Research Article

Performance assessment of data driven water models using water quality parameters of Wangchu river, Bhutan

Yeshi Choden[1] · Sonam Chokden[1] · Tenzin Rabten[1] · Nimesh Chhetri[1,2] · Komal Raj Aryan[3] · Khameis Mohamed Al Abdouli[3]

Received: 24 June 2022 / Accepted: 27 September 2022
Published online: 04 October 2022

Abstract

Multifarious anthropogenic activities triggered by rapid urbanization has led to contamination of water sources at unprecedented rate, with less surveillance, investigation and mitigation. The use of artificial intelligence (AI) in tracking and predicting water quality parameters has surpassed the use of other conventional methods. This study presents the assessment of three main models: adaptive neuro fuzzy inference system (ANFIS), artificial neural network (ANN) and multiple linear regression (MLR) on water quality parameters of Wangchu river located at capital city of Bhutan. The performance and predictive ability of these models are compared and the optimal model for predicting the parameters are recommended based on the coefficient correlation (CC), root mean square error (RMSE) and Nash–Sutcliffe efficiency (NSE) evaluation criteria. Overall NSE and RMSE, the ANN model predicted parameters with maximum efficiency of 97.3 percent and minimum error of 8.57. The efficiency of MLR and ANFIS model are 95.9 percent and 94.1 percent respectively. The overall error generated by MLR and ANFIS are 10.64 and 12.693 respectively. From the analysis made, the ANN is recommended as the most suitable model in predicting the water quality parameters of Wangchu river. From the six-training function of ANN, trainBR (Bayesian Regularization) achieved the CC of 99.8%, NSE of 99.3% and RMSE of 9.822 for next year data prediction. For next location prediction, trainBR achieved CC of 99.2%, NSE of 98.4% and RMSE of 6.485, which is the higher correlation and maximum efficiency with less error compared to rest of the training functions. The study represents first attempt in assessing water quality using AI technology in Bhutan and the results showed a positive conclusion that the traditional means of experiments to check the quality of river water can be substituted with this reliable and realistic data driven water models.

Article highlights

- Total dissolved solids (TDS), electrical conductivity (EC), potential of hydrogen (pH) and dissolved oxygen (DO) are selected as main water quality parameters as data for modeling.
- Artificial neural network model gives highest efficiency and accuracy compared to MLR and ANFIS model.
- Use of artificial intelligence shows better performance to provide water quality and future predictions over conventional methods leading to conservation of water resources and sustainability.

Keywords Wangchu river · Water quality parameters · Data-driven models · Artificial intelligence

✉ Nimesh Chhetri, nimeshchhetri.cst@rub.edu.bt | [1]Civil Engineering and Geology Department, College of Science and Technology, Royal University of Bhutan, Phuentsholing, Bhutan. [2]Centre of Disaster Risk Reduction and Community Development Studies, Royal University of Bhutan, Thimphu, Bhutan. [3]Faculty of Resilience, Rabdan Academy, Abu Dhabi, United Arab Emirates.

1 Introduction

Human health and ecosystems depend on how well water resources are preserved. With the exponential growth in industrialization and continuous increase in population, there is a huge demand and pressure towards water resources [1]. Water quality in rivers is deteriorating due unmanaged disposal of industrial, medical and municipal sewage wastes, and agricultural runoff etc. Due to anthropogenic activities, studies show that water resources even in mountainous area have high content of pollutants such as calcium (Ca), magnesium (Mg), copper (Cu), methane (NH4), nitrogen dioxide (NO2), turbidity (NTU), chemical oxygen demand (COD), total solids, microbes and bacteria [2, 3]. Therefore, it has become vital to assess and simulate water quality and its parameters to ascertain the suitability for various uses [4]. Identifying the water's quality parameter level ensures its suitability for a variety of applications, such as irrigation, drinking and cooking, hydropower generation, and recreational activities. Hence, suitable mitigation measures can be timely implemented to avoid the deterioration of water quality.

Some important indicators of water quality that need to be considered in the current research are the electrical conductivity (EC) and the total dissolved solids (TDS). These parameters if present in higher concentration, it is considered undesirable for consuming [5, 6]. Also, direct assessments of parameters such as EC and TDS are considered time consuming and costly. Therefore, suitable, cost-effective, time saving efficient and consistent methods are desirable for their assessments and predictions [7, 8]. Although various other feasible and important water quality parameters are there that should be evaluated such as dissolved oxygen (DO), pH and biochemical oxygen demand (BOD), although these parameters are fundamentally affected by EC and TDS [9, 10]. Recently, in context related scientific community, the use of mathematical model [11] and data-driven models, such as ANFIS, ANNs and MLR, have become practical alternatives in most studies. In most of the water related studies, artificial intelligence (AI) has been used and found out useful in water modeling and management [10, 12]. Also, helping in making make better decisions while enhancing service delivery and reducing costs [13].

In Abu Ziriq of Iraq, the study was made using different types of artificial intelligence techniques to calculate and predict TDS and EC. Amongst all, ANFIS model outperformed the prediction giving the best fit with the observed data compared to other models [14]. Literatures shows that use of AI models has resulted to more precise results and substantial in resolving the model simulation and prediction of nonlinear interface [15]. The physicochemical test on various sources of water samples suggested that the assessment of water quality parameters as well as conservation management should be carried out periodically to protect the water resources [16].Wen et al., estimated the DO values of Heihe River in northwestern China by developing ANN model. The performance of ANN model observed accurate to estimate DO concentrations [17]. Monstaseri et al. [18] used same model with success in predicting TDS at Iran water resources over a stretch of 20 years. Ay and Kişi [19] used ANN and ANFIS to estimate DO concentration, which was compared with the multiple linear regressions. The models are compared among one other and results indicated that the ANN model was close to accuracy to determine monthly mean DO concentration, thus making artificial intelligence suitable to study water resources [20].

The rapid urbanization, infrastructure development and increased rural urban migration has aggravated the quality of water resources in Bhutan. Moreover, technological methods to model, predict and forecast the quality parameters of water in Bhutan remain unexploited in Bhutan [21]. This study aims to determine the best fit model to assess the water quality parameters in Wangchu river which is located in the capital city of the country. Thus, ANFIS, ANN and MLR models were selected for simulation of water quality parameters including pH, DO, TDS and EC. The CC, RMSE and NSE are determined to see the performance of the models, performing experiment to compare the modeled output with the experimental data and to recommend the suitable model and predict the water quality parameter for the year 2022 and location. Therefore, such advancement of a methodology considering fewer parameters but giving a practical result with higher percentage of accuracy reduces the cost of water quality monitoring. Thus, the machine learning methods in predicting water quality has resulted efficient choice for water planner to improve sustainable management of water resources [22].

The main features of this study consist of six main sections, starting with an introduction which provides detailed literature and specifies the aim and objectives of the study. Section two explains the primary and secondary data considered to carry out the research work and methodology adopted describing various AI models associated regarding water quality modeling. Section three introduces the study area, Wangchu river of Thimphu, Bhutan and water sampling points has been located and presented for the experimental process. Section four describes the criteria of evaluation of selected AI model in detail, in terms of CC, NSE and RMSE. Section five presents all the performance results of ANFIS, MLR and ANN model on selected water quality parameters and best model with higher efficiency with less percentage error is selected for modeling water quality of Wangchu river. Finally, study concludes presenting the significance of how AI models can be used as a reliable and efficient method for assessing water quality and also predict future pollution.

