Uncertainty analysis for the data-driven model using Monte Carlo simulations to predict sodium adsorption rate: A case study, Aras, Sepid-Rud, and Karun Rivers in Iran

Elham Rahnama

Hormozgan University

Omolbanin Bazrafshan

Hormozgan University

Gholamreza Asadollahfardi (✉ asadollahfardi@yahoo.com)

Kharazmi University Faculty of Engineering <https://orcid.org/0000-0002-7867-8757>

Seyed Yaser Samadi

Southern Illinois University Carbondale

Research Article

Keywords: Monte Carlo Simulation, SAR, ANFIS-FCMC, water quality

Posted Date: November 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1058278/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Uncertainty analysis for the data-driven model using Monte Carlo simulations to predict sodium adsorption rate: A case study, Aras, Sepid-Rud, and Karun Rivers in Iran

Elham Rahnama¹, Omolbanin Bazrafshan¹, Gholamreza Asadollahfardi^{2*}, Seyed Yaser Samadi³

1-Departement of Natural Resources Engineering, Faculty of Agricultural Engineering and Natural Resources, University of Horproject

mozgan Bandar Abbas, Iran, rahnamae57@gmail.com,
O.bazrafshan@hormozgan.ac.ir

2- Civil Engineering Department, Faculty of Engineering, Kharazmi University, Tehran, Iran,
fardi@khu.ac.ir, asadollahfardi@yahoo.com

3-Departement of mathematic, school of mathematical and statistical sciences, Southern Illinois University, USA ysamadi@siu.edu

Abstract: Water quality management requires a profound understating of future variations of surface and groundwater qualities for assessment and planning for human consumption, industrial, and irrigation purposes. In this regard, mathematical models, such as Box-Jenkins time series models, Bayesian time series models, and data-driven models are available for future prediction of water quality. However, the uncertainty associated with forecasting is one of the main problems of using these models towards water quality and future planning. In the present work, the uncertainty of the Adaptive Neuro-Fuzzy Inference System, based on Fuzzy c-means clustering, (ANFIS-FCMC) (genfis 3) model is quantified to analyze and predict Sodium Adsorption Rate(SAR) of water of Aras, Sepid-Rud, and Karun Rivers by using Monte Carlo simulations. The results indicate the combined standard and the expanded uncertainty simulated for SAR of Aras River water are 0.58 and 1.16, respectively, and the gap is 2.412 ± 1.1622 . Also, the combined standard and the expanded uncertainty simulated for SAR of Spid-Rud River water were 1.11 and 2.22, respectively, and the gap is equal to 2.235 ± 2.22 . Furthermore, the combined standard and the expanded uncertainty simulated for SAR of Aras River water are 2.063, and 4.126, respectively, and the gap is 4.79 ± 4.126 . Finally, the minimum uncertainty happened to predict SAR of Aras River using ANFIS-FCMC (genfis3) model and maximum SAR uncertainty belong to Karun River.

Keywords: Monte Carlo Simulation, SAR, ANFIS-FCMC, water quality

1. Introduction

For water resources management, forecasting water quality parameters for future planning and management of surface and groundwater is vital. Several statistical models are available for use in water quality modelings such as regression analysis, Box-Jenkins time series, Bayesian time series, and data-driven techniques. Uncertainty emerges from data collection of water quality monitoring and developed models. Uncertainty of the mentioned models is one of the challenges for application in practice. Some methods for analyzing uncertainty exist such as Expression of Uncertainty in Measurement (GUM) and Monte Carlo Simulations (MCS). In the present study, an adaptive network-based fuzzy inference system (ANFIS) is used to predict the sodium adsorption rate(SAR) of three rivers, including Aras, Spid-Rud, and Karun Rivers. However, analyzing the uncertainty of these models is necessary for practice. Some researchers have worked on the uncertainty of data-driven models that we briefly discussed in the following paragraphs.

Zou et al. (2002) used a neural network embedded Monte Carlo (NNMC) method to determine the uncertainty in water quality modeling. By embedding a NN into the conventional Monte Carlo simulation, the proposed method significantly improves upon the conventional method. They applied uncertainty and risk analyses of a phosphorus index for the Triadelphia Reservoir in Maryland. They indicated that the NNMC method had the potential for effective uncertainty analysis of water quality modeling. Nakane and Haidary(2010) developed two linkage models of multiple regression approach in twenty-one river basins in the Chugoku district of west Japan for forecasting the total nitrogen ($r^2=0.70$, $p<0.01$) and the total phosphorus ($r^2=0.47$, $p<0.01$) concentrations. They also applied Monte Carlo method-based sensitivity analysis and their results described that the total nitrogen (TN) regression model would be able to forecast stream water TN concentration between 0.4-3.2 mg/L.

