Prediction of runoff in the upper Yangtze River based on CEEMDAN-NAR model
Xianqi Zhang, Zhiwen Zheng and Kai Wang

ABSTRACT
Scientific and accurate prediction of river runoff is important for river flood control and sustainable use of water resources. This study evaluates the ability of a Nonlinear Auto Regressive model (NAR) in predicting runoff volume. Using the Cuntan Hydrological Station in the upper reaches of the Yangtze River as the research object, the model was established based on the runoff characteristics from 1951 to 2020 and tested by NAR. To improve the prediction efficiency, complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) preprocessing technique is used to decompose the data. The results show that the coupled CEEMDAN-NAR model has better predictive ability than the single model, with a coupled model deterministic coefficient (DC) of 0.93 and a prediction accuracy of Class A.

Key words | CEEMDAN, NAR, runoff projections, upper Yangtze River

HIGHLIGHTS
- Proposed a runoff prediction model using CEEMDAN-NAR hybrid approach.
- Decomposition of runoff data using CEEMDAN preprocessing techniques to improve model prediction accuracy.
- Using models to predict runoff in the upper Yangtze River from 2005 to 2020 and verifying their accuracy
- The CEEMDAN-NAR model provides better predictions than the GRU, NAR and EEMD-NAR prediction models.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (http://creativecommons.org/licenses/by/4.0/).
doi: 10.2166/ws.2021.121
INTRODUCTION

The changes in runoff elements affect the whole hydrological and water resources system, and the change pattern is intricate and complex, with nonlinear, sudden change, randomness and other complex characteristics. 

Improving the accuracy of medium-term and long-term forecasts of runoff in the upper reaches of the Yangtze River is important for the rational planning and allocation of water resources in the Yangtze River basin. At present, the commonly used medium-term and long-term runoff forecasting models include the multiple regression model and the fuzzy pattern recognition model, but the solution function of the multiple regression model is too complicated. Fuzzy systems do not have adaptive learning capabilities, and it is also a tricky problem to automatically generate and adjust affiliation functions and fuzzy rules. For this reason, the introduction of artificial neural networks with better nonlinear and self-learning capabilities for runoff prediction has become a new topic of research for many scholars. However, due to the complex characteristics of hydrological data, such as nonlinearity and mutability, there is some error in using a single neural network to predict long series data. Reducing the nonstationarity of runoff series by building coupled prediction models is a new way to improve the accuracy of runoff prediction.

Roushangar et al. (2021a, 2021b, 2021c) used wavelet transform (WT) and ensemble empirical mode decomposition (EEMD) preprocessing techniques combined with nonlinear neural ensemble methods to predict precipitation index
series for the period 1978–2017 at four stations in northwestern Iran, and the results showed that data preprocessing can effectively improve model accuracy. Roushangar et al. (2021a, 2021b, 2021c) evaluated the predictive capability of kernel extreme learning machine (KELM) for daily suspended sediment concentration (SSC) and discharge (SSD) of rivers and processed the data by WT and EEMD preprocessing techniques to improve the model prediction efficiency. Roushangar et al. (2021a, 2021b, 2021c) evaluated the predictive capability of GPR on daily river stagnation and processed the data by WT and EEMD preprocessing techniques. The results showed that WT-GPR and EEMD-GPR improved the model prediction capability by 25–40%.

Numerous studies have shown that the prediction accuracy of the model can be greatly improved by preprocessing the data compared with the traditional neural network model. Complete ensemble empirical model decomposition with adaptive noise (CEEMDAN) described by Torres et al. (2016) is a method of signal decomposition based on the time-scale characteristics of the data themselves, which can accurately reconstruct the original signal and improve the decomposition efficiency. CEEMDAN is used to decompose the runoff data, and the decomposed series have good stability. Combined with the Nonlinear Auto Regressive model (NAR), which has strong autonomous learning adaptation and generalization ability, a CEEMDAN-NAR model for annual runoff prediction of the upper Yangtze River was constructed, which provided a new way for runoff prediction. The purpose of this study is to evaluate the predictive capability of the NAR model in terms of runoff. To improve the prediction accuracy, the data were preprocessed using CEEMDAN to compare the predictive capability of the coupled CEEMDAN-NAR model with that of a single model. It also provides some reference for the rational planning and allocation of water resources in the Yangtze River basin.