2 Materials and methods

Three different models named ANFIS, MLR and ANN were selected as most suitable model through various literatures for assessing water quality parameter and its prediction. Their performance was analyzed based on assessment criteria such as CC, RMSE and NSC using the water quality parameters of Wangchuk River, Bhutan. The water quality data for the analysis was gathered in two ways. The primary data was collected from the National Environment Commission of Bhutan. Total dissolved solids (TDS), electrical conductivity (EC), potential of hydrogen (pH), temperature, and dissolved oxygen (DO) are some of the parameters measured in the Wangchu River based on standard procedures. The secondary data is the experimental data collected at the same time of year and in the same place. Water samples were taken from five different places along the river's length. After performance assessment of the selected three models which is evaluated based on CC, NSE and RMSE, the most suitable model is recommended. Validation of models is performed using experimental data and next the selected model is trained with its various training functions for prediction of next year and location where the results showed higher CC, RMSE with less error. Finally using the results of models, the condition of Wangchu river quality is ascertained and mitigation measures for maintaining the required quality of water for various purposes are suggested. The details of water model and study are as follow:

2.1 Adaptive neuro fuzzy inference system model

ANFIS is an integrated multilayer feed advancing network that uses neural network algorithms and fuzzy logic to an input data to an output data and this system can be used to predict and model any kinds of input–output data series other than water quality parameters [23–26].

Figure 1 displays a typical characteristic ANFIS structure. Every node in each has distinct role. Layer 1 is an adaptive node with a node function where Gaussian membership function is implemented during analyzation. Layer 2 signifies the strength of each rule. Layer 3 is a fixed node which represents the normalized strength of each rule. Layer 4 is an adaptive node with a node function. Layer 5 is a fixed node which is labeled as Σ, representing the overall output (*F*) as the total of all received inward signals.

2.2 Multiple linear regression model

It is a numerical procedure that considers many variables to forecast the possible result of response variable (refer Fig. 2). It involves more than two variable and depending upon which parameter to be determined, the dependent and independent variable can vary. The idea of regression was first coined in nineteenth century by Francis Galton and since the model can tackle more than two variables, it is term as multiple linear regression [24].

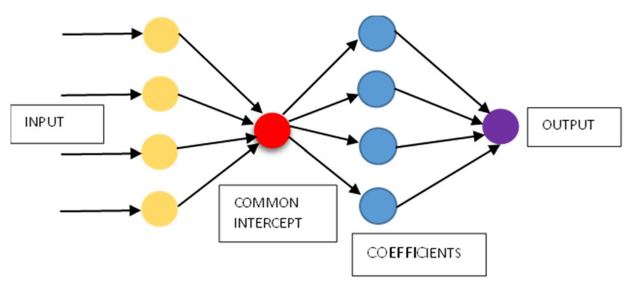


Fig. 2 Typical architecture of MLR model

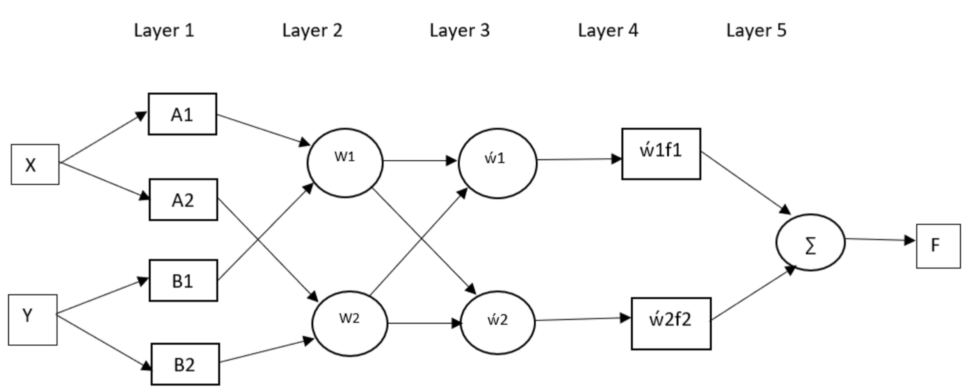


Fig. 1 Typical architecture of ANFIS model

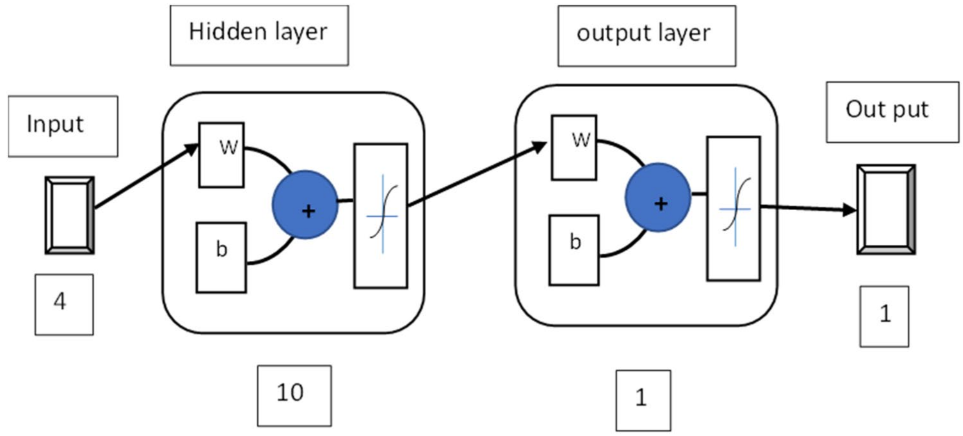


Fig. 3 Typical architecture of ANN model

2.3 Artificial neural network (ANN model)

The ANN model as shown in Fig. 3 is based on human brain which has an ability such as immense parallelism, dispersed illustration and computation, learning and simplification ability, adaptivity, data processing, liability tolerance, and low energy consumption. A neural network comprises of an interrelated group of artificial neurons, and it processes information using a connectionist method to computation [27–29]. ANN mechanically learns the concept from examples which makes them stimulating instead of following rules made by expertise proving its major advantage over traditional expert systems [30, 31].

3 Study area and data

3.1 Wangchuk river

Bhutan has the four main largest rivers named Manas et al. [32]. River basin of Bhutan makes around area of 580,000 km^2, out of which only 8% lies in Bhutan, rest lies under China (50%), India (34%) and Bangladesh (8%) [33]. The site selected for the study is Wangchu river, Thimphu (Fig. 4a). The Wangchu River originates in the high Himalayan glaciers, flows through the country's capital, and eventually flows into India's great oceans in the south, it runs 370 km. People use the river along the way for a variety of purposes, including drinking, sanitation, washing, agricultural purposes, recreation, and hydroelectric power generation [34]. In recent times with the growing population residing along the riverside especially in Thimphu, the more quantity of waste generated being discharged directly into the flowing river [35]. The domestic sewage, agricultural runoff, solid wastes and industrial wastes pollute the river water changing the quality parameters of water. Every current and coming future generations should have secure access to adequate, safe and affordable water, therefore, quality management of rivers should be considered and preserved through alternative solutions [36].

3.2 Water sampling and data collection

Five locations along the stretch of study area from Dodeyna, Pangrizampa, Hejo, Babesa and Khasadrapchu were selected for the water sample collection (Fig. 4b). The data collection and experiment were performed at the same location and same time as that of the secondary data. Data such as pH, EC, DO, water temperature and TDS are collected in two forms: secondary data and experimental data. The first form of data was obtained from the National Environment Commission (NEC) of Bhutan and the latter one was obtained through experiment. Water samples were taken from five different places along the river's length as shown in Fig. 4b. Those data (Table 1) were fed as an input data for the water models. Based on three evaluation criteria, the comparison was made among three models and the suitable model was recommended.

