Nature Environment and Pollution Technology
An International Quarterly Scientific Journal

p-ISSN: 0972-6268
(Print copies up to 2016)
e-ISSN: 2395-3454

Vol. 21 | No. 3 | pp. 1259-1267 | 2022

Original Research Paper

https://doi.org/10.46488/NEPT.2022.v21i03.032

Open Access Journal

Survival Study on Different Water Quality Prediction Methods Using Machine Learning

K. Kalaivanan and J. Vellingiri†
School of Information Technology and Engineering, Vellore Institute of Technology, Vellore-632014, Tamilnadu, India
†Corresponding author: J. Vellingiri; vellingiri.j@vit.ac.in

Nat. Env. & Poll. Tech.
Website: www.neptjournal.com

Received: 04-09-2021
Revised: 24-10-2021
Accepted: 26-10-2021

Key Words:
Water quality prediction
Machine learning
ANN
SVM

ABSTRACT

Water quality analysis is an emergency approach in today's world because people cannot survive without it. As a result of urbanization, industrialization, agricultural practices, and human behavior, water quality analysis have numerous issues in today's world. Manually visiting the water collection station, collecting water samples, analyzing in the lab, feeding data into a database, and so on are all challenges in the water quality analysis processing. Artificial learning model technologies will be used to tackle these challenges. The variety of machine learning approaches to water quality analysis has resulted in a diversity of creation and implementation methods. The study examines artificial intelligence's advancement in water quality prediction from different angles ANN, FUZZY, SVM, and other AI models. The review investigated 40 articles between 2008 and 2020. Groundwater, ponds, lakes, and rivers all water resources were all included in the survey method. The findings of the survey will be used to guide the future study.

INTRODUCTION

Research Background

Water resources span around 70% of the earth's surface (Mishra & Dubey 2015). The water sources are split into two types such as surface water and groundwater. The rivers, lakes, reservoirs, and coastal regions are examples of surface water, whereas infiltration galleries and springs are examples of groundwater (Mustafa et al. 2017). River and groundwater are critical for environmental, social health, and economic growth (Pandhiani et al. 2020, Siebert et al.2010). Rivers provide 65 percent of the water used in agriculture, with the remainder used for drinking, industrial, and other human needs. Human influences such as sewage, urbanization, agricultural, and industrial waste have an impact on river water quality (Bhatti et al. 2019).

However, these problems are caused by changes in the chemical properties of the water and the inability to drink or irrigate (Sakai et al. 2018). The chemical composition of water is related to its physical, chemical, and biological features which are used to determine the condition of water (Bordalo et al. 2006). Some of the water quality variables used to quantify water quality includes Electrical Conductivity, Dissolved Oxygen, Total Dissolved Solid, Chemical Oxygen Demand, Turbidity, Temperature, and pH (Tchobanoglous et al. 1985, Nikoo et al. 2013). Contaminated water is the most dangerous to people's health in underdeveloped countries, accounting for 80% of all health complications (Moore et al. 2003).

The National Sanitation Foundation of the United States proposed and adopted Water Quality Index in a global manner (Brown et al. 1970). The water quality index is another often-used indicator, which is required for massive data calculations, mathematical formulae, time, and effort. WQI categorizes WQ into excellent, poor, and worst according to standards established by regulatory agencies in the research field; results are scientific, yet they are presented in an easy-to-understand fashion (Fernández et al., 2004). Water quality analysis is fundamental in the River WQ analysis, monitoring, and control. Machine learning techniques for predicting water quality are currently based on training and testing techniques (Tung & Yaseen 2020). One of the several purposes of the model used to estimate water quality is to predict how it will change over time (Chen et al. 2018). Therefore, water quality forecasting is crucial for environmental monitoring, ecological sustainability, and human health (Fijani et al. 2019).

Problem Statement

Because water is such an important natural resource for all living organisms, it's necessary to ensure that it's always safe to drink and use. Many environmental variables play a vital role in predicting water quality. The state of the water is changed due to the variation in the effects on the environment, leading to industrial pollution, sewage, wastewater, human overuse, low levels of water, and over-utilization of land and sea resources (Dinka 2018). Water quality analysis was split into physical, chemical, and biological analysis methods. The water quality is predicted depending on turbidity level, moisture content, and water flows (Omer 2019).

Water quality analysis has become a difficult task due to global warming and the increase in population. The population growth leads to another concern water scarcity due to the lack of adequate infrastructure. Water consumption rises in tandem with population growth, affecting people all around the world. This reason leads to people consuming contaminated water which has been connected to the spread of water-borne diseases such as cholera, diarrhea, dysentery, and hepatitis. Twenty-seven waterborne diseases have been identified by the World Health Organization (WHO). Drinking water safety is acknowledged as a threat. This is an issue that worries both developed and developing countries across the world (Jury & Vaux 2007).