Jiang et al. (2013) proposed a method that integrates an ANN into the MCS to improve the computational efficiency of conventional risk assessment. They used an ANN model as a replacement for the iterative finite element method (FEM) runs. The number of water quality parameters decreased. They studied chemical oxygen demand(COD) in the Lanzhou section of the Yellow River in China as a case study and compared their method with the conventional risk assessment approach. They showed that the ANN-MCS-based technique can save much computational effort without a loss of accuracy. Cuesta Antanasijevic et al. (2014) performed the training, validation, testing, and uncertainty analysis of general regression neural network (GRNN) models for the prediction of dissolved oxygen (DO) in the Danube River. They determined optimal data normalization and input choice methods, the determination of the relative significance of uncertainty in different input parameters, in addition to the uncertainty analysis of model outcomes applying the Monte Carlo Simulation (MCS) method. They applied the GRNN models on monthly water quality parameters of 17 stations for 9 years. The finest outcome was gained applying min-max normalized data and the input selection according to the correlation between DO and dependent parameters, which provided the most precise GRNN model, and in combination the smallest number of inputs: pH, temperature, HCO_3^- , sulfate, nitrate, hardness, Na, Cl^- , Conductivity, and Alkalinity.

Cordoba et al. (2014) studied chlorine concentration in a water distribution system by applying a combination of ANN and Monte-Carlo. They examined the model on one specific location applying the hydraulic and water quality parameters for example flow, pH, and temperature. The model permits projecting chlorine concentration at chosen nodes of the water supply system.

Šiljić et al.(2015) developed an initial general regression neural network (GRNN) using 20 inputs to predict biological oxygen demand. Then, they decreased inputs to 15 applying the Monte Carlo simulation technique as the input selection method. The best outcome was obtained with the GRNN model using 25 % fewer inputs than the initial model. Karimi et al. (2018) applied land use/land cover, water quality data, land suitability map, and geological map of catchments discharging to hydrometric stations situated in the western part of the Caspian Sea to develop Multiple Linear Regression Stepwise technique when the model uncertainty analysis was performed using Monte Carlo simulation. The consequences of the Monte Carlo simulation described that though some models had high R^2 , the possibility of negative outputs ($Pr < 0$) produced by the model was similarly high (only positive amounts make sense regarding the water quality variables). Furthermore, the visual assessment of the cumulative distribution function(CDF) curves implied that models of Ca^{2+} , Mg^{2+} and electrical conductivity(EC) had higher uncertainty in predictability compared to Na^+ . Godo-Pla et al.(2019) developed an ANN to forecast the potassium permanganate demand at the influent of a full-scale Treatment Plant of drinking Water. For forecasting water quality and uncertainty of ANN, model parameters estimation, uncertainty, and sensitivity analysis were performed. Monte Carlo simulation was used to determine uncertainty. The chosen architecture included a feed-forward multilayer perceptron(MLP) with four inputs and one node in the hidden layer with a sigmoid transfer function. They achieved suitable ANN with acceptable uncertainty.

Wang et al. (2019) used a fruit fly optimization algorithm (FFOA)-based water quality mechanism modeling technique and investigated a eutrophication risk assessment procedure of lakes and reservoirs based on the mechanism model. They estimated unknown water quality parameters using FFOA. After that, they applied Monte Carlo simulation to forecast water quality evolution and PDFs of water quality indicators at forecasted time indexes were gained to reach water quality forecasting. The simulation outcome indicated that the planned technique can effectively estimate the unknown parameters of the water quality mechanism model, and forecast water quality. Mitrović et al. (2019) developed single-output and concurrent ANNs for the spatial interpolation of 18 water quality parameters at single- and multi-inactive monitoring locations on the Danube River course through Serbia. The variable selection and sensitivity

analysis in the case of simultaneous models were carried out applying a modified method according to Monte Carlo Simulations (MCS). They proposed based on the performance of developed models (mean average percentage error b 10%): a concurrent multi-target model for pH, alkalinity, conductivity, hardness, dissolved oxygen, HCO_3^- , SO_4^{2-} , and Ca, single-output multi-target models for temperature and Cl^- , concurrent single-target models for Mg and CO_2 , single-output single target models for NO_3^- .

The Sodium Adsorption Rate(SAR) formulation is:

$$\text{SAR} = \text{Na}^+ / [(\text{Ca}^{2+} + \text{Mg}^{2+})/2]^{0.5}$$

Where Na^+ , Ca^{+2} , and Mg^{+2} are ions of sodium, calcium, and magnesium, respectively.

To describe the importance of SAR in agricultural purposes, some investigators have studied the consequence of SAR associate with calcium and magnesium (Abukarima and Al-Sulaiman 2018). Shainberg and Letely (1984) stated that rises in sodium levels will decrease the value of infiltration. Similarly, some other researches have shown that a rise of SAR reduced infiltration in the soil (El-Morsy,1991; Emdad et al.2004; Rahnama et al. 2020). Therefore, predicting the SAR of the water river is one of the significant in agriculture.

Considering mentioned literature review, the number of studies to analyze the uncertainty of ANFIS-FCMC using Monte Carlo analysis are few. Also, none of the studies are related to the uncertainty of predicting sodium adsorption rate(SAR) in three rivers, including Aras, Spid-Rud, and Karun River in Iran.

The objective of the present study was to determine the uncertainty of ANFIS-FCMC to forecast the Sodium Adsorption rate (SAR) of water in Aras, Sefid-Rud, and Karun River using Monte Carlo analysis.