4 Performance measures: criteria of evaluation

4.1 Coefficient of correlation (CC)

The CC is a statistical indicator that represents the strength of an association between two variables where the value lies between -1.0 to 1.0. The values greater than 1 or less than -1 indicates the error in the correlation measurement. A correlation of -1 signifies a perfect negative correlation, and a perfect positive correlation by 1. A correlation of 0 specifies that there is no relationship between the two variables. The formula to calculate the coefficient correlation is shown in Eq. (1).

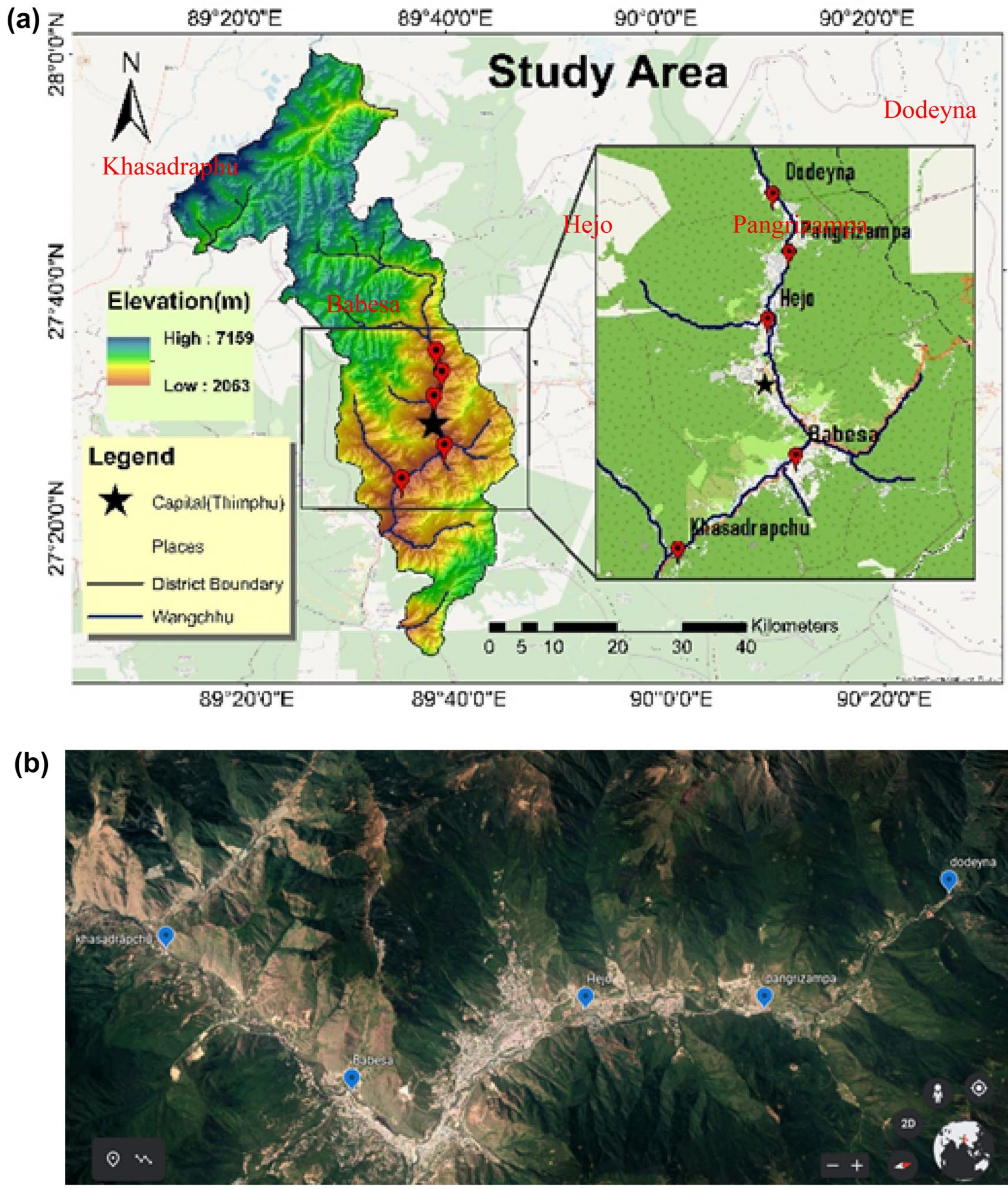


Fig. 4 **a** Map of study area. **b** Wangchu River basin and sampling points

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{\left[n\Sigma x^2 - (\Sigma x)^2\right]\left[n\Sigma y^2 - (\Sigma y)^2\right]}} \tag{1}$$

4.2 Nash–Sutcliffe efficiency (NSE)

The NSE is a regular method that calculates the comparative magnitude of the outstanding variance associated to the measured data variance. It specifies how well the plot of experimental data versus predicted data flits the 1:1 line. NSE equal to 1, means a perfect match of the model to the experimental data. NSE equals to 0 specifies that the model prediction are as accurate as the mean of the observed data (Eq. 2).

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n}(\mathrm{OBS}_i - \mathrm{SIM}_i)^2}{\sum_{i=1}^{n}(\mathrm{OBS}_i - \overline{\mathrm{OBS}})^2}2 \tag{2}$$

Table 1 Water quality data

Sl. no.	Year	Location	Observed avg. pH	Observed avg. EC	Observed avg. DO	Observed avg. TDS
1	2014	Dodeyna	7.2	89.9	9.8	53
2		Pangrizampa	7.25	80	9.5	48.5
3		Hejo	7.3	105	8.5	59.8
4		Babesa	7.5	105.9	7.5	59.8
5		Khasadrapchu	7.5	92	7.2	54.8
6	2015	Dodeyna	7.3	91.1	10.1	57.1
7		Pangrizampa	7.3	90	10	53.2
8		Hejo	7.3	104.8	9	60.8
9		Babesa	7.5	109.1	8	60.5
10		Khasadrapchu	7.4	91	8.1	55.2
11	2016	Dodeyna	8.1	166	8.5	53.3
12		Pangrizampa	8.2	117	8.5	75.3
13		Hejo	8.1	138.2	8.3	71.3
14		Babesa	7.5	143.7	8.1	53.9
15		Khasadrapchu	7.4	139.6	7.3	63.7
16	2017	Dodeyna	8	112.9	8.1	53.3
17		Pangrizampa	11.2	158.8	6.7	75.3
18		Hejo	8.5	150.5	8.2	71.3
19		Babesa	8.5	114.1	8.1	53.9
20		Khasadrapchu	8.3	142.2	7.9	67.3
21	2018	Dodeyna	8.5	168.7	8.5	80.1
22		Pangrizampa	8.8	159.5	9.4	75.6
23		Hejo	9	153.2	8.2	72.6
24		Babesa	9.3	151.4	8	71.9
25		Khasadrapchu	9.7	135	8.7	63.9
26	2019	Dodeyna	8.2	112.9	9.3	59.6
27		Pangrizampa	7.8	144.5	9.7	72.6
28		Hejo	8.2	142.9	8.5	75.1
29		Babesa	8.3	123.8	10.2	88.3
30		Khasadrapchu	8.5	158.8	9.5	67.87
31	2020	Dodeyna	7.1	120.1	9.7	60.7
32		Pangrizampa	8.3	135.2	8.2	62.3
33		Hejo	9.1	130.2	9.5	65.1
34		Babesa	9.4	140.6	8.7	70.2
35		Khasadrapchu	10.2	146.3	8.8	86.9
36	2021	Dodeyna	7.3	157	9.31	90
37		Pangrizampa	7.5	170	10.15	72.5
38		Hejo	7.2	153	8.18	82.5
39		Babesa	7.3	143	9.11	87.5
40		Khasadrapchu	7.6	121	7	97.25

Secondary data from 2014 till 2020 and experimental data for 2021

where OBS_i is the observed value, SIM_i is the predicted value, $\overline{OBS}$ is the average of the observed values and n is the number of data samples.

