ISSN (Print): 2476-8316
ISSN (Online): 2635-3490
Dutse Journal of Pure and Applied Sciences (DUJOPAS), *Vol. 6 No. 3 September 2020*

Modeling of Water Treatment Plant Performance using Artificial Neural Network: Case Study Tamburawa Kano-Nigeria

S.I. Abba[1*], A.G Usman[2], Y.A Danmaraya[3], A.G Usman[4], H.U Abdullahi[5]

[1]Department of Physical Planning Development,
Yusuf Maitama Sule University Kano,
Nigeria

[2]Department of Analytical Chemistry,
Faculty of Pharmacy, Near East University, 99138 Nicosia,
Turkish Republic of Northern Cyprus

[3]Department of Chemistry,
Yusuf Maitama Sule University
Kano, Nigeria

[4] Raw Material and Development Councils (RMRDC)
NO 17, AguiyiIronsi Street, Maitama District,
FCT Abuja,

[5]Department of Chemistry,
Faculty of Science,
Federal University Dutse, Jigawa Nigeria

Email: saniisaabba86@gmail.com

Abstract
Potential implementation of intelligent tool is quite crucial in any complex system like water treatment plant (WTP). In this paper, the application of artificial neural network (ANN) and comparison of three different activation function viz: Logsig, Purelin and Tansig to simulate the performance of Tamburawa WTP in terms of pH and Turbidity were employed. The recorded weekly data set from the Tamburawa WTP for the year 2015, are measured raw and treated pH (pH_R and pH_T), Turbidity ($Turb_R$ and $Turb_T$) (µs/cm), Suspended Solid (SS_R and SS_T) (mg/L), Hardness ($Hardness_R$ and $Hardness_T$) (mg/L) and Chloride (Cl_R and Cl_T) (mg/L) and performance accuracy of the models including determination coefficient (R^2), root mean square error (RMSE) and mean absolute error (MAE) were evaluated. The predictive results demonstrated that Tansig (R^2 = 0.9775, RMSE=0.0027 and MAE=0.0067) outperformed Logsig and Purelinin term of pH prediction while for Turbidity, Logsig proved to be the best activation function with the predictive accuracy of (R^2 = 0.9077,

**Author for Correspondence*

RMSE=0.0283 and MAE=0.0067). The overall outcomes of models depicted the reliability and satisfactory performance of ANN in modeling the performance efficiency of Tamburawa WTP.

Keywords: Artificial neural network, Activation function, water treatment plant, Kano-Nigeria

INTRODUCTION

Water treatment plant (WTP) are process that remove the contaminants from the untreated domestic wastewater with the goal of safeguarding the public health and natural environment (Nourani et al., 2018; Falah Nezhad et al., 2016). Wastewater management is important to protect our environment from deteriorating as well as improving the water scarcity which exist in a place where the water is insufficient to meet or satisfy requirements demands (Dan'azumi and Bichi (2010).). WTP is extremely complex and dynamic process due to its intricacy of the treatment method., Appropriate action, maintenance and control of WTPs is very vital for monitoring the environmental and ecological health (Ogwueleka, 2009; Gaya et al., 2014). There are certain key descriptions of variables which can be used to assess the water treatment plant performance in which pH and Turbidity are among the most important variables. For example, Abba and Elkiran, (2017) assessed the wastewater treatment plant (WWTP)interm of chemical oxygen demand (COD) using artificial neural network (ANN). Tumer and Edebali, (2015) employed ANN to predict the performance of WWTP in term of total suspended solid (TSS). Gaya et al., (2017) analyzed the performance of WTP in term of Turbidity. Similarly, Hamed et al. (2004) predicted biological oxygen demand (BOD) and suspended solid (SS) concentrations based on ANN to determine the performance of WWTP(Gaya *et al.*, 2017).

Based on the established WTP, linear and conventional regression tools have been widely used but they have been generally associated with low accuracy levels, giving room to the development of the artificial intelligence (AI) (e.g. ANN) methods which are considered as accurate and non-linear hydrologic tools (Nourani 2019). Meanwhile, several researchers have established different types of intelligence techniques such asAI which have been gradually applied for modelling and estimation in various discipline of hydrology and environmental engineering in order to rescue the existing traditional models (Kim and Seo, (2015); Abba et al., 2017; Nourani et al., 2018;Nourani et al., 2019; Elkiran et al., 2018; Zhu et al., 2019; Samsudin, et al 2014). The main aim of this paper is to employ the application of artificial neural network for assessing the performance of Tamburawa WTP in terms of treated pH and Turbidity using different activation functions.

