Available online at www.CivileJournal.org

Civil Engineering Journal

(E-ISSN: 2476-3055; ISSN: 2676-6957)

Vol. 10, No. 07, July, 2024

Artificial Intelligence for Application in Water Engineering: The Use of ANN to Determine Water Quality Index in Rivers

Rabah Ismail [1], Adnan Rawashdeh [2], Hashem Al-Mattarneh [3*], Randa Hatamleh [3], Dua'a B. Telfah [3], Aiman Jaradat [3]

[1] *Department of Civil Engineering, Jadara University, Irbid 21110, Jordan.*

[2] *Faculty of Computer Sciences and Information Technology, Yarmouk University, Irbid, 21163, Jordan.*

[3] *Department of Civil Engineering, Yarmouk University, Irbid 21163, Jordan.*

Received 12 April 2024; Revised 19 June 2024; Accepted 24 June 2024; Published 01 July 2024

Abstract

To improve water quality, total daily loads must be established, and this requires determining the quality of the water in rivers, storage tanks, ponds, and coastal areas. Current methods to evaluate water quality involve the collection of water samples for subsequent laboratory analysis. Although these technologies offer precise measurements for a specific location and time, they are expensive, time-consuming, and do not provide the continuous, temporal, or spatial conditions of water quality that are required for managing, assessing, and monitoring water quality. In order to calculate the water quality, the water quality index is modeled using artificial neural network models that incorporate feedforward neural network backpropagation neural networks and radial neural networks. The water quality index of Malaysia's Klang River was determined by training the artificial network using six major sub-quality parameters. Compared to the current method, the artificial neural network simplifies and expedites the computation of the water quality index. The artificial neural network method could provide a significant saving in terms of money and time while offering a robust assessment of water quality. The proposed method could also be used as an early warning system for pollution of water bodies. The best artificial neural network was the feedforward neural network with one hidden layer containing 5 neurons. Furthermore, conventional approaches for calculating the water quality index rely on empirical equations, often introducing a high degree of approximation and uncertainty into the results. Moreover, these equations cannot be applied when some parameters are not measured. In contrast, the artificial neural network methods and technique offer an efficient and straightforward process for estimating and creating prediction models for water quality index.

Keywords: Artificial Intelligence; Artificial Neural Network; Water Quality Index; River; Water Quality Class.

1. Introduction

Water plays a crucial role in sustaining human life and is considered the primary resource for various human activities, including industrial and agricultural endeavors [1]. Therefore, the assessment, monitoring, and preservation of water resources is of utmost importance, particularly in developing nations. Furthermore, water resource assessment is essential for effective river basin and water resources management [2]. Assessing water quality is a critical step in controlling water bodies, with data categorization, modeling, and analysis being key components of this process [3]. Given the significance of surface water for aquatic life support, water supply, recreation, fisheries, and transportation, managing surface water resources is paramount. Rivers, in particular, are

* Corresponding author: hashem.mattarneh@yu.edu.jo

http://dx.doi.org/10.28991/CEJ-2024-010-07-012

vital for agricultural, irrigation, industrial, residential, and other human activities [4]. Various factors impact the quality of river water, with biological and physico-chemical parameters commonly utilized in previous studies for assessing water quality [5]. The needs for determination of water quality, along with modeling, prediction, and analysis of river water quality, are increasing worldwide.

A study conducted by a team of scientists from Wageningen University and Research in the Netherlands predicts that by the 2050s, one-third of the world's rivers could face either water scarcity or pollution due to nitrogen. Their research examines the impact of ongoing climate change on the availability of clean water, highlighting how intensified agriculture and untreated sewage could contribute to water pollution and limit clean water access. These projections for 2050 underscore the urgent need for further research to develop a standardized water quality index (WQI) for rivers [6]. Consistent monitoring and enforcement of water quality standards are essential to hold polluters accountable and safeguard rivers. This involves regular testing of water quality, monitoring discharge levels, and enforcing regulations to restrict the release of pollutants into waterways. Ensuring that companies and individuals are held responsible for the consequences of their actions on rivers can incentivize them to embrace more sustainable practices [7].

The WQI is a method used to assess the quality of rivers, assigning a numerical value between 1 and 100. A WQI of 100 indicates pure water, while a value of 0 indicates heavily polluted water. This index is calculated by considering various sub-parameters such as the pH level, dissolved oxygen, biochemical oxygen demand, and several others [8–10]. This method is implemented in several countries worldwide, including Malaysia, Germany, Greece, Canada, China, and many others. However, the number of sub-parameters used to compute the river water quality index varies from one country to another. This process is time-consuming and requires numerous water samples and laboratory testing. In the subsequent step, the sub-parameters of water quality are normalized to convert them into dimensionless values ranging from 0 to 100. This normalization process involves the use of various empirical and statistical formulas, which necessitates extensive computational processing, which demands significant time, cost, and expertise [11, 12]. Subsequently, these normalized sub-parameters are utilized to determine the WQI of rivers using a multi-linear formula that assigns weights to each sub-parameter. These weights are determined by water quality experts. To address the drawbacks of this method, several researchers have explored alternative techniques. Artificial Intelligence (AI), particularly Artificial Neural Networks (ANN), has shown promise in resolving these issues. ANN has been successfully used to predict and determine similar factors in numerous engineering problems and has been applied in various studies in the field of water resources and environmental applications [13, 14].

This paper explores the potential application of ANN in predicting and determining the WQI of rivers. The study utilizes available water quality data for various sub-parameters to train several ANN models, aiming to identify the most effective model for determining the WQI of rivers. The selected ANN model will then undergo validation and testing using another dataset to assess its accuracy and quality of fit in determining the WQI of rivers. To demonstrate the applicability of the approach, a case study of the Klang River, the largest river in Malaysia, is conducted. The results obtained from this study can be extrapolated to benefit other rivers in Malaysia and worldwide. The following two sections will delve deeper into the existing literature concerning the WQI of rivers and the application of ANN.

2. Water Quality Index (WQI)

Statistical tools and electromagnetic sensors are used in many areas to estimate and predict material properties and quality [15–21]. They are also used to determine water quality in general, and in rivers in particular [15, 22-29]. Water quality is typically assessed by measuring and evaluating a wide range of parameters such as temperature, pH, electrical conductivity, turbidity, and concentrations of a variety of pollutants, including pathogens, nutrients, organic matter, and heavy metals. Different approaches are employed to assess river water quality, with one such method being the determination of WQI. WQI serves as a statistical tool to condense extensive water quality information into a single numerical value [15]. This index provides a concise representation of complex data collected from water bodies, with values ranging between 0 and 100. A higher WQI value signifies good water quality, whereas a lower value indicates poor water quality [30].

Six key factors of water quality, namely biochemical oxygen demand (BOD), dissolved oxygen (DO), suspended solids (SS), chemical oxygen demand (COD), ammoniacal nitrogen (AN), and pH, are typically utilized to determine WQI [31]. These six parameters are converted to a dimensionless unit known as the sub water quality index (Sub-WQI). This Sub-WQI normalized to be a numerical value between 0 and 100. A high value indicates high water quality. The Department of Environment in Malaysia developed the Sub-WQI in 2005. The Sub-WQI in Malaysia is shown in Table 1.