4.3 Root mean square error (RMSE)

RMSE is a standard way to quantify the error of a model in predicting data, given as following Eq. 3:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(M_i - P_i)^2}{N}} \quad (3)$$

where M_i is the observed value, P_i is the predicted value and N is the number of data set.

5 Results and discussion

5.1 ANFIS analysis

The consistency parameters of river water were used as input data for the ANFIS modeling. The 7-year data was split into two parts: training sets and testing sets, with 70 percent and 30 percent of the data going to each. The training stops only when the two datasets match closely and with minimal error did the training come to an end. The projected values were exported and assessed as shown in Fig. 5.

By ANFIS analysis, the following observations were made:

a. The coefficient correlation in predicting the pH, DO, EC and TDS are 0.86, 0.71, 0.67 and 0.56 respectively.
b. The Nash–Sutcliffe efficiency in predicting TDS, pH, DO and EC is 0.513, 0.252, 0.172 and 0.151 respectively.
c. The error generated by the model in predicting EC, TDS, pH and DO are 22.917, 10.875, 0.878 and 0.619 respectively.

5.2 MLR analysis

Similarly, using the MLR model the prediction of data was done and the graphs were plotted (Fig. 6) to work out the different assessment criteria.

The observations made through MLR analysis are:

a. The coefficient correlation in predicting the EC, DO, TDS and pH are 0.653, 0.62, 0.569 and 0.54 respectively.
b. The Nash–Sutcliffe efficiency in predicting EC, DO, TDS and pH are 0.427, 0.396, 0.321 and 0.289 respectively.
c. The error generated by the model in predicting EC, TDS, pH and DO are 18.833, 9.85, 0.767 and 0.69 respectively.

5.3 ANN analysis

The training set is used to build up the neural network model, and the target set is used to check the model performance at several stages of training and to decide when to stop training to avert the over-fitting.

The observations made through ANN analysis are (refer Fig. 7):

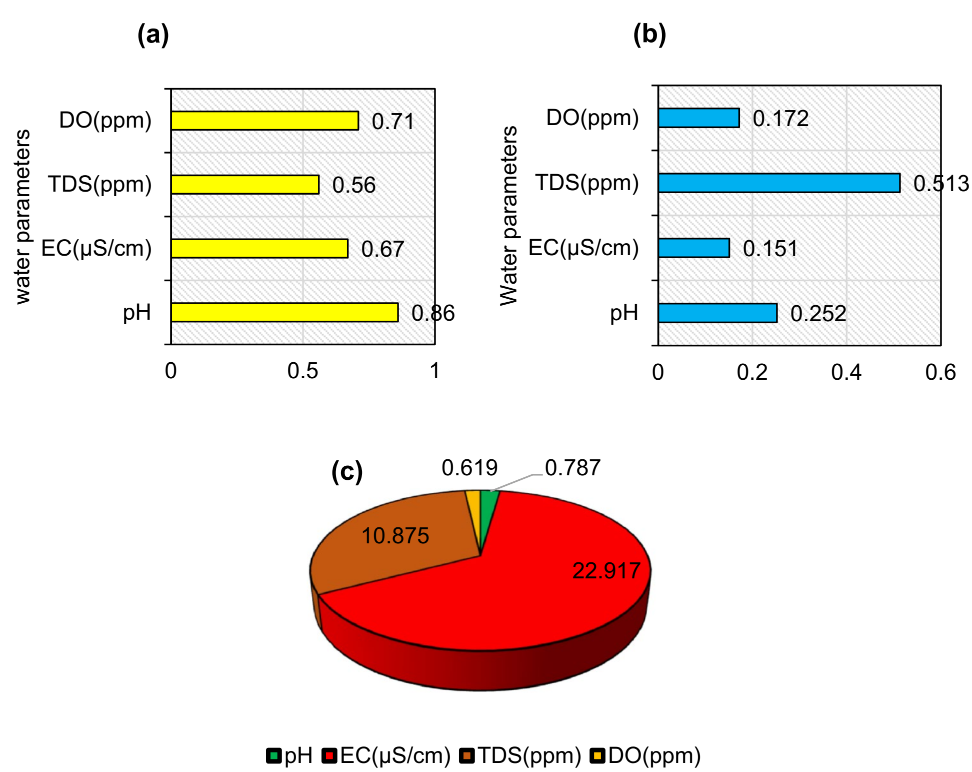


Fig. 5 **a** Coefficient correlation; **b** Nash–Sutcliff efficiency; **c** root mean square error by ANFIS model

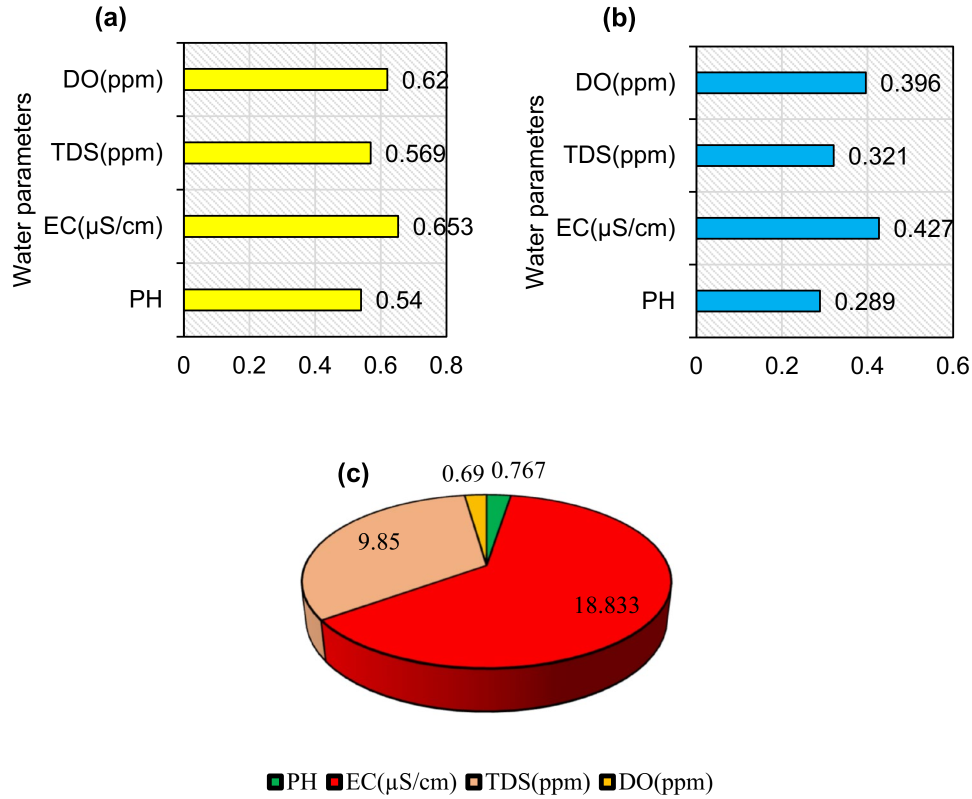


Fig. 6 **a** Coefficient correlation; **b** Nash–Sutcliff efficiency; **c** root mean square error by MLR.