MODELS AND METHODS

CEEMDAN model

CEEMDAN is an improvement of EMD (Empirical Mode Decomposition) by Huang et al. (1998) and EEMD (Ensemble Empirical Mode Decomposition) from Wu & Huang (2009). EMD and EEMD are often subject to modal mixing and data reconstruction errors that affect the accuracy of the decomposition. CEEMDAN does this by adding a limited number of times of adaptive white noise to each process. The reduction in the number of averaging times presupposes the same reduction in the reconstruction error of the signal. CEEMDAN decomposes complex time series signals into a series of simple Intrinsic Mode Functions (IMFs) and a residual term (Res).

The IMF obtained by CEEMDAN decomposition was denoted as $IMFn$, defining the operator $E_c(\cdot)$ as the $c$th order eigenmode component generated by the EMD algorithm. $\omega_i$ is the white noise signal satisfying $N(0,1)$ distribution and $x(t)$ is the original signal. Then the CEEMDAN method can be described as follows:

(1) Decompose using the EMD algorithm to obtain first-order eigenmode components:

$$IMF_1(t) = \frac{1}{M} \sum_{i=1}^{M} IMFi_1(t) = IMF_1(t)$$

(2) Calculate the residuals after first order eigenmode decomposition:

$$r_1(t) = x(t) - IMF_1(t)$$

(3) Perform $i$ experiments ($i-1, \cdots, M$) and decompose the signal $r_1(t) + \epsilon E_1(\omega(t))$ until the eigenmode components of the first-order EMD are obtained. On this basis, continue to calculate the second-order eigenmode component:

$$IMF_2(t) = \frac{1}{M} \sum_{i=1}^{M} E_1(r_1(t) + \epsilon E_1(\omega_i(t))$$

(4) For each remaining stage, that is $n = 2, \cdots, N$ calculate the signal of the nth residual:

$$r_n(t) = r_{n-1}(t) - IMF_n(t)$$
(5) Repeat the steps of (3), such that the intrinsic modal components of order $n + 1$ is:

$$IMF_{k+1}(t) = \frac{1}{N} \sum_{i=1}^{N} E_1(t_i) + \varepsilon_k E_k(w(t))$$  \hspace{1cm} (5)

(6) Repeat (4) and (5) until the residual signal satisfies the termination condition of the decomposition, and finally $K$ modal components are obtained. The final residual signal of the decomposition is:

$$R(t) = X(t) - \sum_{i=1}^{k} IMF_i(t)$$  \hspace{1cm} (6)

The final original signal can be decomposed into:

$$X(t) = R(t) + \sum_{i=1}^{k} IMF_i(t)$$  \hspace{1cm} (7)

**NAR neural networks**

The NAR Ma et al. (2020) neural network uses itself as the regression variable, and it represents the random variables at a subsequent time by a linear combination of random variables over time. As a dynamic neural network based on a time series, its output is not only a static mapping, but also a comprehensive utilization of previous dynamic results, with feedback and memory functions Farhan & Rajiv (2020).

The NAR dynamic neural network consists of a regression neural network with an input layer, an implicit layer, an input delay layer, and an output layer Yuan (2016). Different from the BP neural network, the NAR neural network adds an input delay layer to the hidden layer, which is used as the delay number to record the previous data, so as to realize the dynamic memory of the system. The NAR neural network has good memory and stability, and the network has been used in a wide range of applications Huang & Lu (2016). A schematic diagram of the NAR network structure is shown in Figure 1.

In Figure 1, $y(t)$ on the left represents input data, $y(t)$ on the right represents output data, $1.5$ represents the input and output delay order, $W$ represents the connection power, $b$ represents the threshold value, and $20$ represents the number of implied layers. The NAR common network model expression is as follows:

$$y(t) = f(y(t - 1), y(t - 2), y(t - 3) \ldots y(t - n))$$  \hspace{1cm} (8)

In the formula, $n$ represents the order of the delay layer. It can be seen from this formula that the output $y(t + 1)$ at the next moment depends on the $y(t)$ at the previous $n$ moments, which means that the model has a delay, and the past value is used to infer the current value.

The delay function structure is shown in Figure 2.

This structure diagram shows that the output is time $x(t + 1)$, and the result depends on the recording of the delay function, that is, the input value $x(t), x(t - 1), \cdots, x(t - T)$ of the neural network, so that the NAR neural network model has the ability to record previous data.

