Science of the Total Environment 721 (2020) 137612

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Improving prediction of water quality indices using novel hybrid machine-learning algorithms

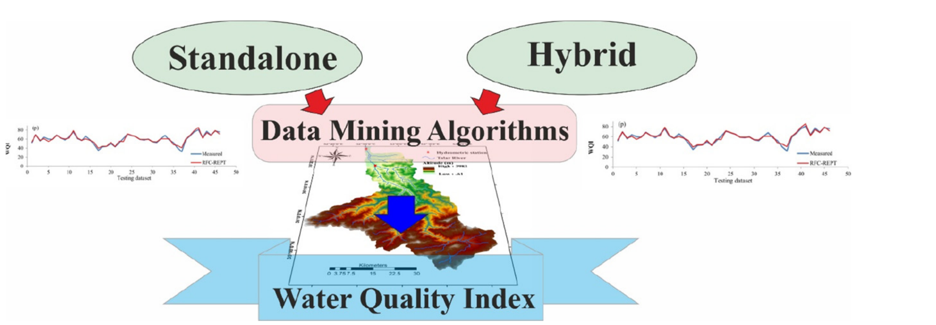
Duie Tien Bui [a,b], Khabat Khosravi [c], John Tiefenbacher [d], Hoang Nguyen [e,*], Nerantzis Kazakis [f,*]

[a] Geographic Information Science Research Group, Ton Duc Thang University, Ho Chi Minh City, Viet Nam
[b] Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City, Viet Nam
[c] School of Engineering, University of Guelph, Guelph, Canada
[d] Department of Geography, Texas State University, San Marcos, TX 78666, USA
[e] Institute of Research and Development, Duy Tan University, Da Nang 550000, Viet Nam
[f] Aristotle University of Thessaloniki, Department of Geology, Lab. of Engineering Geology & Hydrogeology, 54124 Thessaloniki, Greece

HIGHLIGHTS

- 16 novel hybrid data mining algorithm applied for WQI prediction
- BA-RT algorithm outperformed while RFC-RT has the lowest prediction power.
- Fecal coliform was the most effective predictor on WQI estimation.
- The best input combination is not the same for all models.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 23 January 2020
Received in revised form 26 February 2020
Accepted 26 February 2020
Available online 3 March 2020

Editor: Damia Barcelo

Keywords:
Water quality index
Prediction
Data mining
Novel hybrid algorithms

ABSTRACT

River water quality assessment is one of the most important tasks to enhance water resources management plans. A water quality index (WQI) considers several water quality variables simultaneously. Traditionally WQI calculations consume time and are often fraught with errors during derivations of sub-indices. In this study, 4 standalone (random forest (RF), M5P, random tree (RT), and reduced error pruning tree (REPT)) and 12 hybrid data-mining algorithms (combinations of standalones with bagging (BA), CV parameter selection (CVPS) and randomizable filtered classification (RFC)) were used to create Iran WQI ($IRWQI_{sc}$) predictions. Six years (2012 to 2018) of monthly data from two water quality monitoring stations within the Talar catchment were compiled. Using Pearson correlation coefficients, 10 different input combinations were constructed. The data were divided into two groups (ratio 70:30) for model building (training dataset) and model validation (testing dataset) using a 10-fold cross-validation technique. The models were evaluated using several statistical and visual evaluation metrics. Result show that fecal coliform (FC) and total solids (TS) had the greatest and least effect on the prediction of $IRWQI_{sc}$. The best input combinations varied among the algorithms; generally variables with very low correlations displayed weaker performance. Hybrid algorithms improved the prediction power of several of the standalone models, but not all. Hybrid BA-RT outperformed the other models ($R^2 = 0.941$, RMSE = 2.71, MAE = 1.87, NSE = 0.941, PBIAS = 0.500). PBIAS indicated that all algorithms, with the exceptions of RT, BA-RT and CVPS-REPT, overestimated WQI values.

* Corresponding authors.
E-mail addresses: buitiendieu@tdtu.edu.vn (D.T. Bui), kkhosrav@uoguelph.ca (K. Khosravi), tief@txstate.edu (J. Tiefenbacher), nguyenhoang23@duytan.edu.vn (H. Nguyen), kazakis@geo.auth.gr (N. Kazakis).

https://doi.org/10.1016/j.scitotenv.2020.137612
0048-9697/

1. Introduction

Pollution of rivers from both point and non-point sources is a growing and increasingly difficult challenge. Deterioration of water quality significantly effects not only aquatic ecosystems but also reduces the supplies of safe fresh water for human consumption and irrigation. Developing countries often pass through periods of fast economic expansion and every development project has the prospect of generating negative impacts on the environment. Due to rapidly growing wealth and population, demands for greater food production increase pressures on soils' natural fecundity, often over-extracting nutrients leading to the need for artificial fertilizers. Excess fertilizer is usually transported to groundwater and into rivers. Rivers carry contaminants to lakes and oceans, often causing irreparable damage to environments and human health. Monitoring and assessment of water quality is essential for protection of human and environmental health and effective, sustainable water management.

A water quality index (WQI) is a non-dimensional index calculated from selected water quality parameters. Indices allow a categorical assessment of the past and present water quality of water bodies. The variables that are used in WQIs are chemical oxygen demand (COD), biological oxygen demand (BOD), total suspended solids (TSS), pH, temperature, ammoniacal-nitrogen (AN), dissolved oxygen (DO) (Hameed et al., 2016). Groundwater quality indices (GQIs) are usually predicted by measures of common variables like Ca^{2+}, Mg^{2+}, NO_3^-, and others (Bournaris et al., 2015; Vadiati et al., 2016; Rufino et al., 2019). The WQI provides a meaningful value to guide decision-makers' policies and actions. However, the calculation of WQI is not straight forward because sub-indices are calculated within WQI equations.

There are various approaches to calculations of WQI around the world. These include the Interim National Water Quality Standards for Malaysia (INWQS), United States National Sanitation Foundation Water Quality Index (NSFWQI), the Florida Stream Water Quality Index (FWQI), the Canadian Water Quality index (CWQI), the British Columbia Water Quality Index (BCWQI), the Oregon Water Quality Index (OWQI), and others. Iran has also developed a preferred index that focuses on river conditions and characteristics, but Iran is an extensive country with a diversity of climatic conditions that range from arid to humid. Application of this equation to specific sites is therefore filled with great uncertainty.

The disadvantages of WQI calculations are: a) require a lot of time, b) are lengthy, c) the process is complex and d) have inconsistencies as WQIs often use different equations. It may be apparent from this discussion, but there is no universal WQI method. To overcome these issues, some researchers have opted for a non-physical approach, successfully predicting WQI using artificial intelligence (AI) (Yaseen et al., 2018; Iticescu et al., 2019; Leong et al., 2019). AI-based modeling eliminates sub-index calculations and promptly produces a WQI value. Attention to AI algorithms is growing due to advantages that include their non-linear structures, ability to predict complex phenomena, ability to handle big datasets comprised of data at different scales, and insensitivity to missing data. The prediction power of AI algorithms relies heavily on the methods and accuracy of data collection and analysis.

Successful predictions of WQI in rivers have been achieved using AI algorithms like support vector machine (SVM) (Hamzeh Haghibi et al., 2018), least square SVM (LSVM) (Leong et al., 2019), and artificial neural network (ANN) (Khuan et al., 2002; Gazzaz et al., 2012a, 2012b; Sakizadeh, 2016; Hameed et al., 2016; Machiwal et al., 2018). Although SVM achieved high prediction accuracy, its value is reduced by needing to test 4 kernel functions to determine the best. It also requires many parameters for which optimum values need to be determined. And though ANN is the most widely used AI algorithm it has some weaknesses like weak prediction power that occurs when the testing data range is outside of the range of training data and when datasets are small (Khosravi et al., 2018; Choubin et al., 2018). To solve these problems, a hybrid of the ANN model and a fuzzy logic and adaptive neuro-fuzzy inference system (ANFIS) was developed. ANFIS algorithms have been used to map floods (Bui et al., 2018a, 2018b), to model groundwater (Termeh et al., 2019) and to predict water quality (Ahmed and Shah, 2015; Chen and Liu, 2015). ANFIS, even though it is a powerful algorithm, is weakened by internal parameters and the need for accurate determination of weights for membership in a fuzzy rule (Bui et al., 2016). Some have proposed to hybridize ANFIS with a meta-heuristic algorithm to solve this. Meta-heuristic algorithms compute the optimum weight automatically (Kisi and Zounemat-Kermani, 2014; Yaseen et al., 2018; Khosravi et al., 2018; Chen et al., 2019a, 2019b; Khosravi et al., 2019). Compared to standalone algorithms, hybrids can more easily identify the non-linearity of input and output variables because these models are flexible and become more robust with noisy data. Though solving this weakness, hybridization increases the model's complexity and consumption of time, and requires a complicated search for the best meta-heuristic algorithm from a large array of meta-heuristic models with different structures. Researchers continue to explore new, simple, robust, flexible, and reliable soft-computing algorithms.

Recently a new type of AI algorithm, data mining, has been developed to solve regression problems and to reduce the disadvantages of AI. New algorithms like random tree (RT), random forest (RF), M5P, reduced error pruning tree (REPTree), random committee (RC), bagging, and instance-based k-nearest neighbours (IBK) are being used to study issues in hydrology, climatology, and hydraulics to quantify suspended sediment transport (Khosravi et al., 2018), estimate reference evaporation (Khosravi et al., 2019), simulate solar radiation (Sharafati et al., 2019), and predict shear stress in rivers (Khozani et al., 2019). The lack of hidden layers and modeling transparency in the decision tree-based algorithms (i.e. M5P, RF, RT, REPT, and others) enable better modeling performance than achieved by ANN and ANFIS (Kisi et al., 2012).

This main goal of this study is to predict the WQI of the Talar River in Iran using four standalone (M5P, RF, RT, REPT) and 12 novel hybrid data-mining algorithms (bagging, CV parameter selection, and randomizable filtered classifier which are integrated with the four standalone algorithms). Although the stand-alone decision tree algorithm may satisfactorily predict WQI and though its predictive powers have been proven in applications to other hydrological phenomenon, integration with classifier algorithms may improve prediction accuracy and eliminate the inherent shortcomings of each model.