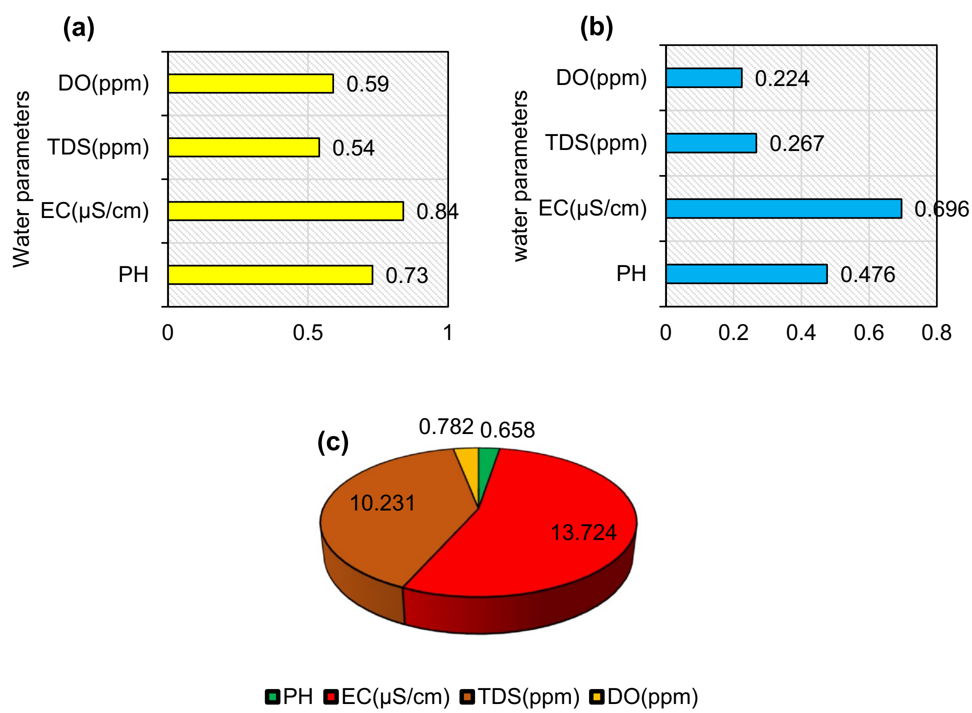


Fig. 7 Coefficient correlation; **b** Nash–Sutcliff efficiency; **c** root mean square error by ANN

a. The coefficient correlation in predicting the EC, pH, DO, and TDS are 0.84, 0.73, 0.59 and 0.54 respectively.
b. The Nash–Sutcliffe efficiency in predicting EC, pH, TDS and DO are 0.696, 0.476, 0.267 and 0.224 respectively.
c. The error generated by the model in predicting EC, TDS, DO and pH are 13.724, 10.231 0.782 and 0.658 s respectively.

5.4 Comparative analysis of models

From Fig. 8a, by comparing the CC values, pH and DO were predicted with strong correlation by ANFIS model and ANN model predicted EC with high correlation. The prediction made by the MLR model gave weak correlation for most of the parameters. Based on all the literature review, model giving strong correlation for different parameters should be adopted for performing the analysis of the particular water parameter [15, 23, 37]. From Fig. 8b, it is evident that ANN model achieved higher efficiency in predicting pH and EC while ANFIS predicted TDS efficiently. From Fig. 8c, the prediction made by ANN model gave minimum error in predicting most of the parameters outperforming the other two models. Thus, the model giving minimum error in the analysis is recommended to be used for predicting the specific parameter.

Looking into the results obtained, it is evident that ANN model performed better than the other two models in predicting most of the water parameters. In addition to that, while looking into the overall NSE and RMSE, the ANN model predicted parameters with maximum efficiency of 97.3 percent and minimum error of 8.57 (Fig. 9a). The efficiency of MLR and ANFIS models are 95.9 percent and 94.1 percent respectively. The overall error generated by MLR and ANFIS are 10.64 and

Fig. 8 Comparison among the models: **a** CC, **b** NSE and **c** RMSE

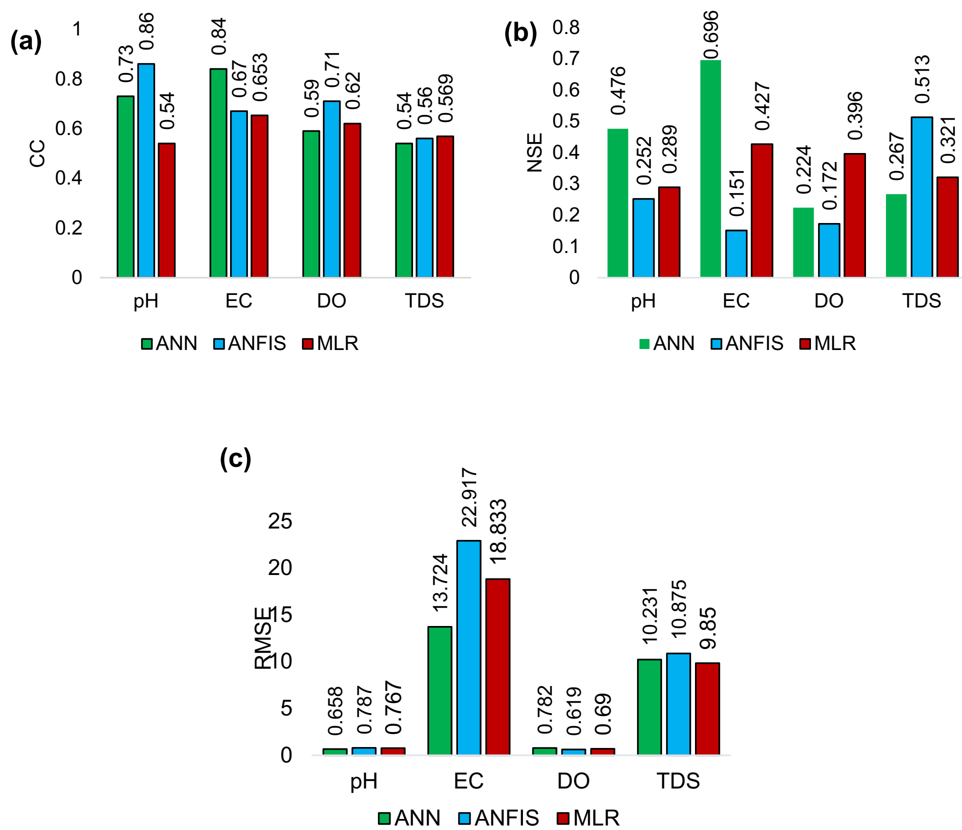
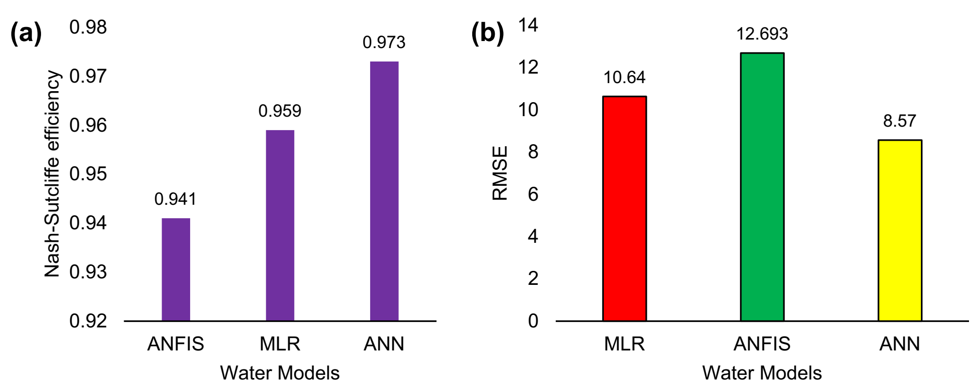


Fig. 9 **a** Overall Nash–Sutcliff efficiency; **b** overall root mean square error by the models.