Table 1. The sub-index calculation formula for the local WQI [32, 33]

Parameter of water quality	Value	Parameter of water quality using WQP_{SI}
DO (%Saturation) = [DO (mg/L)*12.795]-0.05	$X_1 \leq 8$	$DO_{SI} = 0.0$
	$8 < X_1 < 92$	$DO_{SI} = -0.395+0.03X_1^2-0.0002X_1^3$
	$X_1 \geq 92$	$DO_{SI} = 100$
BOD (mg/L)	$X_2 \leq 5$	$BOD_{SI} = 100.4-4.23X_2$
	$X_2 > 5$	$BOD_{SI} = (108e^{-0.055\ X2})-0.1X_2$
COD (mg/L)	$X_3 \leq 20$	$COD_{SI} = 99.1-1.33X_3$
	$X_3 > 20$	$COD_{SI} = (103e^{-0.0157\ X3})-0.04X_3$
AN (NH_3-N) (mg/L)	$X_4 \leq 0.3$	$AN_{SI} = 100.5-105X_4$
	$0.3< X_4 < 4$	$AN_{SI} = (94e^{-0.0573X4})-5\lvert X_4-2\rvert$
	$X_4 \geq 4$	$AN_{SI} = 0$
SS (mg/L)	$X_5 \leq 100$	$SS_{SI} = (97.5e^{-0.00676\ X5})+0.05X_5$
	$100< X_5 < 1000$	$SS_{SI} = (71e^{-0.0016\ X5})-0.015$
	$X_5 \geq 1000$	$SS_{SI} = 0$
pH	$X_6 < 5.5$	$pH_{SI} = 17.2-17.2X_6^2+5.02X_6^3$
	$5.5\leq X_6 < 7$	$pH_{SI} = -242+95.5X_6^2-6.67X_6^3$
	$7\leq X_6 < 8.75$	$pH_{SI} = -181+82.4X_6^2-6.05X_6^3$
	$X_6 \geq 8.75$	$pH_{SI} = 536-77X_6^2+2.76X_6^3$

Over all water quality can be determined using several Sub-WQI and determine a WQI using weight to each WQP_{SI}. WQI can be calculated using Equation 1.

$$WQI = \sum_{i=1}^{k} w_i (WQP_{SI})_i \tag{1}$$

where w is the weight between 0 and 1 developed using experts, WQP_{SI} is a single water quality parameter such as SS and COD. Many countries utilize WQI methods to assess the overall status of their rivers. While the underlying concepts are similar, these indices may vary from one country to another [32, 33]. Where w is the weight between 0 and 1 developed using experts, Sub-WQI is a single water quality parameter such as DO and pH. WQI methods have been used in several countries around the globe such as USA [34], Turkey [35], South Korea [36], Spain [37], Taiwan [38]. In addition, WQI varies from country to country depending on local climate and other factors.

The Malaysian Department of Environment (DOE) initiated the adoption of the WQI to classify the level of pollution of Malaysian rivers in 1974 [38]. This is a set of water quality guidelines, which classify water into water quality categories depending on intended use, such as human consumption, recreational purposes, domestic uses and aquatic life. and so on. The DOE-WQI is the equivalent of an opinion poll. A group of experts were consulted on the selection of criteria, and the weight to be assigned to each parameter. Equation 2 describes the formula used to calculate WQI. All parameters used in Equation 2 are Sub-WQI and it is dimensionless unit not in its original units.

$$WQI = 0.22(DO) + 0.19(BOD) + 0.16(COD) + 0.15(AN + 0.16(SS) + 0.12(pH) \tag{2}$$

Based on the value of WQI the water could be classified into water quality class (WQC). Malaysia water quality class is given in Table 2 [32, 38, 39].

Table 2. Water quality classification according DOE in Malaysia

Water quality class (WQC)	Water condition	Range WQI	Water use
Class-I	Very good	Above 93%	Preservation of the natural world Water-Supply-I: Almost no treatment is required. Fishery-I: Extremely delicate aquatic species IIA: Water-Supply II- Traditional treatment is necessary
Class-II	Good	From 77% to 93%	Fishery-II: Delicate aquatic species IIB: Usage for recreation that involves physical contact Water-Supply-III: Long-term care and treatment is necessary
Class-III	Average	From 52% to 77%	Fishery- III: Adaptable, common, and valuable specie. Drinking livestock
Class-IV	Polluted	From 31% to 52%	Irrigation
Class-V	Very Polluted	Less than 31%	None of the aforementioned

3. Artificial Neural Network (ANN)

Recently, AI techniques have gained recognition as effective tools for modeling complex nonlinear phenomena in water systems and hydrology. Researchers have increasingly utilized ANN to model various engineering problems in several areas including material characterization, prediction of material properties like concrete. The application of ANN in water resource variables applications is also promising [40, 41]. Additionally, machine learning has been employed to investigate the impact of river discharge on the quality of water bodies and coastal water [42-47]. The use of AI models in water engineering management has grown in popularity, utilizing a variety of methods including neural networks, adaptive neuro-fuzzy inference techniques, and hybrid models [48-56]. ANNs, in particular, are information-driven modeling techniques with a flexible statistical structure capable of capturing complex and nonlinear relationships between input and output datasets without prior knowledge of the underlying phenomena. These networks typically consist of three or more layers: an input layer, hidden layers, and an output layer. The input layer solely transmits input information to the neurons of the primary hidden layer [46, 49].

The output layer plays a crucial role in generating outputs corresponding to specified inputs. Meanwhile, the hidden layers, which can range from one to multiple layers, serve as sets of feature detectors. Determining the appropriate network framework is a critical and complex task in system modeling [45]. Figure 1 illustrates a schematic representation of a general 3-layer ANNs model. There are various approaches to applying ANNs, and identifying the best solution poses a challenge, as it requires systematic testing of numerous possibilities. Various frameworks and types of ANNs exist, including Probabilistic Neural Networks, Regression Neural Network, Radial Functions Network (RFN), Multilayer Perceptron Networks, Back Propagation Network (BPN), and Feedforward Neural Network (FNN) [46]. Among these, FNN, BPN and RFN are the most commonly used [49, 52].

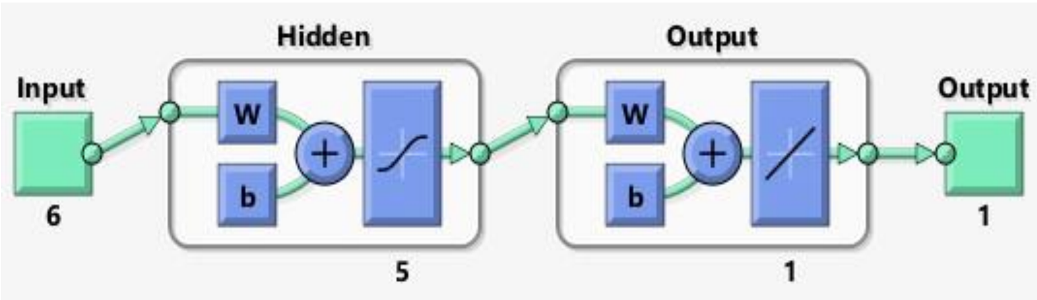


Figure 1. A schematic representation of a general 3-layer ANNs model

The main objective of this study is to establish an Artificial Neural Network (ANN) model to explore the relationship between the Water Quality Index (WQI) and the six key parameters needed to calculate WQI and WQC Understanding the correlation between WQI and these six parameters is pivotal for developing a comprehensive model aimed at managing river water quality and evaluating the overall water quality status. Furthermore, the study aims to devise an input approach for the new model that enables direct computation of WQI from input parameters, thereby eliminating the necessity for parameter indices when certain parameters are not available. Furthermore, conventional approaches for calculating WQI rely on empirical equations, often introducing a high degree of approximation and uncertainty into the results. Moreover, these equations cannot be applied when some parameters are not measured. In contrast, the ANN technique offers an efficient and straightforward method for calculate and create mods for WQI. This research provides a significant contribution to the water engineering fields by creating a model to elucidate the relationship between quality parameters and WQI in the Klang River employing ANNs.

4. Research Methodology

The research methodology used in this study including the study area, water samples, water quality tests, determination of statistical empirical WQI, and modeling of three different ANN techniques are presented in this section. A flowchart of all methods and tests conducted in this study are shown in Figure 2.