1.1 Study area

Figure 1 illustrates the locations of three rivers and water quality monitoring stations. The three rivers are Aras, Sefid-Rud, and Karun. They are located in the north, northwest, southwest of Iran.

Figure 1 the locations of three rivers and water quality monitoring stations including Aras, Sefid-rud, and Karun.

The Aras river originates from Turkey, and it flows along the borders of Turkey, Armenia, the Nakhichevan area of Azerbaijan, and Iran. Finally, it flows to the Kura River in Azerbaijan. The total length of the river is 1,072 km with a basin of 102,000 km². Sefid-Rud River is the longest river in Iran, and it flows into the Caspian.

Its main tributary is the Qezel Owzan with 1,000 km long and drains 56,200 km². The mean annual water flow is 130 m³/s, but it is subject to significant fluctuation from 600 m³/s in the Spring when the melting of the snow and intensive precipitation. In Winter, it decreases to 70-80 m³/s. The mean temperature is between 11 degrees centigrade in January and 26 in July. Rainfall is around 1000 mm /year between Elburz mountain and the Caspian Sea, and precipitation in the southern part is between 200 and 400mm. The Karun River is situated in the southwestern of Iran covering an area of around 67500km². The longitude of the catchment area is 48° to 52° east and the latitude of the basin is 30° to 34° 15' north. The annual precipitation is nearby 550 mm (Keikhosravi et al. 2017).

2. Material and methods

Fars water authority, Iran, has collected water quality data of the three rivers. Table 1 displays the descriptive statistics of the monthly river's water quality variables for 15 years of the three rivers.

Table1 the statistical summary of the three river's water quality

Aras river	SAR	SO ₄ ²⁻ (mg/L)	Cl ⁻ (mg/L)	HCO ₃ ⁻ (mg/L)	pH	EC (μs/cm)	(m ³ /s)
Number	180	180	180	180	180	180	180
Maximum	3.72	10	7	6.50	9.10	2170	596.40
Minimum	0.98	0.55	0.4	1.50	6.70	414	0.00
Mean	2.36	3.04	3.111	4.89	7.97	1112.92	109.87
SD (σ)	0.48	1.25	0.97	0.99	0.38	240.04	99.23
Sefid-Rud river	SAR	SO ₄ ²⁻ (mg/L)	Cl ⁻ (mg/L)	HCO ₃ ⁻ (mg/L)	pH	EC (μs/cm)	Q (m ³ /s)
Number	180	180	180	180	180	180	180
Maximum	7.46	6.20	10.9	6.7	9.63	2018	6056
Minimum	0.14	0.20	0.15	1.3	0.74	247	1.31
Mean	2.31	2.00	4.22	3.31	7.62	1004.38	100.28
SD (σ)	1.29	1.10	2.49	0.90	0.66	399.97	458.43
Karon river	SAR	SO ₄ ²⁻ (mg/L)	Cl ⁻ (mg/L)	HCO ₃ ⁻ (mg/L)	pH	Ec (μs/cm)	Q (m ³ /s)
Number	180	180	180	180	180	180	180
Maximum	9.231	12.68	24.5	4.39	8.7	3570	2835
Minimum	1.69	1.45	2.58	0.64	6.7	740	53.09
Mean	4.78	5.85	9.89	3.00	7.92	1845.65	529.46
SD (σ)	1.51	2.32	3.97	0.63	0.36	577.147	497.34

2.1 Adaptive network-based fuzzy inference system (ANFIS)

The principal of the development of the ANFIS is Takagi-Sugeno fuzzy inference system (FIS) (Güler and Übeyli 2005). The fuzzy system, working based on the membership function, is an effective method

for problem-solving, modeling, data mining, and decreasing the intricacy of data (Jang 1993). The ANFIS applies both the ANN and FIS, which have a successful learning algorithm and a high training rate. Consequently, both nonlinear and complex systems can accurately be modeled using ANFIS (Heddami et al. 2012; Salahi et al. 2017; Asadi et al. 2020). we built the FIS structure using only fuzzy c-means clustering (FCMC), which is a clustering method using the Gaussian kind of membership function (Sivanandam et al. 2007). The ANFIS-FCMC, *Genfis3*, function is available in MATLAB (2018a) defining FCMC method.

Function Genfis3

The *Genfis3* extracts a set of regulations, which simulates data performance. The function requires different sets of input and output data as input arguments. Equation 1 presents a sample command-line for the FCMC (The Math Works 2018):

$$fismat = Genfis3 (datin, datout, type, cluster_n) \quad (1)$$

Where *Genfis3* is a function that builds a fuzzy system due to FCMC; the *type* is either ‘Sugeno’ or ‘Mamdai’, and *cluster_n* is the number of clusters. The Sugeno systems are mathematically well-organized and suitable for math analysis (The Math Works, 2018); consequently, they were used in the present study.

We examined two fuzzies if-then rules based on a first-order Sugeno model to show the ANFIS architecture:

$$Rule\ 1.\ If\ (x\ is\ A_1)\ and\ (y\ is\ B_1)\ then\ (f1=p_1x+q_1y+r_1) \quad (2)$$

$$Rule\ 2.\ If\ (x\ is\ A_2)\ and\ (y\ is\ B_2)\ then\ (f2=p_2x+q_2y+r_2) \quad (3)$$

Where x and y is the inputs. A_i and B_i are the fuzzy sets; $f1$ and $f2$ are the outputs within the fuzzy region quantified by the fuzzy rule; and p_1, q_1, r_1, p_2, q_2 , and r_2 are the design parameters that are gained through the training process.

