.76	0.93	39
	Db7 (3)-M5P	I_3	T_5	3.50	5.07	0.91	0.83	0.95	52	3.30	4.73	0.92	0.84	0.95	50
	Dmey (5)-GEP	I_1	T_2	3.30	4.62	0.93	0.86	0.96	56	3.28	4.40	0.93	0.86	0.96	54
	Dmey (3)-M5P	$\boldsymbol{I_5}$	$\boldsymbol{T_5}$	**2.72**	**4.04**	**0.95**	**0.89**	**0.97**	**61**	**2.69**	**3.67**	**0.95**	**0.90**	**0.97**	**61**
$t+3$	GEP	I_7	T_3	6.67	10.49	0.52	0.26	0.67	–	6.67	9.53	0.54	0.29	0.67	–
	M5P	I_7	T_4	5.72	8.58	0.71	0.50	0.80	18	6.67	8.58	0.69	0.46	0.78	10
	Haar (3)-GEP	I_7	T_2	6.52	9.32	0.66	0.43	0.76	11	6.48	9.57	0.59	0.35	0.70	−0.42
	Haar (3)-M5P	I_4	T_5	4.15	6.04	0.88	0.76	0.92	42	5.58	8.30	0.71	0.50	0.82	13
	Db7 (4)-GEP	I_6	T_3	4.59	6.67	0.84	0.71	0.90	36	4.61	6.68	0.83	0.68	0.89	30
	Db7 (3)-M5P	I_4	T_5	3.69	5.32	0.90	0.82	0.95	49	3.72	5.13	0.90	0.81	0.95	46
	Dmey (5)-GEP	I_3	T_3	4.22	6.23	0.87	0.75	0.93	41	4.10	5.64	0.88	0.77	0.93	41
	Dmey (3)-M5P	$\boldsymbol{I_1}$	$\boldsymbol{T_5}$	**3.07**	**4.83**	**0.92**	**0.85**	**0.96**	**54**	**2.88**	**4.13**	**0.94**	**0.88**	**0.97**	**57**

The bold values show the best results at each lead time.

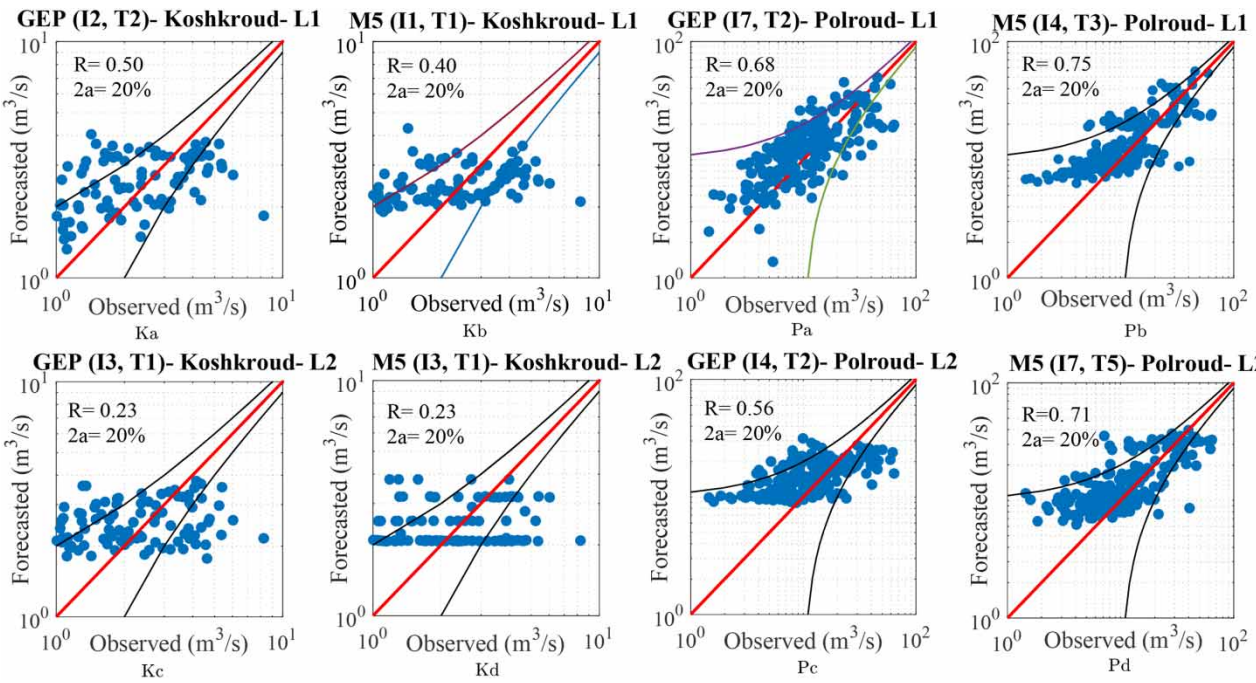


Figure 5 | The validation scatter plots of stand-alone models in logarithmic scale at lead times 1 and 2. The red line shows the best fit (45° line) and the black lines are related to the upper and lower bounds of the confidence level of 80%. K_a and K_b are the best GEP and M5P at lead time 1, K_c and K_d are the best GEP and M5P at lead time 2 for Khoshkroud River. P_a and P_b are the best GEP and M5P at lead time 1, and P_c and P_d are the best GEP and M5P at lead time 2 for Polroud River. Please refer to the online version of this paper to see this figure in color: doi:10.2166/hydro.2021.105.