12.693 respectively as depicted in Fig. 9b. Most of literature indicated that ANN model and ANFIS can be both suitable for modeling of water quality parameter but also indicated that ANN model is slightly better than other two, due to the over estimating and under estimating performance of ANFIS and MLR [15, 17, 19, 38]. Thus, using the ANN model was selected as most suitable for Wangchuk River. Hence, the prediction for the next year and location was performed.

5.5 Prediction for the next year and location

In analysis, the 6 years data from 2014 till 2020 was again used as an input data for the model and predicted the data for 2021 and compared the predicted values for 2021 with experimental values of 2021. Analysis using various training Function of ANN Model were performed for best function with higher efficiency. The different algorithms are incorporated in the backpropagation neural network which have different performances capability in training the data such as (1) Broyden–Fletcher–Goldfarb Quasi-Newton (trainBFG), (2) Bayesian regularization (trainBR), (3) conjugate gradient backpropagation (trainCGB), (4) conjugate gradient Fletcher–Reeves (trainCGF), (5) gradient descent with momentum (trainGDM), (6) gradient descent with adaptive learning rate (trainGDA) [39].

While forecasting data for the next year using various training functions in ANN model, trainBR could achieve higher correlation and maximum efficiency with less error as shown in Table 2 and Fig. 10.

The predicted values were compared with the experimental data of 2021 to validate the result. The results shown in Table 3 validates that the model could predict the next year data with strong correlation between observed and predicted data, and could achieve optimum efficiency and minimum error in predicting the data.

5.6 Prediction for the next location

Similarly, while forecasting data for the location using various training functions in ANN model, trainBR could achieve higher correlation and maximum efficiency with less error as shown in Table 4 and Fig. 11. In this, the data from four locations such as Dodeyna, Pangrizampa, Hejo and Babesa were used as input data and predicted for Khasadrapchu. The predicted value was compared with the experimented value of Khasadrapchu for the validation and the values were found similar with negligible variation. The results shown in Table 5 indicates that the predictions made by the model for the next location are reasonable with strong correlation, high efficiency and minimum error.

6 Conclusion

This study evaluated three different types of Artificial Intelligence, ANFIS, ANN and MLR models to calculate and predict TDS, pH, DO and EC for Wangchu River, Thimphu, Bhutan. Results gave the insights to use water models in predicting and forecasting the water quality parameters at regional and global scale. For the monitoring and management of water resources in a sustainable manner, the suitable model recommended in this study can be used by the individual for various rivers and lakes around the world. The traditional means of experiments to check the quality of river water can be substituted with reliable and realistic water models which are the findings of this study.

The experimental data for the selected parameters were found within the acceptable range by comparing with the standards of the National Environment Standards of Bhutan 2020, but by looking into the past data and comparing it was observed that the quality of the water has deteriorated. As a result, some of the mitigation measures to improve the quality suggested are to prevent the pollution from major sources such as

Table 2 CC and RMSE using different training functions for next year data prediction

Sl. no.	Parameter	Training functions/criteria	trainBFG	trainBR	trainCGB	trainCGF	trainGDM	trainGDA
1	pH	CC	0.602	0.262	0.817	0.400	0.205	0.512
		RMSE	0.147	0.148	0.093	0.177	0.161	0.150
2	EC (μs/cm)	CC	0.332	0.954	0.017	0.530	0.480	0.774
		RMSE	21.844	5.766	18.918	14.369	14.793	14.571
3	DO (mg/L)	CC	0.772	0.611	0.925	0.637	0.404	0.740
		RMSE	1.370	0.891	0.495	1.648	1.092	0.948
4	TDS (mg/L)	CC	0.608	0.783	0.196	0.667	0.235	0.442
		RMSE	8.083	7.678	11.041	7.553	10.512	9.256

Fig. 10 Overall CC, RMSE and NSE using different training functions for next year data prediction

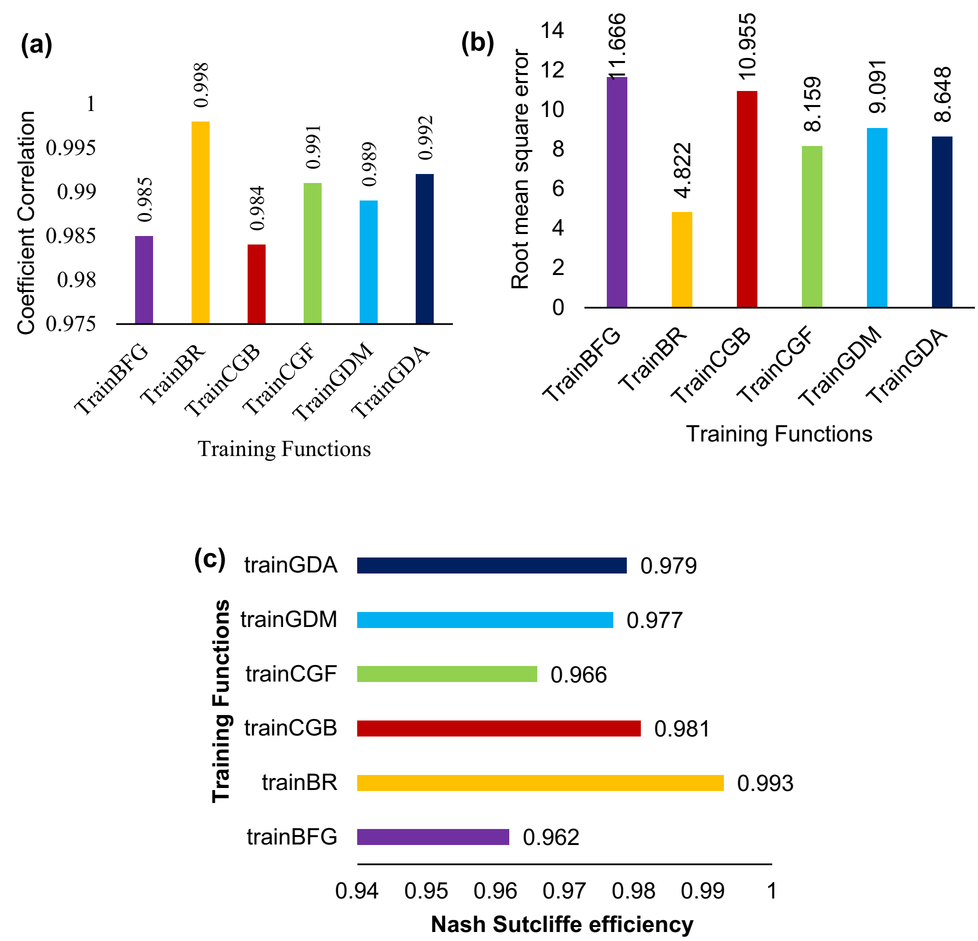


Table 3 ANN model performance for the next year data prediction

Sl. no.	Parameter/criteria	CC	NSE	RMSE
1	PH	0.5	0.201	0.132
2	EC (µS/cm)	0.812	0.658	9.58
3	TDS (ppm)	0.921	0.63	5.01
4	DO (ppm)	0.87	0.662	0.63

the domestic waste should not be discharged directly into the river. As Thimphu being the largest city in the country with largest number of automobiles in the city. The waste water from the car wash and workshops is observed being discharged into the tributaries of Wangchu River which ultimately contaminates the river. A change in water chemistry caused by surface water

