

Impact of Normalization and Input on ARMAX-ANN Model Performance in Suspended Sediment Load Prediction

Hamid Moeeni¹ · Hossein Bonakdari¹ 

Received: 22 March 2017 / Accepted: 24 October 2017 /
Published online: 9 November 2017
© Springer Science+Business Media B.V. 2017

Abstract The suspended sediment load in rivers is an important parameter in watershed planning and management. Since daily suspended sediment time series contain linear and nonlinear components, existing prediction models are associated with limitations. Therefore, this study introduces a new hybrid model comprising two commonly used stochastic and nonlinear models. The sediment load is first modeled by an autoregressive-moving average with exogenous terms (ARMAX) model. Subsequently, the ARMAX residuals are modeled with an artificial neural network (ANN). For this purpose, discharge (Q) and sediment (S) are considered as model input parameters. Three modeling scenarios are defined to investigate the impact of data normalization on the hybrid model. The exponential and Box-Cox transformation methods are combined into a new data normalization method called mixed transformation. The performance of these methods is then compared. In addition, the impact of the type and number of input combinations on ARMAX-ANN model accuracy is evaluated. To this end, 12 input combinations and 1331 ARMAX and ANN models are verified. The ARMAX model inputs include S , Q and the white noise disturbance term (e), while the ANN model inputs include the ARMAX model results and residuals. Moreover, the hybrid model's accuracy is compared with the ARMAX and ANN models.

Keywords Sediment load · River · Stochastic · Hybrid model · Data normalization · Transformation

1 Introduction

Predicting the suspended sediment load of rivers accurately is among the most significant factors in hydraulic structure design. The suspended sediment load of a river can be considered a function of hydrological and meteorological parameters, determining which directly affects

✉ Hossein Bonakdari
bonakdari@yahoo.com

¹ Department of Civil Engineering, Razi University, Kermanshah, Iran

suspended sediment load and is a complicated and costly process. Therefore, applying a classical hydromechanics approach cannot produce reliable results (Alp and Cigizoglu 2007). Hence, in modeling the suspended sediment load in a river, optimum combinations of the most effective and significant parameters should be used. Alp and Cigizoglu (2007) demonstrated that river discharge is the main hydrological factor. Furthermore, other studies have indicated that utilizing measured sediment load along with discharge data contributes to optimum prediction of suspended sediment load (Afan et al. 2015). In some cases, the river discharge parameter is measured and is more suitable for use in modeling and prediction.

Traditional statistical models including autoregressive moving average (ARMA) and autoregressive-moving average with exogenous terms (ARMAX) are applicable for modeling and predicting sediment load. ARMA predictions only involve the impact of sediment, while ARMAX uses other effective modeling parameters such as flow characteristics. However, these models are not adequate for nonlinear hydrological problems (Moeeni and Bonakdari 2016), because as the suspended sediment load system of a river exhibits more complex behaviour, statistical models cannot produce suitable functionality modeling results. Models based on computational intelligence are another means of modeling nonlinear systems; hence, they are an appropriate alternative to statistical models. Rahim and Akif (2015) and Mustafa et al. (2012) evaluated Artificial Neural Network (ANN) model performance in predicting suspended sediment load. Researchers have also studied and compared the performance of ANN models with regression models from physical and statistical perspectives (Demirci et al. 2015; Tiwari and Rai 2015). In addition to ANN, other computational intelligence models may be used to predict suspended sediment. Comparing these models has also attracted much interest (Kisi et al. 2012; Lafdani et al. 2013; Alizdeh et al. 2015; Kumar et al. 2016). Despite the respective studies presenting different results, it appears that ANN models perform reasonably.

However, the ANN model considers the time series under study as nonlinear, whereas time series may not always be purely nonlinear (Moeeni et al. 2017). In practical cases, it is challenging to understand whether a time series is created from a linear or nonlinear underlying process or whether stochastic methods are more effective than ANN methods with out-of-sample prediction data. To overcome the limitations of statistical and ANN models, Zhang (2003) presented a hybrid model combining both ARMA and ANN models (ARMA-ANN). Other researchers have applied this hybrid model for various purposes (Faruk 2010; Nourani et al. 2011; Liu et al. 2012). Moeeni and Bonakdari (2016) improved ARMA-ANN accuracy of forecasting time series with extreme seasonal variation and presented the four-step seasonal autoregressive integrated moving average (SARIMA)-artificial neural network (SARIMA-ANN) model. They used the new model to predict the monthly inflow to a dam reservoir with high and irregular seasonal changes.

Besides modeling, data analysis and pre-processing can also affect result accuracy. Cigizoglu and Kisi (2006) demonstrated that the initial statistical analysis of flow and sediment affect the most suitable ANN model inputs. In the present study, the daily suspended sediment load at upstream and downstream stations on the Cumberland River is modeled. Because both sediment load and discharge impact the results, discharge is also used as model input. However, no study has been found that employs a combination of stochastic and nonlinear models to predict the suspended sediment load of rivers. Therefore, the ARMAX-ANN hybrid model is proposed in this study. The effect of discharge normalization and sediment on modeling accuracy is evaluated. The impact of data normalization on the ARMAX-ANN

model is investigated in three different scenarios. A new data normalization method called mixed transformation is introduced, which is based on both exponential and Box-Cox transformation methods. Subsequently, the performance of the three transformations is compared. The impact of the number and type of ANN model inputs on ARMAX-ANN result accuracy is investigated. The proposed hybrid model results are also compared with the ARMAX and ANN models.

2 Materials and Method

2.1 Data Used

Daily water discharge (Q) and daily suspended sediment load (S) data from two stations located on the Cumberland River, USA, were used in this study. Each time series includes a 10-year period from October 1, 1979 to September 30, 1989. Data for seven years (from October 1, 1979 to September 30, 1986) were used for model training, and the remaining data were used for testing. These sites include time series data with highly irregular behavior, which can be used to evaluate the performance of different models. Since the aim is to identify an efficient model for predicting such data, these stations were selected. Table 1 presents the statistical characteristics of the data. In this table, x_{\min} , \bar{x} , x_{\max} , S_d , S_k , K_u , C_v , M_d , q_1 , q_3 are the minimum, average, maximum, standard deviation, skewness, kurtosis, coefficient of variation, median, first quarter and third quarter of discharge and suspended sediment load series. These quantities represent drastic fluctuation in this time series.