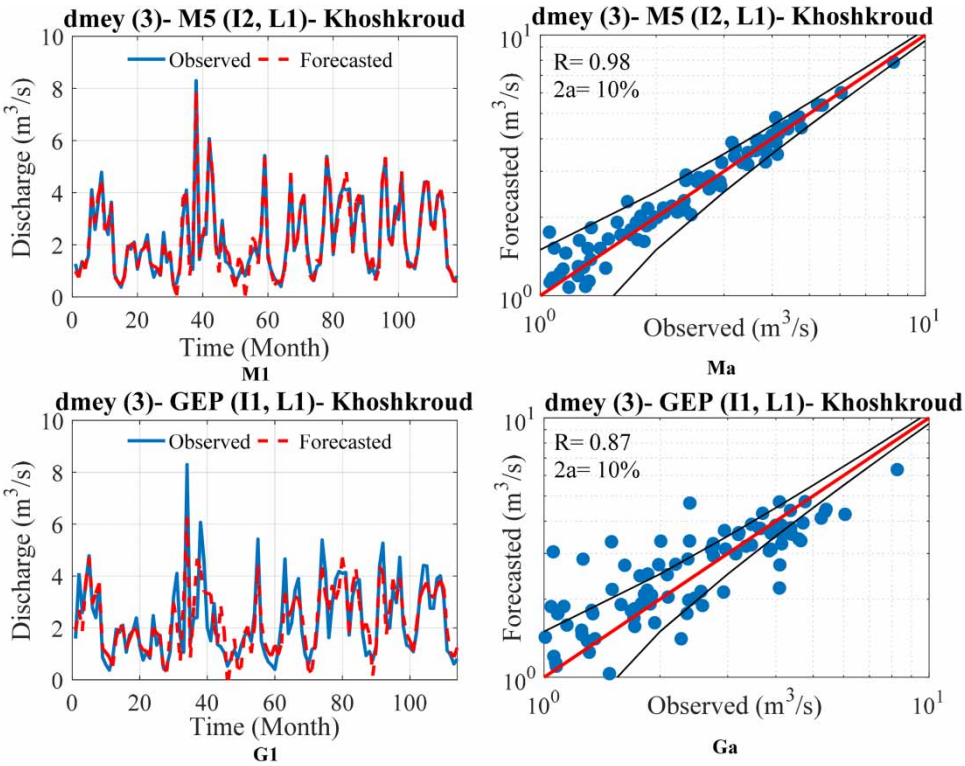


Figure 6 | Comparisons of the observed and forecasted monthly Khoshkroud River flow data for the superior hybrid AI models for testing data, at lead time 1. M_1 and G_1 are time-discharge diagrams. M_a and G_a are scatter plots in logarithmic scale in which the red line shows the best fit (45° line) and the black lines are related to the upper and lower bounds of the confidence level of 90%. With regard to lower efficiency of AI models compared with the hybrid AI models, a small number of outputs associated with AI models put inside 90% confidence level. Hence, this confidence level is statistically more suitable for hybrid AI techniques. Please refer to the online version of this paper to see this figure in color: doi:10.2166/hydro.2021.105.

Figure 7 illustrates that higher confidence of WM5P model associated to the Polroud River performed in 5% confident area (for 1-month-ahead), and additionally, this issue has been efficiently approved by the WGEP hybrid model. The second wavelet property is the decomposition level. M5P and GEP techniques have reacted to variations in the decomposition level values, and mother wavelets were highly sensitive to scales for these two-time series, notably Polroud River. Figures 8 and 9 illustrate variations of statistical criteria for the best hybrid AI models versus decomposition levels and scenarios of input combinations for both Khoshkroud and Polroud Rivers, respectively. From Figures 8 and 9, it can be inferred that in a short lead time (1-month-ahead), the hybrid AI models provided satisfying efficiency for low values of decomposition level, whereas for the higher values of lead time (3-month-ahead), higher values of decomposition level (5 and 6) are required to boost the accuracy level of AI models.