METHODOLOGY

Artificial Neural Network

The ANNs are mathematical modeling tools that are especially useful in the field of prediction and forecasting in complex settings. Historically, there were meant to operate through simulating, at a simplified level, the activity of the human brain(Abba, Usman and

Işik, 2020). The ANN accomplishes this through a large number of highly interconnected processing elements (neurons), working in unison to solve specific problems, such as forecasting and pattern recognition(Elkiran *et al.*, 2018). Each neuron is connected to certain of its neighbors with varying coefficients or weights that represent the relative influence of the different neuron inputs to other neurons(Abba *et al.*, 2019).There is a two-stage operation mode of artificial neural networks. One of them is training the other testing stage. Once it must be trained to use an artificial neural network. The training is carried out using some of the inputs and outputs data set. ANN makes a generalization of these data. Artificial neural networks consist of three layers, including inputs, output and hidden layers, and there are many neurons in each layer as shown (Figure.1).

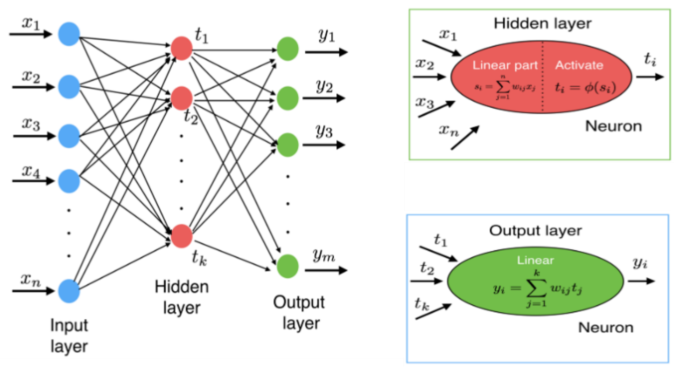


Fig. 1 A schematic diagram of typical three-layer ANN(Nasr *et al.*, 2012)

A neuron is consisting of five main elements including; inputs, weights, summing function, activation (transfer) function and output. Neurons can have multiple inputs, but it can only have one output. Inputs information comes from outside networks or other neurons. In some cases, neurons can create self-input with feedback. A neuron usually receives many simultaneous inputs. Each input has its own relative weight which gives the input the impact that it needs on the processing element's summation function(Elkiran, Nourani and Abba, 2019). Weights are adaptive coefficients within the network that determine the intensity of the input signal as registered by the artificial neuron. They are a measure of an input's connection strength (Tumer and Edebali, 2015). These strengths can be modified in response to various training sets and according to a network's specific topology or through its learning rules. The first step in a processing element's operation is to compute the weighted sum of all of the inputs. Summing function operation sums all of the inputs from the processing elements.

Model development

In model development, various steps were carried out which includes data collection and preprocessing, model design, model training and testing and lastly, model execution (see, Fig 2). Prior to model development of any ANNs approach, the significant input selection is quite significant to be employed to different models(Pham *et al.*, 2019). Therefore, the inputs employed in this modeling were the most significant according to pre-analysis(Abba *et al.*, 2020).

Data Processing.