**Gated recurrent unit (GRU) models**

Hochreiter & Schmidhuber (1997) proposed Long Short-Term Memory (LSTM) neural networks, a deformation model of recurrent neural networks, to overcome the problem of long dependencies in recurrent neural networks. However, the
LSTM neural network model has a more complex form and takes a long time to train. To address this problem, Cho et al. (2014) improved the LSTM model and proposed the GRU neural network model. The GRU neural network consists of a reset gate and an update gate. The reset gate determines how the new input information is combined with the previous memory, and the update gate defines the amount of previous memory saved to the current time step. The GRU model has simplified the structure and reduced the number of parameters while retaining the function of long and short-term memory network units, thus significantly improving the training speed and overcoming the long-term dependency problem of traditional neural networks, which makes it stand out in the problem of predictive analysis of long series data. The GRU principle is shown in Figure 3.

The GRU model can accurately extract the structural features of the data, has a fast learning speed and a low number of iterations, with good predictive performance, and is suitable for the analysis of nonlinear data.

The prediction results of this model are compared with those of the CEEMDAN-NAR model to compare and analyse whether the prediction results of the coupled CEEMDAN-NAR model are good.

The specific expressions and structural diagram of the GRU model are shown below:

\[ z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \]  
(9) 

\[ r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \]  
(10) 

\[ \tilde{h}_t = \tanh(W \cdot [r_t \cdot h_{t-1}, x_t]) \]  
(11) 

\[ h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \]  
(12)

CEEMDAN-NAR models

In order to improve the prediction accuracy of runoff, the CEEMDAN-NAR model is proposed, which is a ‘decomposition–prediction–reconstruction’ process, using CEEMDAN to decompose the runoff data into a finite number of subseries with high to low frequencies, which include local features and residual terms of the original data at different time scales and are relatively linear and steady-state. Using NAR to predict subseries with different frequencies, the predicted data contain more information and the prediction results are more consistent with the real situation. The specific steps of CEEMDAN-NAR coupling model are as follows:

1. CEEMDAN decomposition: MATLAB is used to perform CEEMDAN decomposition on the original data to obtain the IMF components and residual terms of the time series.
2. Divide training data and validation data: The runoff data from 1951 to 2004 is used as the training data of the NAR network, and the data from 2005 to 2020 is used as the validation data.
3. NAR neural network prediction: The NAR network is used to repeatedly debug the training data to make the prediction reach the best effect.
4. Analysis of prediction results: Finally, the predicted IMF components and residual terms are cumulatively reduced and compared with the original data.

The technical route of the CEEMDAN-NAR coupling prediction model is shown in Figure 4.

Statistical evaluation indicators

In order to more clearly reflect the error of the model prediction results and the prediction accuracy, four indicators, the deterministic coefficient (DC), the mean absolute error (MAE), the mean absolute percentage error (MAPE) and the root mean square error (RMSE), were used for the analysis, and the equations were calculated as follows:

\[ DC = 1 - \frac{\sum_{i=1}^{n} [y_i(i) - \bar{y}_i(i)]^2}{\sum_{i=1}^{n} [\bar{y}_i(i) - \bar{y}][i]} \quad i = 1, 2, \ldots, n \]  
(13)
In the equation, $y_c(i)$ is the predicted value, $y_0(i)$ is the measured value, $\bar{y}_0$ is the mean value of the measured value, and $n$ is the sequence length.

According to the Specification of Hydrological Intelligence Forecasting (SL250-2000), the closer the DC value is to 1, the better is the model prediction. When $DC \geq 0.9$, the prediction accuracy is Grade A; when $0.7 \leq DC < 0.9$, the prediction accuracy is Grade B; when $0.5 \leq DC < 0.7$, the prediction accuracy is Grade C; when $DC < 0.5$, the prediction result is unreliable.

**CASE STUDIES**

**Data sources**

The Cuntan Hydrological Station on the upper reaches of the Yangtze River is located in Chongqing, 7.5 km downstream of the confluence of the Yangtze River and the Jialing River, with a catchment area of 866,600 km². It controls the basic water regime after the Minjiang, Tuojiang, Jialing and Chishui Rivers merge into the Yangtze River. An important reservoir control station for the Three Gorges Project and an important hydrologic control station for the upper reaches of the Yangtze River. Cuntan Hydrological Station was built in 1936, and is a national level hydrological station and has complete and continuous measurement information. The largest flood observed since records began at the station occurred in 1981, when the water level reached 191.45 meters. The main measurement of water level, flow, sediment, water quality, etc. Cuntan hydrological station controls more than 60% of the water upstream of the Yangtze River, providing countless important information for leaders at all levels who are responsible for flood control.