Furthermore, hybrid AI techniques have the potential of modeling time-series flow, and especially applying the DWT could improve the performance of hybrid AI models significantly at 3-month-ahead for both rivers ($RI_{(WM5P,\ Khoshkroud,\ L3)} = 0.48$, $RI_{(WM5P,\ Polroud,\ L3)} = 0.57$). For both basins, WM5 has yielded an equation that stood relatively at the same accuracy level provided by the WGEP model. In the case of extreme values of streamflow, WM5P provided more accurate results than the WGEP technique. Figure 10 demonstrates that hybrid AI approaches estimated discharges for 3-month-ahead with R values greater than 70%. According to the above-mentioned results, although meteorological variables might suffer from proper localization (i.e., meteorological stations are not at the near vicinity of hydrometric station) and noises of time series, they could provide a significant improvement in the efficiency level of stand-alone AI techniques (M5P and GEP). For instance, the M5P technique had the worst performance for both case study rivers ($NSE_{Khoshkroud,\ M5P,\ I5,\ L3} = -0.06$, $NSE_{Khoshkroud,\ M5P,\ I1,\ L3} = -0.07$, $NSE_{Polroud,\ M5P,\ I7,\ L3} = 0.46$, $NSE_{Polroud,\ M5P,\ I1,\ L3} = 0.05$) when M5P is not fed by meteorological data. In this way, the effect of wavelet function usage on the performance of AI techniques is more efficient than that of typical datasets (i.e., M and L). Based on the RI index, the largest improvement in

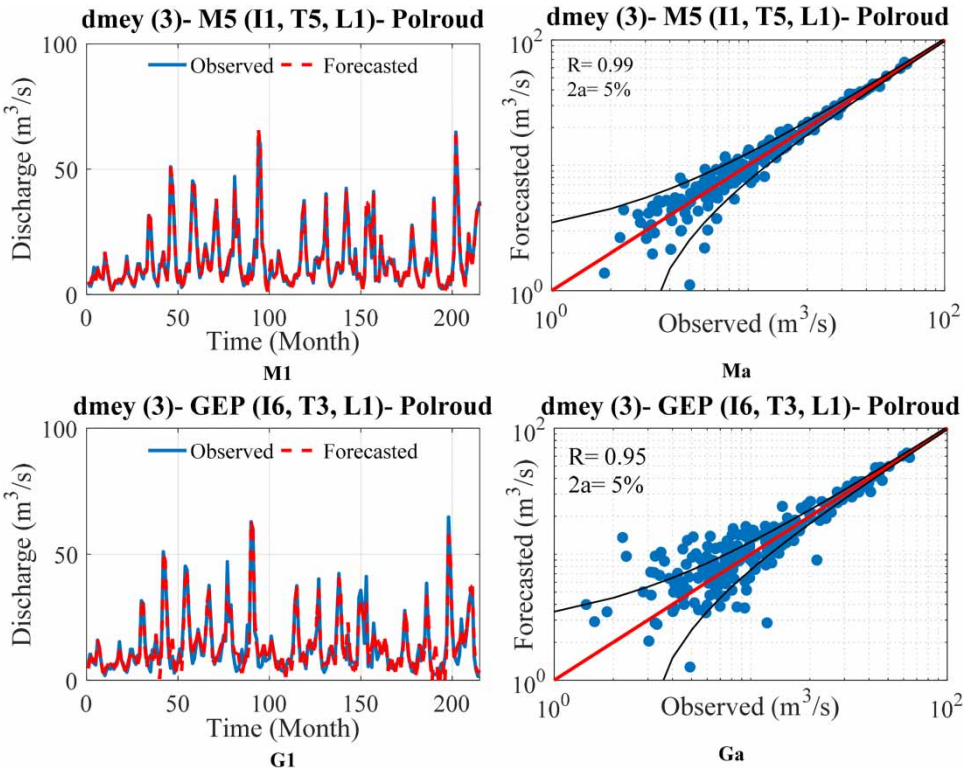


Figure 7 | Comparisons between the observed and forecasted monthly Polroud River flow data for the superior hybrid AI models for testing data at lead time 1. M_1 and G_1 are time-discharge diagrams. M_a and G_a are scatter plots in logarithmic scale in which the red line shows the best fit (45° line) and the black lines are related to the upper and lower bounds of the confidence level of 95%. Please refer to the online version of this paper to see this figure in color: doi:10.2166/hydro.2021.105.