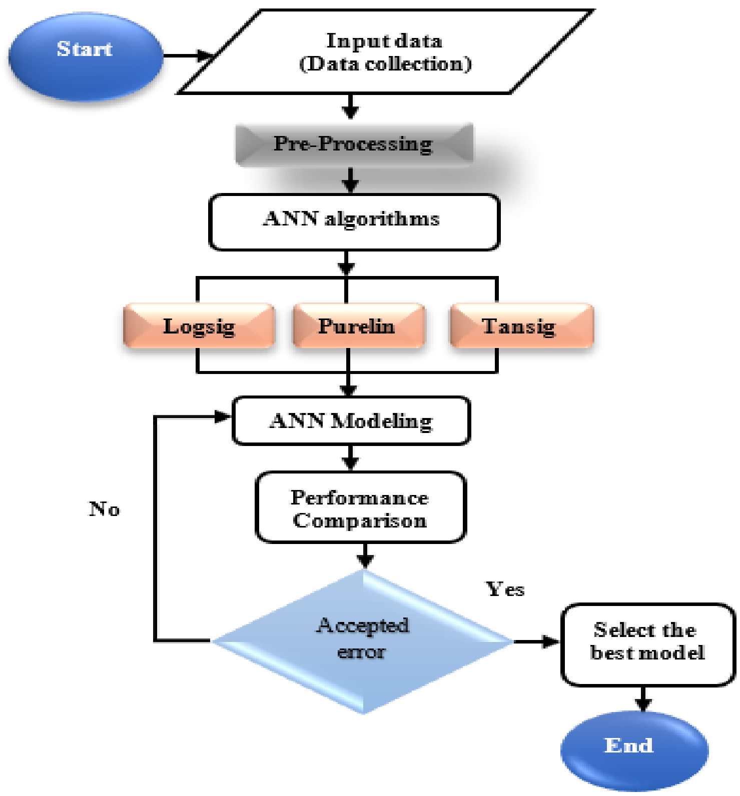
The initial weight are normally assigned randomly during the simulation process, the larger the data both in quality and quantity the better the training, as lack of enough data may create over fitting (over learning), which is one of the major problem occur in learning, in addition lack of suitable number of hidden neurons and inappropriate epoch (iteration) can also lead to over fitting, small nodes can also result in under fitting while many nodes can lead to over fitting in the hidden layers. The suitable architecture is the one which produced the minimal error term in both training and testing data. In this work, a supervised training, backpropagation algorithm is selected(Selin and Abba, 2020). The backpropagation algorithm minimizes the MSE between the observed and the predicted output in the output layer, through two phases. In the forward phase, the external input information signals at the input neurons are propagated forward to compute the output information signal at the output neuron. In the backward phase, modifications to the connection strengths are made based on the basis of the difference in the predicted and observed information signals at the output neuron. The structure that resulted in minimum errors was the one selected(Abba, Hadi and Abdullahi, 2017). The model uses the trainlm - Levenberg–Marquardt function for training. Learngdm (gradient descent with momentum weight and bias learning function) function was used as the adaption learning function and R^2, MAE and RMSE (see, equation 2,3, and 4) was determined as the performance function(Abdullahi, Usman and Abba, 2020). Owing to the overfitting problems, satisfactory training performance does not always be in agreement with the testing performance(Ghali *et al.*, 2020). As such, validation process needs to be conducted, different types of validation approach can be applied including cross-validation which is called k-fold cross-validation, others are holdout, leave one out and so on, this work employed cross-validation approach(Elkiran *et al.*, 2018).

$$R^2 = 1 - \frac{\sum_{j=1}^{N}\left[(Y)_{obs,j} - (Y)_{com,j}\right]^2}{\sum_{j=1}^{N}\left[(Y)_{obs,j} - \overline{(Y)}_{obs,j}\right]^2} \qquad (1)$$

$$MAE = \frac{1}{N}\sum_{j=1}^{N}\left|(Y)_{obs,j} - (Y)_{com,j}\right| \qquad (2)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left((Y)_{obs,j} - (Y)_{com,j}\right)^2} \qquad (3)$$

Where N, Y_{obs}, Y_{com} and $\overline{Y}$ indicate the number of samples, observed data, computed data and mean of observed values, respectively.



t

Description of the study area

Kano as one of the populace state in Nigeria deserved a treated water from rivers for consumption and other domestic chores. This purification should atleast fulfill the minimum requirement by world health organization through the use of the conventional treatment process (see, Fig 3a). The Tamburawa water treatment plant (TWTP) with a capacity of producing 150L portable water per day to fulfill the communities need in Kano city and the surroundings. The raw water from the source is pump via pump station which enters preliminary treatment unit where grits and some of the suspended solids are removed to avoid pump wear and pipe deterioration. Fig.3b shows the schematic flow chart of the important operational process(S. I. Abba, 2019).