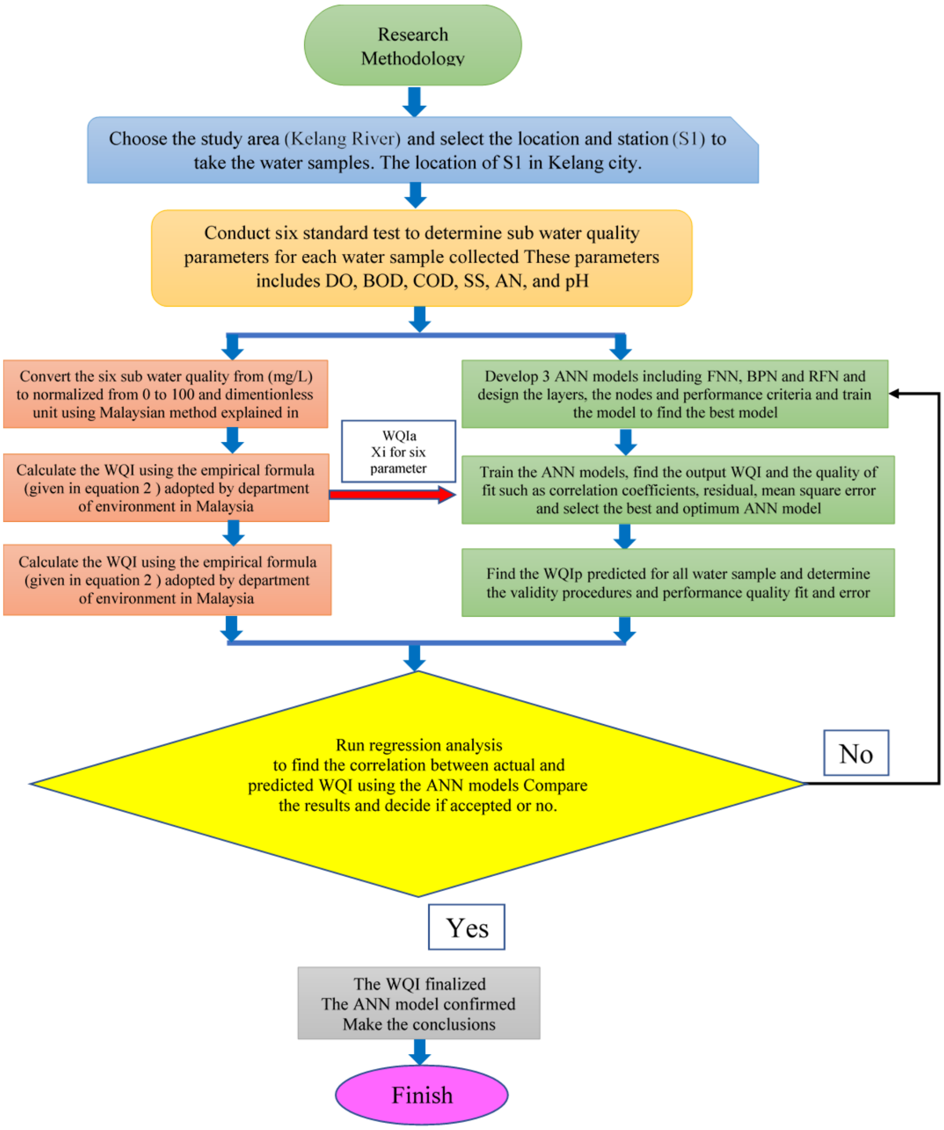


Figure 2. Flowchart explains the methods and tests conducted and used in this study

4.1. Location of Study

This study was undertaken to explore the application of ANN for modeling and predicting the WQI of the Klang River in Malaysia. Situated in the Selangor state of Malaysia, the Klang River is the focal point of this investigation. Figure 3 provides an overview of the study area, depicting the location of the river within the map of Malaysia. Additionally, Figure 3 highlights the specific stations within Klang city where water samples were collected and subjected to testing and analysis.

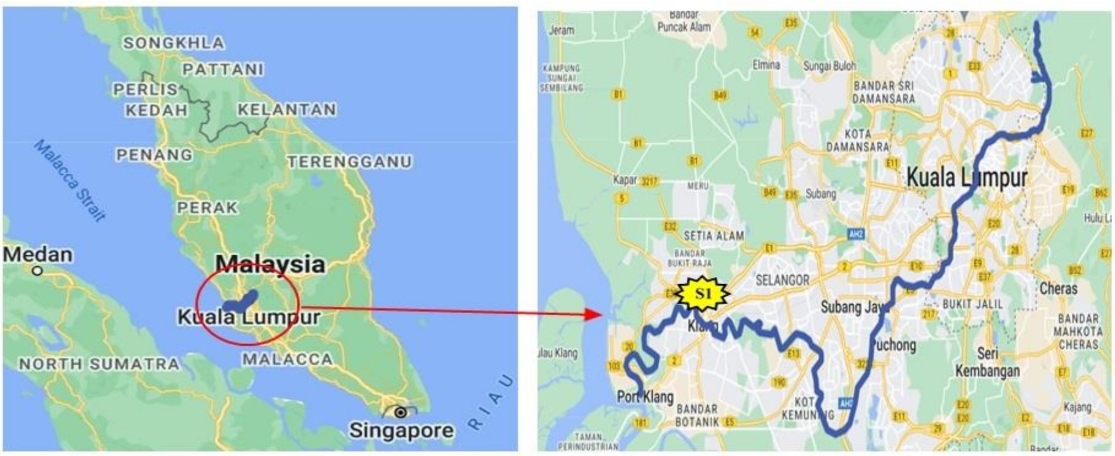


Figure 3. Kelang River location in Malaysia and station S1 where water samples were collected

4.2. Water Quality Tests of Collected Samples

All water samples were collected from station S1 as located in Figure 3. After collected the water samples several standard tests were conducted to determine the six water quality parameters including DO, pH, SS, BOD, COD and AN. Furthermore, the six sub index parameters were calculated and normalized to be from 0 to 100 and dimensionless as explained in Table 1. Subsequently, the WQI was calculated using Equation 2. The actual WQI calculated using Equation 2 for each water sample are presented in Table 3.

Table 3. Water quality sub index and WQI of Kelang River calculated using method of Malaysian DOE

Water sample Code and Number	DO (mg/L)	BOD (mg/L)	COD (mg/L)	SS (mg/L)	AN (mg/L)	pH	Actual WQI
WS 1	2.45	27	12	36	0.83	6.30	6.30
WS 2	2.35	25	11	45	0.83	6.27	6.27
WS 3	2.33	29	13	184	0.83	6.27	6.27
WS 4	2.31	24	11	27	0.78	6.09	6.09
WS 5	2.31	25	11	34	0.92	6.24	6.24
WS 6	2.30	25	11	26	0.85	6.26	6.26
WS 7	2.31	32	14	247	0.85	6.28	6.28
WS 8	2.28	38	17	243	0.83	6.32	6.32
WS 9	2.22	26	12	306	0.83	6.26	6.26
WS 10	2.24	26	12	36	0.83	6.26	6.26
WS 11	2.25	24	11	180	0.83	6.32	6.32
WS 12	2.25	32	14	211	0.83	6.32	6.32
WS 13	2.26	25	11	27	0.84	6.27	6.27
WS 14	2.78	33	15	143	0.94	6.08	6.08
WS 15	2.36	37	17	380	0.84	6.28	6.28
WS 16	2.33	30	13	162	0.83	6.28	6.28
WS 17	2.34	28	13	65	0.83	6.29	6.29
WS 18	2.32	24	11	27	0.78	6.12	6.12
WS 19	2.29	40	18	180	0.83	6.27	6.27
WS 20	2.34	32	14	253	0.78	6.09	6.09
WS 21	2.34	38	17	146	0.90	6.09	6.09
WS 22	2.30	49	22	238	0.78	6.12	6.12

4.3. ANN Model which will be Used to Predict WQI

Three ANN tools were used to model the WQI these tools including radial functions network (RFN), back propagation network (BPN), and Feedforward Neural Network (FNN). Several trials were conducted, and it was found that the best ANN method was FNN. For this reason, the results of the other two ANN tools will not be presented and the discussion will focus on FNN only. The FNN were developed using MATLAB environment.

4.4. Performance Criteria for Selecting the Best ANN Models

To determine the best Artificial Neural Network (ANN) model for predicting Water Quality Index (WQI), several performance criteria were taken into account. These included the number of trials during learning, correlation coefficients, mean square error, and residual error. Additionally, various activation functions were employed in the proposed Feedforward Neural Network (FNN) models. These functions included linear, hyperbolic tangent (tanh), logistic or sigmoid, and Gaussian functions (see Figure 4). The primary objective of an activation function is to introduce non-linearity to a neuron's output. It achieves this by computing the weighted sum and then applying bias, determining whether a neuron should be activated or not. The water sample data was divided into three groups for training, validation, and testing purposes. Specifically, the first group comprised 70% of the samples and was used for training the FNN models. The second group, accounting for 15% of the total water samples, served as a validation set for evaluating the performance of the FNN models. The remaining 15% of the samples were allocated to the third group, which was used to assess the degree of fitting of the FNN models. This systematic division allowed for robust evaluation and optimization of the ANN models for WQI prediction.