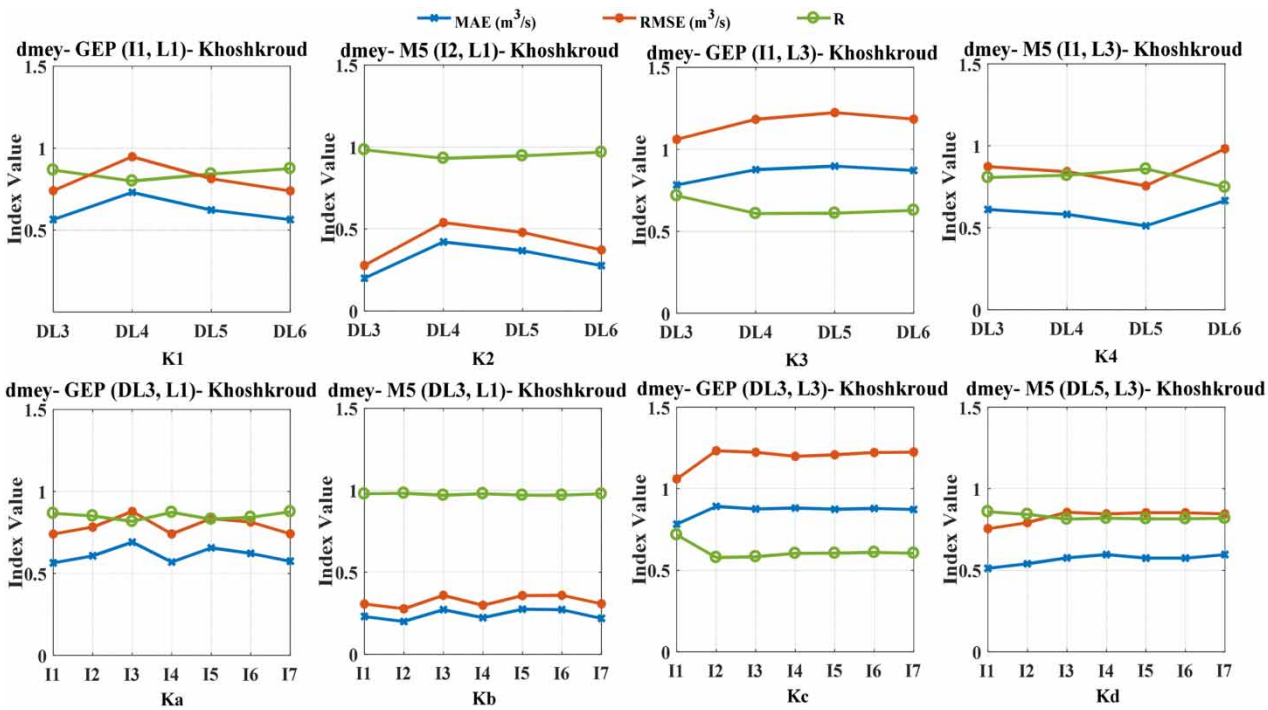


Figure 8 | Performance of statistical indices versus (K_1–K_4) decomposition level and (K_a–K_d) input type for the superior hybrid models for Khoshkroud River at lead times 1 and 3.