2. Study area

The Talar catchment is located in a region between 52°40′ and 53°20′ E and between 35°40′ and 36°25′ N in Mazandaran and Semnan provinces in northern Iran (Fig. 1). The Talar catchment includes the main river (the Talar) and 5 main tributaries. The length of the main river is about 151 km. It flows from the south to the north and drains into the Caspian Sea. The catchment area is approximately 2900 km^2. The average slope is 54.32% and the average elevation is 1699 m. The region has a humid climate and a mean annual precipitation total of 730 mm. The river's average discharge rate is 72 m^3/s at the Shirgah hydrometric station. Geologically, approximately 19.5% (about 565 km^2) of the region is sandstone and shale of the Shemshak formation. Grasslands and forest are the two most common landcovers in the basin.

3. Methodology

3.1. Data collection and preparation

Six years (2012 to 2018) of monthly water quality data were compiled at 2 strategically-placed water quality monitoring stations within the catchment. These data included measures of BOD, COD, DO, pH, total

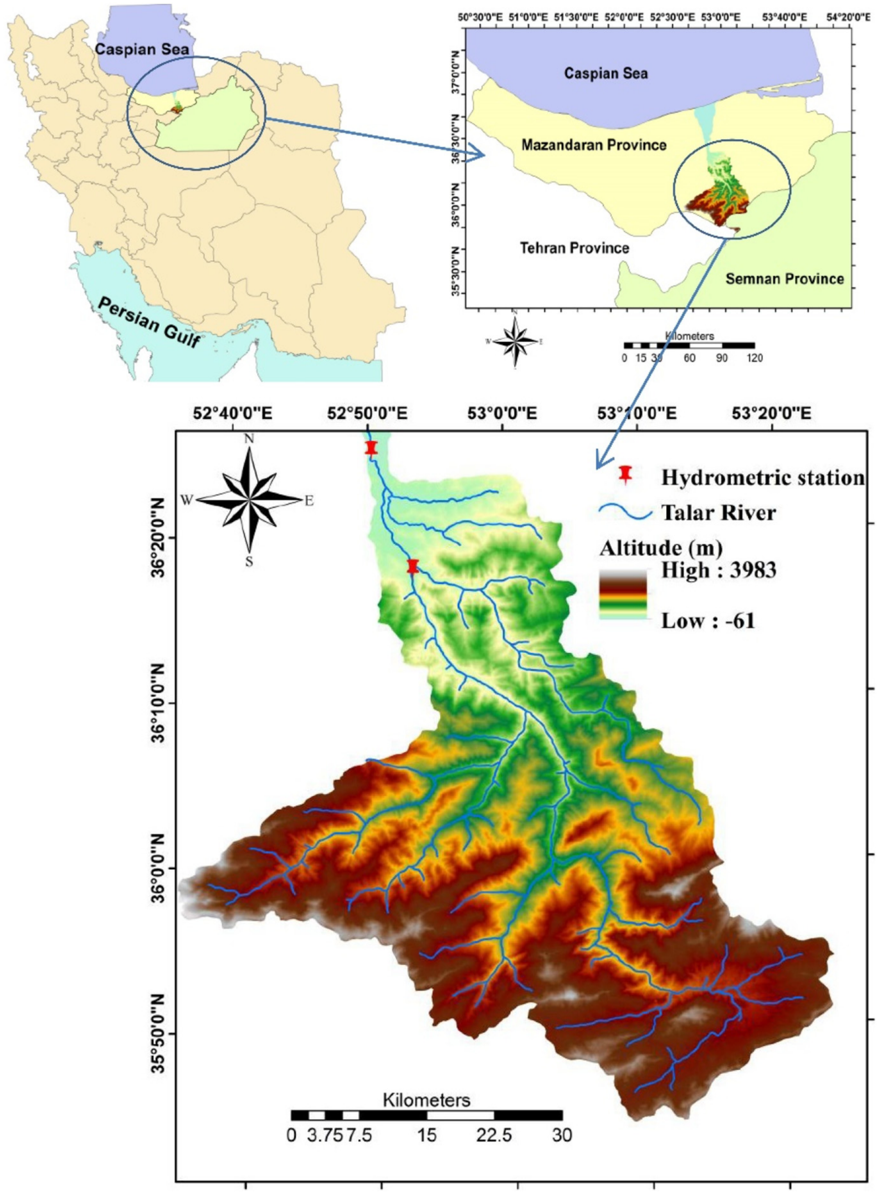


Fig. 1. Talar River catchment water quality monitoring stations.

solids (TS), fecal coliform (FC), phosphate (PO_4^{-2}), nitrate (NO_3^-), turbidity (TU), and electrical conductivity (EC). These were compiled and used to compute the $IRWQI_{sc}$. Sampling, preservation, and analysis protocols followed standard methods in the Mazandaran environmental protection organization laboratory. The dataset was divided into two sub-sets for model training and model testing (70:30) using the 10-fold cross-validation technique. Though there is no universal guideline, this ratio is the most common modeling approach for both spatial (Pham et al., 2018; Gholami et al., 2019; Khosravi et al., 2019) and temporal prediction (Khosravi et al., 2018; Sharafati et al., 2019; Khozani et al., 2019; Khosravi et al., 2019).

The NSFWQI equation was used to develop the $IRWQI_{sc}$. The greater the IRWQI, the cleaner the river (a WQI above 80 indicates that a river is clean) (Leong et al., 2019). The $IRWQI_{sc}$ is calculated:

$$IRWQIsc = \sum_{i=1}^{n} W_i \times SI_i \qquad (1)$$

where $IRWQI_{sc}$ is the water quality index of Iran (0–100) (Table 1), W_i is the weight of variable i (between 0 and 1), and SI_i is the sub-index resulting from the quality-index curve (0–100). The methods of calculation are consistent with the NSFWQI; these are described in detail by Misaghi et al. (2017) and Fathi et al. (2018).

The descriptive statistics (Table 2) indicate that the WQI ranges from 15.40 to 91.65 (maximum value is 100); water quality varied from low quality to very good (Wan Mohtar et al., 2019). The water in the upper catchment is high quality, apparently containing minimal pollution generated from protected forestlands. The water quality quickly

Table 1
$IRWQI_{sc}$ ranges and its qualitative description.

Index	Range	Quality
$IRWQI_{sc}$	<15	Very low
	15–29.9	Low
	30–44.9	Approximately low
	45–55	Moderate
	55.1–70	Approximately good
	70.1–85	Good
	>85	Very good

Table 2
Descriptive statistics of the variables in the training and testing datasets.

Variables	Training dataset				Testing dataset			
	Min	Max	Mean	Std. deviation	Min	Max	Mean	Std. deviation
BOD (mg/L)	1.00	99.00	9.381	11.48	1.00	45.00	9.37	8.42
COD (mg/L)	1.00	250.00	40.23	32.41	11.00	105.00	43.20	22.94
TS (mg/L)	0.90	1850.00	91.36	162.32	17.00	293.00	83.00	71.37
DO (mg/L)	2.36	10.93	5.69	2.29	0.78	7.20	4.53	1.57
FC (No/100 cm^3)	250	1320	892.24	785.23	268	1295	864.97	765.48
pH	6.10	9.23	7.70	0.37	6.23	7.87	7.20	0.36
PO_4 (mg/L)	0.023	0.071	0.058	0.025	0.029	0.070	0.055	0.023
NO_3 (mg/L)	1.98	9.36	6.50	2.19	2.10	9.15	6.37	2.01
Turbidity (NTU)	12.5	56.5	40.36	10.68	13.5	50.45	38.46	9.39
EC (μS/cm)	0.302	900	320.12	105.32	165	350	289.32	210.36
WQI	15.40	91.65	58.23	12.34	32.00	81.00	59.88	11.29

deteriorates as the system enters urbanized area due to untreated and uncontrolled discharges into the Talar River (Othman et al., 2012). Following common practice, the data were normalized (x_i') to a 0-to-1 range to improve prediction power using Eq. (2):

$$x_i' = (x_i - x_{\min})/(x_{\max} - x_{\min}) \tag{2}$$

where, x_i' is the normalized value of a variable's (i.e., BOD, COD, etc.), x_i, value at a location, and x_{min}, and x_{max} are the minimum and maximum values of that variable.

3.2. Constructing input combinations

Before modeling, the best input combination must be determined and the optimum value for each model's operator must be found. Ten variables were considered as potential inputs and finally based on correlation coefficients (CCs) between input and WQI (Table 3), ten input combinations were constructed (Table 4).

The variable with the highest CC was FC and was the first variable introduced into the model. This is the most effective variable as it can provide the best estimate of the WQI on its own. Each variable with the next highest CC (i.e., BOD, then NO_3, then DO, etc.) was added to the previous combination until the last variable with lowest CC is included (i.e., pH and combination 10). All 10 input combinations using fixed input variable values (or default values) in each model to find the most effective (i.e., most predictive) combination. Root mean-square error (RMSE) criteria were used for evaluation in the testing phase.

3.3. Determining optimum values

After determining the best input variables, optimum values for each model's operator were determined by trial and error (Khosravi et al., 2018; Choubin et al., 2018; Sharafati et al., 2019; Khosravi et al., 2019). As there is no global optimum value for each operator (values differ from study to study), an assortment of values should be tested to find the effective value. To do this, each model was run with default values. According to these results, higher and lower values were entered arbitrarily until the optimum value was determined (Table 5).

4. Descriptions of the models

As stated above, this study applies sixteen artificial intelligence models for predicting WQI with two main groups, specifically:

- Group 1 (decision-tree algorithms): M5P; random forest (RF); random tree (RT); and reduced error pruning tree (REPT);
- Group 2 (meta-classifier or hybrid Algorithms): Bagging (BA); CV parameter selection (CVPS); and randomizable filtered classifier (RFC); including BA-M5P; BA-RF; BA-RT; BA-REPT; CVPS-M5P; CVPS-RF; CVPS-RT; CVPS-REPT; RFC-M5P; RFC-RF; RFC-RT; and RFC-REPT.