2.2 ARMAX Model

ARMAX is capable of modeling sediment load using water discharge as a linear input. This model is denoted by ARMAX (n_a, n_b, n_c) with the following equation:

$$(1 - a_1q - \dots - a_{n_a}q^{n_a})S_t = (1 - b_1q - \dots - b_{n_b}q^{n_b})Q_{t-k} + (1 - c_1q - \dots - c_{n_c}q^{n_c})e_t \quad (1)$$

Table 1 Statistical characteristics of the data employed

	Barbourville station						Pineville station					
	Discharge (Q)			Sediment load (S)			Discharge (Q)			Sediment load (S)		
	Train	Test	Whole	Train	Test	Whole	Train	Test	Whole	Train	Test	Whole
x_{\min}	1.44	2.15	1.44	0.83	0.14	0.14	1.42	1.53	1.42	0.50	0.14	0.14
\bar{x}	43.9	47	44.8	1289	1355	1306	34.5	38.5	35.7	1151	1091	1133
x_{\max}	1178	835	1178	182,000	123,000	182,000	892	966	966	244,000	135,000	244,000
S_d	71.5	73.2	72	7061	6108	6831	56.6	61.3	58.1	7751	5936	7253
S_k	5.7	4.1	5.2	14.8	10.8	14	6.7	5.8	6.4	18.4	13.8	18
K_u	59	25	48	300	167	278	72	60	68	461	261	456
C_v	1.6	1.6	1.6	5.5	4.5	5.2	1.6	1.6	1.6	6.7	5.4	6.4
M_d	21.1	25	22.3	56	40	50	17.3	20.7	18.3	48	37	44
q_1	7.5	6.4	7.4	15	8.7	12	6.7	6.3	6.6	13	8.7	11
q_3	50.4	52.1	51	292	286.8	290	40.2	44.5	42.2	187	195.3	189

Where S_t and Q_{t-k} are the suspended sediment load and discharge time series, e_t is the white noise disturbance series, $(a_1, a_2, \dots, a_{n_a})$ is the vector of autoregressive coefficients, $(b_1, b_2, \dots, b_{n_b})$ is the vector of exogenous input coefficients, $(c_1, c_2, \dots, c_{n_c})$ is the vector of moving average coefficients, n_a , n_b and n_c are orders of the autoregressive, exogenous input and moving average components, k is the dead time in the system (here it is equal to zero) and q is the delay operator.

To identify the best input combination and achieve the most accurate results, each value for n_a , n_b , and n_c was considered between 0 and 10 in the form of the set $[0, 1, 2, \dots, 10]$. Hence, three sets of 11 memberships were considered. The parameters of the $11^3 = 1331$ ARMAX model were estimated and the best input combination was identified based on the evaluation criteria. The model based on this combination is considered the most accurate of all models.

2.3 ANN Model

One of the known structures of this model is the multilayer perceptron neural network (MLPNN). In this structure, a multi-input vector such as discharge and sediment load is used to model daily suspended sediment load. MLPNN includes three layers: input, hidden and output layers. The used MLPNN structure contains two inputs and one hidden layer with 4 nodes. The following relationship is related to a network with a hidden layer of S and Q inputs:

$$\hat{S}_t = a + \sum_{j=1}^g a_j f \left(b_j + \sum_{i=1}^h b_{ij} S_{t-i} + \sum_{i=1}^h c_{ij} Q_{t-i} \right) \tag{2}$$

where \hat{S}_{t+1} is the predicted discharge, c_{ij} and b_{ij} are model parameters (connection weights) in the hidden layer, a_i is a model parameter in the output layer, b_j and a are bias components in the hidden and output layers, g is the number of neurons in the hidden layer, h is the number of inputs and f is the transfer function. For modeling, the present perceptron structure with a hidden layer contains sigmoid and linear transfer functions in the hidden and output layers, respectively. The number of hidden layer neurons considered is between 1 and 20. These models were trained by the Levenberg-Marquardt (LM) algorithm, after which the best model was selected based on error criteria.

$$f(t) = \frac{1}{1 + \exp(-t)} \tag{3}$$

2.4 Time Series Normalization

Time series normality is an essential assumption in statistical models (Salas et al. 1988; Marco et al. 2012). Various research works present different interpretations of series normality. More precisely, it means that time series data follow a normal probability distribution. In most cases, series with extreme fluctuations in discharge and daily sediment for instance, do not follow a normal distribution, as seen in Fig. 1. This figure presents the discharge time series and daily sediment for both upstream and downstream stations as normal probability plots. It is observed that each time series is long with a normal line. Therefore, none of these series is normal.

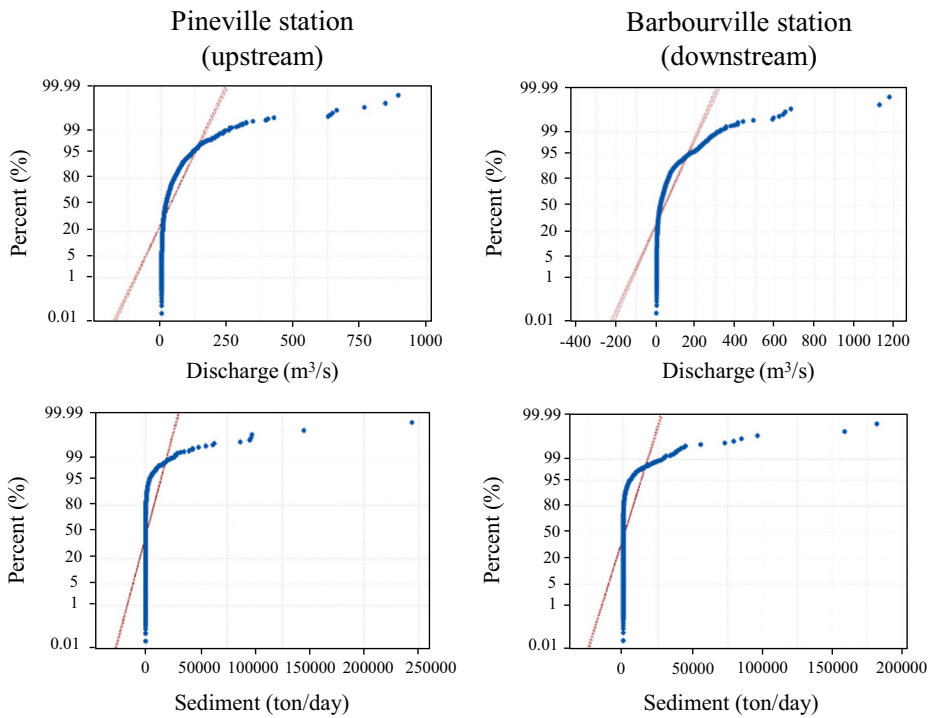


Fig. 1 Normal probability plots of original series

Data normalization has always had an important role in improving the result accuracy of statistical models. The effect of this practice is investigated on the result accuracy of the proposed hybrid model, which is based on a statistical model and a computational intelligence model. The exponential and Box-Cox transformations were used for normalization in this study. In addition, the transformation notion served as a basis for presenting a mixed transformation. The discharge and daily sediment time series for both stations were normalized using each of the three transformations. Their relationships are presented below:

$$\text{Exponential function transformation : } y = x^a \tag{4}$$

$$\text{Box-Cox transformation : } y = \begin{cases} \frac{(x + b)^c - 1}{c} & c \neq 0 \\ \text{Ln}(x + b) & c = 0 \end{cases} \tag{5}$$

$$\text{Mixed transformation : } y = \begin{cases} \frac{(x^{a'} + b')^{c'} - 1}{c'} & c' \neq 0 \\ \text{Log}(x + b') & c' = 0 \end{cases} \tag{6}$$

where x is the original series, y is the normalized series, a is the exponential transformation coefficient, b and c are the Box-Cox transformation coefficients and a' , b' and c' are the mixed

transformation coefficients. There is usually no time series that completely follows a normal distribution after normalization, but will approach the distribution as much as possible. The normality results of each mentioned transformation were compared using the Jarque-Bera, Doornick chi-squared and Anderson-Darling tests. The statistics related to these tests are provided below:

$$JB = N \left(\frac{S_k^2}{6} + \frac{(K_u - 3)^2}{24} \right) \tag{7}$$

$$DCS = Z_1^2 + Z_2^2 \sim \chi^2(2) \tag{8}$$

$$AD = -N - \frac{1}{N} \sum_{t=0}^N (2t-1) (\ln(F(Y_t)) + \ln(1-F(Y_{N+1}))) \tag{9}$$

where *JB*, *DCS* and *AD* are the Jarque-Bera, Doornick chi-squared and Anderson-Darling test statistics, respectively; *S_k* is skewness, *K_u* is kurtosis, *Z₁²* and *Z₂²* are the transformed skewness and kurtosis, $\chi^2(2)$ is the chi-square distribution with two freedom degrees, *F(x_t)* is the cumulative distribution function of the standard normal distribution and *Y_t* is the ordered and standardized data. A smaller quantity of each statistic leads to a perfectly normal time series.