The location of the study area is shown in Figure 5.

The runoff data from 1951–2020 from Cuntan Hydrological Station were used for the study data, In which runoff data from 1951 to 2004 were selected as simulation training for the model and data from 2005 to 2020 were used for validation. The runoff changes at the Cuntan Hydrological Station from 1951 to 2020 are shown in Figure 6.

**CEEMDAN decomposition**

CEEMDAN was used to decompose the runoff data from 1951 to 2020 at the Cuntan Hydrological Station on the upper reaches of the Yangtze River. After repeated experiments, the annual runoff was decomposed into four IMF components and one remaining term. The results are shown in Figure 7.

As can be seen from the Figure 7, the IMF component exhibits variations in frequency, amplitude and wavelength of the runoff time series. Among them, the IMF1 component...
has the highest frequency, the largest amplitude, and the shortest wavelength. From IMF1–IMF4, the amplitude of the IMF component gradually decreases, the frequency gradually decreases, and the wavelength gradually becomes longer. The remaining items represent the overall change trend of runoff at Cuntan Station from 1951 to 2020. It can be seen that the runoff from 1951 to 2020 has a downward trend.

**Annual runoff projections**

The CEEMDAN-NAR model predicts the decomposed IMF1-Res for 2005–2020 at the Cuntan Hydrographic Station, using the decomposed data from 1951 to 2004 as the model training sample to predict the data for 2005–2020, and comparing the predicted data with the true values to show the absolute error between the predicted and true values of IMF1-Res. The delay

---

**Figure 5** | Location map of the study area.

**Figure 6** | Change process of runoff at Cuntan Station from 1951 to 2020.
order is 1:5, and the number of hidden layer nodes is 20. The decomposed four IMF components are predicted with the residual term, and the absolute error between the predicted and true values of each component is displayed. The prediction results are shown in Figure 8.

It can be seen from Figure 8 that after the CEEMDAN decomposition process, the nonlinearity and nonsmoothness of the annual runoff time series are greatly reduced, and the IMF1–IMF2 data have the greatest volatility and the worst fitting effect, and the absolute error produced is also larger, fluctuating within ±200. The predicted values from IMF3-Res fit better and better with the true values, and the absolute errors show a gradually decreasing trend.

As can be seen from Table 1, the deterministic coefficient of the CEEMDAN-NAR model for predicting the annual runoff of each component at Cuntan Station becomes better gradually, and the deterministic coefficient of IMF1 is 0.81. According to the Specification of Hydrological Intelligence Forecasting (SL250-2000), the prediction accuracy of IMF1 is Grade B, and from IMF2-Res, the deterministic coefficients the coefficients from IMF2-Res are all >0.90, and the prediction accuracy is Class A.

In order to study the annual runoff of Cuntan hydrological station in a more visual and detailed way, the predicted components were superimposed, and the reconstructed data were organized into adult runoff displays to compare the absolute error between the real and predicted values of runoff. The specific results are shown in Figure 9. The DC, MAE, RMSE, and MAPE of the runoff from 2005 to 2020 at Cuntan Station were calculated, as shown in Table 2.

It can be seen from Figure 9 and Table 2 that the CEEMDAN-NAR coupling prediction model has a good prediction effect. The real value of runoff and the predicted value have the same trend, and the fitting effect is good. The difference between the real value and the predicted value in 2006. The maximum absolute error is 169.15 × 10^8 m^3, and the absolute error is the smallest in 2010. The model has a certainty coefficient of 0.94 for runoff prediction at Cuntan Station from 2005 to 2020, which reaches the Class A accuracy specified in Hydrological Intelligence Forecasting (SL250-2000).
Figure 8 | IMF1-Res annual runoff prediction effect map of Cuntan Station.
Comparative analysis

To verify the superiority of the CEEMDAN-NAR coupled model, the GRU model, NAR model, EEMD-NAR model, and CEEMDAN-NAR coupled model were compared. The prediction results of the CEEMDAN-NAR model and the other models are shown in Figure 10.