The traditional method of water quality forecasting is a manual approach, such as raw data collected at intervals and analyzed in the lab. This approach leads to a time-consuming process and a risky policy-making process related to water. The previous data collection approaches resulted in a dataset that was noisy and unbalanced. For the above reason, researchers spend a lot of time pre-processing and cleaning data (Pratt & Loizos 1992). Designing water-related data is challenging because of its nonlinearity, nonstationarity, and complexity (Chang et al. 2016). Because of the size of the data, traditional methods cannot meet current demand and require the use of artificial intelligence (AI) models and their development.

METHOD

This survey paper focuses on a review of water quality prediction using AI. The initial step is to collect water quality papers from Springer, Elsevier, and IEEE and possible resources. The second strategy was to search for articles from 2008 to 2020 using terms such as water quality, river, lake, machine learning, and deep learning research. The contribution of this survey study is:

- To conduct a thorough literature study to determine the current machine learning approaches for water quality prediction.
- To draw attention to the flaws and limits of present approaches.
- To compare various sources of surface water with the Common AI Approach.
- To recommend future research directions.

MACHINE LEARNING

Machine learning is the latest trending technology in water quality prediction. It is a technique that allows computers to learn automatically from previous results. Machine learning is a predictive analytics technique that makes predictions based on past data. Data and algorithms were incredibly important in machine learning (Kelleher 2019). In machine learning, most of the data is used in training, and less data is used in testing. It is an embrace of four of learning such as ANN, SVM, Fuzzy, and other AI models.

Artificial Neural Network (ANN) Model

The ANN algorithm was introduced by McCulloch in the year of 1943 (McCulloch & Pitts 1943). It is a simple method used for nonlinear data and solves complex problems. The Artificial Neural Network type is analogous to a human brain. It has neurons that are interconnected in different layers of networks (Chen et al. 2020). ANN is composed of three layers such as entry, hidden, and exit. The user will be offered input in the extreme form of formats adopted by the input layer. There may be a hidden layer between the input and the output layer. It is capable of carrying out a wide range of tasks. The output layer provides a result based on the predicted response (Najah et al. 2014). Fig.1 represents ANN Architecture.

Fuzzy Based Model

The Fuzzy based model was introduced by Zadeh (1965). Fuzzy Logic Systems (FLS) accepted incomplete, unclear, skewed input and produce the exact output. Fuzzylogic (FL) is a method similar to human thinking. The FL approach simulates human decision-making. Some logic blocks allow the device to recognize the number of inputs and outputs defined as true or false (Wang et al. 2003). The structure of Fuzzy Logic is depicted in Fig. 2.

Support Vector Machine (SVM) Model

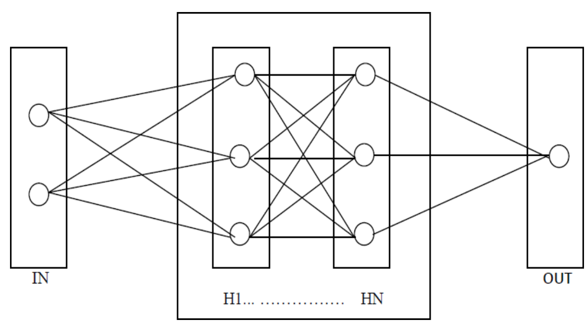

Fig. 1: ANN Architecture.

SVM is among the most effective classification and regression algorithms. The goal of the SVM algorithm is to make the simplest call purpose for bifurcate. A hyperplane is one of the best decision boundaries. It has two types such as linear and non-linear. Fig. 3 represents SVM Architecture.

Others AI Models

Some other models are also used to predict water quality such as hybrid and other models.

KEY METHODS AND SYSTEMS FOR WATER QUALITY PREDICTION

This part will cover water quality predictions such as inner relation analysis in water and pollution, various parameter predictions, anomaly detection methods in datasets, and dimension reduction in the feature selection process.

Analysis of Inner Relationship in Water Quality Parameters

Wu et al. (2019) introduced the adaptive frequency analysis method (ADP-FA) used to address data problems through the details of the intelligence frequency domain for internal relationships and personal discovery. The scalability characteristics were determined from the indicator, geography, and time domain. Prasad et al.(2020) developed different machine learning models that deal with binary and multi-class classification strategies used in water quality in lakes. They used the confusion matrix to test their work. The LSTM algorithm was found to provide the best level of accuracy and precision.