Accordingly, the individual models in the group 1 were used to predict and evaluate the WQI. Subsequently, ensemble models were developed based on the individual models in groups 1 to investigate the accuracy of the WQI prediction compared to standalone algorithms. Finally, sixteen models in two groups are compared and assessed to find out the best model for predicting WQI. The details of the models can be found in the Supplementary material.

5. Models evaluation and comparison

Five statistical metrics were used to evaluate the models quantitatively. They included RMSE, coefficient of determination (R^2), mean absolute error (MAE), Nash-Sutcliffe efficiency (NSE), and the percentage of bias (PBIAS) and percent of relative error index (PREI). These metrics are calculated as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(WQI_{predicted} - WQI_{measured}\right)^2} \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(WQI_{measured} - WQI_{predicted}\right)^2}{\sum_{i=1}^{n}\left(WQI_{measured} - \overline{WQI}_{measured}\right)^2} \tag{10}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left| WQI_{measured} - WQI_{predicted} \right| \tag{11}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(WQI_{measured} - WQI_{predicted}\right)^2}{\sum_{i=1}^{n}\left(WQI_{measured} - \overline{WQI}_{measured}\right)^2} \tag{12}$$

$$PBIAS = \left(\frac{\sum_{i=1}^{n}\left(WQI_{measured} - WQI_{predicted}\right)}{\sum_{i=1}^{n} WQI_{predicted}}\right) \cdot 100 \tag{13}$$

$$PREI = \left(\frac{WQI_{Measured} - WQI_{\ Predicted}}{WQI_{Measured}}\right) * 100$$

where $\overline{WQI}_{predicted}$ and $\overline{WQI}_{measured}$ are mean values of predicted and measured WQI, respectively.

In addition, visual comparisons were used for model evaluation. These techniques included scatter plots and box plots. Scatter plots are

Table 3
Pearson correlation coefficient between each input variable and WQI.

Variables	BOD	COD	TS	DO	FC	pH	PO_4	NO_3	Turbidity	EC
Correlation (r)	−0.78	−0.56	−0.11	0.69	−0.83	0.05	−0.23	−0.72	−0.12	−0.58

often used to analyze the distribution of the datasets used, as well as to assess model performance (Kahng et al., 1998; Touchette et al., 1985). Data structure is also analyzed with scatter plots to examine density and distribution, for example. Box plots are standard devices to evaluate the distribution and density of the datasets and results. In general, the range of the datasets or results is divided into four groups, each consisting of 25% of the observations (quartiles). Extreme values (minimums and maximums), medians, and first (upper) and third (lower) quartile predictions can be analyzed.

6. Results and analysis

6.1. Best input combination

Several water quality variables were used to construct different input combinations based on CCs (Table 3). Khuan et al. (2002) presented an equation to estimate WQI and revealed that pH was the least important predictor of WQI. This is in accordance with our analysis and agrees with Yaseen et al. (2018). However, in the study by Mohammadpour et al. (2015) pH was the most important predictor of WQI. The 10 input combinations described earlier were constructed and were used to train the 16 algorithms. They were evaluated with a testing dataset, and the most effective was selected for modeling and further analysis (Table 6). As all models were built with the training dataset, this result only indicates how well the models fit with the training dataset. They were not used to evaluate the model. Evaluation was done with the testing data (Khosravi et al., 2016).

The combinations of variables numbered 4 (FC, BOD, NO3, and DO), 5 (FC, BOD, NO3, DO, and EC), and 6 (FC, BOD, NO3, DO, EC, and COD) were determined by testing to be the optimal combinations for specific standalone and hybrid algorithms. They were best able to estimate WQI and achieved the lowest RMSE values as there was non-linearity among the input variables and between the inputs and outputs. Variables with high (direct or inverse) CCs to the WQI can add prediction power to a point, but at some point, additional variables (e.g., PO_4, TU, TS and pH) tend to weaken the prediction power. Input combination number 4 was the best input for the RFC hybrids (RFC-M5P, RFC-RF, RFC-RT, and RFC-FEPT). Combination 5 was best for three standalone (RF, RT, and REPT) algorithms, for the bagging hybrids (BA-RF, BA-RT, and BA-REPT), and the CVPS hybrids (CVPS-RF and CVPS-REPT). And combination 6 was optimal for modeling with the standalone M5P and three hybrid algorithms (BA-M5P, CVPS-M5P and CVPS-RT).

Table 4
Different input combinations.

No	Different input combinations
1	FC
2	FC, BOD
3	FC, BOD, NO_3
4	FC, BOD, NO_3, DO
5	FC, BOD, NO_3, DO, EC
6	FC, BOD, NO_3, DO, EC, COD
7	FC, BOD, NO_3, DO, EC, COD, PO_4
8	FC, BOD, NO_3, DO, EC, COD, PO_4, TU
9	FC, BOD, NO_3, DO, EC, COD, PO_4, TU, TS
10	FC, BOD, NO_3, DO, EC, COD, PO_4, TU, TS, pH

6.2. Model performance and validation

The sixteen algorithms were validated (Figs. 2 and 3). According to the results, all models performed well. However, BA-RT, BA-RF, BA-M5P, CVPS-RF, and RF have the highest prediction power among the models.

Table 5
Optimum values for each model's operator during testing phase.

Model's operators	Optimum value						
	M5P	RF	RT	REPT	Bagging	CVPS	RFC
Batch-size	100	100	100	100	100	100	100
Batch-size percent	–	100	–	–	100	–	–
Minimum number of Instance	2	–	1	2	–	–	–
Number of decimal places	3	2	3	2	2	2	3
Built regression tree	No	–	–	–	–	–	–
Do not check capabilities	No	No	No	No	No	No	No
Unpruned	No	–	–	No		–	–
Debug	Yes	Yes	No	No	No	No	No
Maximum depth of tree	–	0	2	−1	–	–	–
Number of features	–	0	–	–		–	–
Number of iteration	–	100	–	–	10	–	–
Number of execution slots	–	1	–	–	1	–	–
Store out of bag predictions	–	No	–	–	No	–	–
Break ties randomly	–	No	No	–	–	–	–
Calculate out of bag	–	No	–	–	–	–	–
K-value	–	–	0	–	–	–	–
Minimum variance proportion	–	–	0.001	0.001	–	–	–
Number of fold	–	–	0	3	–	10	–
Initial count	–	–	–	0	–	–	–
Classifier	–	–	–	–	Decision tree algorithms	Decision tree algorithms	Decision tree algorithms
Filter	–	–	–	–	–	–	Random projection

Table 6
Identification of the optimal input combination based on the RMSE.

Models	RMSE of Various Input Combinations										
	Phase	1	2	3	4	5	6	7	8	9	10
M5P	Training	8.27	5.36	4.1	3.05	2.16	2.12	2.15	2.18	2.45	2.93
	Testing	9.21	6.37	4.75	3.95	3.84	**3.65**	3.87	3.99	4.12	4.31
RF	Training	5.54	2.21	1.62	1.2	0.94	0.93	0.98	1.15	1.43	1.52
	Testing	10.81	6.28	4.7	3.53	**2.76**	2.81	2.82	2.97	3.12	3.24
RT	Training	5.038	0.39	0.13	0.13	0.18	0.14	0.15	0.17	0.18	0.21
	Testing	12	7.87	6.42	4.71	**3.9**	4.43	4.79	4.93	5.21	5.29
REPT	Training	8.03	4.72	3.62	3.07	2.95	2.95	3.10	3.26	3.44	3.65
	Testing	9.26	6.33	5.29	4.49	**4.28**	4.42	4.51	4.55	4.71	4.83
BA-M5P	Training	8.34	5.28	3.91	2.91	2.23	2.26	2.49	2.73	2.84	2.93
	Testing	9.13	6.08	4.66	4.33	4.23	**4.00**	4.16	4.26	4.41	4.49
BA-RF	Training	6.12	3.13	2.33	1.72	1.35	1.36	1.38	1.46	1.51	1.58
	Testing	10.54	6.2	4.64	3.5	**2.78**	2.84	2.89	2.96	3.13	3.36
BA-RT	Training	5.74	2.52	1.89	1.42	1.21	1.25	1.36	1.45	1.49	1.58
	Testing	10.99	6.52	4.98	3.66	**2.96**	3.14	3.22	3.29	3.38	4.25
BA-REPT	Training	7.2	4.36	3.22	2.3	2.1	2.14	2.28	2.35	2.46	2.67
	Testing	9.55	6.08	4.81	3.67	**3.15**	3.23	3.37	3.48	3.58	3.71
CVPS-M5P	Training	8.27	5.36	4	3.06	2.16	2.12	2.28	2.37	2.45	2.64
	Testing	9.21	6.37	4.74	4.35	4.21	**3.95**	4.31	4.39	4.51	4.62
CVPS-RF	Training	5.52	2.2	1.55	1.17	0.89	0.9	0.99	1.15	1.29	1.41
	Testing	10.88	6.32	4.64	3.5	**2.72**	2.76	2.88	3.04	3.21	3.45
CVPS-RT	Training	5.03	0.39	0.14	0.14	0.15	0.16	0.19	0.25	0.31	0.38
	Testing	12	7.87	6.36	4.76	5.01	**4.73**	4.89	5.12	5.29	5.56
CVPS-	Training	7.97	5.19	4.34	3.45	3.04	3.18	3.31	3.42	3.55	3.78
REPT	Testing	9.29	6.9	5.5	4.71	**4.44**	4.49	4.61	4.73	4.84	4.95
RFC-M5P	Training	8.28	5.36	4.07	3.15	4.97	3.56	3.87	4.35	4.78	5.10
	Testing	9.21	6.37	4.67	**3.95**	6.69	6.28	6.57	6.84	6.95	7.13
RFC-RF	Training	5.51	2.2	1.65	1.2	1.97	1.74	1.83	1.98	2.34	2.36
	Testing	10.74	6.34	4.91	**3.69**	5.25	4.74	5.10	5.27	5.62	5.83
RFC-RT	Training	5.03	0.39	0.13	0.14	0.13	0.13	0.16	0.19	0.25	0.31
	Testing	11.97	8	6.47	**5.15**	7.23	6.23	6.62	6.93	7.21	7.36
RFC-REPT	Training	7.95	5.1	3.79	3.09	4.94	4.32	5.21	5.56	5.92	6.31
	Testing	9.42	6.47	5.27	**4.85**	6.77	6.1	6.84	6.93	7.26	7.52

Bold and shading shows the best input combination.