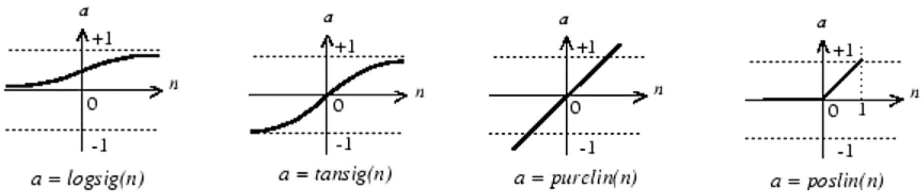


Figure 4. Various activation functions used in ANN

5. Results and Discussion

Water samples collected from Station 1 (S1) underwent comprehensive analysis and testing according to appropriate standards. Six key quality parameters, namely Dissolved Oxygen (DO), Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Suspended Solids (SS), Ammoniacal Nitrogen (AN), and pH, were determined. Subsequently, utilizing Malaysian standard protocols, the sub-quality index parameters (WQP_{SI}) were calculated and detailed in earlier sections. The study then progressed to developing ANN models within the MATLAB environment. All data regarding the WQI for the Klang River and its corresponding sub water quality parameters (WQP_{SI}) were imported into MATLAB for sophisticated modeling and analysis. Three ANN methodologies were explored: radial functions network (RFN), back propagation network (BPN), and Feedforward Neural Network (FNN). Results indicated that the predicted WQI obtained from RFN and BPN models exhibited relatively lower performance compared to those obtained through the FNN. The correlation coefficient (R) between the predicted WQI and the actual WQI ranged from 0.91 to 0.93 for RFN and BPN models. Consequently, this paper primarily focuses on the outcomes derived from the feedforward neural network (FNN). This finding is supported by two recent studies conducted in the Johor River in southern Malaysia [11] and the Klang river basin [53].

Figure 5 illustrates the FNN model utilized in this study to predict WQI in rivers, employing AI tools. The FNN model comprises three layers: an input layer, a hidden layer, and an output layer. The input layer consists of six nodes, each corresponding to one of the six sub-index water quality parameters. These parameters are used as input data for the neural network. The hidden layer, containing five neurons in this specific configuration, serves as an intermediary processing stage. The number of neurons in the hidden layer can be adjusted to optimize the model's performance, experimenting with different quantities until the optimal configuration is determined. Finally, the output layer consists of one node, representing the predicted WQI for the river. The output is calculated through the neural network's processing of the input data, providing an estimation of the overall WQI based on the given parameters.

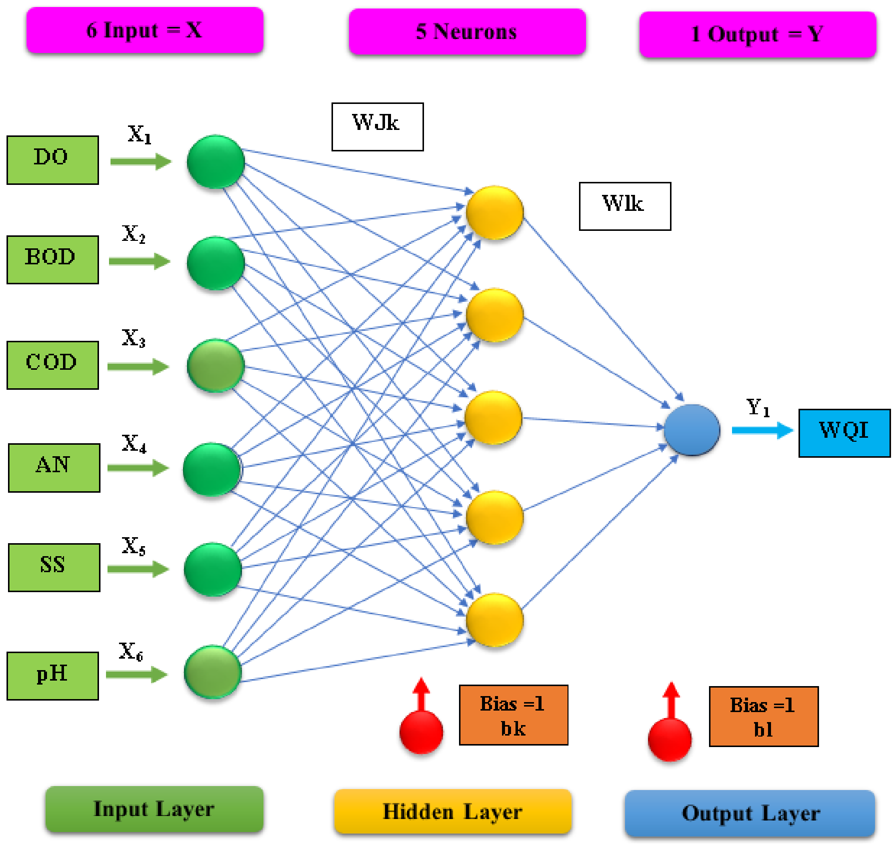


Figure 5. Best feedforward neural network (FNN) contains input layer, hidden layer with 5 neuron

5.1. ANN using FNN Model with Three Layers and 5 Neurons in the Hidden Layer

Several trials were investigated to determine the best feedforward neural network model. These trials use different numbers of hidden layers, different numbers of neurons in the hidden layer, and one output layer has a single outcome represented the WQI. In addition, several activation functions also used in the proposed FNN models including linear, hyperbolic tangent (tanh), Logistic or sigmoid and Gaussian functions. The best FNN model has one hidden layer contains 5 neurons. The activation function in the hidden layer was Logistic or sigmoid and the activation function in the output layer was linear. The next good quality FNN model was with 9 neurons. The results of FNN model with 5 neurons are presented in Figures 6 to 8.

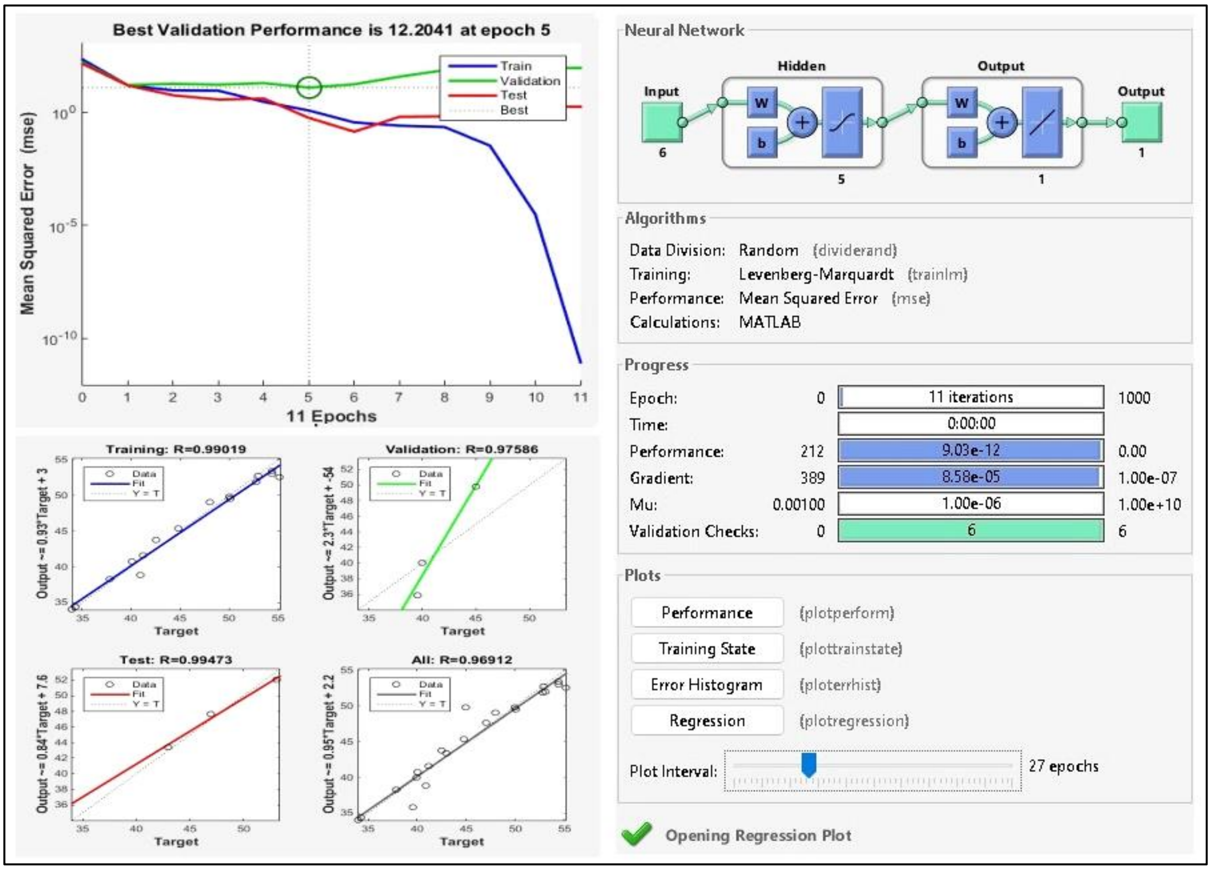


Figure 6. Performance of feedforward neural network model with 5 neurons in hidden layer (first run)