2.5 ARMAX-ANN Hybrid Model

The ARMAX-ANN hybrid model presented in this study is used to predict the daily suspended sediment load as follows:

$$\hat{S}_t = (a_1q + \dots a_{n_a}q^{n_a})S_t + (1-b_1q - \dots - b_{n_b}q^{n_b})Q_{t-k} - (c_1q + \dots + c_{n_c}q^{n_c})e_t + \left(a + \sum_{j=1}^g a_j f \left(b_j + \sum_{i=1}^h b_{ij} e_{t-i} \right) \right) \tag{10}$$

The parameters in this equation are the parameters defined for Eqs. 1 and 2. If the Box-Cox conversion (Eq. 5) is used to normalize the discharge time series and sediment load, the equation provided for the hybrid model is changed as follows:

$$\hat{S}_t = \begin{cases} \left(c \left((a_1q + \dots a_{n_a}q^{n_a}) \left(\frac{(S_t + b)^c - 1}{c} \right) + (1-b_1q - \dots - b_{n_b}q^{n_b}) \left(\frac{(Q_{t-k} + b)^c - 1}{c} \right) \right) + 1 \right)^{\frac{1}{c}} \\ \quad + (1-c_1q - \dots - c_{n_c}q^{n_c})e_{n(t)} & c \neq 0 \\ -b + \left(a + \sum_{j=1}^g a_j f \left(b_j + \sum_{i=1}^h b_{ij} e_{t-i} \right) \right) & \\ \quad \left(\exp \left((a_1q + \dots a_{n_a}q^{n_a}) \text{Ln}(S_t + b) + (1-b_1q - \dots - b_{n_b}q^{n_b}) \text{Ln}(Q_{t-k} + b) \right) - b \right) \\ \quad + \left(a + \sum_{j=1}^g a_j f \left(b_j + \sum_{i=1}^h b_{ij} e_{t-i} \right) \right) & c = 0 \end{cases} \tag{11}$$

Following linear component normalization and data modeling, the normal residuals are modeled with the nonlinear component. The hybrid model equation with the Box-Cox conversion is as follows:

$$\hat{S}_t = \begin{cases} \left(\left((a_1 q + \dots + a_{n_a} q^{n_a}) \left(\frac{(S_t + b)^c - 1}{c} \right) + (1 - b_1 q - \dots - b_{n_b} q^{n_b}) \left(\frac{(Q_{t-k} + b)^c - 1}{c} \right) \right) + 1 \right)^{\frac{1}{c}} - b & c \neq 0 \\ \left(\exp \left((a_1 q + \dots + a_{n_a} q^{n_a}) \text{Ln}(S_t + b) + (1 - b_1 q - \dots - b_{n_b} q^{n_b}) \text{Ln}(Q_{t-k} + b) \right) + (1 - c_1 q - \dots - c_{n_c} q^{n_c}) e_{n(t)} + \left(a + \sum_{j=1}^g a_j f \left(b_j + \sum_{i=1}^h b_{ij} e_{n(t-i)} \right) \right) \right) - b & c = 0 \end{cases} \quad (12)$$

In order to study the impact of normalization on the results, three modeling scenarios were defined for the ARMAX-ANN model. Figure 2 shows the hybrid model steps based on the three scenarios. The process in each scenario is as follows:

- **Scenario 1:** The data does not undergo normalization. The daily suspended sediment load series ($S(t)$) is initially modeled by ARMAX. The model inputs are $S(t)$ and $Q(t)$, while the output is the linear term of sediment load (\hat{S}_L). Then the ARMAX model residual ($e(t)$) is modeled by ANN and the ANN model output is the nonlinear term of sediment load ($\hat{e}(t)$). Finally, the linear and nonlinear components are summed, leading to the modeled sediment load (\hat{S}) in scenario 1.
- **Scenario 2:** Here, $S(t)$ and $Q(t)$ are normalized first. The normalized sediment load series ($S_n(t)$) and normalized discharge series ($Q_n(t)$) are obtained. $S_n(t)$ is modeled by ARMAX and the normalized linear term of sediment load (\hat{S}_{nL}) is obtained. With component denormalization, the linear term of sediment load (\hat{S}_L) is computed. Then the difference between $S(t)$ and \hat{S}_L is calculated, which is the ARMAX model error after denormalization ($e(t)$). In the next step, the $e(t)$ series is modeled by ANN and the nonlinear term of sediment load ($\hat{e}(t)$) is obtained. Finally, the modeled sediment load (\hat{S}) is the calculated sum of the linear and nonlinear components.
- **Scenario 3:** In this scenario, $S(t)$ and $Q(t)$ are normalized. Then $S_n(t)$ is modeled by ARMAX and the normalized linear term of sediment load (\hat{S}_{nL}) is obtained. The ARIMA model residual from the normalized series ($e_n(t)$) is modeled by ANN. Therefore, the normalized nonlinear term ($\hat{e}_n(t)$) is obtained. The sum of $\hat{e}_n(t)$ and \hat{S}_{nL} is the normalized modeled sediment load (\hat{S}_n). With series denormalization, the modeled sediment load (\hat{S}) is obtained.

In this hybrid model, the ANN model inputs are selected by two approaches. Overall, 12 input combinations are used. Group 1 (e_{t-1} (Model 1), e_{t-1}, e_{t-2} (Model 2), $e_{t-1}, e_{t-2}, e_{t-3}$ (Model 3), $e_{t-1}, e_{t-2}, e_{t-3}, e_{t-4}$ (Model 4)) contains inputs consisting only of ARMAX model residuals. In addition to the ARMAX model residuals, group 2 (S_{t-1}, e_{t-1} (Model 5), $S_{t-1}, e_{t-1}, e_{t-2}$ (Model 6), $S_{t-1}, e_{t-1}, e_{t-2}, e_{t-3}$ (Model 7), $S_{t-1}, e_{t-1}, e_{t-2}, e_{t-3}, e_{t-4}$ (Model 8)) includes this model's results in a time step before (S_{t-1}) as well. In addition to S_{t-1} , group 3 ($S_{t-1}, S_{t-2}, e_{t-1}$ (Model 5), $S_{t-1}, S_{t-2}, e_{t-1}, e_{t-2}$ (Model 6), $S_{t-1}, S_{t-2}, e_{t-1}, e_{t-2}, e_{t-3}$ (Model 7), $S_{t-1}, S_{t-2}, e_{t-1}, e_{t-2}, e_{t-3}, e_{t-4}$ (Model 8)) contains S_{t-1} as an input besides other inputs.