To further analyze the performance, PREI, which examines the performance of models based on their tendency to over- or underestimate the WQI, was applied (Fig. 4). Among the standalone algorithms, the lowest error belonged to the RF model. M5P also had a very low error rate, but there was a large error (about −40) which reduced its performance. Although the error ranges for RT and REPT were between ±10, these algorithms didn't predict each value properly (i.e., each error value was far from 0 and is dispersed). Hybrid models, especially the bagging algorithm, enhanced the predictive power of the standalone algorithms (i.e., compare Fig. 4a with e, c with g and d and h).

The box plots of the measured and predicted WQI values showed that the RFC-RT, RFC-M5P, CVPS-M5P, M5P, and BA-M5P models are highly accurate, predicting the maximum values of WQI well. Only RFC-RF predicted the lower values accurately (Fig. 5).

The weakness of visually comparing the models' predictions to evaluate their performance is that those with higher prediction capabilities are easy to recognize, but identifying the best algorithm and the rank of success is difficult. Therefore, some quantitative information that provides better evidence of the performance of each algorithm is necessary (Table 7). Although all of the algorithms performed well ($R^2 > 0.75$), the hybrid BA-RT ($R^2 = 0.941$) had the highest prediction success and the CVPS-REPT ($R^2 = 0.823$) had the lowest. Similarly, the BA-RT model had the lowest RMSE (2.71) and its MAE (1.87) was also the best. When 0.75 < NSE ≤ 1, a model has very good prediction power (Moriasi et al., 2007). Thus, all showed very good performance, but BA-RT outperformed the others (NSE = 0.941). The PBIAS metric indicated that all the algorithms except RT, BA-RT, and CVPS-REPT overestimated WQI. The rank (best to worst) of the algorithms based on performance is: BA-RT, BA-RF, BA-M5P, CVPS-RF, RF, RFC-RF, BA-

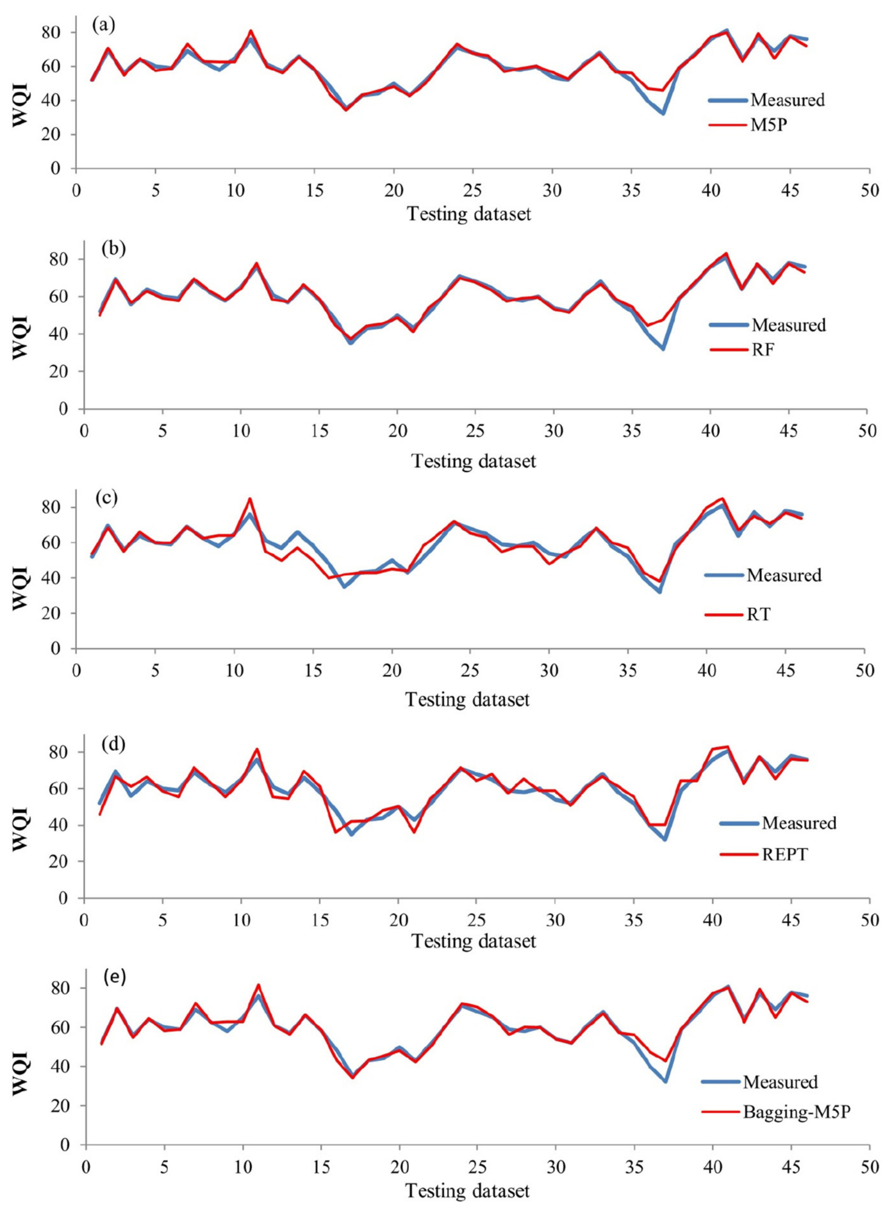


Fig. 2. Time variation graph of predicted and measured value during testing phase.

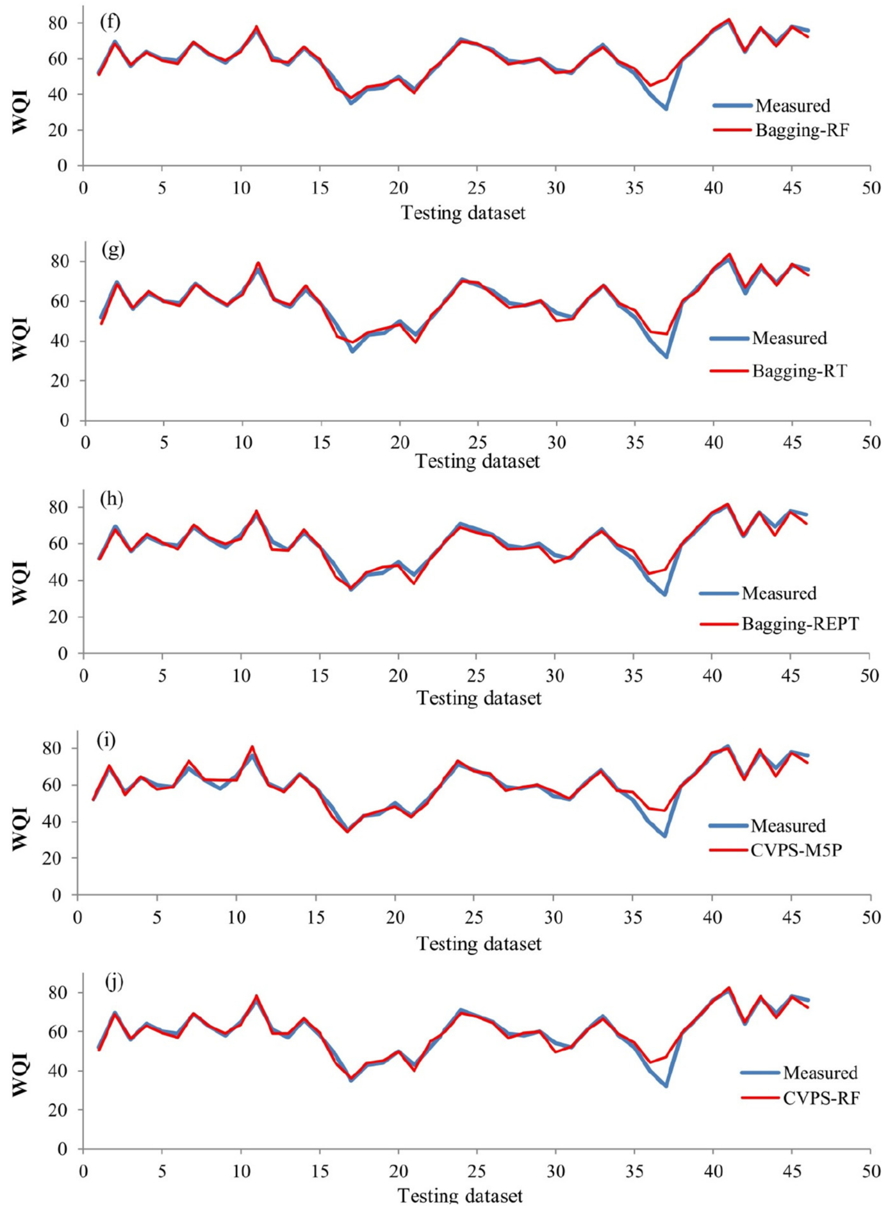


Fig. 2 (*continued*).

REPT, M5P = CVPS-M5P, RFC-REPT, RFC-M5P, REPT, RT, RFC-RT, CVPS-RT, CVPS-REPT.

7. Discussion

Streamflow is a vital factor in water quality sampling (Harmel et al., 2006). The seasonality of rainfall affects runoff volumes and influences the water quality (Yunus and Nakagoshi, 2004). Additionally, river water quality can influence also influence groundwater quality due to direct percolation (Kapetas et al., 2019). Additionally, using river water for irrigation purposes could also influence groundwater resources and hence the application of river water irrigation should be performed in zones where aquifers are protected (Busico et al., 2020).

Northern Iran, with a humid climate, receives approximately 730 mm of precipitation annually. Daily streamflow in its rivers fluctuates and may receive surges of discharge. Measured BOD, COD, and other characteristics vary by season, with the highest BOD and COD (transported through runoff) observed during wet seasons. The variability and complexity of precipitation over a watershed requires a robust and flexible WQI model to minimise the effects of nonlinearity and to increase prediction accuracy.