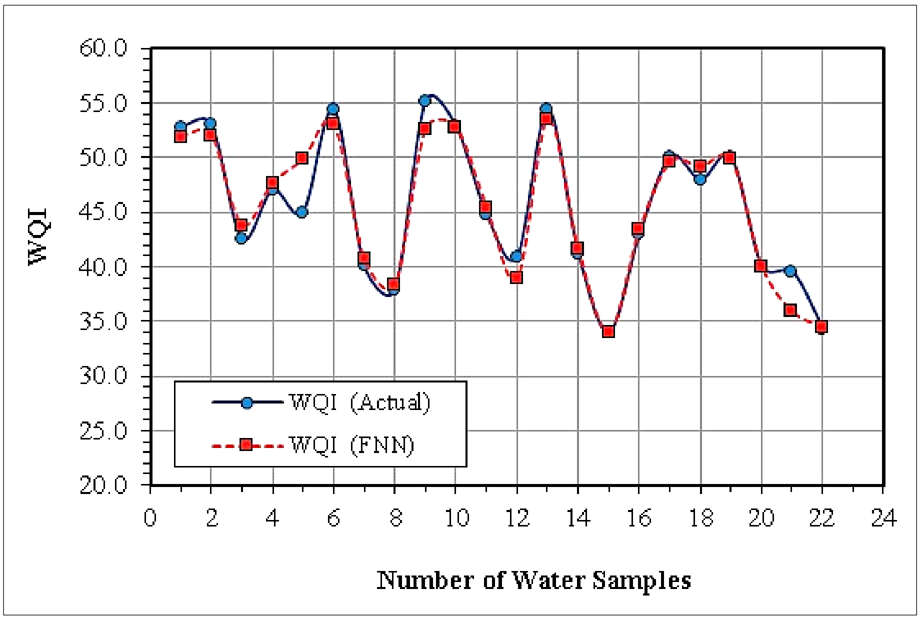
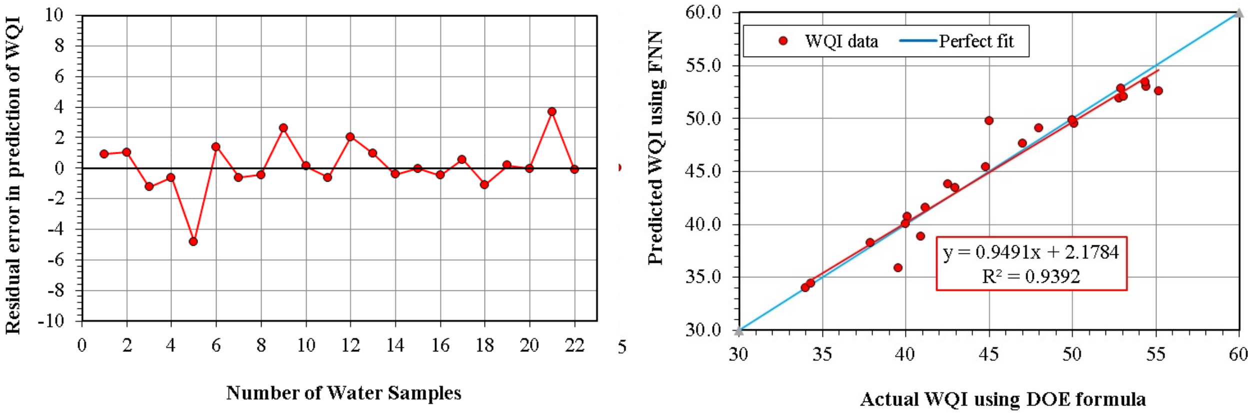


Figure 7. Quality of fit and performance of FNN model with 5 neurons in hidden layer (first run)



(a) Residual error (b) Predicted WQI versus actual

Figure 8. Performance of feedforward neural network model with 5 neurons in hidden layer (first run)

The best Feedforward Neural Network model, utilizing 5 neurons in the hidden layer, successfully achieved the target with an impressive precision of 1.0×10^{-7}. This model was trained using a dataset comprising 16 water samples, and it reached the target within 11 epochs of training. Notably, the correlation coefficient between the predicted WQI and the actual WQI of the water samples used for training was exceptionally high, at 0.99, indicating a nearly perfect fit of the model to the training data. Furthermore, when validating the model, the correlation coefficients between the predicted and actual WQI for the validation samples and the test samples were also notably high, at 0.97 and 0.99, respectively. This suggests strong performance and generalization capability of the model beyond the training dataset. Moreover, the parameter quality performance evaluation of the model revealed no significant difference between the predicted WQI and the actual WQI for the 6 samples used to test the model's predictability, further confirming the robustness and accuracy of the FNN model in predicting water quality indices.

The predictability assessment conducted using the best model for 22 water samples, as illustrated in Figure 7 and 8, demonstrates that the predicted WQI closely aligns with the actual WQI, with minimal non-significant differences observed. Additionally, the residual errors in the predicted values of WQI for the water samples are predominantly less than 3%, indicating a high degree of accuracy in the model's predictions. Furthermore, the quality fit correlation between the predicted and actual WQI was calculated to be 0.969, with a square correlation coefficient of 0.9392. These metrics signify a strong relationship between the predicted and actual values, further validating the effectiveness of the ANN technique in computing and modeling WQI. Moreover, employing the ANN technique offers an efficient and straightforward approach to computing and modeling WQI compared to the existing method developed by the Malaysian Department of Environment. The ANN simplifies and accelerates the computation process, potentially streamlining water quality assessment procedures.

Another investigation was conducted to assess the sensitivity of the best FNN model by varying the samples used for training the ANN model. The results, depicted in Figure 9, yielded similar quality and performance parameters to those observed previously. Notably, the model achieved the minimum target error of 1.0×10^{-7} at epoch 12, with a correlation coefficient (R) of 0.937. Furthermore, the FNN model demonstrated its utility in predicting the WQI of water samples with missing or unavailable data for some sub-quality index parameters, such as pH or BOD. Even in cases where certain data were missing, the FNN model exhibited the capability to predict the WQI of these samples at an acceptable level compared to the actual values. This capability may prove invaluable to practicing engineers who encounter situations where some data are missing or unavailable for calculating the WQI using conventional methods and formulas.

5.2. ANN using FNN Model with Three Layers and 9 Neurons in the Hidden Layer

Figure 10 depicts the FNN model with 9 neurons in its hidden layer. This model exhibits good predictability of WQI, although it is slightly less accurate compared to the FNN model with 5 neurons. The results obtained from this model indicate that the targeted level of performance was achieved at epoch 7, with a correlation coefficient of 0.954.