The results show that the prediction results of the CEEMDAN-NAR model are in line with expectations, and the period and trend are highly consistent with the sequence of the measured data. Table 3 shows that the errors of CEEMDAN-NAR and EEMD-NAR models, which are ‘decomposition–prediction–reconstruction’ models, are significantly smaller than those of the single-prediction models. The reason is that the smoothness of the time series after decomposition is greatly improved, which improves the prediction accuracy. The CEEMDAN decomposition solves the problems of modal confusion and residual noise in the reconstruction sequence in the EEMD decomposition process. The decomposition process is complete and has no reconstruction errors. The errors of the prediction results of the CEEMDAN-NAR model are smaller than those of the EEMD-NAR model. Among them, compared with the GRU model, NAR model, and EEMD-NAR model, the MAE is reduced by 52.31%, 29.47%, and 14.13%, respectively, the MAPE is reduced by 2.72%, 1.00%, and 0.39%, and the RMSE is reduced by 51.11%, 23.02%, and 11.13%.

The nonstationary and nonlinear characteristics of the runoff data will affect the training process of the single prediction model, resulting in low accuracy of the prediction results. The CEEMDAN method effectively reduces the number of iterations, improves the reconstruction accuracy, and is more suitable for the analysis of nonlinear signals. CEEMDAN-NAR coupled with NAR prediction model is applied to annual runoff prediction, which has obvious advantages compared with the traditional single-prediction model and EEMD decomposition.

CONCLUSION

(1) Establish CEEMDAN-NAR to establish a coupling model, and use this model in the runoff prediction of the Cuntan Hydrometric Station in the upper reaches of the Yangtze River. The results show that the runoff of the Cuntan Hydrological Station in the upper reaches of the Yangtze River in 1951–2020 exhibited a

| Table 1 | Analysis of the deterministic index of annual runoff for each component at Cuntan Station |
|-----------|---------------------------------|
| Component | IMF1 | IMF2 | IMF3 | IMF4 | Res | overall |
| DC        | 0.81 | 0.92 | 0.96 | 0.99 | 0.99 | 0.92 |

Figure 9 | 2005–2020 runoff forecast effect map of Cuntan Station.
downward trend. The change cycle and trend of the model prediction results are consistent with the real data, the prediction error is small, and the deterministic coefficient (DC) is 0.94, reaching the Class A accuracy standard, and the prediction results are reliable.

(2) The coupled CEEMDAN-NAR model outperforms the GRU model, NAR model, and EEMD-NAR model in prediction. Taking the annual runoff data decomposed by CEEMDAN as the input data of the model can effectively improve the prediction accuracy of the NAR prediction model. The coupled prediction model based on CEEMDAN can obtain more accurate and stable prediction results, which is helpful for the study of hydrological time series prediction. The use of this model can provide a technical reference for the development and utilization of water resources, water allocation and optimal reservoir operation in the Yangtze River basin.

(3) The CEEMDAN-NAR coupled model has an effective decomposition algorithm and a powerful, fast and stable prediction tool. The use of the CEEMDAN method to decompose the runoff sequence is helpful to understand the characteristics of the periodic change of runoff, and it can transform the nonlinear and nonstationary time series into a stationary time series, which helps to improve the accuracy of prediction and solves the problem of using the NAR model alone. Forecasting is difficult to reveal the characteristics of runoff changes and the problems of low prediction accuracy, which has broad application prospects. The prediction algorithm based on the CEEMDAN-NAR model can not only be used for annual runoff prediction, but also for the prediction of other time series such as sediment, rainfall, inventory, and meteorological factors. In addition, based on the time series analysis, the physical mechanism of runoff and long-term prediction are not considered, and how to conduct a more comprehensive analysis and improve the prediction accuracy will be the focus of the next research.

**DATA AVAILABILITY STATEMENT**

All relevant data are included in the paper or its Supplementary Information.

**REFERENCES**

Alizadeh, M. J., Kavianpour, M. R., Kisi, O. & Nourani, V. 2017 A new approach for simulating and forecasting the rainfall-runoff process within the next two months. *Journal of Hydrology* **548**, 588–597.

Bai, J. Z., Shi, B., Feng, M. Q. & Zhou, L. K. 2011 Medium- and long-term runoff prediction based on an adaptively tuned ant colony-RBF neural network model. *Journal of Natural Resources* **26** (06), 1065–1074.

Chen, X., Wu, J. L. & Wang, L. 2005 Artificial neural network models predict climate change impacts on runoff from the Bosten Lake basin. *Lake Science* **17** (003), 207–212.
Cho, K., Merrienboer, B. V. & Gulcehre, C. 2014 Learning phrase representations using RNN encoder-decoder for statistical machine translation. Computer Science 9 (4), 30–37.

Coulibaly, P., Anctil, F., Rasmussen, P. & Bernard, B. 2000 A recurrent neural networks approach using indices of low-frequency climatic variability to predict regional annual runoff. *Cghu Special Issue of Hydrological Processes* 14 (15), 2755–2777.