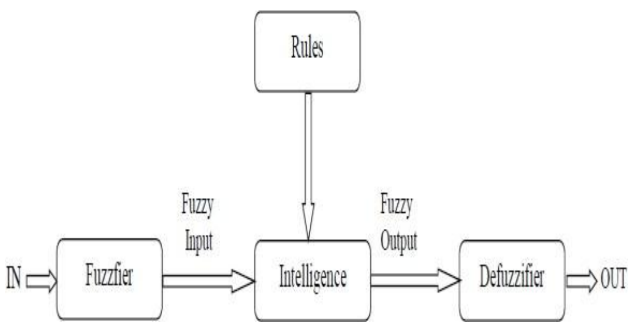

Fig. 2: Fuzzy Logic Structure.

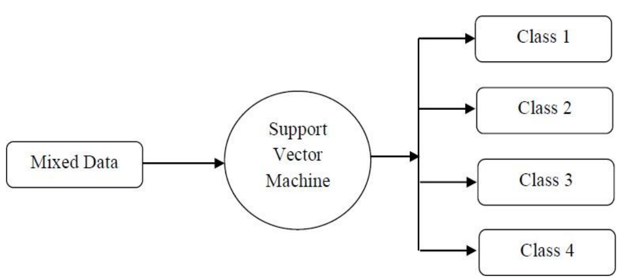

Fig. 3: Support Vector Model Architecture.

Total Dissolved Solids (TDS) Prediction

Niroobakhsh et al. (2012) compared two ANN models to predict TDS such as multi-layer perceptions (MLP) and the random forest (RF). The RF results can handle an enormous amount of information and can accurately forecast TDS levels. Tarke et al. (2016) An ANN model was used to predict the number of TDS within rivers. The backpropagation technique and the Levenberg-Marquardt optimization procedure are used to improve the ANN model's performance.

Dissolved Oxygen Prediction

Ahmed & Shah (2017) designed an ANFIS-based model for estimating BOD levels in the river. ANFIS model predicted the BOD whose performance was assessed by Mean Squared Error (MSE), Mean Absolute Error (MSE), Correlation coefficient (R), and Nash model efficiency (E). As a result, the adaptive neuro-fuzzy inference approach can accurately predict biochemical oxygen demand. Chen et al. (2018) introduced a new model BPNN combined with the artificial bee colony (ABC) method for DO forecasting. According to the findings, enhanced ABC–BPNN beats regular BPNN in terms of accuracy and generalization in local searches.

pH Prediction

Rajaee et al. (2018) tested ANN, WNN, MLR, and WMLR models for pH predicting. The WNN models used a wavelet transformation algorithm to predict pH levels in an advanced manner. This technique penetrates the mother wavelet while keeping its length constant. The WNN outperformed the others by eliminating the noise caused by the pH change.

Salinity Prediction

Melesse et al. (2020) used a hybrid machine learning model for salinity forecasting. M5-Prime, RF, and eight new hybrid algorithms were employed in this investigation. The water quality parameters such as pH, HCO, Cl, SO_4, Na, Mg, Ca, and Total Dissolved Solid were chosen. The validation results M5P-based hybrid algorithms outperformed RF-based hybrid algorithms. Barzegar et al. (2017) applied ANN, ANFIS, WNN, and WANFIS models to predict saltiness. The dataset was preprocessed using a discrete wavelet to improve accuracy. As a result, the WANFIS and WNN models outperformed other models.

Anomaly Detection Prediction

Deng et al. (2015) introduced a new model hybrid Fuzzy

Time-series for reducing noise in the dataset. The Gaussian cloud algorithm and the heuristic GCT method were employed by FTS. The performance of results extensively compared ARIMA, RBFNN, NAR, SVM, ANN–GCT, and OSM models. The FTS surpassed other models. Muharemi et al. (2019) proposed a machine learning method for the detection of changes or anomalies found in water quality over time. The information was collected from a German public water firm. Logistic regression, linear discriminate analysis, SVM, ANN, DNN, RNN, and LSTM were introduced to improve the data quality. F-scoring, accuracy, and recall methods are used to evaluate efficiency. As a result, the LSTM performed better than the LR model.

Dimension Reduction

Techniques for minimizing the number of parameters in a dataset are defined as dimensionality reduction. The dimension reduction technique increases machine learning performance at a high level. The ANN coupled discriminate analysis method was used by Voza & Vuković(2018). This method was used to extract two parameters such as the water temperature and electrical conductivity. Singh et al. (2011) implemented the SVM method for dimension reduction purposes. This process is compared with kernel-based discriminate analysis (DA), kernel-based partial least square (KPLS), standard DA, and PLS methods. The findings of SVM modeling show promise in terms of managing huge amounts of WQ data for optimization, categorization, and prediction.

COMPARISON OF DIFFERENT WATER QUALITY PREDICTION METHODS

In this review, machine learning approaches were used to simulate lake water quality, river water quality, and groundwater quality. This section discusses the present state of the water quality prediction technique comparative methodology in three categories of (1) ANFIS-FA with lake water (ii) ANFIS-WDT with river water (iii) ANFIS with groundwater is discussed in detail.