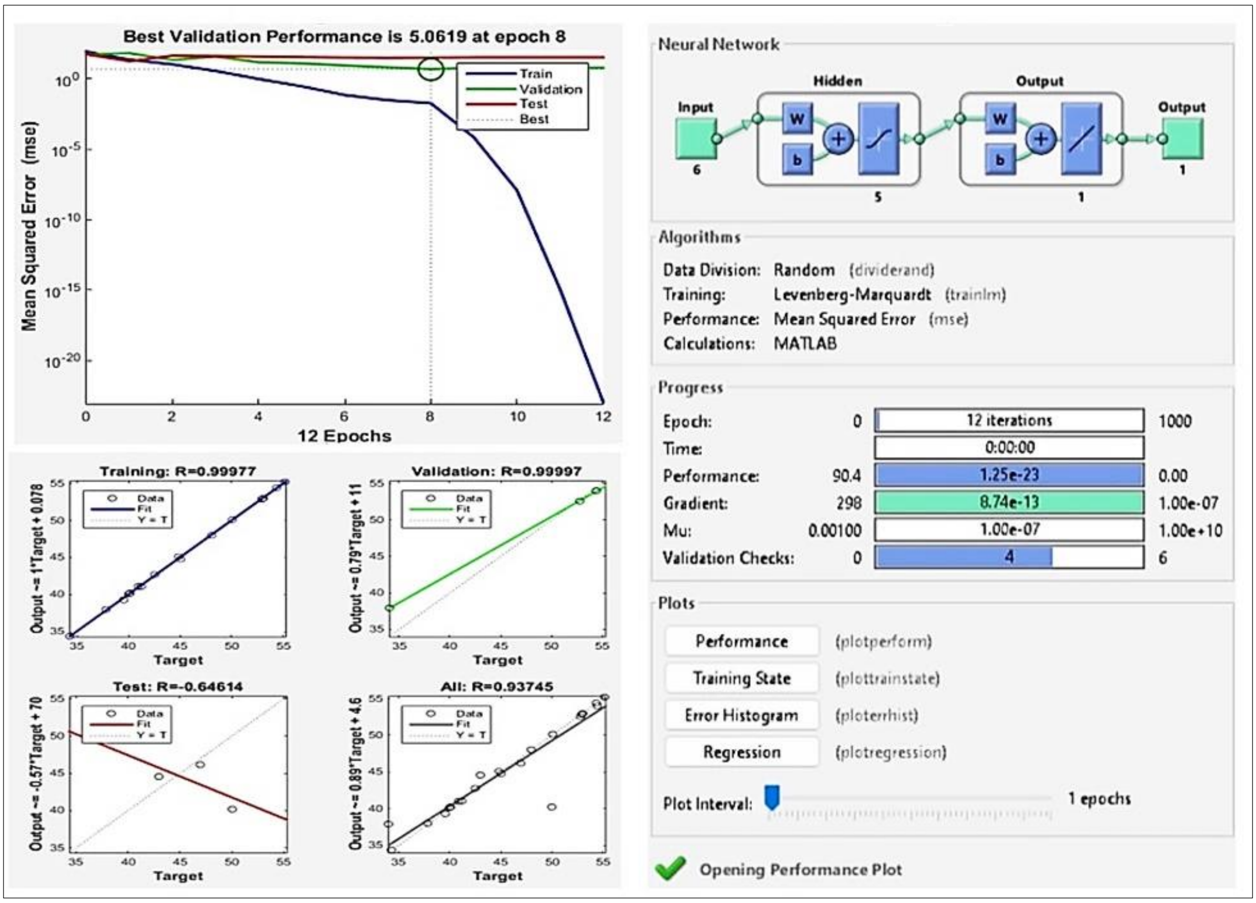


Figure 9. Performance of feedforward neural network model with 5 neurons in hidden layer (second run)

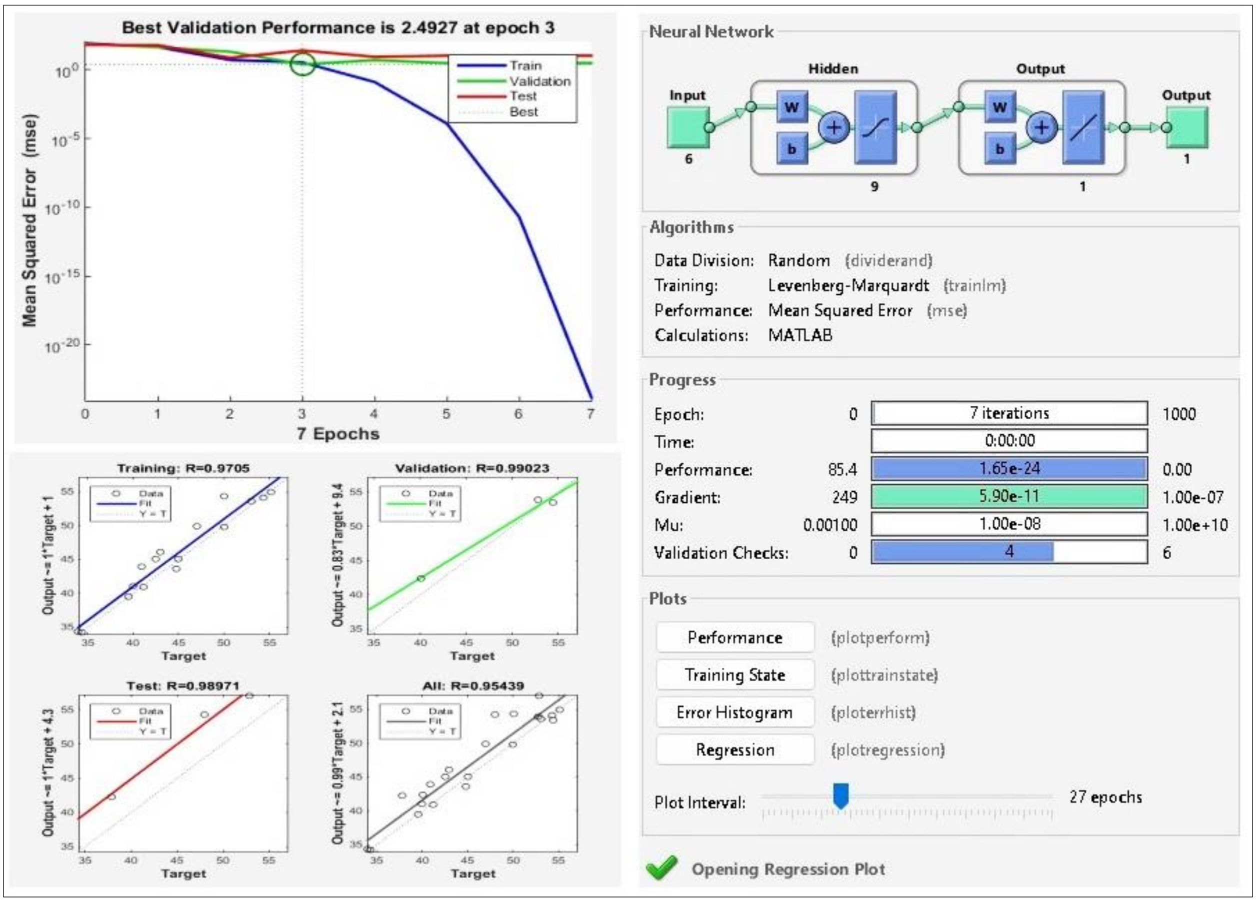


Figure 10. Performance of feedforward neural network model with 9 neurons in hidden layer

Moreover, the model successfully predicted the WQI of 4 out of the 6 water samples used to test the model without any significant difference between the actual and predicted values. For the remaining two samples, although the quality of the predicted WQI was slightly lower, it still fell within acceptable limits.

Overall, while the FNN model with 9 neurons may not exhibit the same level of precision as the model with 5 neurons, it still demonstrates commendable performance in predicting WQI, providing valuable insights for water quality assessment.

6. Conclusion

This study employed various artificial neural network methods, including backpropagation networks (BPN), radial neural networks (RNN), and feedforward neural networks, to predict and model the water quality index for rivers. All three neural network architectures exhibited satisfactory performance in determining WQI in rivers. Among the three artificial neural network (ANN) models examined in this study, the feedforward neural network (FNN) demonstrated the highest performance with a correlation coefficient of 0.99. In contrast, the correlation coefficients for the other two ANN models, namely, the Bayesian neural network (BPN) and the recurrent neural network (RNN), were 0.91 and 0.93, respectively. Therefore, the FNN emerges as the most suitable model for estimating the Water Quality Index (WQI) in rivers.

However, the feedforward neural network (FNN) outperformed the others in terms of learning time, strong correlations, and lower residual error in predicting WQI values. The most favorable results for WQI in rivers were achieved using two FNN models. The first had a hidden layer containing 5 neurons, while the other contained 9 neurons. Notably, these outcomes were based on limited data obtained from a single station along the Klang River in Klang City. To bolster the application of ANN for WQI assessment in rivers across Malaysia and other regions, it is imperative to gather more comprehensive water quality data from multiple stations and different time points to train the ANN models effectively. Additionally, incorporating data from other rivers would contribute to enhancing the accuracy and predictive capability of these models.

7. Declarations

7.1. Author Contributions

Conceptualization, R.I., A.R., H.A., and A.J.; methodology, R.I., A.R., and H.A.; software, A.R.; validation, H.A., R.I., R.H., and D.B.T.; formal analysis, R.I., D.B.T., R.H., and A.J.; data curation, R.I.; writing—original draft preparation, R.I., A.R., H.A., R.H., D.B.T., and A.J.; writing—review and editing, R.I., A.R., H.A., R.H., D.B.T., and A.J. All authors have read and agreed to the published version of the manuscript.