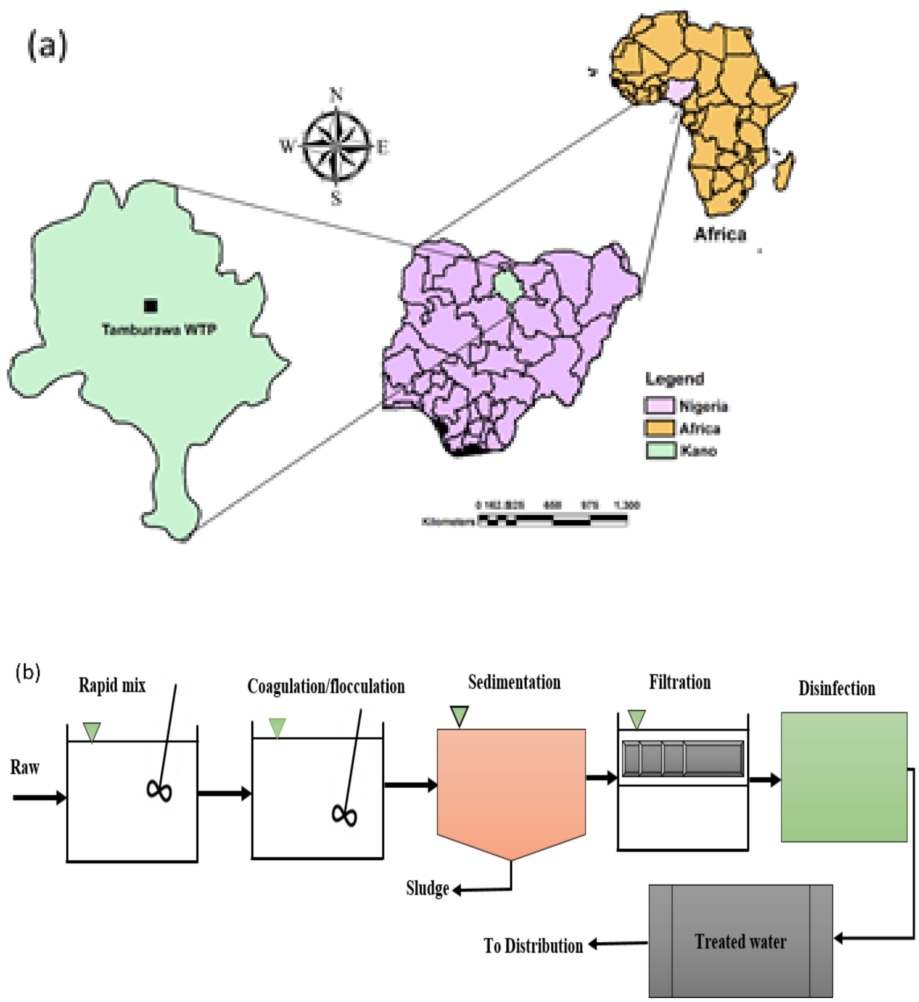


Fig. 3. Tamburawa WTP and its important processes

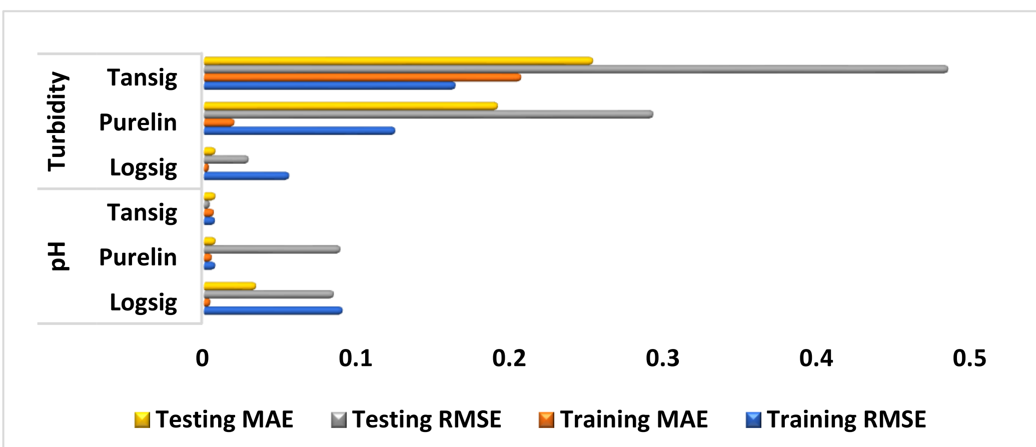
RESULTS AND DISCUSSION

The major motivation of this study is engaging the potential of highly powerful and the most common artificial intelligence for the prediction and modeling the performance of Tamburawa WTP.Iin order to overcome the weakness of conventional mathematical approach in simulating the parameters. In this study, all the modeling aspect were carried out in MATLAB program (2018). The model uses the trainlm - Levenberg–Marquardt function for training. At defaults, the MATLAB program randomly divides input variables and target variables into three sets. But the overfitting problems of ANN, the data were subjected in to validation process using cross-validation of k-fold. As such, the data was portioned into two different part 75% for training and 25% for testing phase. Many different functions such as ‘‘Logsig”, “Tansig”, and “Purelin” were tested as transfer function for hidden layer and “purelin” function was used as the output layer transfer function. The suitable training algorithm at different layers, the number of hidden layers, the number of neurons, determination of the transfer and training functions are highly responsive parameters in the design of artificial neural networks. Owing to that, different ANN architectures were tried and the best was selected.Table 2 shows the performance results of the three-transfer function.