Quality Risk Analysis For Sustainable Smart Water Supply Using Data Perception

Water source management is used to regulate the supply of drinking water. The control of the water source was essential to enhance the water quality for end-users. An Adaptive Frequency Analysis (Adp-FA) method was introduced for determining data for individual predictions. A practical solution for building a risk analysis method for the local water distribution system. The designed method was employed was used to test the spectrum analyzer, model accuracy, and processing time with ANN and RF techniques. A new framework was introduced for the quality of the water examination with data processing mode. An Adp-FA method was employed for risk classification and prevention depending on water quality data acquired from the water distribution system. The developed technique can measure water quality indicators across many domains, including region and time.

The frequency property relationship between water quality indicators was examined for risk identification, forecast, and appraisal study. It observed a practical water source monitoring method using data obtained directly from industrial processes. It avoids issues such as the dependability of laboratory results and industrial applicability. It was beneficial to the exits water distribution system in the local infrastructural system. It established a link between readily available physical, chemical, and biological features. Pollution control decisions were made based on industrial and residential activities in water supply locations.

Cycle identification was used to find significant cycles for indicator changes in the temporal domain. The maximum value computation was used to monitor and estimate the amounts of the presence of various biological microorganisms. The training set adoption was used to do parameter adjustment. The created model was utilized to forecast the accurate bacteria indicators in tendency and values. The values were assigned different risk modes in accordance with applicable water source management regulations in different countries and locations. The decision support in water treatment plants was employed to forecast risky modes.

Machine Learning Methods For Better Water Quality Prediction

Artificial intelligence (AI) was carried out to develop dynamic mathematical expressions. It can recognize complex and nonlinear correlations involving input and output data. The Johor River Basin has suffered significant deterioration as a result of development and human activities. Different methods like ANFIS, RBF-ANN, and MLP-ANN were introduced. The data collected from monitoring sites and experimental studies were polluted by the noise impulses with random and systematic error. It is difficult to build an accurate prediction when there is noisy data present. Based on previous data on water quality factors, a Neuro-Fuzzy Inference System (WDT-ANFIS) with an augmented wavelet de-noising technique was used. Ammoniacal nitrogen (AN), suspended solids (SS), and pH were among the water quality indicators measured. Three evaluation assessment process methods were introduced. The first evaluation was carried out using neural network connection weight partitioning, which determined the importance of each model parameter in the network. The second and third evaluation processes established the beneficial input to create a single and combination of parameters correspondingly.

Estimation Of Total Dissolved Solids (TDS) Using New Hybrid Machine Learning Models

Groundwater (GW) is the main source of water for agricultural and residential purposes.GW quality was determined in an optimal manner. Decision-makers constructively handled groundwater sources when there was a clear understanding of the quantity and quality of GW. The ANFIS model was used to forecast hydrodynamic parameters and water quality indicators. The ANFIS model included an adaptive Takagaki-Sugeno fuzzy model. The ANFIS model has premises and subsequent parameters. The subsequent and premises parameters were regarded as choice parameters. As the objective function, RMSE was used. The fitness value calculation indicated the quality of the solutions. The optimization technique was used to modify the values of the agents. Because the training approach frequently produces a suboptimal answer when determining ANFIS parameters, the optimization algorithm was used to optimize and quantify the ANFIS parameters. In hybrid ANFIS and optimization methods, divides the first level data into training and test data. The second step involved training ANFIS based on training data.

PERFORMANCE ANALYSIS OF WATER QUALITY PREDICTION TECHNIQUES

To determine the different water quality prediction methods, a number of the data points are considered as input to conduct the testing. Different parameters were analyzed for improving the water quality prediction. For experimental consideration, water quality data is obtained from three separate water sources: lake water, river water, and groundwater. The dataset URL is given as https://tnpcb.gov.in/water-quality.php.The quantitative analysis is compared with different parameters such as,

- Prediction Time
- Prediction Accuracy and
- Error rate

Analysis of Prediction Time

Prediction time is calculated by counting the number of data points and the time required to estimate the water quality of one data point. The prediction time is determined as,

$$P_t = N * \textit{Time consumed for predicting water quality of one data} \quad \ldots(1)$$

From (1), the prediction time (P_t) is determined. 'N' represents the number of data points. The prediction time is expressed in milliseconds. Table 1 shows the prediction time as a function of the number of data points, which ranges from 25 to 250. The prediction accuracy comparison takes place on the existing Adp-FA method, WDT-ANFIS, and ANFIS model. The prediction time utilizing the Adp-FA method is comparatively reduced when compared to the WDT-ANFIS methodology and the ANFIS model, as shown in the table value. A graphical representation of the prediction time is shown in Fig. 4.