In this study, four standalone tree-based algorithms (M5P, RF, RT, and REPT) were employed to predict WQI in the Talar catchment. Twelve additional hybrid algorithms were developed by combining the standalone models with bagging (BA), CV parameter selection (CVPS) and randomizable filter classification (RFC) algorithms. The performances of the sixteen models were compared.

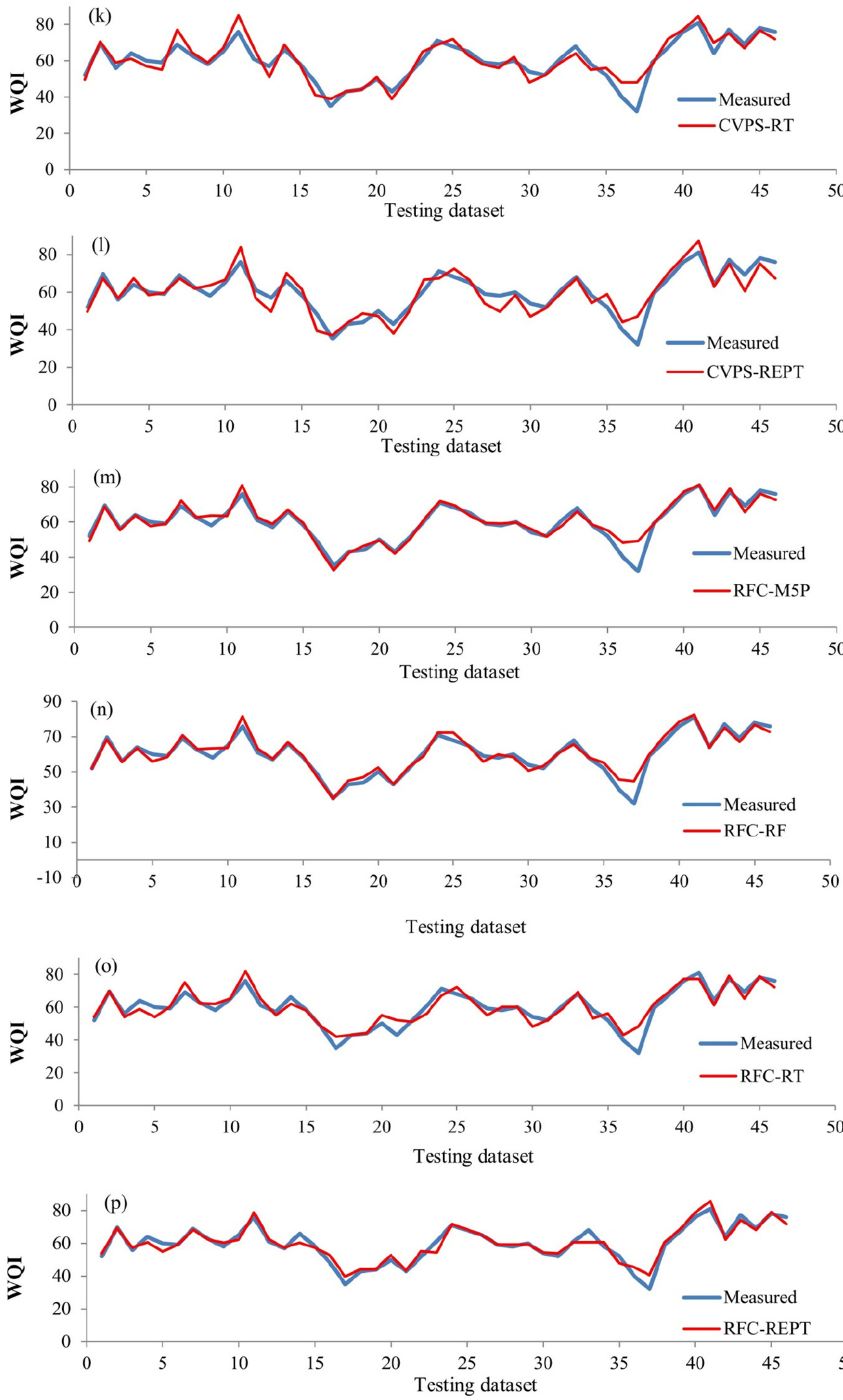


Fig. 2 (*continued*).

Khosravi et al. (2019) compared the prediction power of some standalone tree-based algorithms with the neuron-based algorithm ANFIS as well as ANFIS hybridized with meta-heuristic optimization algorithms. Their findings illustrate that standalone models of neuron-based algorithm have a lower prediction powers due to substantial weaknesses. Hybridization can enhance their predictions significantly. Their results also indicate that the performance of standalone tree-based algorithms is similar to the performance of ANFIS hybridized with meta-heuristic optimization, which achieved higher prediction power than tree-based algorithms. The hybrid models in this study increased the performance of some standalone tree-based models, but not in all cases. Tree-based algorithms have very high prediction powers on their own.

We find that hybrid tree-based models (especially the bagging algorithm) are more robust and flexible than standalone models. They perform better when estimating non-linear phenomenon like the WQI (De'ath and Fabricius, 2000). Our findings are in agreement with the conclusions of Khosravi et al. (2018), Khozani et al. (2019), Sharafati et al. (2019), and Ghorbani et al. (2017). It can be concluded that not all models follow this pattern and some hybrid models may be exceptions to this. The predictive power of hybrid models depends entirely on the base algorithm used. For example, the standalone RF performed better than the BA-REPT when estimating WQI. Among the standalone

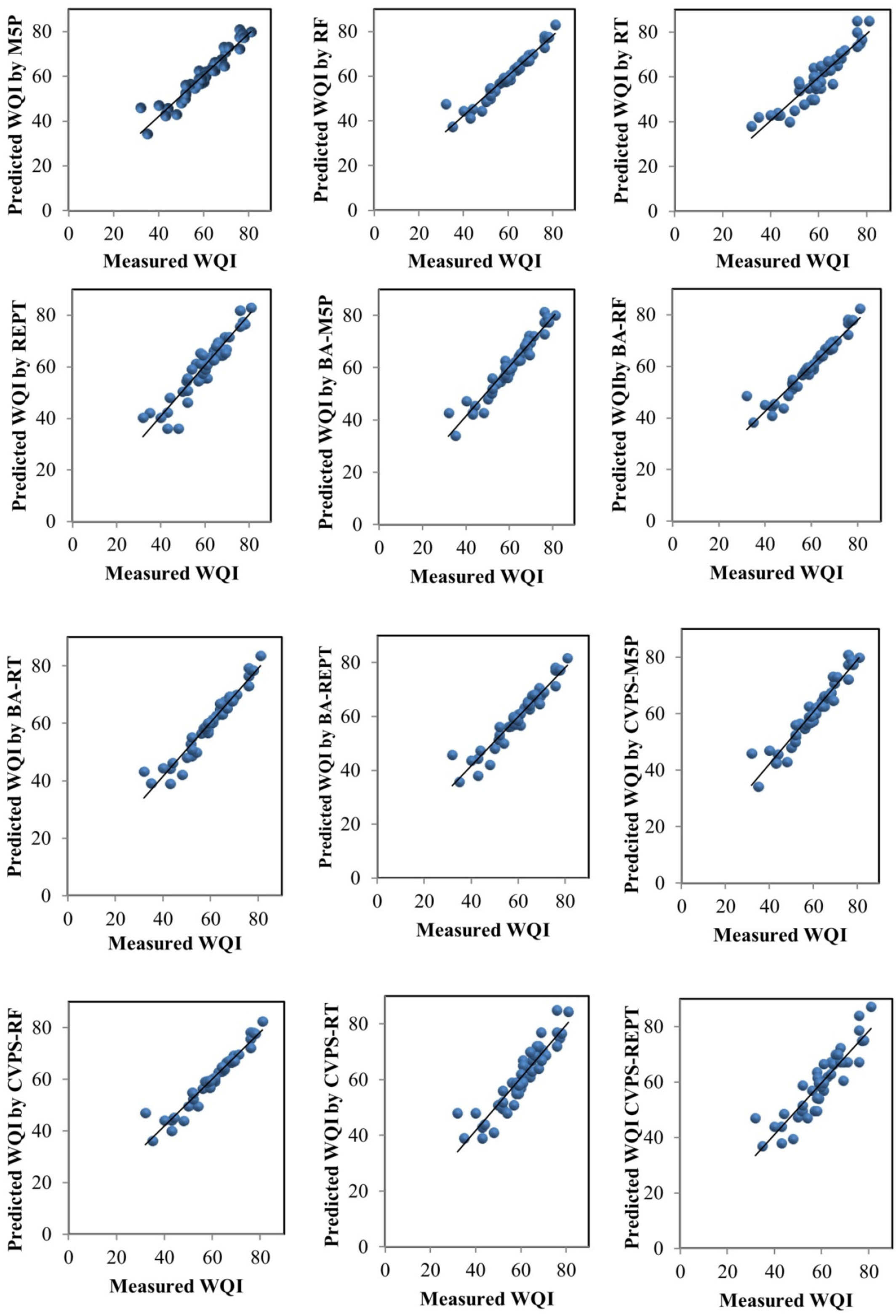


Fig. 3. Scatter plots of predicted and measured values produced by the 16 models during testing.