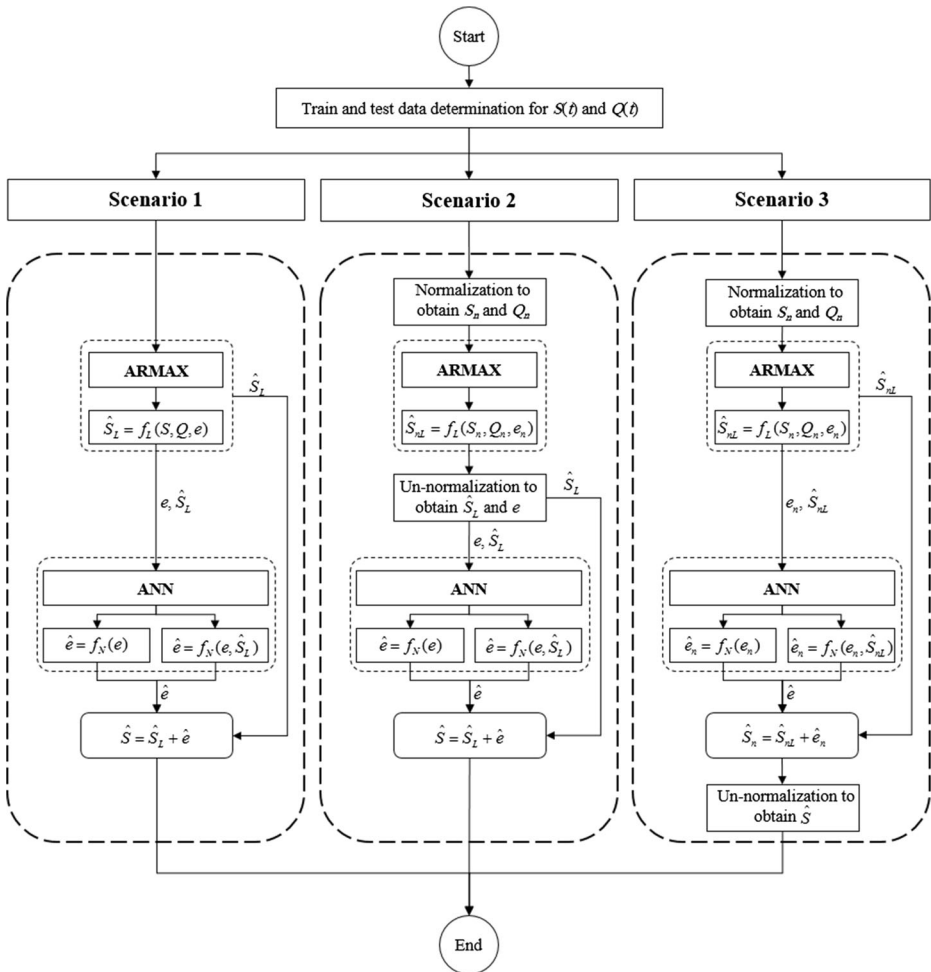


Fig. 2 Hybrid ARMAX-ANN model flowchart with three scenarios

In the first case, the inputs only include ARMAX model error values in past time steps (group 1). This means that for nonlinear component estimation only the residuals are used. Nonetheless, the linear component value can affect the nonlinear component. Therefore, to assess the relationship between these two components in the second case, the results of this model are used in past time steps as input in addition to the ARMAX model errors (groups 2 and 3). In other words, in this case the linear component is used to estimate the nonlinear component. The effects of using linear components as ANN model inputs on ARMAX-ANN model accuracy are examined. Thus, the impact of input type and number on the model results can be evaluated and compared better. The ARMAX-ANN model results are compared based on each of the 12 input combinations in the superior scenario.

2.6 Evaluation Criteria

The performance of the hybrid model in the normalization scenarios with different inputs is compared with individual models using the scatter index (*SI*), mean absolute error (*MAE*),

mean absolute relative error (*MARE*), coefficient of residual mass (*CRM*), variance accounted for (*VAF*), correlation coefficient (R^2) and Akaike's information criterion (*AIC*). The *SI* and *MAE* criteria are based on average absolute error and are sensitive to mean and limited data. The *MARE* criterion, which is based on relative error, is more sensitive to base values (small quantities). The *CRM* criterion indicates the underestimation or overestimation of the predicted discharge. *VAF* shows the difference in error variance compared to the data variance. R^2 indicates the extent of the linear relationship between the actual and predicted values. *AIC* indicates the most parsimonious model using the errors and the number of model parameters. These criteria are as follows:

$$SI = \left(\left(\sum_{t=1}^n (S_t - \hat{S}_t)^2 \right) / n \right)^{0.5} / \bar{S}(t) \quad (13)$$

$$MAE = \sum_{t=1}^n |S_t - \hat{S}_t| / n \quad (14)$$

$$MARE = \sum_{t=1}^n \left| (S_t - \hat{S}_t) / S_t \right| / n \quad (15)$$

$$CRM = \left(\sum_{t=1}^n S_t - \sum_{t=1}^n \hat{S}_t \right) / \sum_{t=1}^n S_t \quad (16)$$

$$VAF = \left(1 - \text{var}(S_t - \hat{S}_t) / \text{var}(S_t) \right) \times 100 \quad (17)$$

$$R^2 = \left(\sum_{t=1}^n (S_t - \bar{S}_t) (\hat{S}_t - \bar{\hat{S}}_t) \right)^2 / \left(\sum_{t=1}^n (S_t - \bar{S}_t)^2 \sum_{t=1}^n (\hat{S}_t - \bar{\hat{S}}_t)^2 \right) \times 100 \quad (18)$$

$$AIC = \frac{2N(k+1)}{N-k-1} + N \ln(\sigma_\varepsilon^2) \quad (19)$$

3 Results and Discussion

3.1 Performance of the Normalization Transformations

Figure 3a and b illustrate normal probability plots of the discharge and suspended sediment load series. Comparing these two figures with Fig. 1 signifies that the transformed series is closer to a normal distribution than the original series. Therefore, all three transformations were effective on normalization. However, comparing the results of these three transformations clarifies which one is the most capable. It is evident from Fig. 3a that at both upstream and downstream stations the exponential transformation showed weaker performance than the other transformations in daily discharge normalization. The Box-Cox transformation results exhibited the greatest proximity to the normal line. Hence, it seems that in daily discharge