From Fig. 4, the prediction time depending on a different number of data points is described. The blue color line represents the prediction time of the Adaptive Frequency Analysis method. The red color and green color lines represent the prediction time of the WDT-ANFIS technique and the ANFIS model. From the above figure, it is clear that prediction time using the Adp-FA method is lesser when compared to the WDT-ANFIS technique and ANFIS model. This is because of introducing ANN and RF for improving the accuracy rate and minimizing the processing time. The modern data analysis methods were employed to tackle the water quality issues problems in management in smart water distribution systems for transferring knowledge across multiple indications, regions, and time domains. Consequently, the prediction time of the Adp-FA method is reduced by 16% when compared to the WDT-ANFIS technique and 40% when compared to the ANFIS model

Analysis of Prediction Accuracy

Prediction accuracy is defined as the ratio of the number of data points properly predicted to the total number of data points considered as input. As a result, the prediction accuracy is calculated as follows.

$$P_A = \left(\frac{\textit{Number of patient data that are correctly predicted the water quality}}{\textit{Number of data points}}\right) * 100 \quad \ldots(2)$$

Table 1: Tabulation of prediction time.

Number of data points	Prediction Time (ms)		
	Adp-FA method	WDT-ANFIS technique	ANFIS model
25	25	29	35
50	28	31	39
75	32	34	44
100	35	39	47
125	37	42	50
150	39	45	54
175	41	48	58
200	43	52	62
225	45	56	66
250	48	59	70

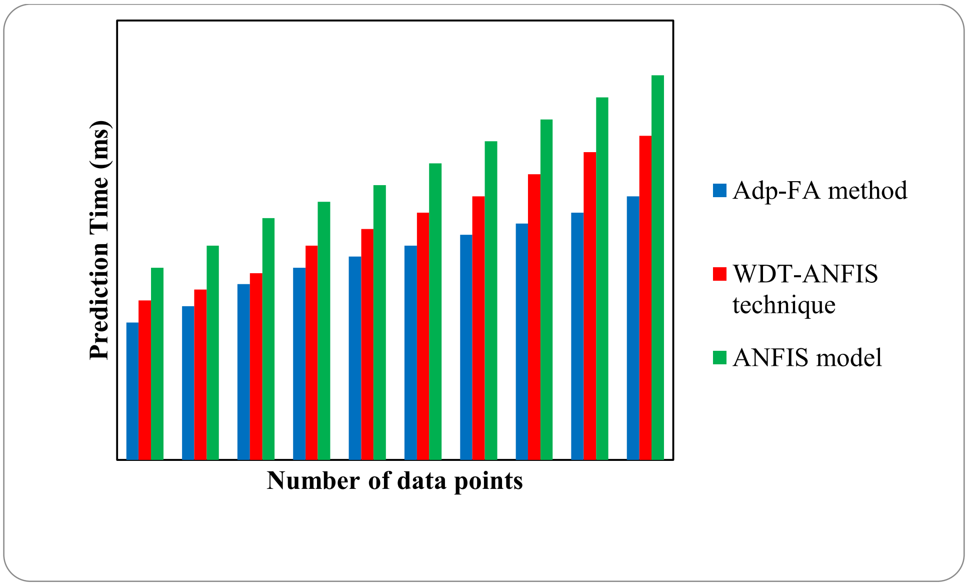

Fig. 4: Measurement of Prediction Time.

From (2), the prediction accuracy (P_A) is determined. The prediction accuracy is expressed as a percentage (%). Table 2 shows the prediction accuracy as a function of the number of data points, which ranges from 25 to 250. The prediction accuracy comparison takes place on the existing Adp-FA method, WDT-ANFIS, and ANFIS model. According to the table values, the prediction accuracy of the WDT-ANFIS methodology is higher than that of the Adp-FA method and the ANFIS model. Fig. 5 depicts a graphical representation of prediction accuracy.

From Fig.5, the prediction accuracy based on different numbers of data points is described. The blue color line represents the prediction accuracy of the Adaptive Frequency Analysis method. The red color and green color lines represent the prediction accuracy of the WDT-ANFIS technique and ANFIS model. It is clear from the statistics that the prediction accuracy using WDT-ANFIS technology is relatively high compared to the ANFIS model and the Adp-FA method. This is due to the application of a wavelet de-noising technique based on past water quality parameter data. Neural network connection load is estimated based on partitioning, which determines the importance of each input parameter in the network. The second and third evaluation processes were ascertained to construct the single and combination of parameters correspondingly. Consequently, the prediction accuracy of the WDT-ANFIS technique is increased by 7% when compared to the Adp-FA method and 11% when compared to the ANFIS model.

Table 2: Tabulation of Prediction Accuracy.