Five distinct layers, which are used to form an ANFIS structure are available (Jang 1993; Wali et al. 2012). A detailed description of ANFIS is accessible in other studies (Yetilmezsoy 2010a). We explained ANFIS shortly which is: The first layer covers input nodes, which present the inputs to the ANFIS model. Each node is an adaptive node indicating a membership grade of a linguistic label. The first layer output suit the input of layer two. The second layer is a fixed node, which carries a prior quantity of MFs, which are selected based on the input value. The nodes in layer two calculate the fuzzy rules and create them for layer three with a suitable degree of activity. The third layer performances are to normalize the degree of activity for all rules. The defuzzification is the fourth layer. The fifth layer comprises the output nodes and computes the sum of all outputs of each rule getting from the preceding layer. Figure 2 describes a schematic flow diagram of ANFIS and MCS for the uncertainty of model predicting of the SAR.

2.2 Model selection criteria

To obtain the usability of the ANFIS, two statistical norms were examined, i.e., root means square error (RMSE) and mean bias error (MBE). The MBE indicates if the model forecasting in the training process overestimate (MBE>0) or underestimates (MBE<0). The MBE equal to zero is the best score. Equations (4) and (5) describes RMSE and MBE (Willmott and Matsuura 2005).

$$RMSE = \sqrt{n^{-1} \sum_{i=1}^n (P_i - O_i)^2} \quad (4)$$

$$MBE = n^{-1} \sum_{i=1}^n P_i - O_i \quad (5)$$

where O_i and P_i are the observed(measured) and the predicted(forecasted) amount of the parameters respectively, the number of data indicated by n.

2.3 Model Efficiency

To confirm the performance of the RBF neural network and the ANFIS, we calculated the coefficient of determination (R^2) according to Equations (6) to (8).

$$R^2 = \left\{ \frac{\sum(O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum(O_i - \bar{O})^2 * \sum(P_i - \bar{P})^2}} \right\}^2 \quad (6)$$

Where, O_i and P_i = the observed (measured) and the predicted data of the parameters, respectively. $\bar{O}$ and $\bar{P}$ = the mean of observed and the projected data. Also, we calculated the index of agreement (IA) to determine how close the predicted data was to the observed data.

$$IA = 1 - \frac{\sum(P_i - O_i)^2}{||P_i - \bar{O}|| + ||O_i - \bar{O}||^2} \quad (7)$$

The index of agreement (IA) varies from 0.0 to 1.0.

The efficiency E, suggested by Nash and Sutcliffe (1970), is obtained as one minus the sum of the absolute squared differences between the projected data and observed data normalized by the variance of the observed amounts in the period under study. The variety of E is between 1.0 (perfect fit) and $-\infty$. Equation 12 describes E (Krause 2005).

$$E = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (8)$$

Rahnama et al. (2020) developed a different ANN model, including the Radial base function(RBF) and Adaptive neuro-fuzzy inference system(ANFIS) to predict SAR of water of four rivers including Aras, Sepid-Rud River, Karun, and Mond. Input variables were SO_4^{2-} (mg/L), Cl^{-1} (mg/L), HCO_3^{-} (mg/L), pH, EC ($\mu\text{s}/\text{cm}$) and Discharge (m^3/s). The number of input data was 144 for training and 36 for testing the RBF and ANFIS models. In this study, the uncertainty of one of the ANFIS neural networks, , genfis3, is analyzed for the SAR of the first three rivers.

2.4 Monti Carlo Analysis(MCA)

Our focus in this study was on evaluating the uncertainty of ANFIS neural network models by using Monte Carlo analysis (MCA). The Anderson–Darling (AD) test (Goktepe et al. 2008) was applied to evaluate the goodness-of-fit and find proper probability density functions (PDF) to each input variable with a 5% significance level.

First, different probability density functions (PDFs) with a 5% significance (such as normal, Logistic, Log logistic, and 3-Parameter Weibull distributions) were allocated to each input parameter using Minitab 18 software. Then, about 1000 random data was generated by MCS for each input parameter including SO_4^{2-} (mg/L), Cl^{-1} (mg/L), HCO_3^{-} (mg/L), pH, EC ($\mu\text{s}/\text{cm}$) and Discharge (m^3/s). After that, we developed ANFIS (genfis3) neural network using the generated 1000 data for each input parameter using MATLAB (2018). Furthermore, we computed the PDFs of the output of genfis3 using input data generated by MCA. Figure 2 illustrates ANFIS-genfis3 and MCA ANFIS-genfis3 of SAR prediction in the rivers.

Figure 2 ANFIS-genfis3 and MCA ANFIS-genfis3 of SAR prediction in the rivers.