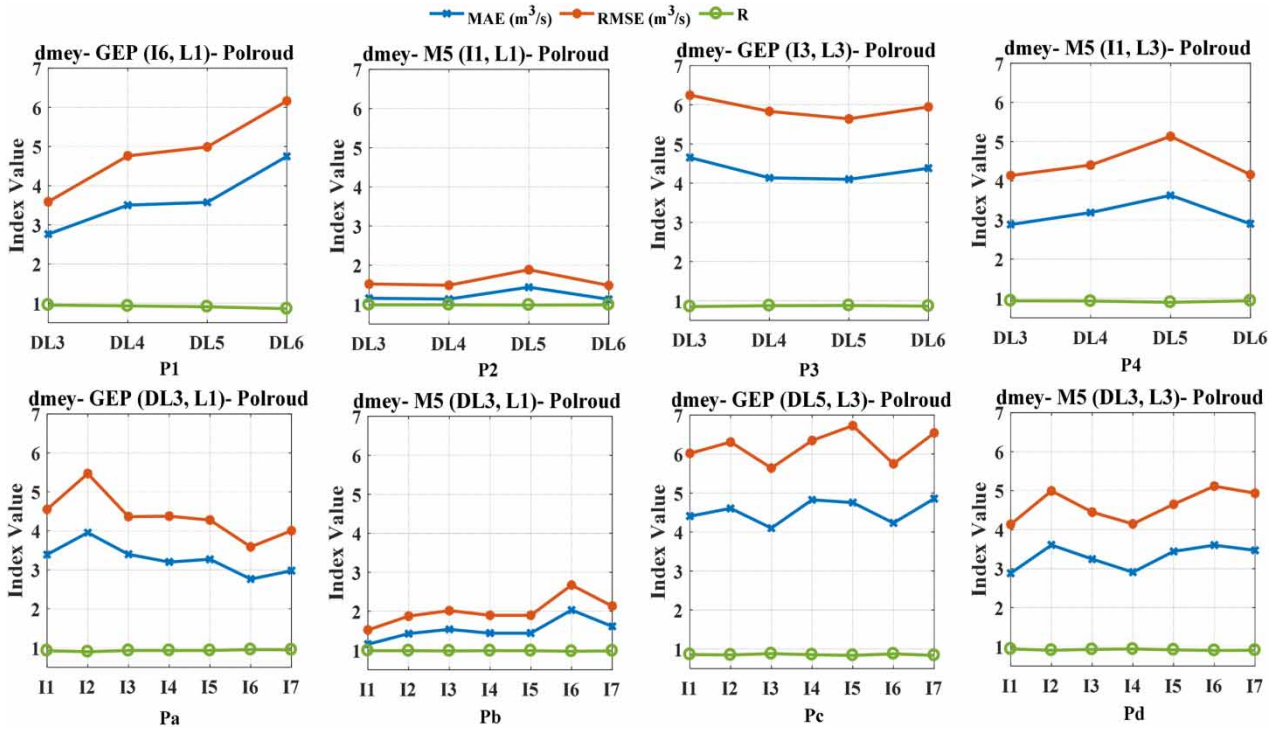


Figure 9 | Performance of statistical indices versus (P_1—P_4) decomposition level and (P_a—P_d) input type for the superior hybrid models for Polroud River at lead times 1 and 3.

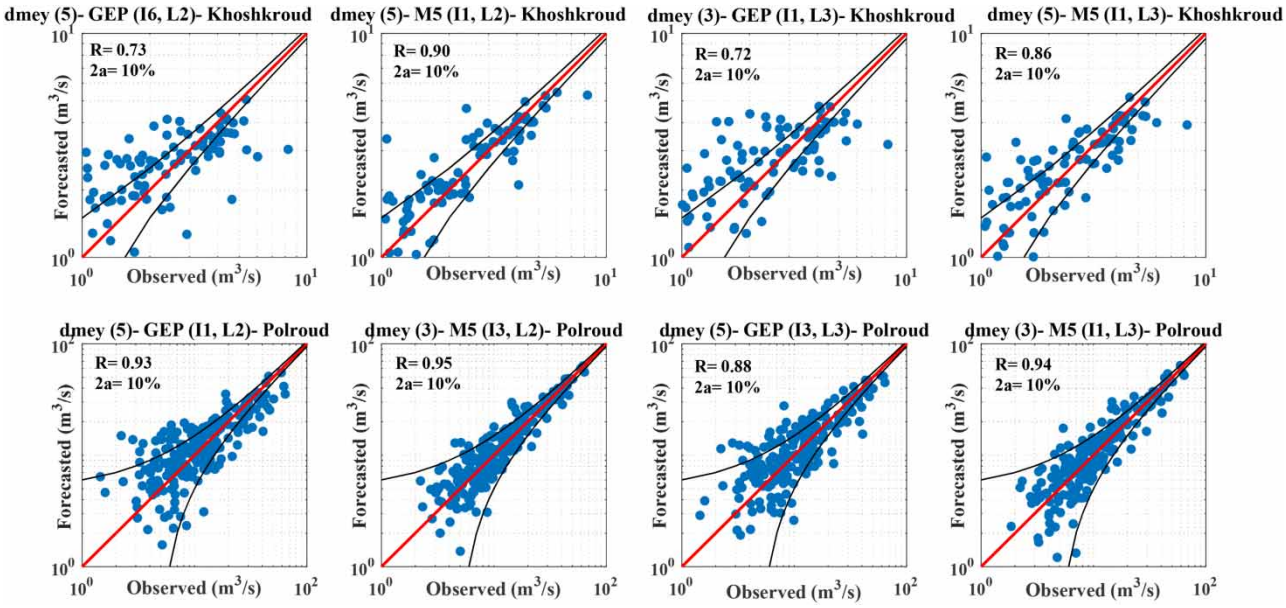


Figure 10 | Validation scatter plots for WM5 and WGEP hybrid models at lead times 2 and 3 in logarithmic scale. The red line shows the best fit and the black are related to the upper and lower bounds of the confidence level of 90%. K_1 (WGEP) and K_2 (WM5P) are at lead time 2, K_3 (WGEP) and K_4 (WM5P) are at lead time 3, for Khoshkroud River. P_a (WGEP) and P_b (WM5P) are at lead time 2, P_c (WGEP) and P_d (WM5P) are at lead time 3, for Polroud River. Please refer to the online version of this paper to see this figure in color: doi:10.2166/hydro.2021.105.

model capability was observed in WM5P-dmey at L1 compared to stand-alone AI approaches (GEP and M5P) for both case study rivers.