Number of data points	Prediction Accuracy (%)		
	Adp-FA method	WDT-ANFIS technique	ANFIS model
25	67	72	64
50	69	74	76
75	71	77	73
100	75	80	75
125	72	78	77
150	70	75	79
175	74	79	82
200	77	82	85
225	80	85	88
250	83	88	89

Analysis of Error rate

The error rate is defined as the ratio of the number of data points that incorrectly predict the water quality of the total data points to be considered as input. Therefore, the error rate is assumed to be as follows.

$$E_R = \left(\frac{\textit{Number of patient data that are incorrectly predicted water quality}}{\textit{Number of data points}}\right) * 100 \quad \ldots(3)$$

From (3), the error rate (E_R) is determined. The error rate is expressed as a percentage (%). Table 3 shows the error rate as a function of the number of data points, which ranges from 25 to 250. The Error rate comparison takes place on the

Table 3: Tabulation of error rate.

Number of data points	Error rate (%)		
	Adp-FA method	WDT-ANFIS technique	ANFIS model
25	21	25	15
50	24	28	17
75	27	31	19
100	22	26	18
125	20	22	16
150	18	20	14
175	22	24	17
200	26	28	20
225	30	33	23
250	33	37	26

Table 4: Merit and demerit.

Method	Merits	Demerits
ADP-FA	The approach is used to identify industrial contamination in waterways with ease.	The anticipated technique was difficult to implement in the absence of significant frequency effects.
WDT-ANFIS	Artificial neural weight vectors are employed in this approach to quickly identify input data.	Time consumption for water quality prediction was not reduced
ANFIS	This approach was used to predict dissolved solids in aquifers in a simple manner.	However, the accuracy of outputs was not improved by using multi-objective MFO.

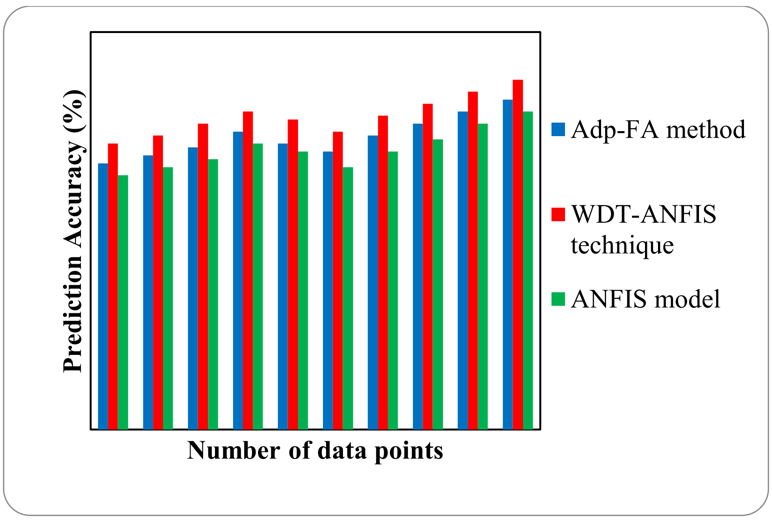

Fig.5: Measurement of prediction accuracy.

existing Adp-FA method, WDT-ANFIS, and ANFIS model. Fig. 6 depicts a graphical representation of the Error rate.

From above Fig. 6, the error rate based on the different numbers of data points is illustrated. The blue color line is also shown the error rate of the Adaptive Frequency Analysis method. The red color and green color lines are also shown in the error rate of the WDT-ANFIS technique and ANFIS model. From the figure, it is clear that the error rate using the ANFIS model is comparatively lesser when compared to the WDT-ANFIS technique and Adp-FA method. This is due to the usage of the Takagaki-Sugeno fuzzy model with the assumption and subsequent parameters, which is flexible.

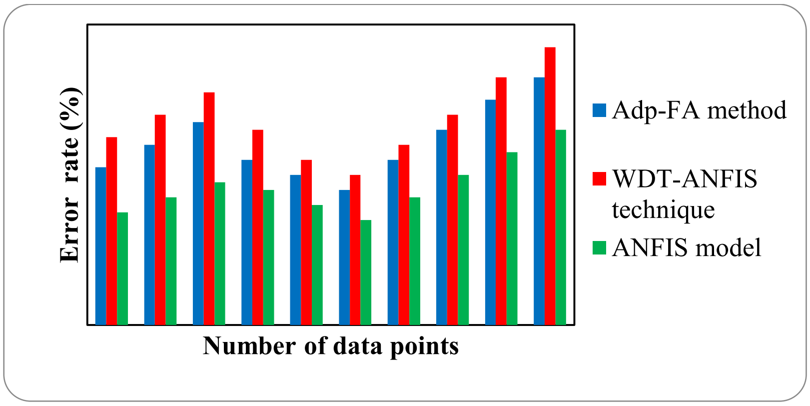

Fig. 6: Measurement of error rate.