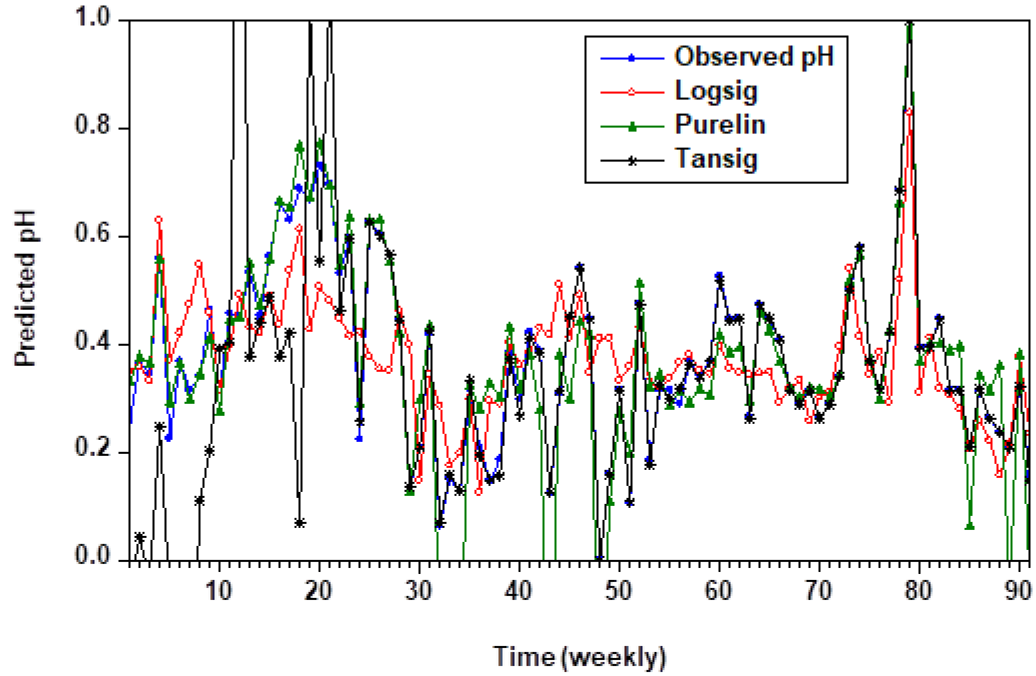
Table 1. Performance analysis results

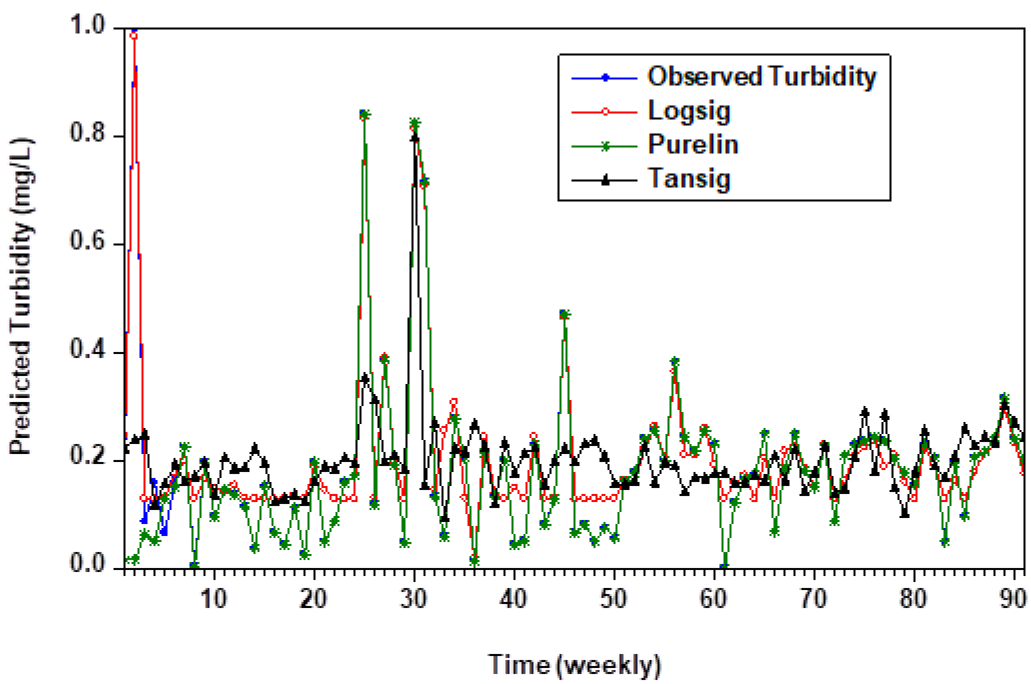
		Training			Testing		
Parameters	Transfer function	R^2	RMSE	MAE	R^2	RMSE	MAE
	Logsig	0.8793	0.0893	0.0032	0.8779	0.0836	0.0331
pH	Purelin	0.8770	0.0066	0.0043	0.8755	0.0879	0.0068
	Tansig	0.9863	0.0064	0.0057	0.9775	0.0027	0.0067
	Logsig	0.9166	0.0546	0.0021	0.9077	0.0283	0.0067
Turbidity	Purelin	0.8739	0.1235	0.0191	0.8706	0.2911	0.1901
	Tansig	0.8761	0.1627	0.2053	0.8724	0.4837	0.2520