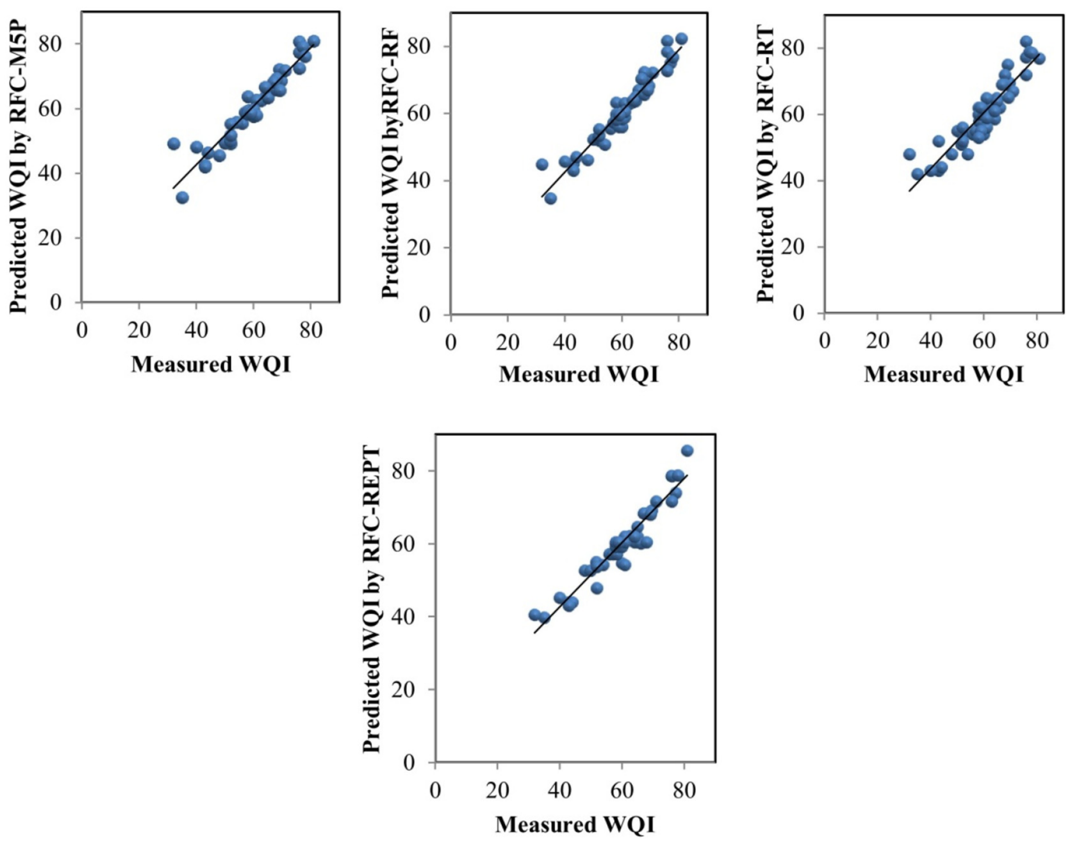


Fig. 3 (*continued*).

models, RF outperformed M5P, REPT, and RT; this is in accordance with Khosravi et al. (2019) but counter to the results of Khozani et al. (2019), who found that M5P performed better than RF when predicting apparent shear stress in open channels. Therefore, there is no “best” and “most robust” algorithm that always performs better in all circumstance. Many algorithms and their hybrids should be tested to find the models that work the best in each situation.

Besides a model's structure, one of the most important influencers of performance is the determination of the best combination of variables to input. The effects of combining specific variables on the result isn't consistent from one catchment to the next due not only to the spatial complexity of processes in each catchment, but also because of the variety of point and non-point sources of pollution that cause nonlinear relationships between variables and water quality. Some research neglected consideration of different variable combinations to determine the best set. Some introduced all variables simultaneously (Hameed et al., 2016). And some have even applied different methods, like multiple linear regression (which depends upon CC), to select the best input variables (Barzegar et al., 2016). This study demonstrates that different input combinations affect the results in diverse ways. Therefore, to improve performance, a variety of combinations of input variables should be tested and the most effective set should be used. The most effective combination is not consistent between models; each model may have its own “best” combination. The results are a function of the structure of each algorithm and the fit of the dataset to the model's structure (data structure and distribution).

Sahoo et al. (2015) used ANFIS and Yaseen et al. (2018) used hybrid ANFIS to model WQI. They proposed three different hybridized ANFIS models with fuzzy c-means (FCM) data clustering, grid partition (GP), and subtractive clustering (SC). Their results showed that Yaseen et al.'s (2018) results ($R^2 = 0.9269$ for ANFIS-FCM) were better than Shoo et al.'s (2015) results ($R^2 = 0.792$ for the standalone ANFIS model). Tiwari et al. (2018) used a hybrid algorithm (ANFIS-SC) and achieved an R^2 value of 0.90 and Li et al. (2019) used a standalone SVR ($R^2 = 0.82$) and a hybrid (SVR-FFA) ($R^2 = 0.90$). In our results, all standalone and hybrid models yielded better WQI prediction than any other previous algorithm tested for WQI prediction. We recommend that future research should test datasets to predict WQI by dividing them into wet- and dry-season datasets and comparing the resulting models to the results for the entire dataset.

8. Conclusion

This study evaluated he effectiveness of four standalone (RF, M5P, RT, and REPT) and 12 hybrid data-mining algorithms (hybrids of the standalones with bagging, CVPS, and RFC) for predicting monthly WQI in a humid environment in northern Iran. Our aim was to develop and propose novel algorithms not only for WQI prediction, but also for other aspect of water science in regions with very thin distributions of water quality gauging stations.

The modeling process revealed that fecal coliform concentration was the most important determinant of WQI. This was followed in order of importance by BOD, NO_3^-, DO, EC, COD, PO_4^{2-}, Turbidity,

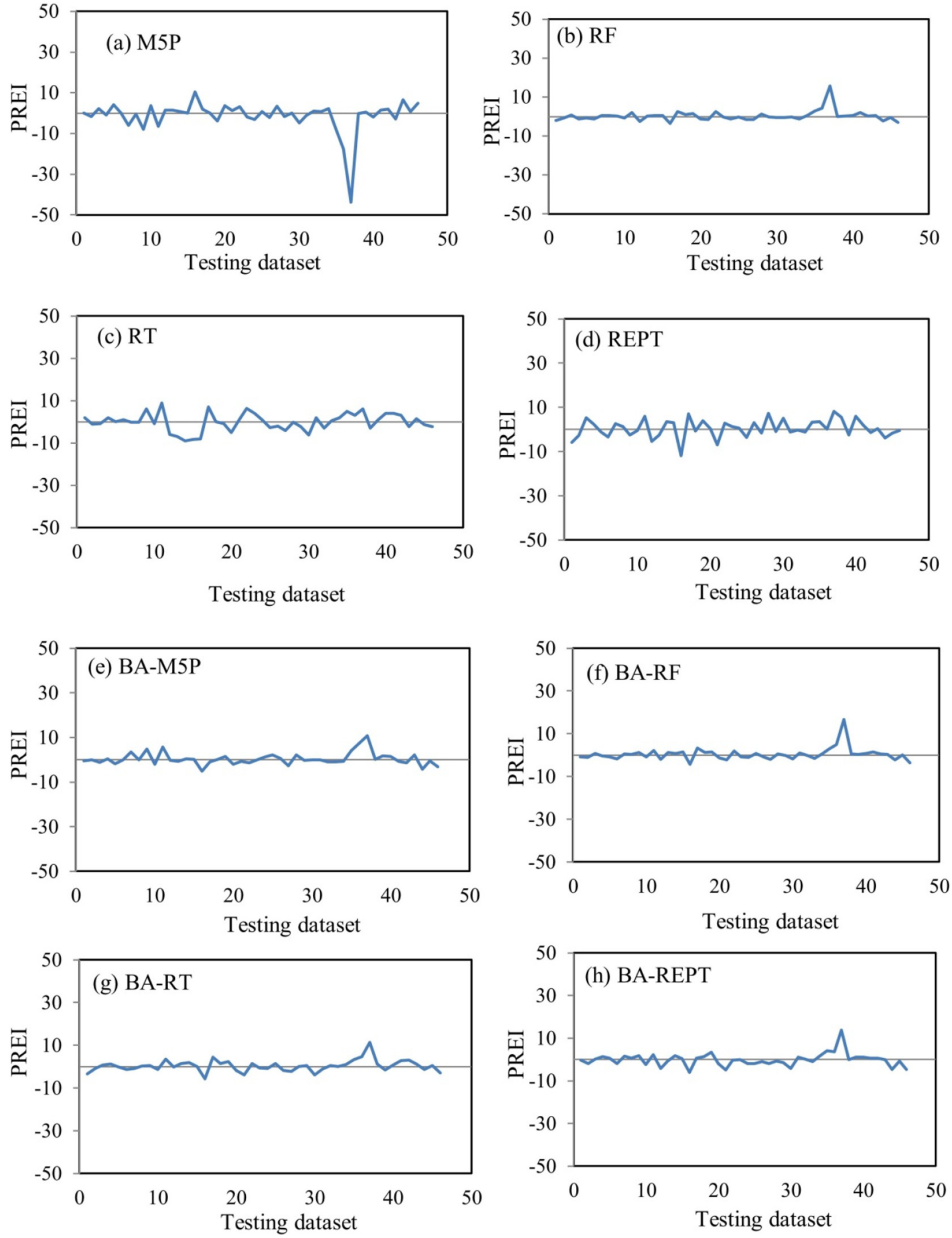


Fig. 4. Error graph of estimated valued compare to measured data in testing phase.

TS, and pH. We also found that using different combinations of variables resulted in different levels of performance of the models and that the effects of changing the inputs on the models for our study area had inconsistent and dissimilar effects on modeling in other catchments, even when using the same variable combinations. The models revealed that prediction power is best when the variables with the highest CCs are used. Variables with very low CCs negatively impact predictive power. The results of the hybrids, compared to the standalone algorithms, showed that they had improved prediction accuracies (i.e., better than those of the standalone algorithms), but may not be as successful in all cases. The level of prediction by the BA-RT model was better than all other models. In order of decreasing performance after BA-RT are RF, bagging-RF, bagging-RT, bagging-REPT, RFC-RF, RT, M5P = CVPS-M5P, RFC-M5P, bagging-M5P, REPT, CVPS-REPT, CVPS-RT, RFC-REPT, and RFC-RT. Although the BA-RT hybrid had the highest performance, it didn't predict extreme WQI values accurately. Nearly all algorithms overestimated WQI values, but RT, BA-RT and CVPS-REPT did not.