The mean, median, variance, and percentiles are computed to guarantee there will be about the same properties if more input and output data are randomly generated and evaluated (Dehghani et al., 2014; Goktepe et al., 2008, Asadi et al. 2020). Generally, a short, wide, and irregular output probability distribution explains a higher uncertainty degree of the model (Arhami et al., 2013; Gao et al., 2018; Asadi et al. 2020).

3. Results

ANFIS-genfis3

Table 2 indicates the results ANFIS- genfis3 for predicting SAR of water in Aras, Sefid-Rud, and Karun Rivers.

Table 2 results of ANFIS (genfis3) neural network uncertainty in predicting SAR of water of Aras River (Rahnama et al. 2020).

Rivers		ANFIS-FCMC					
		MAE	MSE	MBE	R ²	IA	E
		Training			Testing(Model performance)		
Aras	Genfis3	0.19	0.068	1.2E-06	0.61	0.87	0.72
Sefid-Rud	Genfis3	0.10	0.038	0.001	0.88	0.95	0.83
Karun	Genfis3	0.24	0.098	8.3E-06	0.72	0.94	0.79

As indicates in Table 2, training of ANFIS-FCMC(genfis3) for the SAR of the three rivers is acceptable because MBE approaches zero, which shows the adequacy of the training models. However, considering training of genfis3 for SAR forecasting of the three rivers, MBE of Aras is equal to 1.2E-06, which is less than MBE of Spid-Rud and Karun Rivers. It means the model adequacy of training of Aras water is higher than two other rivers. R² and IA is also acceptable for the models' performance of SAR forecasting of the three rivers. However, the model performance of SAR prediction of the Spid-Rud river (R²=0.88) is higher than the two rivers. The main objective of the present work was to analyze the uncertainty of ANFIS-FCMC(genfis3) of the three mentioned rivers.

3.1 Aras River results

Results of ANFIS (genfis3) neural network uncertainty in predicting SAR of water of Aras River

Table 3 indicates the fitting of different probability density function(PDF) distributions on sulfate, chloride, bicarbonate, pH, and flow rate of 180 inputs data to the genfis3 neural network of Aras River water. Results were tested with Anderson-Darling and p-value.

Table 3 fitting of different probability density function(PDF) distributions on sulfate, chloride, bicarbonate, pH, and discharge of 180 inputs data to the genfis3 neural network of Aras River water using

Anderson-Darling test

Parameters	PDF	AD test	p-value
SO ₄ ²⁻ (mg/L)	Log logistic	0.79	0.022
Cl ⁻¹ (mg/L)	3-Parameter Weibull	1.73	<0.05
HCO ₃ ⁻ (mg/L)	3-Parameter Weibull	1.27	<0.05
pH	Normal	0.57	0.135
EC (μs/cm)	Logistic	1.64	<0.05
Discharge (m ³ /s)	3-Parameter Weibull	0.58	0.136

As described in Table 3, different distributions were fitted to the water quality data of the river. The results of this table were the outcome of a test of 15 different PDF distributions by Minitab 18 software. Some results reached a p-value of less than 0.05, which is not a statistically well fit. However, the data collected and analyzed by the Iran water authority, and the authors of the article were not played a role in the uncertainty of water quality sampling and the dispersion of data, and the number of samples. Indeed, the best possible PDF fit on the available data was carried out in this study. This explanation is true for the input parameters to ANFIS (genfis 3) of the Sepid-Rud and Karun rivers. Antanasijevic et al. (2014) stated the same our justification on determining the uncertainty of the Total Dissolved Solid(TDS) of the Danube River using Monte Carlo simulations on neural network.

Figure3 presents the fit of different PDFs on the input parameters data of the ANFIS(genfis3) neural network for Aras River

Figure3 fitting different PDF to the input parameters data of the ANFIS(genfis3) neural network for Aras River

Table 4 describes the results of the Monte Carlo method uncertainty on the results of genfis3 neural network of SAR of Aras river water.

Table4 the results of Monte Carlo simulation(MCS) uncertainty on the results of ANFIS(genfis3)
neural network of SAR of water prediction in Aras river

Descriptive Statistics	ANFIS-GENFIS3	MC
N	180	952
Mean	2.35	2.41
Standard Deviation(SD)	0.51	0.58
2.5 th percentile of the model output distribution	1.11	1.31
97.5 th percentile of the model output distribution	3.19	3.78

Figure 4 indicates the tolerance interval of the SAR of water with a 95% confidence interval and minimum coverage of 95% of the population of Aras River using MCS. According to Figure4 which is almost a normal fit. The combined standard uncertainty simulated for SAR is equal to 0.58. The expanded uncertainty is equal to two times of standard deviations with a 95% confidence interval (Farrance and Frenkel, 2014). Therefore, the expanded uncertainty for SAR of Aras water river is equal to 1.16, and the gap is equal to 2.412 ± 1.1622 .

Figure 4 the tolerance interval of the SAR of water with 95% confidence interval and minimum coverage of 95% of the population of Aras River using MCA

3.2 Sepid-Rud River results

Table 5 illustrates the fitting of different probably density function(PDF) distributions on sulfate, chloride, bicarbonate, pH, and discharge of 180 inputs data to the genfis3 neural network of Sepid-Rud River water using Anderson-Darling and p-value tests.