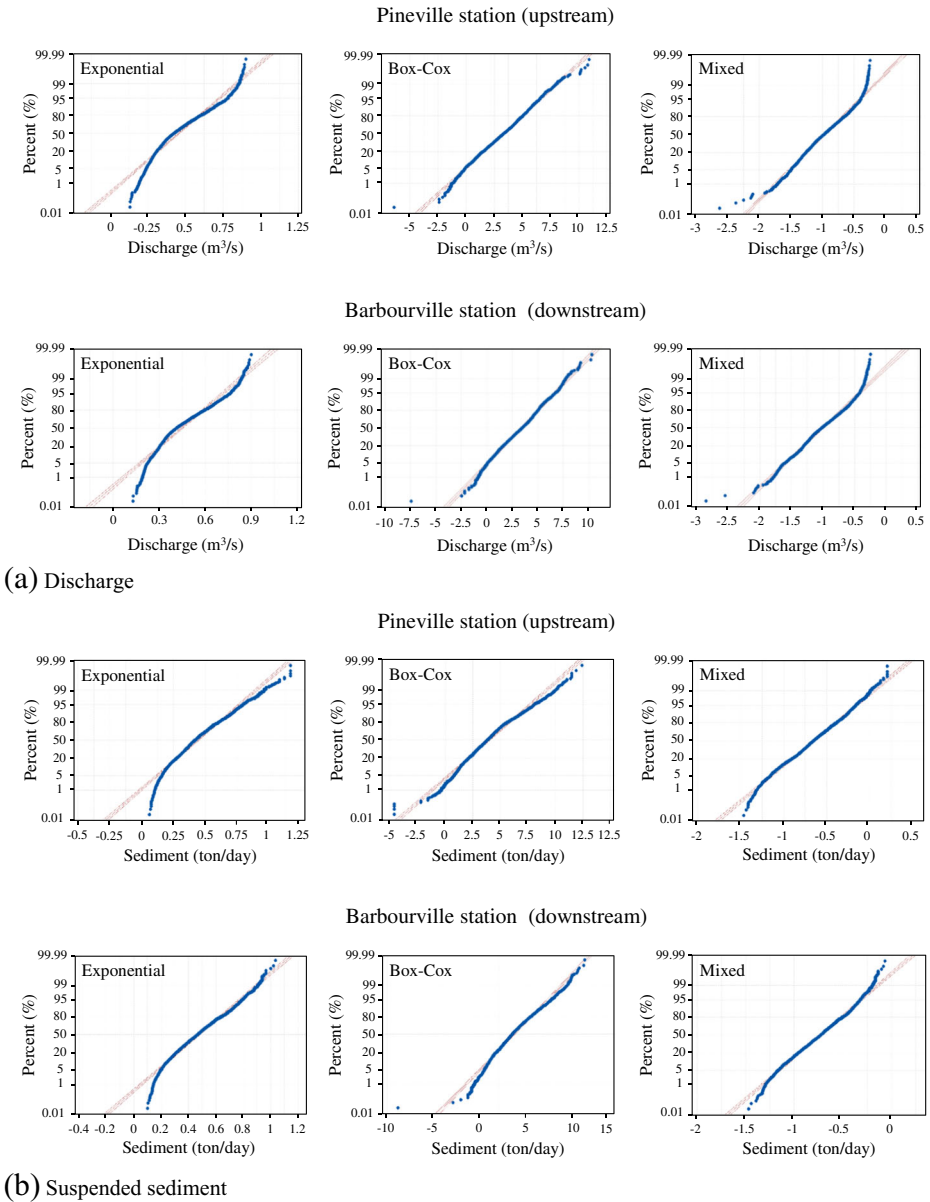


Fig. 3 Normal probability plots of transformed discharge and suspended sediment series

normalization, the Box-Cox transformation is more capable than the others. The comparison of the transformation results in Fig. 3b indicates that the exponential transformation had the weakest performance in sediment load normalization. The intuitive results of the Box-Cox and mixed transformations are almost the same.

In addition to the intuitive investigation, the transformation results based on the normalization tests appeared more accurate and scientific. The tests include the Jarque-Bera (*JB*), Doornick Chi-square (*DCS*) and Anderson-Darling (*AD*) tests. The results of these tests for the

transformed discharge and sediment series are provided in Table 2. In this table, the critical values for all three tests are presented relative to a significance level of 1%. By comparing the computational statistics and these critical values, it can be concluded that in some cases the series was not normalised with any transformation, indicating a limitation in the normalization of daily time series data. However, by comparing the performance of the transformations, it can be determined which transforms bring the discharge and sediment series closer to a normal distribution. For the transformed discharges, the exponential transformation statistic is larger than the other two transformations for both upstream and downstream stations. It was found that the transformation made the original series too proximate to a normal series, but the statistic magnitude indicates that the series was still not normal. As a result, the exponential transformation is not appropriate for normalizing daily discharge. The test statistics for the Box-Cox transformation are very small. As a result, this transformation normalized the daily discharge at both stations well. A comparison of the mixed transformed values with the two others shows that these transformations present much stronger performance than the exponential transformation. However, the Box-Cox transformation is the best for normalizing daily discharge.

The normality evaluation test results for the transformed sediment series (Table 2) are different from those for discharge. For the downstream station, the *JB* and *AD* values for the exponential transformation are smaller than the Box-Cox transformation, but the difference is very small. Meanwhile, the *DCS* value for the Box-Cox transformation is smaller with a large difference. At the upstream station, the values of these three tests for Box-Cox are smaller than the exponential transformation. As a result, the Box-Cox transformation is more capable of normalizing daily sediment load than the exponential transformation. Comparing the mixed transformation results with the two other transformations indicates that the obtained values are significantly smaller (especially for the upstream station). The transformation presented in this study is the best and most powerful for daily sediment load normalization. It should be noted that daily sediment load fluctuations are very intense. Due to these fluctuations, normalizing this non-normal time series is very difficult. However, the proposed transformation can handle this well.

Table 2 Statistics of normal tests for the discharge and suspended sediment series

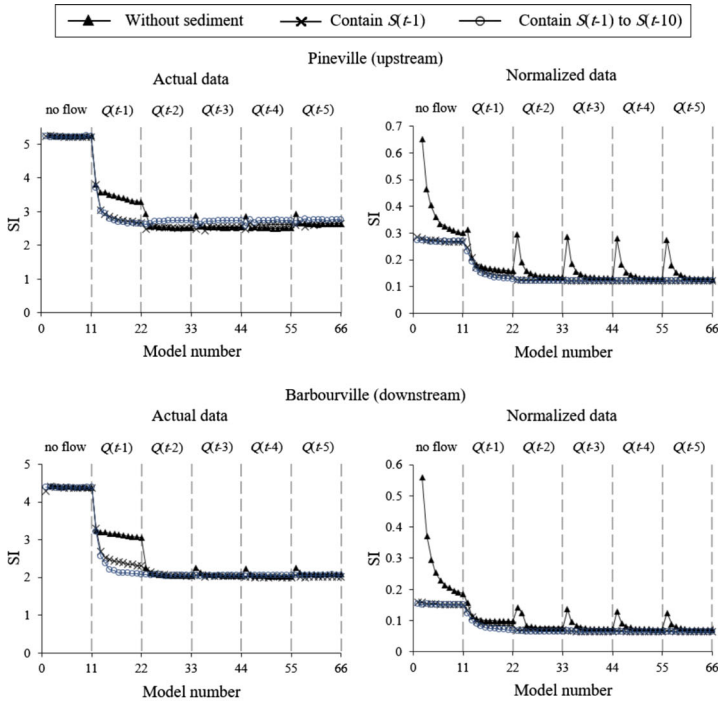
	Transformation	Downstream			Upstream		
		<i>JB</i>	<i>DCS</i>	<i>AD</i>	<i>JB</i>	<i>DCS</i>	<i>AD</i>
Discharge	Critical value	9.21	9.21	1.1	9.21	9.21	1.1
	Exponential	172	424	35.3	177	435	35.1
	Normal?	No	No	No	No	No	No
	Box-Cox	0.16	0.09	2.1	0.1	0.13	1.1
	Normal?	Yes	Yes	No	Yes	Yes	Yes
	Mixed	14.2	17.6	5.3	8.4	9.7	4.9
	Normal?	No	No	No	yes	Almost yes	No
Suspended sediment	Exponential	69	120	8.3	144	203	13.8
	Normal?	No	No	No	No	No	No
	Box-Cox	70	59	9.8	108	76	11.7
	Normal?	No	No	No	No	No	No
	Mixed	31	39	2.4	13	14	1.4
	Normal?	No	No	No	No	No	Almost yes