Equations (14) and (15) are, respectively, dmey (3)-M5P linear relationship for 1-month-ahead of Polroud and Khoshkroud Rivers:

$$\begin{aligned} Q(t+1)_{\text{Polroud}} = {} & 0.054D_1(t) + 1.62D_2(t) + 2.99D_3(t) + 5.50A_3(t-1) - 3.11D_1(t-1) \\ & - 2.74D_2(t-1) - 4.08D_3(t-1) - 7.299A_3(t-2) - 2.95D_1(t-2) \\ & + 2.23D_2(t-2) + 2.29D_3(t-2) + 2.79A_3(t-3) - 1.87D_1(t-3) \\ & - 1.61D_2(t-3) - 0.70D_1(t-4) + 0.47D_2(t-4) - 0.51D_3(t-4) \end{aligned} \tag{14}$$

$$\begin{aligned} Q(t+1)_{\text{Khoshkroud}} = {} & 0.004 + 1.95A_3(t) - 2.25D_1(t) + 0.72D_2(t) + 2.49D_3(t) + 0.41P(t) \\ & - 3.14D_1(t-1) - 1.25D_2(t-1) - 2.31D_3(t-1) - 0.5P(t-1) - 2.04A_3(t-2) \\ & - 2.93D_1(t-2) - 0.19D_2(t-2) + 0.76D_3(t-2) + 1.1A_3(t-3) \\ & - 1.84D_1(t-3) - 0.64D_1(t-4) - 0.56D_2(t-4) \end{aligned} \tag{15}$$

where A_3, D_1:D_3, P, and t represent the approximation of decomposed streamflow series for third level, details of decomposed streamflow series for the third level, precipitation, and time (month), respectively.

In this study, stand-alone AI models and hybrid AI techniques were performed 336 and 3,528 times, respectively. Essentially, the length of time series is an influential factor in the complexity degree of M5P and GEP models. It can be generally said that the increase of time-series length and number of input variables caused to present a complicated M5P formulation, whereas simpler equations by the GEP model were obtained owing to the existence of nonlinear mathematical expression in GEP formulation. Furthermore, meteorological data have had a substantial effect on the efficiency of AI models, which can be treated as a piece of axillary information especially during long lead time. The confidence level is dependent on the type of mother wavelet and is not a constraint to the decomposition level extremely. In fact, the wavelet mother function extracted important properties of time series such as decomposition level of time series, scale, and delimited frequencies. Widening the band of wavelet mother caused to boost the capability of preprocessing in time series and extraction of main properties. As a result, there exists high sensitivity to the variations in the decomposition level and scenarios of inputs combination. This issue is the reason why dmey wavelet is more robust than other typical wavelet functions. According to Tables 5 and 6, the performance of preprocessing time series decreased both the number of rules applied in the M5P and the complexity of the linear equation by M5P; therefore, the accuracy level of Equations (14) and (15) booted. In general, it can be said that the increase in the number of meteorological variables and preprocessing the time series led to bring the forecasting results for both distinctive basins at the same level of precision. On the contrary, once AI models were not fed by meteorological data, the results of AI models for Polroud River were more accurate than those for Khoshkroud River.

CONCLUSION

Accurate forecasting of streamflow is a vital issue in hydrology and water resources management. Intelligence computing is a satisfactory solution to increase confidence in river flow modeling. Thus, in the current research, the monthly streamflow of two distinctive rivers (Khoshkroud and Polroud) was forecasted by WGEP, WM5P model tree, and their stand-alone AI models (GEP and M5P). To understand the influence of preprocessing on the performance of hybrid AI models, meteorological (M) variables and long lead time (L) have been considered within several input scenarios. Results indicated that the accuracy level of all the AI techniques was in close connection with properties of time series such as being long-term persistence and having extreme values. In fact, these characterizations of time series mean comprehensiveness of streamflow datasets. WT functions were capable of eradicating limitations of time series; then, the results of hybrid AI models were accurate. More specifically, although time series associated with Khoshkroud River had lower comprehensiveness in terms of missing information and length of time series compared with Polroud River, the accuracy level of hybrid AI models fed by Khoshkroud River was approximately equal to the those obtained by Poulroud River. Additionally, meteorological information had the capability to cover deficits of input time series, although effects of meteorological data on the performance of stand-alone AI technique was more tangible than those of hybrid AI models. Due to the existence of more complexity and unknown patterns of time series in longer lead times, the AI models benefited from climatic variables and higher decomposition levels. WM5P with piece-wise linear structure has been considered as the most accurate hybrid

AI model owing to its higher compatibility of hybrid M5P technique to various basins. Although WM5P has a simpler structure than the WGEP model, the WM5P technique has the higher capability ($NSE_{(Polroud,\ WM5P,\ H,\ L3)} = 0.88$ and $NSE_{(Khoshkroud,\ WM5P,\ H,\ L3} = 0.74$) in the simulation of river flow even for extreme values and long lead times. Ultimately, the most significant improvement in the AI models' capability was related to the WM5p-dmey3 in comparison with the stand-alone GEP approach. One of the limitations of this study is the lack of meteorological stations in the location of the hydrometric stations. Also, since the modeling period is monthly, another limitation is accurate forecasting for long-term periods (more than 3 months). It is suggested that the effect of catchment area on the performance of the proposed models be evaluated, and additionally, the proposed hybrid AI models can be combined with other intelligent methods or time-series models such as ARMA and ARIMA for the streamflow forecasting.