Table 5 the results of fitting of different PDF distributions to sulfate, chloride, bicarbonate, pH, and discharge of input data to genfis3 neural network of Sepid-Rud River using Anderson-Darling test

Parameters	PDF	AD test	p-value
SO ₄ ²⁻ (mg/L)	Weibull	0.60	0.125
Cl ⁻¹ (mg/L)	Weibull	1.17	<0.05
HCO ₃ ⁻ (mg/L)	Normal	0.97	0.014
pH	Weibull	0.27	0.250
EC (μs/cm)	Normal	1.43	<0.05
Discharge (m ³ /s)	2-Parameter	1.45	0.033
	Exponential		

Figure5 presents the fit of different PDF to the input parameters data (SO_4^{2-} (mg/L)

, Cl^- (mg/L), HCO_3^- (mg/L), pH, EC ($\mu\text{S}/\text{cm}$), and Discharge (m^3/s)) of the ANFIS(genfis3) neural network for Sepid-Rud River

Figure5 fitting different PDF distributions to the input parameters data of the ANFIS(genfis3) neural network for Sepid-Rud River

Table 6 the results of Monte Carlo simulation(MCS) uncertainty on the results of ANFIS(genfis3)neural network of SAR of water prediction in Sepid-Rud River

Descriptive Statistics	ANFIS-GENFIS3	MC
N	180	892
Mean	2.26	2.23
SD	1.24	1.11
2.5 th percentile of the model output distribution	0.09	0.61
97.5 th percentile of the model output distribution	5.95	5.48

Figure 6 illustrates the tolerance interval of the SAR of water with a 95% confidence interval and minimum coverage of 95% of the population of Sepid-Rud River using MCS. According to Figure 6 which is approximately a normal fit. The combined standard uncertainty simulated for SAR is equal to 1.11. The expanded uncertainty is equal to two times of standard deviations with a 95% confidence interval (Farrance and Frenkel, 2014). Therefore, the expanded uncertainty for SAR of Sepid-Rud River water is equal to 2.22, and the gap is equal to 2.235 ± 2.22 . When we compare the uncertainty of SAR in Aras and Sepid-Rud River water, it is clear that Sepid-Rud River has higher uncertainty than Aras.

Figure 6 the tolerance interval of the SAR of water with 95% confidence interval and minimum coverage of 95% of the population of Sepid-Rud River using MCS

3.3 Karun River results

Table 7 describes the fitting of different probably density function(PDF) distributions on sulfate, chloride, bicarbonate, pH, and discharge of 180 inputs data to the genfis3 neural network of Karun River water using Anderson-Darling and p-value tests.

Table 7 the results of fitting of different PDF distributions to sulfate, chloride, bicarbonate, pH, and discharge of input data to genfis3 neural network of Karun River using Anderson-Darling test

Parameters	PDF	AD test	p-value
SO_4^{2-} (mg/L)	Normal	0.43	0.311
Cl^{-1} (mg/L)	Normal	0.86	0.027
HCO_3^{-} (mg/L)	Normal	0.50	0.207
pH	Weibull	0.84	0.012
EC ($\mu\text{s}/\text{cm}$)	Normal	0.72	0.061
Discharge (m^3/s)	Lognormal	0.77	0.044

Figure5 illustrates the fit of different PDF to the input parameters data (SO_4^{2-} (mg/L)

, Cl^- (mg/L), HCO_3^- (mg/L), pH, EC ($\mu\text{s}/\text{cm}$), and Discharge (m^3/s) of the ANFIS(genfis3) neural network for Karun River

Figure7 fitting different PDF distribution to the input parameters data of the ANFIS(genfis3) neural network for Karun River

Table 8 shows the results of Monte Carlo uncertainty analysis of genfis3 neural network in predicting SAR of water on Karun river. As indicated in Table 8, the standard deviation (SD) of MCA is higher than ANFIS-FCMC (genfis 3).

Table 8 the results of Monte Carlo simulation(MCS) uncertainty on the results of genfis3 neural network of SAR of water prediction in Karun River

Descriptive Statistics	ANFIS-GENFIS3	MC
N	180	950
Mean	4.82	4.79
SD	1.45	2.06
2.5 th percentile of the model output distribution	1.97	0.87
97.5 th percentile of the model output distribution	8.29	10.1

Figure8 describes the tolerance interval of the SAR of water with a 95% confidence interval and minimum coverage of 95% of the population of Karun River using MCS.

Figure 8 the tolerance interval of the SAR of water with 95% confidence interval and minimum coverage of 95% of the population of Karun River using MCS

Figure 8 describes the tolerance interval of the SAR of water with 95% confidence interval and minimum coverage of 95% of the population of Karun River using MCS. Based on Figure 6, which is roughly a normal fit. The combined standard uncertainty simulated for SAR is equal to 2.063. The expanded uncertainty for SAR of Sepid-Rud River water is equal to 4.126, and the gap is equal to 4.79 ± 4.126 . Table 9 describes the difference between the mean and standard deviation of ANFIS(Grnfis3) models and Monte Carlo analysis.