Cui, D. W. 2013 Application of multi-hidden layer BP neural network models in runoff prediction. *Hydrology* 53 (01), 68–73.

Cui, D. W. & Jin, B. 2016 Application of the bird swarm algorithm-projection tracing regression model to multivariate annual runoff prediction. *People’s Pearl River* 37 (11), 26–30.

Dong, L. H., Xiong, L. H. & Yu, K. X. 2012 Research advances in effects of climate change and human activities on hydrology. *Advances in Water Science* 23 (2), 278–285.

Farhan, M. K. & Rajiv, G. 2020 ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India. *Journal of Safety Sciences and Resilience* 36 (05), 1146–1155.

Guo, S. L., Guo, J. L., Hou, Y. K., Hou, L. H. & Hong, X. J. 2015 Projecting future runoff changes in the Yangtze River basin based on Budyko’s hypothesis. *Advances in Water Science* 26 (02), 151–160.

Hochreiter, S. & Schmidhuber, J. 1997 Long short-term memory. *Neural Computation* 9 (8), 1735–1780.

Huang, J. F. & Lu, W. C. 2016 Wavelet-NAR neural network-based weather element time series forecasting and weather index rainbow option valuation. *Systems Engineering Theory and Practice* 36 (05), 1146–1155.

Huang, N. E., Shen, Z. & Long, S. R. 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A* 454, 903–995.

Komasi, M. & Sharghi, S. 2016 Hybrid wavelet-support vector machine approach for modelling rainfall-runoff process. *Water Science & Technology* 73 (8), 1937–1953.

Koradia, A. K., Bhalala, A. D. & Tiwari, M. K. 2020 Rainfall-runoff simulation modelling using artificial neural networks in semi-arid middle Gujarat region. *Indian Journal of Soil Conservation* 47 (3), 231–238.

Ma, F. H., Jin, Y. C. & Sun, C. Y. 2020 Short-time prediction model for metro passenger flow based on EMD-optimized NAR dynamic neural network. *Journal of Applied Sciences* 38 (06), 936–943.

Mahabir, C., Hicks, F. E. & Fayek, A. R. 2020 Application of fuzzy logic to predict seasonal runoff. *Hydrological Processes* 17 (18), 3749–3762.

Nourani, V., Andalib, G., Sadikoglu, F. & Sharghi, E. 2018 Cascade-based multi-scale AI approach for modeling rainfall-runoff process. *Hydrology Research* 49 (4), 1191–1207.

Roushangar, K., Aghajani, N., Ghasempour, R. & Alizadeh, F. 2021a The potential of ensemble WT-EEMD-kernel extreme learning machine techniques for prediction suspended sediment concentration in successive points of river. *Journal of Hydroinformatics*. https://doi.org/10.2166/hydro.2021.146 (accessed 27 April 2021).

Roushangar, K., Chamani, M., Ghasempour, R., Azamathulla, H. M. & Alizadeh, F. 2021b A comparative study of wavelet- and empirical mode decomposition-based GPR models for river discharge relationship modeling at consecutive hydrometric stations. *Water Supply*. https://doi.org/10.2166/ws.2021.073 (accessed 27 April 2021).

Roushangar, K., Ghasempour, R. & Nourani, V. 2020 The potential of integrated hybrid pre-post-processing techniques for short-to-long term drought forecasting. *Journal of Hydroinformatics* 25 (1), 117–135.

Torres, M. E., Colominas, M. A., Schlotthauer, G. & Flandrin, P. 2016 Complete ensemble empirical mode decomposition with adaptive noise. In ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing-Proceedings, pp. 4144–4147.

Wan, X. & Tang, T. G. 2009 Chaotic network model for runoff and its application to the prediction study of Yangtze River Three Gorges. *China Rural Water and Hydropower* 14 (04), 1–3.

Wu, Z. & Huang, N. E. 2009 Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis* 01 (01), 1–41.

Yuan, L. S. 2016 *Vehicle Speed Prediction Based on NAR Neural Networks and Applications*, PhD Thesis. Dalian University of Technology, Dalian, China.

Zhu, Z. H., Zhang, H. W. & Liu, X. 2014 Mine power hazard prediction based on fuzzy pattern recognition. *Journal of Natural Hazards* 23 (04), 19–25.

First received 4 March 2021; accepted in revised form 12 April 2021. Available online 26 April 2021