3.2 ARMAX Model Result Evaluation

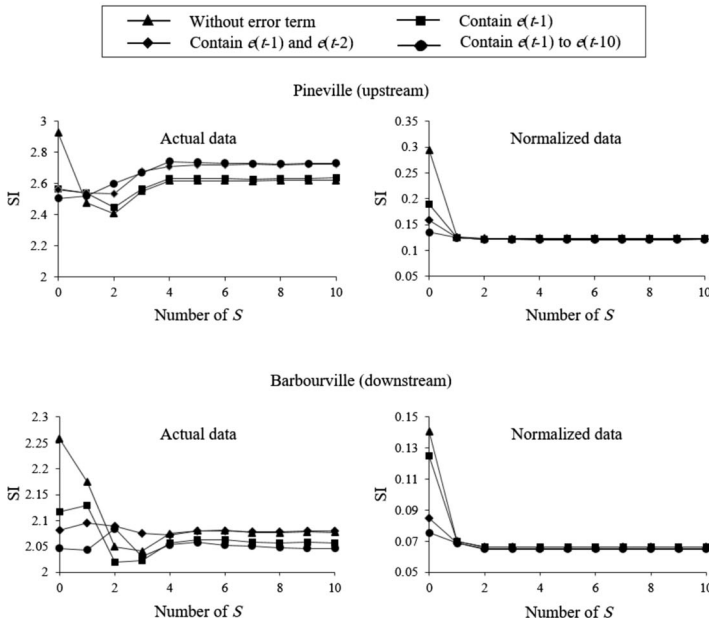
As ARMAX is one of the components of the hybrid model presented, its results can affect the hybrid model results. Hence, to achieve the best results, the ARMAX model inputs were considered between 0 and 10. The inputs include discharge (Q), sediment (S) and white noise disturbance (e). Therefore, 1331 models were defined for each station. All these models were obtained once based on the original series and normalized series. It should be noted that according to the results mentioned in the previous section, the Box-Cox and mixed transformations were used to normalize Q and S respectively. To identify the best combination of inputs and most accurate model, Fig. 4a is provided. This figure displays the SI criterion in the testing period for both stations and in the actual and normalized modes. Due to the multitude of models, the results are provided for some cases. It is observed that in both modeling cases, models in which the discharge is not part of the inputs had greater prediction error than models which include this factor. With the addition of $Q(t-1)$ as an input, the prediction error was reduced at once. Adding $Q(t-2)$ also led to a drop in estimated values, but its impact was less. It is very clear that adding more discharge data as input had no impact on improving the outcome. Because the graphs are almost horizontal, no effective error reduction is seen. To achieve the best results with the ARMAX model it is sufficient to use the daily discharge for one and two days before. Another point is that when only the previous day's discharge was used, increasing the number of white noise disturbance components had a significant impact on outcome improvement. The error reduction of models 11 to 22 in each diagram demonstrates this fact. However, when the discharge from two days before was used, the impact of white noise disturbance components on model error reduction diminished. It can be seen in each diagram that models 22 to 33 are not very different from each other. Thus, it seems that the white noise disturbance component may not be required in modeling.

The appropriate number of input data (discharge) was detected previously, and the best number of other input data (sediment) will be discussed subsequently. Figure 4b shows the SI criterion changes in test mode compared to the number of sediment input data for the case where $Q(t-1)$ and $Q(t-2)$ are also considered inputs. It is observed that for both stations and modes, with increasing the number of sediment input data to 2, the diagrams display a decline. Subsequently, the diagrams become horizontal as normalized data are used. In the case where actual data are used, the diagrams initially rise (up to 4 sediment input data) and then become horizontal. This suggests that using sediment input data at more than two steps before is either ineffective in normal mode or increases the error in real mode. Another point is that in models that lack sediment input data, increasing the turbulence component in the inputs improves the results. However, this component is ineffective in models that include sediment input data in normal mode.

In real mode, when using two or more sediment input data, increasing the number of turbulence components in the inputs shows the reverse trend and higher model error. As a result, it is not recommended to use these components in a state when discharge and sediment are present as inputs. In summary, using $S(t-1)$ and $S(t-2)$ with $Q(t-1)$ and $Q(t-2)$ as ARMAX model inputs is sufficient. The model that includes these inputs produced the best results for both stations and in both real and normal modes. Hence, ARMAX(2, 2, 0) is introduced as the best of all models.



(a) Discharge



(b) Suspended sediment

Fig. 4 Scatter index values for the ARMAX model based on the number of discharge and suspended sediment input data

3.3 The Impact of Normalization on ARMAX-ANN Model Results

In this section, the results from the three scenarios defined for the ARMAX-ANN hybrid model are compared. Figure 5 shows the *MAE* and *MARE* values in test mode for each scenario and all 12 input combinations. It can be seen that the highest error values for all input combinations are related to scenario 1. The values obtained for both criteria (especially *MARE*) are significantly higher than other scenarios. Unlike the two other scenarios, there was no normalization done in scenario 1. The effective role of normalization in improving the ARMAX-ANN hybrid model outcomes is evident. However, for all 12 cases, scenario 3 exhibited better results than scenario 2. Thus, it can be concluded that if denormalization is done in the last step, normalization has the greatest impact on increasing hybrid model accuracy.

3.4 The Impact of Input Type and Number on Hybrid Model Accuracy

In the previous section, scenario 3 was introduced as the most accurate and effective scenario for the proposed hybrid model. Therefore, to evaluate the effect of the input type and number on the accuracy of this model, the results of scenario 3 are used in this section. Figure 6 shows the *MAE*, *SI*, *CRM* and *VAF* values for each of the 12 input combinations compared with each other. The 12 cases were divided into three