The enhancement algorithm reached the optimal solution to determine the best ANFIS parameters. Consequently, the error rate of the ANFIS model is reduced by 32% when compared to the Adp-FA method and 49% when compared to the WDT-ANFIS technique.

RESULTS AND LIMITATIONS OF THIS STUDY

The results section explains the comparative research of the aforesaid literature review approaches.

The comparison is based on the ANFIS model, which incorporates frequency analysis as well as the wavelet denoising approach. The assessment measures employed in this study are prediction accuracy, prediction time, and error rate. Table 4 shows the merit and demerit of the models.

CONCLUSION AND FUTURE WORK

This paper presents a variety of survey techniques for water quality analysis using a machine learning approach. Many machine learning models are now being integrated with Big data and IoT to increase model performance. The performance in water quality prediction depends upon the accuracy and latency of detection, which is connected to the computing time of the model. According to the review, the time consumption of the machine learning models is not decreasing in estimating the quality of water.

Simultaneously, machine learning models lack certain information that allows them to handle real-time input.

The following list of research gaps and possibilities can be further explored in detail.

1. To introduce new hybrid ML models that can accurately forecast water quality.
2. To introduce new DL models for predicting water quality.
3. To improve optimization methods by adding new algorithms to forecast the water quality.

It is hoped that this work will contribute to the analysis of water quality in general and to an overview of the problem of predicting water quality as needed by people around the world.

REFERENCES

Ahmed, A.M. and Shah, S.M.A. 2017. Application of adaptive neuro-fuzzy inference system (ANFIS) to estimate the biochemical oxygen demand (BOD) of Surma River. J. King Saud Univ. Eng. Sci., 29(3): 237-243.

Barzegar, R., Adamowski, J. and Moghaddam, A.A. 2016. Application of wavelet-artificial intelligence hybrid models for water quality prediction: a case study in Aji-Chay River, Iran. Stochastic Environ. Res. Risk Assess., 30(7): 1797-1819.

Bhatti, N.B., Siyal, A.A., Qureshi, A.L. and Bhatti, I.A. 2019.Socio-economic impact assessment of small dams based on T-paired sample test using SPSS software. Civil Eng. J., 5(1): 153-164.

Bordalo, A.A.,Teixeira, R. and Wiebe,W.J. 2006. A water quality index was applied to an international shared river basin: the case of the Douro River. Environ. Manag., 38(6): 910-920

Brown, R.M., McClelland, N.I., Deininger, R.A. and Tozer, R.G. 1970. A water quality index-do we dare. Water Sewage Works, 117(10): 651

Chang, F.J., Chen, P.A., Chang, L.C. and Tsai, Y.H. 2016. Estimating spatio-temporal dynamics of stream total phosphate concentration by soft computing techniques. Sci. Total Environ., 562: 228-236.

Chen, S., Fang, G., Huang, X. and Zhang, Y. 2018.Water quality prediction model of a water diversion project based on the improved artificial bee colony–backpropagation neural network. Water, 10(6): 806.

Chen, Y., Song, L., Liu, Y., Yang, L. and Li, D. 2020. A review of the artificial neural network models for water quality prediction. Appl. Sci., 10(17): 5776.

Deng, W., Wang, G. and Zhang, X.2015. A novel hybrid water quality time series prediction method based on cloud model and fuzzy forecasting. Chemometr. Intell. Lab. Syst., 149:39-49.

Dinka, M.O. 2018. Safe Drinking Water: Concepts, Benefits, Principles and Standards. Water Challenges of an Urbanizing World, IntechOpen, London, pp.163-181.

Fernández, N., Ramírez, A. and Solano, F. 2004. Physico-chemical water quality indices-a comparative review. Bistua: Revista de la Facultad de Ciencias Básicas, 2(1): 19-30.

Fijani, E., Barzegar, R., Deo, R., Tziritis, E. and Skordas, K. 2019. Design and implementation of a hybrid model based on a two-layer decomposition method coupled with extreme learning machines to support real-time environmental monitoring of water quality parameters. Sci. Total Environ., 648: 839-853.

Jury, W.A. and Vaux, H.J. 2007. The emerging global water crisis: managing scarcity and conflict between water users. Adv. Agro., 95: 1-76.

Kelleher, J.D., 2019. Deep Learning. MIT Press, Cambridge, MA.

McCulloch, W.S. and Pitts, W.1943. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys., 5(4): 115-133.

Melesse, A.M., Khosravi, K.,Tiefenbacher, J.P. and Heddam, S. 2020. River water salinity prediction using hybrid machine learning models. Water, 12(10): 2951.