It should be noted that when these algorithms produce reliable results using a dataset covering a short time period, they will be much

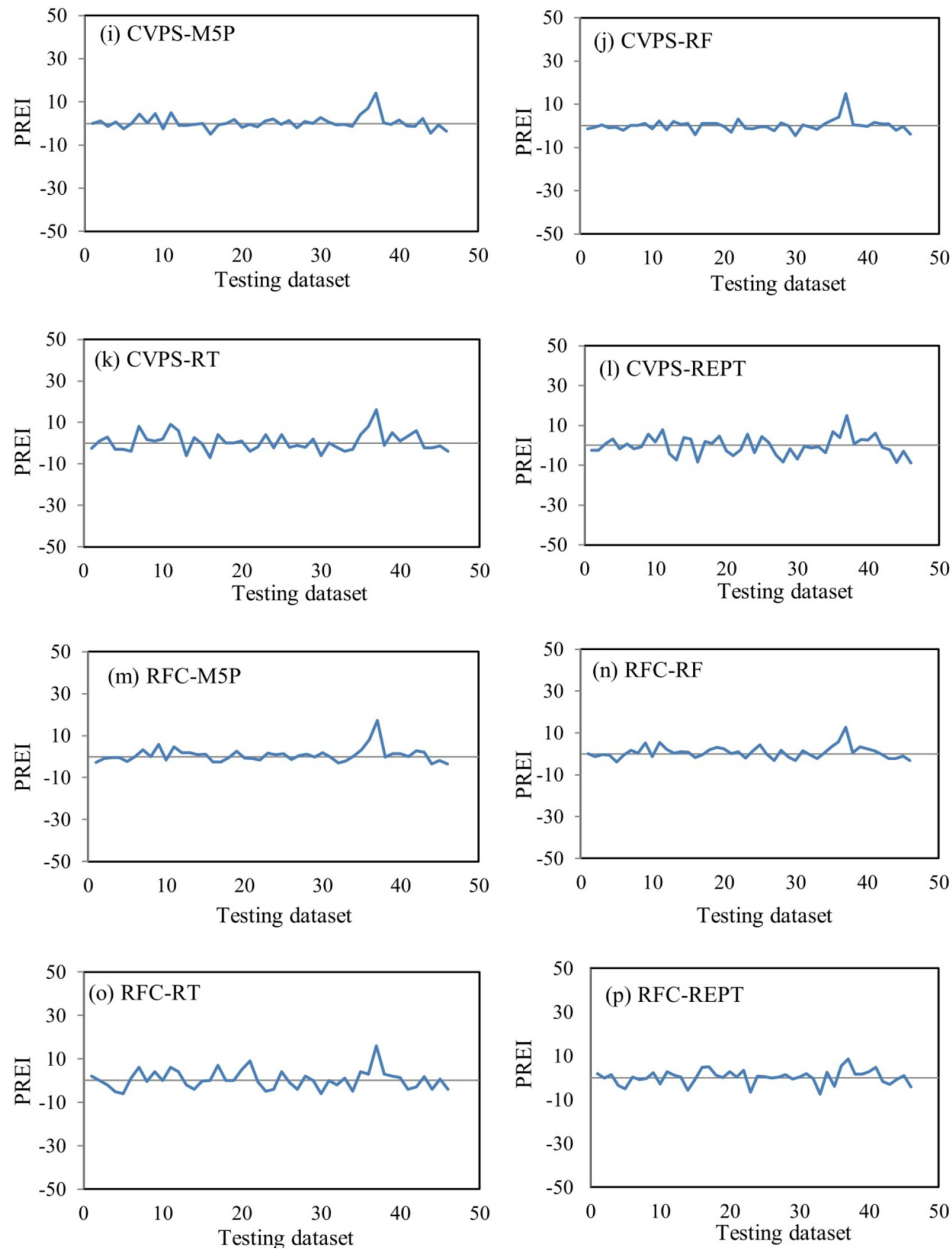


Fig. 4 (*continued*).

more robust for datasets covering longer periods. Therefore, these algorithms may be particularly useful for developing regions that have more limited gauging networks and where gauging networks have been established more recently. Our results suggest that the proposed BA-RT algorithm could be a reliable and cost-effective algorithm to improve groundwater quality management in humid regions in northern Iran. This model is likely to be much more useful in developing countries where the costs of measuring some water-quality parameters are high and may be generally prohibitive. These results cannot be generalized and applied to other study areas or equated to other hydrological data, however. BA-RT would undoubtedly be a powerful algorithm, but it may not be the best (i.e., most accurate) in all places and in all circumstances.

CRediT authorship contribution statement

Duie Tien Bui: Conceptualization. **Khabat Khosravi:** Investigation, Formal analysis, Conceptualization, Writing - original draft. **John**

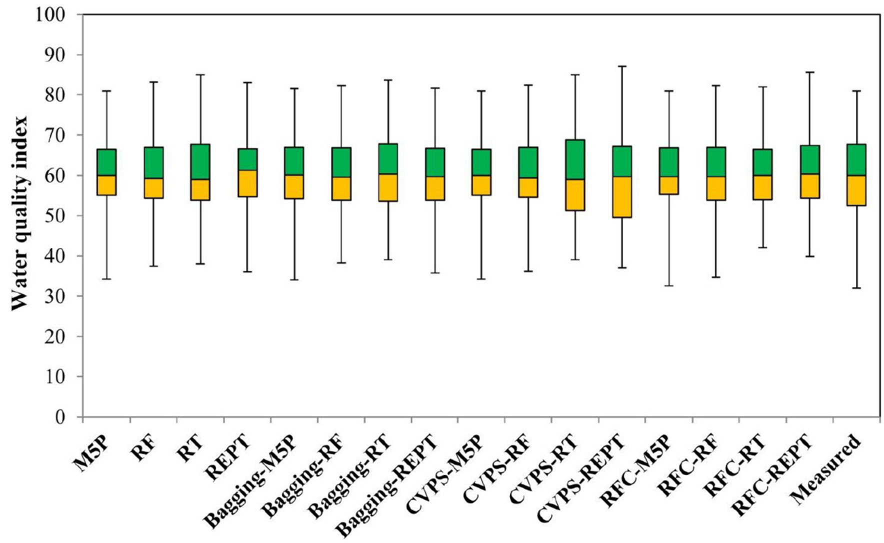


Fig. 5. Box plot of applied algorithms for model performance.

Table 7
Identification of the best algorithm for prediction of the WQI prediction.

Algorithms	R^2	RMSE	MAE	NSE	PBIAS	Rank Order
M5P	0.923	3.11	1.96	0.922	-0.754	8
RF	0.932	2.97	1.67	0.930	-0.600	5
RT	0.873	4.13	3.22	0.863	0.227	12
REPT	0.880	4.05	3.19	0.868	-0.774	11
BA-M5P	0.938	2.79	1.79	0.937	-0.650	3
BA-RF	0.940	2.78	1.51	0.938	-0.467	2
BA-RT	**0.941**	**2.71**	**1.87**	**0.941**	**0.500**	1
BA-REPT	0.925	3.055	2.01	0.925	0.204	7
CVPS-M5P	0.923	3.11	1.96	0.922	-0.754	8
CVPS-RF	0.937	2.82	1.67	0.936	-0.237	4
CVPS-RT	0.853	4.47	3.41	0.839	-1.040	14
CVPS-REPT	0.823	4.90	3.90	0.807	0.419	15
RFC-M5P	0.907	3.47	2.14	0.903	-1.19	10
RFC-RF	0.931	3.01	2.09	0.927	-1.15	6
RFC-RT	0.854	4.30	3.16	0.851	-0.72	13
RFC-REPT	0.920	3.18	2.37	0.918	-0.385	9

Shading shows the most powerful algorithm with highest prediction power.

Tiefenbacher: Writing - review & editing. **Hoang Nguyen:** Methodology. **Nerantzis Kazakis:** Conceptualization, Writing - original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2020.137612.

References

Ahmed, A.A.M., Shah, S.M.A., 2015. Application of adaptive neuro-fuzzy inference system (ANFIS) to estimate the biochemical oxygen demand (BOD) of Surma River. Journal of King Saud University – Engineering Sciences https://doi.org/10.1016/j.jksues.2015.02.001.

Barzegar, R., Moghaddam, A.A., Adamowski, J., Fijani, E., 2016. Comparison of machine learning models for predicting fluoride contamination in groundwater. Stoch. Env. Res. Risk A. 31 (10), 2705–2718.

Bournaris, Th., Papathanasiou, J., Manos, B., Kazakis, N., Voudouris, K., 2015. Support of irrigation water use and eco-friendly decision process in agricultural production planning. Oper. Res. 15, 289–306.

Bui, D., Pradhan, B., Nampak, H., Bui, T.Q., Tran, Q.A., Nguyen, Q.P., 2016. Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibility modelling in a high-frequency tropical area. J. Hydrol. 540, 317–330.

Bui, D., Khosravi, K., Li, S., Shahabi, H., Panahi, M., Singh, V., Chapi, k., et al., 2018a. New hybrids of anfis with several optimization algorithms for flood susceptibility modeling. Water 10 (9), 1210.

Bui, D., Panahi, M., Shahabi, H., Singh, V.P., Shirzadi, A., Chapi, K., Khosravi, K., et al., 2018b. Novel hybrid evolutionary algorithms for spatial prediction of floods. Sci. Rep. 8 (1), 15364.

Busico, G., Kazakis, N., Cuoco, E., Colombani, N., Tedesco, D., Voudouris, K., Mastrocicco, M., 2020. A novel hybrid method of specific vulnerability to anthropogenic pollution using multivariate statistical and regression analyses. Water Res. 171, 115386.

Chen, Wei-Bo, Liu, Wen-Cheng, 2015. Water quality modeling in reservoirs using multivariate linear regression and two neural network models. Advances in Artificial Neural Systems https://doi.org/10.1155/2015/521721.

Chen, W., Panahi, M., Khosravi, K., Pourghasemi, H.R., Rezaie, F., 2019a. Spatial prediction of groundwater potentiality using ANFIS ensembled with teaching-learning-based and biogeography-based optimization. J. Hydrol. 572, 435–448.