When we compare the uncertainty of SAR in Karun and Sepid-Rud River water, it is clear that Karun River models (Genfis 3) have higher uncertainty than Aras in predicting SAR (Table 9). The uncertainty of SAR in Aras River water is the lowest compared to Sfid-Rud River and Karun River. Uncertainty emerges from confirmation of model and water quality data analysis. Considering the uncertainty of SAR of Aras River water (Table 9), which is low. It may state model configuration (ANFIS-FCMC; genfis 3)

and data measurement was precise. However, in Karun River, the SAR model from genfis 3 has higher uncertainty than Sifid-Rud River and Aras River, which come from model configuration and water quality collection.

Coulthard et al. (2013) stated, “uncertainty in numerical models has many origins: input data, model simplifications, algorithm structure, calibration process, calibration and validation data, as well as equifinality.” However, according to the study of Kamali et al. (2017), most uncertainties emerge from input data, and less uncertainty belongs to model configuration. The input data was collected by Fars Water Authority in Iran.

Table 9 different between the mean and standard deviation of ANFIS(Grnfis3) models and Monte Carlo analysis

Name of River	Mean of genfis3	Mean of MCA	The standard deviation of genfis 3	The standard deviation of MCA
Aras River	2.35	2.41	0.51	0.58
Sefid-Rud River	2.26	2.23	1.24	1.1
Karun River	4.82	4.79	1.45	2.06

Considering the literature review in present wor, none of the uncertainty studies predict SAR of water in Rivers using ANFIS-FCMC and MCA. Therefore, there was no opportunely to confirm the results with other researches.

4. Conclusions

Considering using Monti Carlo analysis(MCA) to evaluate the uncertainty of ANFIS-FCMC(genfis3) model for predicting Sodium Adsorption Rate(SAR) of water of Aras, Sepid-Rud, and Karun Rivers, the following conclusion can be reached.

- 1- The ANFIS-FCMC (genfis3) model performance of SAR prediction of water of the Spid-Rud River ($R^2=0.88$) is higher than the Aras and Karun Rivers.
- 2- The combined standard and the expanded uncertainty simulated for SAR of Aras River water using *ANFIS-FCMC* (genfis3) and MCS is equal to 0.58 and 1.16, respectively, and the gap is equal to 2.412 ± 1.1622 .
- 3- The combined standard and the expanded uncertainty simulated for SAR of Spid-Rud River water applying *ANFIS-FCMC* (genfis3) and MCA is equal to 1.11 and 2.22, respectively, and the gap is equal to 2.235 ± 2.22 .
- 4- The combined standard and the expanded uncertainty simulated for SAR of Aras River water using *ANFIS-FCMC*(genfis3) and MCA is equal to 2.063 and 4.126, respectively, and the gap is equal to 4.79 ± 4.126 .
- 5- The least uncertainty (using MCA) belongs to forecasting of SAR of Aras River using ANFIS-FCMC (genfis3) model and highest SAR predicting uncertainty be owned by Karun River.

Conflict of interest statement

We certify that we have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

5. References

Aboukarima A M, Al-Sulaiman MA (2018) Effect of sodium adsorption ratio and electrical conductivity of the applied water infiltration in a sandy-loam soil, *Water SA*,44(2018)1,105-110

Antanasijevic', D. Pocajt, V Peric'-Grujic', A. Ristic', M.(2014) Modelling of dissolved oxygen in the Danube River using artificial neural networks and Monte Carlo Simulation uncertainty analysis, *Journal of Hydrology* 519, 1895–1907.

Arhami, M., Kamali, N., Rajabi, M.M., 2013. Predicting hourly air pollutant levels using artificial neural networks coupled with uncertainty analysis by Monte Carlo simulations. *Environ. Sci. Pollut. Res.* 20, 4777–4789.

Coulthard, T. J. Van De Wiel ,M. J. (2013) *Treatise on Geomorphology*, ISBN:978-0-08-088522-3

Cuesta Cordoba, G.A. Tuhovčáka, L. Tauša, M.(2014) Using artificial neural network models to assess water quality in water distribution networks, *Procedia Engineering* 70, 399 – 408, DOI: 10.1016/j.proeng.2014.02.045.

Emdad, M R, Raine S R, Smith R J, Fardad H (2004) Effect of water quality on soil structure and infiltration under furrow irrigation, *Irrigation Sci.*, 23:55-60.DOI: 10.1007/s00271-004-0093-y

Dehhaghi, M., Tabatabaei, M., Aghbashlo, M., Panahi, H.K.S., Nizami, A.-S., 2019. A state-of-the-art review on the application of nanomaterials for enhancing biogas production. *J. Environ. Manag.* 251, 109597.

El-Morsy EA, Malik M, Letey J (1991) Interaction between water quality and polymer treated on infiltration rate and clay migrations, *Soil techno.* ,493: 221-231. DOI:10.1016/0933-3630(91)90003-6

Farrance, I. Frenkel, R. (2014) *Uncertainty in Measurement: A Review of Monte Carlo Simulation Using Microsoft Excel for the Calculation of Uncertainties Through Functional Relationships, Including Uncertainties in Empirically Derived Constants*, *Clin Biochem Rev* 35 (1) ,37-61.