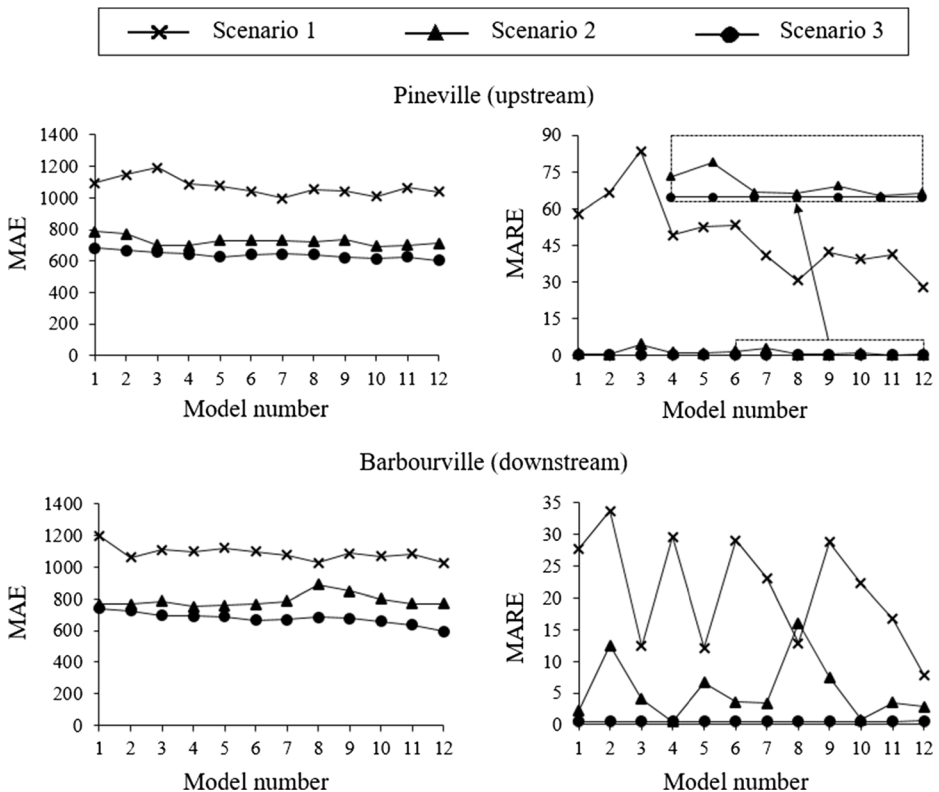


Fig. 5 ARMAX-ANN model results obtained in each scenario

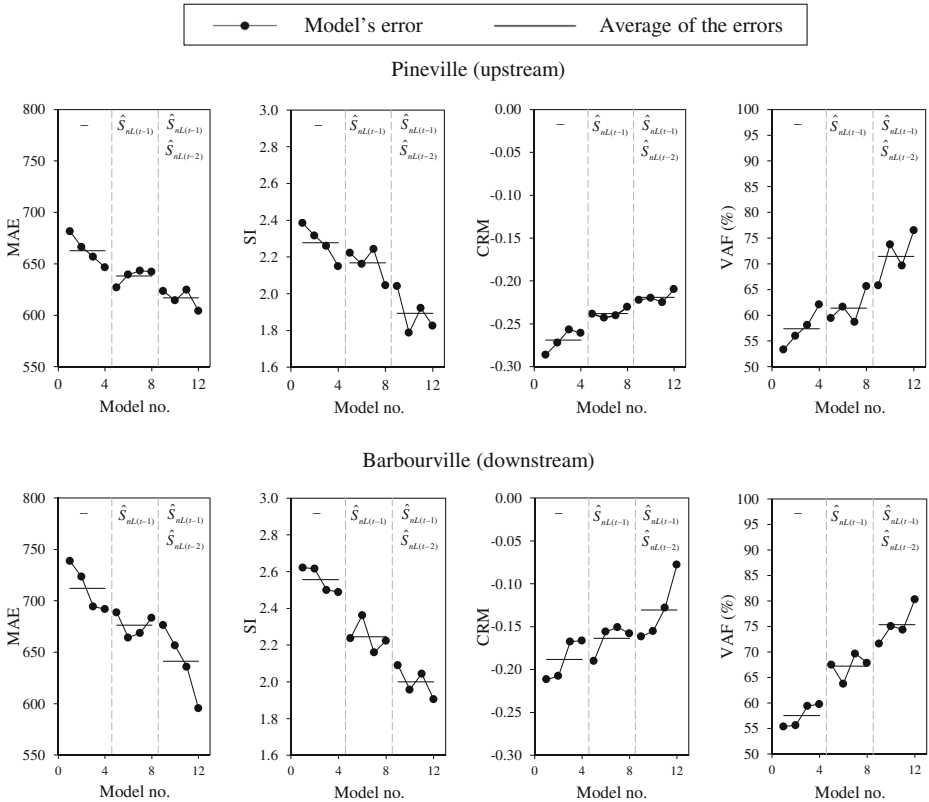


Fig. 6 Effect of input number and type on ARMAX-ANN model accuracy in scenario 3

groups of four. In the first group, the ANN model inputs in the hybrid model were only residuals of the ARMAX ($\hat{e}_{n(t-1)}$ to $\hat{e}_{n(t-4)}$) model. In the second group, the result of the ARMAX model for the day before ($\hat{S}_{nl(t-1)}$) was added to the first group's inputs. In the third group, the ARMAX model results for one and two days before ($\hat{S}_{nl(t-1)}$ and $\hat{S}_{nl(t-2)}$) were added to the first group's inputs. The results of the three groups separately are shown in Fig. 6. This figure also presents the average of each criterion for each group. It can be seen that with no exception, the averages for the second and third groups are better than the first group. These results for all four criteria and both upstream and downstream stations are the same. Thus, it is concluded that adding the ARMAX model result to the ANN model input leads to greater hybrid model accuracy. It is also observed that group three performed better than group two. Therefore, using a higher input number produces better results and the best result was observed for model 12 with the following inputs: $\hat{e}_{n(t-1)}$, $\hat{e}_{n(t-2)}$, $\hat{e}_{n(t-3)}$, $\hat{e}_{n(t-4)}$, $\hat{S}_{nl(t-1)}$ and $\hat{S}_{nl(t-2)}$. The ARMAX model results led to enhanced hybrid model result accuracy. Although in some cases the ARMAX model may have performed well, this model is linear and there is no guarantee of a nonlinear relationship between residuals. The results of this study signify that the ARMAX model has a strong nonlinear relationship with its residuals for daily data. This relationship leads to an increase in nonlinear ANN model accuracy and subsequently ARMAX-ANN model accuracy.

Table 3 Values of AIC criteria and the number of model parameters

Station	Input	1	2	3	4	5	6	7	8	9	10	11	12
Upstream	k	17	17	21	25	29	25	33	29	25	25	21	17
	Part 1	37	37	45	53	62	53	70	62	53	53	45	37
	Part 2	17,525	17,461	17,407	17,298	17,371	17,310	17,392	17,189	17,184	16,895	17,052	16,940
	AIC	17,561	17,497	17,452	17,351	17,433	17,363	17,462	17,251	17,238	16,948	17,097	16,976
Downstream	Average	17,465				17,377				17,065			
	k	17	17	17	21	21	25	25	25	21	21	17	13
	Part 1	37	37	37	45	45	53	53	53	45	45	37	28
	Part 2	17,192	17,101	17,181	17,134	17,162	17,150	17,113	17,103	17,059	17,052	17,081	17,158
Average	AIC	17,229	17,137	17,217	17,179	17,207	17,203	17,167	17,156	17,104	17,097	17,117	17,186
	Average	17,191				17,183				17,126			

In order to determine the more accurate model with fewer parameters, the principle of parsimony is considered. The Akaike Information Criterion (*AIC*) is calculated for the results of scenario 3 and all 12 input combinations. According to Eq. 19, this measure consists of two parts. The first part indicates the effect of the number of parameters and the second part shows the effect of result accuracy on the *AIC* criterion. Table 3 displays these two parts along with the *AIC* criterion and the number of parameters used in the ARMAX-ANN(*k*) model. It is evident from the values in parts 1 and 2 that the effect of the number of model parameters on the *AIC* value is much lower than the effect of the error values. This is explained by the fact that the sediment load has very high daily fluctuations, which thereby leads to an increase in data variance. This large variance remained in the error values after modeling and reduced the effect of the number of model parameters. Therefore, it can be concluded that for the studied data, the most accurate model is the most parsimonious model. Table 3 also demonstrates that the same as the other error criteria and based on *AIC*, the increase in the number of inputs as well as the use of the ARMAX model improved the results accordingly.