Table 4 CC and RMSE using different training functions for next location data prediction

Sl. no.	Parameter	Training functions/ criteria	trainBFG	trainBR	trainCGB	trainCGF	trainGDM	trainGDA
1	pH	CC	0.850	0.970	0.910	0.970	0.640	0.510
		RMSE	0.728	0.245	0.492	0.288	0.855	0.944
2	EC (µs/cm)	CC	0.910	0.970	0.890	0.850	0.630	0.980
		RMSE	18.690	7.230	7.750	18.130	21.270	5.260
3	DO (mg/L)	CC	0.480	0.660	0.530	0.880	0.750	0.770
		RMSE	0.879	0.829	0.867	0.415	0.593	0.587
4	TDS (mg/L)	CC	0.750	0.660	0.800	0.700	0.690	0.370
		RMSE	13.780	10.730	10.390	11.450	13.550	20.130

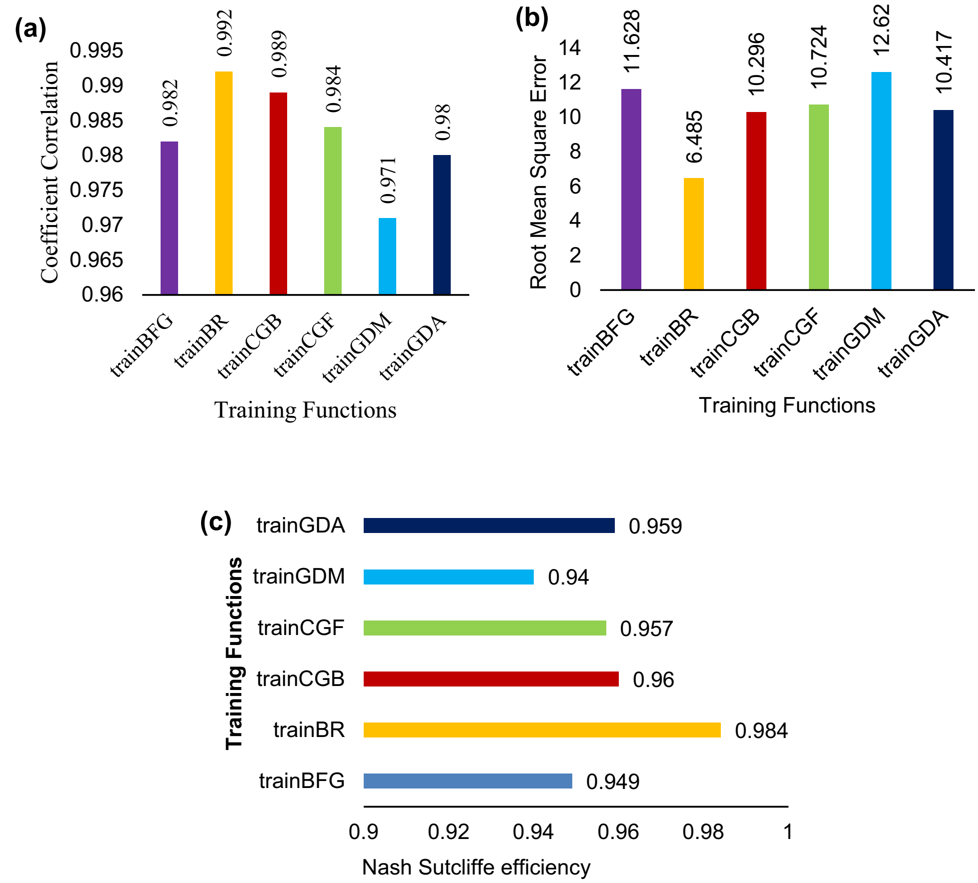


Fig. 11 Overall CC, RMSE and NSE using different training functions for next location data prediction

Table 5 ANN model performance for the next location data prediction

Sl. no.	Parameter/criteria	CC	NSE	RMSE
1	PH	0.776	0.494	0.727
2	EC (µS/cm)	0.88	0.773	11.145
3	TDS (ppm)	0.68	0.336	11.395
4	DO (ppm)	0.52	0.262	0.712

pollution can have a detrimental impact on an environment at all levels. Proper dumping of waste or discharging after treatment of waste water will be effective for water resource conservation. A constant monitoring to fresh water resources should be constantly conducted and with current practices, the prediction by ANN model shows that the water of the capital city of Bhutan will keep deteriorating over time. The utilization of applied methods in this study can be considered in other rivers of the Country as well as around the globe in order to investigate water quality. Furthermore, the models applied in this study could provide a basis for managers, engineers and policy makers for impressive designs, management and decisions making over different rivers in Bhutan and around the world.

Acknowledgements The author is thankful to College of Science and Technology, Royal University of Bhutan, Phuentsholing 21101, for assisting the project through providing essential reviews timely. We exuberantly thank the management of Phuentsholing Thromde (city) for their support during the entire phase of project. Heartful gratitude to all agencies for aiding us with data acquisition. The data is a record of consultancy services carried out by the authors and fund were covered from it. The author appreciates reviewers for their insightful comments on manuscript.

Author contributions Conceptualization, YC and NC.; data collection and field survey, YC and SC.; methodology, KRA. and SC; data curation, NC; writing—original draft preparation, NC and SC; writing—review and editing, KMAA and YC; supervision, TR; all authors have read and agreed to the published version of the manuscript.

Funding The authors did not receive support from any organization for the submitted work.

Declarations

Conflict of interest The authors don't have any relevant financial or non-financial interests to disclose.