7.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

7.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

7.4. Acknowledgements

The authors thank their universities; Yarmouk University, Jordan, University Tenaga National (UNITEN), Malaysia, and Jadara University, Jordan for using the lab facilities.

7.5. Conflicts of Interest

The authors declare no conflict of interest.

8. References

[1] Wan Mohtar, W. H. M., Abdul Maulud, K. N., Muhammad, N. S., Sharil, S., & Yaseen, Z. M. (2019). Spatial and temporal risk quotient-based river assessment for water resources management. Environmental Pollution, 248, 133–144. doi:10.1016/j.envpol.2019.02.011.

[2] Caddis, B., Nielsen, C., Hong, W., Anun Tahir, P., & Yenn Teo, F. (2012). Guidelines for floodplain development – a Malaysian case study. International Journal of River Basin Management, 10(2), 161–170. doi:10.1080/15715124.2012.688750.

[3] Sharma, A., Naidu, M., & Sargaonkar, A. (2013). Development of computer automated decision support system for surface water quality assessment. Computers & Geosciences, 51, 129–134. doi:10.1016/j.cageo.2012.09.007.

[4] Luo, P., He, B., Takara, K., Razafindrabe, B. H. N., Nover, D., & Yamashiki, Y. (2011). Spatiotemporal trend analysis of recent river water quality conditions in Japan. Journal of Environmental Monitoring, 13(10), 2819–2829. doi:10.1039/c1em10339c.

[5] Avvannavar, S. M., & Shrihari, S. (2007). Evaluation of water quality index for drinking purposes for river Netravathi, Mangalore, South India. Environmental Monitoring and Assessment, 143(1–3), 279–290. doi:10.1007/s10661-007-9977-7.

[6] Wang, M., Bodirsky, B. L., Rijneveld, R., Beier, F., Bak, M. P., Batool, M., Droppers, B., Popp, A., van Vliet, M. T. H., & Strokal, M. (2024). A triple increase in global river basins with water scarcity due to future pollution. Nature Communications, 15(1), 880. doi:10.1038/s41467-024-44947-3.

[7] Mohan Raj, K., & Vairavel, K. S. (2024). Quality index metrics with Bi-GRU-based water quality prediction. Energy Sources, Part A: Recovery, Utilization and Environmental Effects, 46(1), 171–187. doi:10.1080/15567036.2023.2272658.

[8] Chidiac, S., El Najjar, P., Ouaini, N., El Rayess, Y., & El Azzi, D. (2023). A comprehensive review of water quality indices (WQIs): history, models, attempts and perspectives. Reviews in Environmental Science and Biotechnology, 22(2), 349–395. doi:10.1007/s11157-023-09650-7.

[9] Murivhami, T., Tartibu, L. K., & Olayode, I. O. (2023). Applications of Artificial Neural Network in the Prediction of Water Quality Index. 2023 14th International Conference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT), Cape Town, South Africa. doi:10.1109/icmimt59138.2023.10199388.

[10] Mamat, N., Mohd Razali, S. F., & Hamzah, F. B. (2023). Enhancement of water quality index prediction using support vector machine with sensitivity analysis. Frontiers in Environmental Science, 10, 1061835. doi:10.3389/fenvs.2022.1061835.

[11] Sidek, L. M., Mohiyaden, H. A., Marufuzzaman, M., Noh, N. S. M., Heddam, S., Ehteram, M., Kisi, O., & Sammen, S. S. (2024). Developing an ensembled machine learning model for predicting water quality index in Johor River Basin. Environmental Sciences Europe, 36(1), 67. doi:10.1186/s12302-024-00897-7.

[12] Diop, M., Mall, I., Sonko, E. H. M., Diop, T., Badji, L., & Mbow, C. (2023). Development and Application of Water Quality Index (WQI) for the Evaluation of the Physico-Chemical Quality of Groundwater in Gold Mining Areas of Southeastern Senegal. Journal of Water Resource and Protection, 15(02), 33–50. doi:10.4236/jwarp.2023.152003.

[13] Muthukumaran, V., & Vinoth kumar, V. (2022). Water Quality Index Process Using Artificial Neural Networks. International Journal of Information Technology, Research and Applications, 1(1), 33–37. doi:10.59461/ijitra.v1i1.12.

[14] Sekarlangit, S., Widodo, C. E., & Tarno, T. (2024). A Comparison of Artificial Neural Network (ANN) and Long Short-Term Memory (LSTM) in River Water Quality Prediction. International Journal of Current Science Research and Review, 7(4), 2000-2005. doi:10.47191/ijcsrr/v7-i4-02.

[15] Štambuk-Giljanović, N. (1999). Water quality evaluation by index in Dalmatia. Water Research, 33(16), 3423–3440. doi:10.1016/S0043-1354(99)00063-9.

[16] Tao, H., Bobaker, A. M., Ramal, M. M., Yaseen, Z. M., Hossain, M. S., & Shahid, S. (2019). Determination of biochemical oxygen demand and dissolved oxygen for semi-arid river environment: application of soft computing models. Environmental Science and Pollution Research, 26(1), 923–937. doi:10.1007/s11356-018-3663-x.

[17] Al-Mattarneh, H., & Alwadie, A. (2016). Development of Low Frequency Dielectric Cell for Water Quality Application. Procedia Engineering, 148, 687–693. doi:10.1016/j.proeng.2016.06.554.

[18] Al-Mattarneh, H., Ismail, R., Nuruddin, M., Shafiq, N., & Dahim, M. (2016). Characterization of Pb and Cd contaminated sandy soil by dielectric means. Civil, Offshore and Environmental Engineering, 327–330.

[19] Fan, X., Lv, S., Xia, C., Ge, D., Liu, C., & Lu, W. (2024). Strength prediction of asphalt mixture under interactive conditions based on BPNN and SVM. Case Studies in Construction Materials, e03489. doi:10.1016/j.cscm.2024.e03489.

[20] Dahim, M., Abuaddous, M., Al-Mattarneh, H., Rawashdeh, A., & Ismail, R. (2021). Enhancement of road pavement material using conventional and nano-crude oil fly ash. Applied Nanoscience (Switzerland), 11(10), 2517–2524. doi:10.1007/s13204-021-02103-z.

[21] Abdullahi, M., Al-Mattarneh, H. M. A., & Mohammed, B. S. (2009). Statistical modeling of lightweight concrete mixtures. European Journal of Scientific Research, 31(1), 124–131.

[22] Hameed, M., Sharqi, S. S., Yaseen, Z. M., Afan, H. A., Hussain, A., & Elshafie, A. (2017). Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia. Neural Computing and Applications, 28(1), 893–905. doi:10.1007/s00521-016-2404-7.

[23] Tiyasha, Tung, T. M., & Yaseen, Z. M. (2020). A survey on river water quality modelling using artificial intelligence models: 2000–2020. Journal of Hydrology, 585, 124670. doi:10.1016/j.jhydrol.2020.124670.

[24] Ho, J. Y., Afan, H. A., El-Shafie, A. H., Koting, S. B., Mohd, N. S., Jaafar, W. Z. B., Lai Sai, H., Malek, M. A., Ahmed, A. N., Mohtar, W. H. M. W., Elshorbagy, A., & El-Shafie, A. (2019). Towards a time and cost-effective approach to water quality index class prediction. Journal of Hydrology, 575, 148–165. doi:10.1016/j.jhydrol.2019.05.016.

[25] Ismail, R., Dahim, M., Jaradat, A., Hatamleh, R., Telfah, D., Abuaddous, M., & Al-Mattarneh, H. (2021). Field Dielectric Sensor for Soil Pollution Application. IOP Conference Series: Earth and Environmental Science, 801(1), 012003. doi:10.1088/1755-1315/801/1/012003.

[26] Zain, M. F. M., Karim, M. R., Islam, M. N., Hossain, M. M., Jamil, M., & Al-Mattarneh, H. M. A. (2015). Prediction of strength and slump of silica fume incorporated high-performance concrete. Asian Journal of Scientific Research, 8(3), 264–277. doi:10.3923/ajsr.2015.264.277.

[27] Abdullahi, M., Al-Mattarneh, H. M. A., & Mohammed, B. S. (2009). A matlab program for diagnosis and adjustment of mix proportions of structural lightweight concrete. European Journal of Scientific Research, 31(1), 106–123.