3.5 Comparison of Models

Table 4 shows the R^2 , *MAE*, *CRM* and *VAF* results for these models in test mode for both stations. The ANN model results were extracted from Kisi et al.'s (2012) study. This table also provides the ARMAX model results for both actual and normalized data. A comparison of the ARMAX model results in these cases shows that for the two stations, the *MAE* and *CRM* criteria in normalization mode are much better. However, the R^2 and *VAF* criteria for these two cases exhibit very little difference from each other. Thus, it can be said that the normalization enhanced the ARMAX model results. By comparing the ANN model with the ARMAX (normal) model, one cannot be deemed superior to the other. The reason may be the diverse nature of these two models. The ARMAX model is linear and based on probability and statistics, while the ANN model is nonlinear and based on computational intelligence. By comparing the results in Table 4 it can be concluded that the hybrid model produced the best results. The highest difference between this model and other models is seen in the *MAE* criterion. It should be noted that this criterion determines the error value directly. It is therefore more important than other criteria, which are supplemental to *MAE*. The hybrid model presented in this study benefits from both ARMAX and ANN models' advantages and is able to improve prediction accuracy as well.

Table 4 Model results

Station	Model	R^2	<i>MAE</i>	<i>CRM</i>	<i>VAF</i>
Pineville	Hybrid	0.864	604	-0.210	78.4
	ANN (Kisi et al. 2012)	0.889	818	0.377	77.1
	ARMAX (normal)	0.820	659	-0.274	55.9
	ARMAX	0.860	1272	-0.814	67.8
Barbourville	Hybrid	0.885	596	-0.078	80.3
	ANN (Kisi et al. 2012)	0.848	798	-0.161	81.5
	ARMAX (normal)	0.875	712	-0.185	53.9
	ARMAX	0.803	1135	-0.376	74.8

Bold present the best value for each index for two stations

4 Conclusions

In this study, the daily suspended sediment load at two stations was predicted by the hybrid ARMAX-ANN model. The ARMAX model was selected for modeling the linear part, and the effect of the discharge parameter can be considered in the ARMAX model. ANN was chosen for modeling the nonlinear time series component due to its simplicity and relatively higher speed compared to other nonlinear models. Three different scenarios were defined for normalization with the hybrid model. Moreover, 12 input combinations were identified. A summary of the results is as follows:

- 1- The proposed mixed transformation outperformed the exponential and Box-Cox transformations in daily sediment load normalization. This transformation also performed better than the exponential transformation in daily discharge normalization.
- 2- In addition to the ARMAX model, data normalization led to increased accuracy of the ARMAX-ANN hybrid model results as well. Scenario 3 was the best modeling scenario using the hybrid model.
- 3- Among 1331 specified input combinations for ARMAX modeling, the model with inputs $S(t-1)$, $S(t-2)$, $Q(t-1)$ and $Q(t-2)$ was the most accurate and parsimonious.
- 4- Among 12 defined input combinations, the model with inputs $\hat{e}_{n(t-1)}$, $\hat{e}_{n(t-2)}$, $\hat{e}_{n(t-3)}$, $\hat{e}_{n(t-4)}$, $\hat{S}_{nl(t-1)}$ and $\hat{S}_{nl(t-1)}$ showed the best results.
- 5- The ARMAX-ANN hybrid model exhibited superior performance over each individual ANN and ARMAX model.
- 6- It is suggested to verify other linear and nonlinear models as part of a hybrid model in terms of modeling daily time series with high irregular behavior.

References

- Afan HA, El-Shafie A, Yaseen ZM, Hameed MM, Mohtar WHMW, Hussain A (2015) ANN based sediment prediction model utilizing different input scenarios. *Water Resour Manag* 29:1231–1245
- Alizadeh MJ, Joneyd PM, Motahhari M, Ejlali F, Kiani H (2015) A Wavelet-ANFIS model to estimate sedimentation in dam reservoir. *Int J Comput Appl T* 114:19–25
- Alp M, Cigizoglu HK (2007) Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environ Model Softw* 22:2–13
- Cigizoglu HK, Kisi O (2006) Methods to improve the neural network performance in suspended sediment estimation. *J Hydrol* 317:221–238
- Demirci M, Üneş F, Saydemir S (2015) Suspended sediment estimation using an artificial intelligence approach. In: Heininger P, Cullmann J (eds) *Sediment Matters*. Springer International Publishing, Switzerland, pp 83–95
- Faruk DÖ (2010) A hybrid neural network and ARIMA model for water quality time series prediction. *Eng Appl Artif Intell* 23:586–594
- Kisi O, Dailr AH, Cimen M, Shiri J (2012) Suspended sediment modeling using genetic programming and soft computing techniques. *J Hydrol* 450:48–58
- Kumar D, Pandey A, Sharma N, Flügel WA (2016) Daily suspended sediment simulation using machine learning approach. *Catena* 138:77–90
- Lafdani EK, Nia AM, Ahmadi A (2013) Daily suspended sediment load prediction using artificial neural networks and support vector machines. *J Hydrol* 478:50–62
- Liu H, Tian HQ, Li YF (2012) Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. *Appl Energy* 98:415–424
- Marco JB, Harboe R, Salas JD (2012) *Stochastic hydrology and its use in water resources systems simulation and optimization*. Springer Science & Business Media, Peniscola
- Moeeni H, Bonakdari H (2016) Forecasting monthly inflow with extreme seasonal variation using the hybrid SARIMA-ANN model. *Stoc Environ Res Risk Assess*. <https://doi.org/10.1007/s00477-016-1273-z>

- Moeeni H, Bonakdari H, Ebtehaj I (2017) Integrated SARIMA with neuro-fuzzy systems and neural networks for monthly inflow prediction. *Water Resour Manag* 31:2141–2156
- Mustafa M, Rezaur R, Saiedi S, Isa M (2012) River suspended sediment prediction using various multilayer perceptron neural network training algorithms - a case study in Malaysia. *Water Resour Manag* 26:1879–1897
- Nourani V, Kisi Ö, Komasi M (2011) Two hybrid artificial intelligence approaches for modeling rainfall–runoff process. *J Hydrol* 402:41–59
- Rahim A, Akif A (2015) Optimal artificial neural network modeling of sedimentation yield and runoff in high flow season of Indus River at Besham Qila for Terbela dam. *Int J Sci Res* 4:479–483
- Salas J, Delleur J, Yevjevich V, Lane W (1988) Applied modeling of hydrologic time series. Water Resources Publications, Colorado
- Tiwari H, Rai SP (2015) Discharge and sediment time series, uncertainty analysis using the maximum likelihood estimator and artificial neural network. *J Water Res Environ Eng* 1:1–9
- Zhang GP (2003) Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* 50:159–175