ACKNOWLEDGEMENTS

The authors are grateful to the editor and three anonymous reviewers for their constructive and insightful comments, which help enhance the quality of the manuscript.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES

Abdollahi, S., Raeisi, J., Khalilianpour, M., Ahmadi, F. & Kisi, O. 2017 Daily mean streamflow prediction in perennial and non-perennial rivers using four data driven techniques. *Water Resources Management* **31**, 4855–4874.

Abrahart, R. J. & See, L. 1998 September Neural network vs. ARMA modelling: constructing benchmark case studies of river flow prediction. In: *GeoComputation'98. Proceedings of the Third International Conference on GeoComputation*, University of Bristol, UK, pp. 17–19.

Adamowski, J. F. 2008 River flow forecasting using wavelet and cross-wavelet transform models. *Hydrological Processes* **22** (25), 4877–4891.

Adnan, R. M., Liang, Z., Heddam, S., Zounemat-Kermani, M., Kisi, O. & Li, B. 2019 Least square support vector machine and multivariate adaptive regression splines for streamflow prediction in mountainous basin using hydro-meteorological data as inputs. *Journal of Hydrology* **586**, 124371.

Akcakoca, F. & Apaydin, H. 2020 Modelling of bektas creek daily streamflow with generalized regression neural network method. *International Journal of Advances in Scientific Research and Engineering* **6** (2), 97–103.

Amini, M., Abbaspour, K. C., Khademi, H., Fathianpour, N., Afyuni, M. & Schulin, R. 2005 Neural network models to predict cation exchange capacity in arid regions of Iran. *European Journal of Soil Science* **56** (4), 551–559.

Behnood, A., Behnood, V., Gharehveran, M. M. & Alyamac, K. E. 2017 Prediction of the compressive strength of normal and high-performance concretes using M5P model tree algorithm. *Construction and Building Materials* **142**, 199–207.

Boucher, M. A., Quilty, J. & Adamowski, J. 2020 Data assimilation for streamflow forecasting using extreme learning machines and multilayer perceptrons. *Water Resources Research* **56** (6), e2019WR026226.

Chu, H., Wei, J. & Wu, W. 2020 Streamflow prediction using LASSO-FCM-DBN approach based on hydro-meteorological condition classification. *Journal of Hydrology* **580**, 124253.

Cigizoglu, H. K. 2003 Incorporation of ARMA models into flow forecasting by artificial neural networks. *Environmetrics* **14** (4), 417–427.

Diop, L., Bodian, A., Djaman, K., Yaseen, Z. M., Deo, R. C., El-Shafie, A. & Brown, L. C. 2018 The influence of climatic inputs on stream-flow pattern forecasting: case study of Upper Senegal River. *Environmental Earth Sciences* **77** (5), 1–13.

Fahimi, F., Yaseen, Z. M. & El-shafie, A. 2017 Application of soft computing based hybrid models in hydrological variables modeling: a comprehensive review. *Theoretical and Applied Climatology* **128** (3–4), 875–903.

Ferreira, C. 2001 Gene expression programming: a new adaptive algorithm for solving problems. arXiv preprint cs/0102027.

Ferreira, C. 2006 *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence*, Vol. 21. Springer, Berlin.

Hadi, S. J. & Tombul, M. 2018 Monthly streamflow forecasting using continuous wavelet and multi-gene genetic programming combination. *Journal of Hydrology* **561**, 674–687.

Huang, W., Xu, B. & Chan-Hilton, A. 2004 Forecasting flows in Apalachicola River using neural networks. *Hydrological Processes* **18** (13), 2545–2564.

Hussain, D. & Khan, A. A. 2020 Machine learning techniques for monthly river flow forecasting of Hunza River, Pakistan. *Earth Science Informatics* **13**, 939–949.

Jacquin, A. P. & Shamseldin, A. Y. 2009 Review of the application of fuzzy inference systems in river flow forecasting. *Journal of Hydroinformatics* **11** (3–4), 202–210.

Karimi, S., Shiri, J., Kisi, O. & Shiri, A. A. 2015 Short-term and long-term streamflow prediction by using 'wavelet-gene expression' programming approach. *ISH Journal of Hydraulic Engineering* **22** (2), 148–162.