[28] Dahim, M., Abuaddous, M., Ismail, R., Al-Mattarneh, H., & Jaradat, A. (2020). Using a dielectric capacitance cell to determine the dielectric properties of pure sand artificially contaminated with Pb, Cd, Fe, and Zn. Applied and Environmental Soil Science, 2020(8838054). doi:10.1155/2020/8838054.

[29] Ermeey, A. K., Ghodgaonkar, D. K., & Al-Mattarneh, H. M. A. (2003). Three probe reflectometer algorithms for complex coefficient measurements of water quality at microwave frequencies. Asia-Pacific Conference on Applied Electromagnetics, 2003. APACE 2003. doi:10.1109/apace.2003.1234481.

[30] Lee Yoot Khuan, Hamzah, N., & Jailani, R. (2002). Prediction of water quality index (WQI) based on artificial neural network (ANN). Student Conference on Research and Development. doi:10.1109/scored.2002.1033081.

[31] Almattarneh, H., Alwadie, A., Malkawi, A., & Nuruddin, M. F. (2014). A novel method for monitoring hydration process of cement paste material. Applied Mechanics and Materials, 567, 333–338. doi:10.4028/www.scientific.net/AMM.567.333.

[32] Ministry of Natural Resources and Environment (2006). Malaysia Environmental Quality Report 2006. Department of Environment, Ministry of Natural Resources and Environment, Petaling Jaya, Malaysia.

[33] Gazzaz, N. M., Yusoff, M. K., Aris, A. Z., Juahir, H., & Ramli, M. F. (2012). Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. Marine Pollution Bulletin, 64(11), 2409–2420. doi:10.1016/j.marpolbul.2012.08.005.

[34] Cude, C. G. (2001). Oregon water quality index: A tool for evaluating water quality management effectiveness. Journal of the American Water Resources Association, 37(1), 125–137. doi:10.1111/j.1752-1688.2001.tb05480.x.

[35] Boyacioglu, H. (2007). Development of a water quality index based on a European classification scheme. Water SA, 33(1), 101–106. doi:10.4314/wsa.v33i1.47882.

[36] Song, T., & Kim, K. (2009). Development of a water quality loading index based on water quality modeling. Journal of Environmental Management, 90(3), 1534–1543. doi:10.1016/j.jenvman.2008.11.008.

[37] Debels, P., Figueroa, R., Urrutia, R., Barra, R., & Niell, X. (2005). Evaluation of water quality in the Chillán River (Central Chile) using physicochemical parameters and a modified Water Quality Index. Environmental Monitoring and Assessment, 110(1–3), 301–322. doi:10.1007/s10661-005-8064-1.

[38] Liou, S. M., Lo, S. L., & Wang, S. H. (2004). A generalized water quality index for Taiwan. Environmental Monitoring and Assessment, 96(1–3), 35–52. doi:10.1023/B:EMAS.0000031715.83752.a1.

[39] Naji, A., & Ismail, A. (2012). Metals Fractionation and Evaluation of their Risk Connected with Urban and Industrial Influx in the Klang River Surface Sediments, Malaysia. EnvironmentAsia, 5(1), 17-25.

[40] Malik, A., Kumar, A., Kisi, O., & Shiri, J. (2019). Evaluating the performance of four different heuristic approaches with Gamma test for daily suspended sediment concentration modeling. Environmental Science and Pollution Research, 26(22), 22670–22687. doi:10.1007/s11356-019-05553-9.

[41] Sepahvand, A., Singh, B., Sihag, P., Nazari Samani, A., Ahmadi, H., & Fiz Nia, S. (2019). Assessment of the various soft computing techniques to predict sodium absorption ratio (SAR). ISH Journal of Hydraulic Engineering, 27(Sup1), 124–135. doi:10.1080/09715010.2019.1595185.

[42] Alizadeh, M. J., Kavianpour, M. R., Danesh, M., Adolf, J., Shamshirband, S., & Chau, K. W. (2018). Effect of river flow on the quality of estuarine and coastal waters using machine learning models. Engineering Applications of Computational Fluid Mechanics, 12(1), 810–823. doi:10.1080/19942060.2018.1528480.

[43] Olyaie, E., Banejad, H., Chau, K. W., & Melesse, A. M. (2015). A comparison of various artificial intelligence approaches performance for estimating suspended sediment load of river systems: a case study in United States. Environmental Monitoring and Assessment, 187(4), 189. doi:10.1007/s10661-015-4381-1.

[44] Elzwayie, A., El-shafie, A., Yaseen, Z. M., Afan, H. A., & Allawi, M. F. (2017). RBFNN-based model for heavy metal prediction for different climatic and pollution conditions. Neural Computing and Applications, 28(8), 1991–2003. doi:10.1007/s00521-015-2174-7.

[45] Ghorbani, M. A., Khatibi, R., Karimi, V., Yaseen, Z. M., & Zounemat-Kermani, M. (2018). Learning from Multiple Models Using Artificial Intelligence to Improve Model Prediction Accuracies: Application to River Flows. Water Resources Management, 32(13), 4201–4215. doi:10.1007/s11269-018-2038-x.

[46] Yaseen, Z. M., Sulaiman, S. O., Deo, R. C., & Chau, K. W. (2019). An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. Journal of Hydrology, 569, 387–408. doi:10.1016/j.jhydrol.2018.11.069.

[47] Ismail, R., Alsadi, J., Hatamleh, R., Telfah, D., Jaradat, A., Aljamal, M., Trrad, I., & Al-Mattarneh, H. (2024). Employing CNN and black widow optimization for sustainable wastewater management in an environmental engineering context. Asian Journal of Civil Engineering, 25(5), 3973–3988. doi:10.1007/s42107-024-01024-w.

[48] Al-Mattarneh, H., & Dahim, M. (2018). Determination of Water Quality Parameters Using Microwave Nondestructive Method. International Journal of Engineering & Technology, 7(3.32), 182-185.

[49] Othman, F., Alaaeldin, M. E., Seyam, M., Ahmed, A. N., Teo, F. Y., Ming Fai, C., Afan, H. A., Sherif, M., Sefelnasr, A., & El-Shafie, A. (2020). Efficient river water quality index prediction considering minimal number of inputs variables. Engineering Applications of Computational Fluid Mechanics, 14(1), 751–763. doi:10.1080/19942060.2020.1760942.

[50] Ismail, R., Aljamal, M., Al-Mattarneh, H., Das, A., & Pawar, R. K. (2023). Sustainable Infrastructure Development: Integrating Environmental and Social Factors in Civil Engineering". Journal of Advanced Zoology, 44(S2), 4730–4737. doi:10.53555/jaz.v44iS2.2205.

[51] Dahim, M., Al-Mattarneh, H., & Ismail, R. (2018). Simple capacitor dielectric sensors for determination of water content in transformer oil. International Journal of Engineering & Technology, 7(3), 157–160. www.sciencepubco.com/index.php/IJET

[52] Ismail, R. (2024). Improving wastewater treatment plant performance: an ANN-based predictive model for managing average daily overflow and resource allocation optimization using Tabu search. Asian Journal of Civil Engineering, 25(2), 1427–1441. doi:10.1007/s42107-023-00853-5.

[53] Kasmuri, N., Hadi, N. A., & Omar, M. (2017). Water Quality Assessment in Klang River Basin, Malaysia. Proceedings of the 37th IAHR World Congress, 6865, 270–273, 13–18 August, 2017, Kuala Lumpur, Malaysia.

[54] Telfah, D. B., Jaradat, A. Q., & Ismail, R. (2024). Examining the Long-Run and Short-Run Relationship between Water Demand and Socio-Economic Explanatory Variables: Evidence from Amman. Sustainability (Switzerland), 16(6), 2315. doi:10.3390/su16062315.

[55] Ismail, R. M. A., Enemose, E. A., Al-Jamal, M., Trrad, I., Al-Mattarneh, H., Tripathi, V., & Patil, P. Y. (2022). Study on Cu Based Organic Linkers for Effective Emission Control over Diesel Engine Effuluents. Key Engineering Materials, 928, 61–67. doi:10.4028/p-4nz3up.

[56] Zain, M. F. M., Al-Mattarneh, H., Rama, S. N., & Mahmud, H. B. (2008). A Knowledge-based system for mix design of concrete containing pozzolanic materials. Proceedings of the AEI 2008 Conference - AEI 2008: Building Integration Solutions, 328. doi:10.1061/41002(328)57.