Chen, W., Pradhan, B., Li, S., Shahabi, H., Rizeei, H.M., Hou, E., Wang, S., 2019b. Novel hybrid integration approach of bagging-based fisher's linear discriminant function for groundwater potential analysis. Nat. Resour. Res. 28, 1239–1258. https://doi.org/10.1007/s11053-019-09465-w.

Choubin, B., Darabi, H., Rahmati, O., Sajedi-Hosseini, F., Kløve, B., 2018. River suspended sediment modelling using the CART model: a comparative study of machine learning

techniques. Sci. Total Environ. 615, 272–281. https://doi.org/10.1016/j.scitotenv.2017.09.293.
De'ath, G., Fabricius, K.E., 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology https://doi.org/10.1890/0012- 9658(2000)081[3178:CARTAP]2.0.CO;2.
Fathi, E., Zamani Ahmadmahmoodi, R., Bidaki, R., 2018. Water quality evaluation using water quality index and multivariate methods, Beheshtabad River, Iran. Appl Water Sci 8, 210. https://doi.org/10.1007/s13201-018-0859-7.
Gazzaz, N.M., Yusoff, M.K., Aris, A.Z., Juahir, H., Ramli, M.F., 2012a. Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. Mar. Pollut. Bull. 64, 2409–2420.
Gazzaz, N.M., Yusoff, M.K., Zarin Aris, A., Juahir, H., Firuz, M., 2012b. Artificial neural network modeling of the water quality index for Kinta River using water quality variables as predictors. Mar. Pollut. Bull. 64 (11), 2409–2420.
Gholami, M., Ghachkanlu, E.N., Khosravi, K., Pirasteh, S., 2019. Landslide prediction capability by comparison of frequency ratio, fuzzy gamma and landslide index method. Journal of Earth System Science 128 (2), 42.
Ghorbani, M.A., Deo, R.C., Karimi, V., Yaseen, Z.M., Terzi, O., 2017. Implementation of a hybrid MLP-FFA model for water level prediction of Lake Egirdir, Turkey. Stoch. Environ. Res. Risk Assess., 1–15 https://doi.org/10.1007/s00477-017-1474-0.
Hameed, M., Sharqi, S.S., Yaseen, Z.M., Afan, H.A., Hussain, A., Elshafie, A., 2016. Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia. Neural Computing and Application 28, 893–905. https://doi.org/10.1007/s00521-016-2404-7.
Hamzeh Haghibi, A., Nasrolahi, A., Parsaie, A., 2018. Water quality prediction using machine learning methods. Water Quality Research Journal 53 (1), 3–13.
Harmel, R.D., Cooper, R.J., Slade, R.M., Haney, R.L., Arnold, J.G., 2006. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. Transactions of the American Society of Agricultural and Biological Engineers 49 (3), 689–701.
Iticescu, C., Georgescu, L.P., Murariu, G., Topa, C., Timofti, M., Pintilie, V., Arseni, M., 2019. Lower Danube water quality quantified through WQI and multivariate analysis. Water 11, 1305.
Kahng, S., Iwata, B.A., Fischer, S.M., Page, T.J., Treadwell, K.R., Williams, D.E., Smith, R.G., 1998. Temporal distributions of problem behavior based on scatter plot analysis. J. Appl. Behav. Anal. 31, 593–604.
Kapetas, L., Kazakis, N., Voudouris, K., McNicoll, D., 2019. Water allocation coordination and governance in multi-stakeholder environments: insights from Axios Delta, Greece. Sci. Total Environ. 695, 133831.
Khosravi, K., Nohani, E., Maroufinia, E., Pourghasemi, H.R., 2016. A GIS-based flood susceptibility assessment and its mapping in Iran: a comparison between frequency ratio and weights-of-evidence bivariate statistical models with multi-criteria method. Nat. Hazards 83 (2), 1–41.
Khosravi, K., Mao, L., Kisi, O., Yaseen, Z.M., Shahid, S., 2018. Quantifying hourly suspended sediment load using data mining models: case study of a glacierized Andean catchment in Chile. J. Hydrol. 567, 165–179.
Khosravi, K., Daggupati, P., Alami, M.T., Awadh, S.M., Ghareb, M.I., Panahi, M., et al., 2019a. Meteorological data mining and hybrid data-intelligence models for reference evaporation simulation: a case study in Iraq. Comput. Electron. Agric. 167, 105041.
Khosravi, K., Barzegar, R., Miraki, S., Adamowski, J., Daggupati, P., Pham, B., 2019b. Stochastic modeling of groundwater fluoride contamination: introducing lazy learners. Groundwater (In press).
Khosravi, K., Shahabi, H., Pham, B.T., Adamowski, J., Shirzadi, A., Pradhan, B., et al., 2019c. A comparative assessment of flood susceptibility modeling using Multi-Criteria Decision-Making Analysis and Machine Learning Methods. J. Hydrol. 573, 311–323.
Khozani, Z.S., Khosravi, K., Pham, B.T., Kløve, B., Mohtar, W., Yaseen, Z., 2019. Determination of compound channel apparent shear stress: application of novel data mining models. J. Hydroinf. 21 (5), 798–811.
Khuan, L.Y., Hamzah, N., Jailani, R., 2002. Prediction of water quality index (WQI) based on artificial neural network (ANN). Student Conference on Research and Development, Shah Alam, Malaysia, Malaysia.
Kisi, O., Zounemat-Kermani, M., 2014. Comparison of two different adaptive neuro-fuzzy inference systems in modelling daily reference evapotranspiration. Water Resour. Manag. 28 (9), 2655–2675.
Kisi, O., Dailr, A.H., Cimen, M., Shiri, J., 2012. Suspended sediment modeling using genetic programming and soft computing techniques. J. Hydrol. 450, 48–58.
Leong, W.C., Bahadori, A., Zhang, J., Ahmad, H., 2019. Prediction of water quality index (WQI) using support vector machine (SVM) and least square-support vector machine (LS-SVM). International Journal of River Basin Management https://doi.org/10.1080/15715124.2019.1628030 (In Press).
Li, J., Ali Abdulmohsin, H., Hasan, S., Kaiming, L., Al-Khateeb, B., Ghareb, M.I., 2019. Hybrid soft computing approach for determining quality indicator: Euphrates River. Neural Comput. & Applic. 31 (3), 827–837.
Machiwal, D., Cloutier, V., Güler, C., Kazakis, N., 2018. A review of GIS-integrated statistical techniques for groundwater quality evaluation and protection. Environmental Earth Science 77, 681.
Misaghi, F., Delgosha, F., Razzaghmanesh, M., Myers, B., 2017. Introducing a water quality index for assessing water for irrigation purposes: a case study of the Ghezel Ozan River. Sci. Total Environ. 589 (2017), 107–116.
Mohammadpour, R., Shaharuddin, S., Chang, C.K., et al., 2015. Prediction of water quality index in constructed wetlands using support vector machine. Environ. Sci. Pollut. Res. 2015 (22), 6208. https://doi.org/10.1007/s11356-014-3806-7.
Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Binger, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50, 885–900. https://doi.org/10.13031/2013.23153.
Othman, F., M E, A.E., Mohamed, I., 2012. Trend analysis of a tropical urban river water quality in Malaysia. J. Environ. Monit. 14, 3164–3173.
Pham, B.T., Prakash, I., Khosravi, K., Chapi, K., Trinh, P.T., Ngo, T.Q., Hosseini, S.V., et al., 2018. A comparison of Support Vector Machines and Bayesian algorithms for landslide susceptibility modelling. Geocarto International 34 (13), 1385–1407.
Rufino, F., Busico, G., Cuoco, E., Darrah, T.H., Tedesco, D., 2019. Evaluating the suitability of urban groundwater resources for drinking water and irrigation purposes: an integrated approach in the Agro-Aversano area of Southern Italy. Environ. Monit. Assess. 191, 768. https://doi.org/10.1007/s10661-019-7978-y.
Sahoo, M., Patra, K.C., Khatua, K.K., 2015. Inference of Water Quality Index Using ANFIA and PCA. Aquatic Procedia 4, 1099–1106.
Sakizadeh, M. Model, 2016. Artificial intelligence for the prediction of water quality index in groundwater systems. Earth Syst. Environ. 2016 (2), 8. https://doi.org/10.1007/s40808-015-0063-9.
Sharafati, A., Khosravi, K., Khosravinia, P., Ahmed, K., Salman, S.A., Yaseen, Z.M., 2019. The potential of novel data mining models for global solar radiation prediction. Int. J. Environ. Sci. Technol. 1–18.
Termeh, S.V.R., Khosravi, K., Sartaj, M., Keesstra, S.D., Tsai, F.T.C., Dijksma, R., Pham, B., 2019. Optimization of an adaptive neuro-fuzzy inference system for groundwater potential mapping. Hydrogeol. J. 27 (7), 2511–2534.
Tiwari, S., Babbar, R., Kaur, G., 2018. Performance evaluation of two ANFIS models for predicting water quality index of river Satluj (India). Advances in civil engineering https://doi.org/10.1155/2018/8971079.
Touchette, P.E., MacDonald, R.F., Langer, S.N., 1985. A scatter plot for identifying stimulus control of problem behavior. J. Appl. Behav. Anal. 18, 343–351.
Vadiati, M., Asghari-Moghadam, A., Nakhaei, M., Adamowski, J., Akbarzadeh, A.H., 2016. A fuzzy-logic based decision-making approach for identification of groundwater quality based on groundwater quality indices. J. Environ. Manag. 184 (2), 255–270.
Wan Mohtar, W.H.M., Abdul Maulud, K.N., Muhammad, N.S., Sharil, S., Yaseen, Z.M., 2019. Spatial and temporal risk quotient based river assessment for water resources management. Environ. Pollut. 248 (2019), 133–144.
Yaseen, Z.M., Ramal, M.M., Diop, L., Jaafar, O., Demir, V., Kisi, O., 2018. Hybrid adaptive neuro-fuzzu models for water quality index estimation. Water Resour. Manag. 32, 2227–2245.
Yunus, A.J.M., Nakagoshi, N., 2004. Effects of seasonality on streamflow and water quality of the Pinang River in Penang Island, Malaysia. Chin. Geogr. Sci. 14 (2), 153–161.