Karimi, S., Shiri, J., Kisi, O. & Xu, T. 2017 Forecasting daily streamflow values: assessing heuristic models. *Hydrology Research* **49** (3), 658–669.

Karimi, S., Shiri, J., Kisi, O. & Xu, T. 2018 Forecasting daily streamflow values: assessing heuristic models. *Hydrology Research* **49** (3), 658–669.

Khairuddin, N., Aris, A. Z., Elshafie, A., Sheikhy Narany, T., Ishak, M. Y. & Isa, N. M. 2019 Efficient forecasting model technique for river streamflow in tropical environment. *Urban Water Journal* **16** (3), 183–192.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K. & Herrnegger, M. 2018 Rainfall–runoff modelling using long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences* **22** (11), 6005–6022.

Mallat, S. G. 1989 A theory for multi resolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11** (7), 674–693.

Martí, P., Shiri, J., Duran-Ros, M., Arbat, G., De Cartagena, F. R. & Puig-Bargués, J. 2013 Artificial neural networks vs. gene expression programming for estimating outlet dissolved oxygen in micro-irrigation sand filters fed with effluents. *Computers and Electronics in Agriculture* **99**, 176–185.

Najafzadeh, M. & Saberi-Movahed, F. 2019 GMDH-GEP to predict free span expansion rates below pipelines under waves. *Marine Georesources & Geotechnology* **37** (3), 375–392.

Najafzadeh, M., Saberi-Movahed, F. & Sarkamaryan, S. 2018 NF-GMDH-based self-organized systems to predict bridge pier scour depth under debris flow effects. *Marine Georesources & Geotechnology* **36** (5), 589–602.

Nourani, V., Davanlou Tajbakhsh, A., Molajou, A. & Gokcekus, H. 2019 Hybrid wavelet-M5 model tree for rainfall-runoff modeling. *Journal of Hydrologic Engineering* **24** (5), 04019012.

Parasuraman, K. & Elshorbagy, A. 2005 Wavelet networks: an alternative to classical neural networks. In: *Proceedings of IEEE International Joint Conference on Neural Networks 2005*, Vol. 5, pp. 2674–2679.

Quinlan, J. R. 1992 Learning with continuous classes. In: *5th Australian Joint Conference on Artificial Intelligence*, Vol. 92, pp. 343–348.

Saberi-Movahed, F., Najafzadeh, M. & Mehrpooya, A. 2020 Receiving more accurate predictions for longitudinal dispersion coefficients in water pipelines: training group method of data handling using extreme learning machine conceptions. *Water Resources Management* **34** (2), 529–561.

Shahabi, S., Khanjani, M. J. & Kermani, M. R. H. 2017 Significant wave height modelling using a hybrid wavelet-genetic programming approach. *KSCE Journal of Civil Engineering* **21**, 1–10.

Shoaib, M., Shamseldin, A. Y., Melville, B. W. & Khan, M. M. 2015 Runoff forecasting using hybrid wavelet gene expression programming (WGEP) approach. *Journal of Hydrology* **527**, 236–344.

Solomatine, D. P. & Dulal, K. N. 2003 Model trees as an alternative to neural networks in rainfall – runoff modelling. *Hydrological Sciences Journal* **48** (3), 399–411.

Solomatine, D. P. & Ostfeld, A. 2008 Data-driven modelling: some past experiences and new approaches. *Journal of Hydroinformatics* **10** (1), 3–22.

Solomatine, D. P. & Xue, Y. 2004 M5 model trees compared to neural networks: application to flood forecasting in the upper reach of the Huai River in China. *ASCE Journal of Hydrologic Engineering* **9** (6), 491–501.

Tyralis, H., Papacharalampous, G. & Langousis, A. 2019 Super learning for daily streamflow forecasting: large-scale demonstration and comparison with multiple machine learning algorithms. arXiv preprint arXiv:1909.04131.

Wang, Y. & Witten, I. H. 1996 *Induction of Model Trees for Predicting Continuous Classes*.

Willmott, C. J. 1981 On the validation of models. *Physical Geography* **2** (2), 184–194.

Wu, C. L. & Chau, K. W. 2010 Data-driven models for monthly streamflow time series prediction. *Engineering Applications of Artificial Intelligence* **23** (8), 1350–1367.

Yarar, A. 2014 A hybrid wavelet and neuro-fuzzy model for forecasting the monthly streamflow data. *Water Resources Management* **28** (2), 553–565.

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. *Journal of Hydrology* **530**, 829–844.

Zakhrouf, M., Bouchelkia, H., Stamboul, M. & Kim, S. 2020 Novel hybrid approaches based on evolutionary strategy for streamflow forecasting in the Chellif River, Algeria. *Acta Geophysica* **68**, 167–180.

First received 16 June 2020; accepted in revised form 21 September 2021. Available online 5 October 2021