References

1. Akhtar N, Syakir Ishak MI, Bhawani SA, Umar K (2021) Various natural and anthropogenic factors responsible for water quality degradation: a review. Water (Switzerland) 13:19. https://doi.org/10.3390/w13192660
2. Rosca OM, Dippong T, Marian M, Mihali C, Mihalescu L, Hoaghia MA, Jelea M (2020) Impact of anthropogenic activities on water quality parameters of glacial lakes from Rodnei mountains, Romania. Environ Res 182:109136
3. Dippong T, Hoaghia MA, Mihali C, Cical E, Calugaru M (2020) Human health risk assessment of some bottled waters from Romania. Environ Pollut 267:115409
4. Ma X, Wang L, Yang H, Li N, Gong C (2020) Spatiotemporal analysis of water quality using multivariate statistical techniques and the water quality identification index for the Qinhuai river basin, east china. Water (Switzerland) 1210:1–19. https://doi.org/10.3390/w12102764
5. Omar NH (2012) Water quality paramter. IntechOpen book chapter. http://dx.doi.org/https://doi.org/10.5772/intechopen.89657
6. Meride Y, Ayenew B (2016) Drinking water quality assessment and its effects on residents health in Wondo genet campus, Ethiopia. Environ Syst Res 51:1–7. https://doi.org/10.1186/s40068-016-0053-6
7. Sattari MT, Joudi AR, Kusiak A (2016) Estimation of water quality parameters with data-driven model. J Am Water Works Assoc 1084:232–239. https://doi.org/10.5942/jawwa.2016.108.0012
8. Ahmed U, Mumtaz R, Anwar H, Mumtaz S, Qamar AM (2020) Water quality monitoring: from conventional to emerging technologies. Water Sci Technol Water Supply 20(1):28–45
9. Bhateria R, Jain D (2016) Water quality assessment of lake water: a review. Sustain Water Resour Manag 2(2):161–173. https://doi.org/10.1007/s40899-015-0014-7
10. Rahmanian N, Hajar S, Ali B, Homayoonfard M, Ali NJ, Rehan M, Sadef Y, Nizami AS (2015) Analysis of physiochemical parameters to evaluate the drinking water quality in the state of Perak, Malaysia. J Chem. https://doi.org/10.2166/ws.2019.144
11. Dippong T, Mihali C, Năsui D, Berinde Z, Butean C (2018) Assessment of water physicochemical parameters in the Strîmtori-Firiza reservoir in Northwest Romania. Water Environ Res 90(3):220–233
12. Qiu Y, Li J, Huang X, Shi H (2018) A feasible data-driven mining system to optimize wastewater treatment process design and operation. Water (Switzerland). https://doi.org/10.3390/w10101342
13. Jenny H, Wang Y, Alonso EG, Minguez R (2020) Using Artificial Intelligence for Smart Water Management Systems 55(143)
14. Mammri AMM, Kantoush SA, Kobayashi S, Sumi T, Saber M (2019) Real-time measurement of flash-flood in a wadi area by LSPIV and STIV. Hydrology 6(1):27
15. Abba SI, Hadi SJ, Abdullahi J (2017) River water modelling prediction using multi-linear regression, artificial neural network, and adaptive neuro-fuzzy inference system techniques. Procedia Comput Sci 120:75–82. https://doi.org/10.1016/j.procs.2017.11.212
16. Vasudev S, Nagar DB, Choudhary MM (2018) Analysis of drinking water quality parameters a case study of Hanumangarh town. Int J Trend Sci Res Dev 25:75–82. https://doi.org/10.31142/ijtsrd15774
17. Xiaohu W, Jing F, Meina D, Chuangi Z (2013) Artificial neural network modeling of dissolved oxygen in the Heihe River, Northwestern China. Environ Monit Assess 1855:4361–4371
18. Montaseri M, Zaman ZG, Sanikhani H (2018) Water quality variations in different climates of Iran: toward modeling total dissolved solid using soft comMonstaseri, majid Ghavidel, z sarvin Sanikhani, Hadiputing techniques. Stoch Env Res Risk Assess 32:2253–2273
19. Ay M, Kişi O (2017) Estimation of dissolved oxygen by using neural networks and neuro fuzzy computing techniques. KSCE J Civ Eng 215:1631–1639. https://doi.org/10.1007/s12205-016-0728-6
20. Ay M, Özyıldırım S (2018) Artificial intelligence (AI) studies in water resources. Nat Eng Sci 3(2):187–195
21. Tariq M, Wangchuk K, Muttil N (2021) A critical review of water resources and their management in Bhutan. Hydrology. https://doi.org/10.3390/hydrology8010031
22. Kouadri S, Elbeltagi A, Islam ARMT, Kateb S (2021) Performance of machine learning methods in predicting water quality index based on irregular data set: application on Illizi region (Algerian southeast). Applied Water Sci 11(12):190. https://doi.org/10.1007/s13201-021-01528-9
23. Areerachakul S (2012) Comparison of ANFIS and ANN for estimation of biochemical oxygen demand parameter in surface water. Int J Chem Biol Eng 64:286–290
24. Al-Mukhtar M, Al-Yaseen F (2019) Modeling water quality parameters using data-driven models, a case study Abu-Ziriq Marsh in South of Iraq. Hydrology 624:1–17. https://doi.org/10.3390/hydrology6010024
25. Jang JSR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 233:665–685. https://doi.org/10.1109/21.256541
26. Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. IEEE Trans Syst Man Cybern 151:116–132. https://doi.org/10.1109/TSMC.1985.6313399
27. Nemati S, Fazelifard MH, Terzi O, Ghorbani MA (2015) Estimation of dissolved oxygen using data-driven techniques in the Tai Po River, Hong Kong. Environ Earth Sci 745:4065–4073. https://doi.org/10.1007/s12665-015-4450-3
28. Salari M, Salami Shahid E, Afzali SH, Ehteshami M, Conti GO, Derakhshan Z, Sheibani SN (2018) Quality assessment and artificial neural networks modeling for characterization of chemical and physical parameters of potable water. Food Chem Toxicol 118:212–219. https://doi.org/10.1016/j.fct.2018.04.036
29. Asadollahfardi G, Taklify A, Ghanbari A (2012) Application of artificial neural network to predict TDS in Talkheh Rud river. J Irrig Drain Eng 138(4):364–370. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000402
30. Bilgili M, Sahin B, Yasar A (2007) Application of artificial neural networks for the wind speed prediction of target station using reference stations data. Renew Energy 3214:2350–2360. https://doi.org/10.1016/j.renene.2006.12.001
31. Singh KP, Basant A, Malik A, Jain G (2009) Artificial neural network modeling of the river water quality—a case study. Ecol Model 2206:888–895. https://doi.org/10.1016/j.ecolmodel.2009.01.004

32. Garg T (2019) The 4 major rivers of Bhutan—the lifeline of the country. Holidify, Wildlife & Nature. https://www.holidify.com/pages/rivers-of-bhutan-1221.html
33. Sarma JN (2004) An overview of the Brahmaputra river system. Kluwer Academic Publishers, Boston, pp 72–87. https://doi.org/10.1007/978-94-017-0540-0_5
34. Kharat R, Mundra A (2019) Water issues in Bhutan: internal disputes and external tensions. In: Ranjan A (ed) Water issues in Himalayan South Asia. Palgrave Macmillan, Singapore. https://doi.org/10.1007/978-981-32-9614-5_3
35. Tenden Z (2016) Upper Wangchu turning into a garbage dump. The Bhutanese. https://thebhutanese.bt/upper-wangchu-turning-into-a-garbage-dump/. Accessed 24 Mar 2021
36. Bhutan Water Policy (2007) National environment commission. http://www.nec.gov.bt/. Accessed 24 Mar 2021
37. Bisht AK, Singh R, Bhutiani R, Bhatt A, Kumar K (2017) Water quality modelling of the River Ganga using artificial neural network with reference to the various training functions. Environ Conserv J 1812:41–48. https://doi.org/10.36953/ecj.2017.181206
38. Folorunsho JO, Mu'azu MB, Garba S, Obiniyi AA, Ajibade AO (2014) A comparison of ANFIS and ANN–based models in river discharge forecasting. N Ground Res J Phys Sci 1(1):1–6
39. Sharma BK, Venugopalan P (2014) Comparison of neural network training functions for hematoma classification in brain CT images. IOSR J Comput Eng 161:31–35. https://doi.org/10.9790/0661-16123135