Gao, M., Yin, L., Ning, J., 2018. Artificial neural network model for ozone concentration estimation and Monte Carlo analysis. *Atmos. Environ.* 184, 129–139.

- Godo-Pla, L., Emiliano, P., Valero, F., Poch, M., Sin, G., Monclús, H. Predicting the oxidant demand in full-scale drinking water treatment using an artificial neural network: Uncertainty and sensitivity analysis, *Process Safety and Environmental Protection*, 125, 317-327.
- Goktepe, A. B., Altun, S., Altintas, G., & Tan, O. (2008). Shear strength estimation of plastic clays with statistical and neural approaches. *Building and Environment*, 43(5), 849-860.
- Übeyli E D (2005) Adaptive neuro-fuzzy inference system for classification of EEG signals Güler I, using wavelet coefficients. *J. Neurosci. Methods* 148: 113–121. DOI: 10.1016/j.jneumeth.2005.04.013
- Heddami S, Bermad A, Dechemi N (2012) ANFIS-based modeling for coagulant dosage in drinking water treatment plant: a case study. *Environ. Monit. Assess.* 184(4): 1953–1971. DOI: 10.1007/s10661-011-2091-x
- Jang J-S (1993) ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man. Cybern.* 23: 665–685. DOI: 10.1109/21.256541
- Jiang, Y. Nan, Z. Yang, S. (2013) Risk assessment of water quality using Monte Carlo simulation and artificial neural network method, *Journal of Environmental Management*, 122, 130-136.
- Kamali, B., Abbaspour, K.C. and Hong Yang, H. (2017) Assessing the Uncertainty of Multiple Input Datasets in the Prediction of Water Resource Components Water, DOI: 10.3390/w9090709
- Karimi, Sh.2, Jabbarian Amiri, B., Malekian, A. (2018) Modeling the Impact of Watershed Physical Attributes on Surface Water Quality and Uncertainty Assessment of the Models Using the Monte Carlo Simulation, *Iran-Water Resources Research*, 14(3), 304-317.
- Keikhosravi M S, Masoodian, S A, Balling, Jx, R.C. Svoma, B.M (2017) Spatial and temporal variations of snow cover in the Karun River Basin, Iran, 2003-2015, *Water*, 9, 965:1-16. DOI:10.3390/w9120965

Keikhosravi M S, Masoodian, S A Balling, Jx, R.C. Svoma, B.M (2017) Spatial and temporal variations of snow cover in the Karun River Basin, Iran,2003-2015, *Water*,9, 965:1-16.

DOI:10.3390/w9120965

Krause P, Boyle D, Base F (2005) Comparison of different efficiency criteria for hydrological model assessment, *Advances in Geosciences J.* 5:89–97.DOI:10.5194/adgeo-5-89-2005

Linkage Models Using Monte Carlo Method *Int. J. Environ. Res.*, 4(1):121-130.

Mitrović, T. Antanasijević, D. Lazović, S. Perić-Grujić, A. Ristić, M.(2019) Virtual water quality monitoring at inactive monitoring sites using Monte Carlo optimized artificial neural networks: A case study of Danube River (Serbia), *Science of the Total Environment*, 654 (2019) 1000–1009,

<https://doi.org/10.1016/j.scitotenv.2018.11.189>.

Nakane, K. Haidary, A. Sensitivity Analysis of Stream Water Quality and Land Cover

Nash J E, Sutcliffe J V (1970) River flow forecasting through conceptual models Part I-A discussion of principles *J. Hydro.* 27 (3): 282-290. DOI:10.1016/0022-1694(70)90255-6.

Shainberg I, Lete J (1984) The response of soils to sodic and saline conditions, *Hilgardia*,52(2) :2 1-57

DOI:10.3733/hilg. v52n02p057

Šiljić, A. Antanasijević, D. Perić-Grujić, A. Ristić, M. Pocajt, V.(2015) Artificial neural network modeling of biological oxygen demand in rivers at the national level with input selection based on Monte Carlo simulations, *Environmental Science and Pollution Research* **22**, 4230–4241.

The Math Works, I., 2018. Fuzzy Logic Toolbox User’s Guide. Natick, United States. Tufaner, F., Demirci, Y., 2020. Prediction of biogas production rate from the anaerobic hybrid reactor by artificial neural network and nonlinear regressions models. *Clean Technol. Environ. Policy* 1–12

Wang, X. Zhou, Y. Zhao, Z. Li Wang, L. Xu, J. Yu, J. (2019) A novel water quality mechanism modeling and eutrophication risk assessment method of lakes and reservoirs *Nonlinear Dyn.* 96,1037–1053,<https://doi.org/10.1007/s11071-019-04837-6>

Willmott C, Matsuura K (2005) Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance., *Climate Research J.*, 30: 79-82.DOI: 10.3354/cr030079

Zou, R. Lung, W-S. Guol, H.(2002) Neural Network Embedded Monte Carlo Approach for Water Quality Modeling under Input Information Uncertainty *Journal of Computing in Civil Engineering*,16,2 [https://doi.org/10.1061/\(ASCE\)0887-3801\(2002\)16:2\(135\)](https://doi.org/10.1061/(ASCE)0887-3801(2002)16:2(135))