Mishra, R.K. and Dubey, S.C. 2015. Freshwater availability and its global challenge. Int. J. Eng. Sci. Invent. Res. Develop., 2(6): 65-83.

Mohammed, H., Hameed, I.A. and Seidu, R. 2018. Detection of water safety conditions in distribution systems based on artificial neural network and support vector machine. Int. Conf. Adv. Intell. Syst. Inform., 61: 567-576.

Moore, M., Gould, P. and Keary, B.S. 2003. Global urbanization and impact on health. Int. J. Hyg. Environ. Health, 206(5): 269-278.

Muharemi, F., Logofǎtu, D. and Leon, F. 2019. Machine learning approaches for anomaly detection of water quality on a real-world data set. J. Inform. Telecomm., 3(3): 294-307.

Mustafa, A., Sulaiman, O. and Shahooth, S. 2017. Application of QUAL2K for water quality modeling and management in the lower reach of the Diyala river. Iraqi J. Civ. Eng, 11: 66-80.

Najah, A., El-Shafie, A., Karim, O.A. and El-Shafie, A.H. 2013. Application of artificial neural networks for water quality prediction. Neural Comp. Appl., 22(1): 87-201.

Nikoo, M.R., Karimi, A., Kerachian, R., Poorsepahy-Samian, H. and Daneshmand, F. 2013. Rules for optimal operation of reservoir-river-groundwater systems considering water quality targets: Application of M5P model. Water Resour. Manag., 27(8): 2771-2784.

Niroobakhsh, M., Musavi-Jahromi, S. H., Manshouri, M., and Sedghi, H. 2012. Prediction of water quality parameter in Jajrood River basin: application of multi-layer perceptron (MLP) perceptron and radial basis function networks of artificial neural networks (ANNs). African Journal of Agricultural Research, 7(29): 4131-4139.

Omer, N.H., 2019. Water quality parameters. Water quality-science, assessments and policy, p.18.

Pandhiani, S.M., Sihag, P., Shabri, A.B., Singh, B. and Pham, Q.B. 2020. Time-series prediction of streamflows of Malaysian rivers using data-driven techniques. J. Irrig. Drain. Eng., 146(7): 104020.

Prasad, V.V.D., Venkataramana, L.Y., Kumar, P.S. and Prasannamedha, G. 2020.Water quality analysis in a lake using deep learning methodology: prediction and validation. Int. J. Environ. Anal. Chem., 13: 45-56

Pratt, B. and Loizos, P. 1992. Choosing Research Methods: Data Collection for Development Workers (Vol. 7). Oxfam, Oxford, UK.

Rajaee, T., Ravansalar, M., Adamowski, J.F. and Deo, R.C.2018.A new approach to predict daily pH in rivers based on the "à trous" redundant wavelet transform Algorithm. Water, Air, & Soil Pollution, 229(3):1-18

Sakai, N., Mohamad, Z.F., Nasaruddin, A., Abd Kadir, S.N., Salleh,(ed.). 2018. Eco-Heart Index as a tool for community-based water quality monitoring and assessment Ecological Indicators, 91:38-46.

Siebert, S., Burke, J., Faures, J. M., Frenken, K.,(ed.). 2010. Groundwater use for irrigation–a global inventory. Hydrology and earth system sciences, 14(10):1863-1880.

Singh, K.P., Basant, N. and Gupta, S.2011.Support vector machines in water quality management. Analytica chimica acta, 703(2):152-162.

Tarke, P.D., Sarda, P.R. and Sadgir, P.A.2016. Performance of ANNs for prediction of TDS of Godavari river, India. International Journal of Engineering Research, 5(2):115-118.

Tchobanoglous, G. and Schroeder, E.E.1985. Water quality: characteristics, modeling, modification.

Tung, T.M. and Yaseen, Z.M., 2020. A survey on river water quality modelling using artificial intelligence models: 2000-2020. Journal of Hydrology, 585:124670.

Voza, D. and Vuković M.2018. The assessment and prediction of temporal variations in surface water quality-a case study. Environmental monitoring and assessment, 190(7):1-16.

Wang, K. 2003. Intelligent condition monitoring and diagnosis systems: a computational intelligence approach (Vol. 93). IOS press.

Wu, D., Wang, H., Mohammed, H. and Seidu, R. 2019. Quality risk analysis for sustainable smart water supply using data perception. IEEE Transactions on Sustainable Computing, 5(3): 377-388.

Zadeh, L.A. 1996. Fuzzy sets. In Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zade,pp 394-432.

Zhang, J., Zhu, X., Yue, Y. and Wong, P.W. 2017. August. A real-time anomaly detection algorithm/or water quality data using dual time-moving windows. In 2017 Seventh international conference on innovative computing technology,pp. 36-41.