From the results above it can be observed generally that, ANN is a satisfying tool for modeling the effluents characteristic of WTP at Tamburawa. Among the transfer function, *Tansig*outperform the other two in term of performance criterion which served as the best option for selecting activation function in modeling pH, although other activation function provided satisfactory results. Likewise, for Turbidity prediction *Logsig* emerged to outperform the other two and therefore became the best and reliable option for the simulation of Turbidity in Tamburawa WTP. This conclusion can be justified by considering the RMSE and MAE in Fig. 4.



According this graphs it can prove that the performance of both pH and Turbidity are very low with regards to the RMSE and MAE. This error criteria have been used in numerous science and engineering research. Further investigation of the performance results can be indicated in Fig. 5 that depicted the time series plot of both pH and Turbidity. From the plots it can be seen that, there are very good agreement between the observed and predicted values for both the pH and Turbidity which is attributed to the *Tansig* and *Logsig*, respectively. The outcomes also demonstrated that using trial and error procedure is very important in ANN modeling since there is no promising algorithms that is unique and produce the best performance in all the three employed transfer function.





The findings of the results found in this work is in line with the results of the study conducted by (Parmar and Bhardwaj, 2015), (Elkiran, Nourani and Abba, 2019), (Alizadeh and Kavianpour, 2015) and (Yaseen *et al.*, 2018) for determination and prediction of water quality index using various data driven approaches.

CONCLUSION

In this paper, the performance of Tamburawa WTP was assessed in term of pH and Turbidity using the powerful non-linear model called ANN. Three transfer function were used for comparison and the results demonstrated the promising capability of ANN in modeling the effluent water quality of Tamburawa. Therefore, the model developed in this work has a satisfactory result and accuracy. As a result, the ANN modeling could effectively predict the performance of Tamburawa WWTP. For all the modeling, R^2, RMSE, and MAE were used as the predictive performance criteria of the models.It is concluded that, ANN provides an effective analyzing and diagnosing tool to understand and simulate the plant, and is used as a valuable performance assessment tool for plant operators and decision makers. It is suggested that more robust tool such as wavelet, genetic algorithms, extreme learning machine, fuzzy logic etc., should be used in the same treatment plant for comparison as well as enhancement of the prediction results.

ACKNOWLEDGEMENT

The authors would like to acknowledge the effort of Tamburawa WTP for providing us with the data to carry out this research. Secondly, Assoc.Prof. M.S. Gaya for the knowledge guidance in the field of AI-based models and lastly, the management of Buildtrain Nig. Ltd, financial assistance to this publication.

REFERENCES

Abba, S. I. *et al.* (2019) 'Modelling of Uncertain System : A comparison study of Linear and Non-Linear Approaches', *IEEE*.

Abba, S. I. *et al.* (2020) 'Emerging evolutionary algorithm integrated with kernel principal component analysis for modeling the performance of a water treatment plant', *Journal of Water Process Engineering*. Elsevier, 33(October 2019), p. 101081. doi: 10.1016/j.jwpe.2019.101081.

Abba, S. I., Hadi, S. J. and Abdullahi, J. (2017) 'River water modelling prediction using multi-linear regression, artificial neural network, and adaptive neuro-fuzzy inference system techniques', *Procedia Computer Science*. Elsevier B.V., 120, pp. 75–82. doi: 10.1016/j.procs.2017.11.212.

Abba, S. I., Usman, A. G. and Işik, S. (2020) 'Simulation for response surface in the HPLC optimization method development using artificial intelligence models: A data-driven approach', *Chemometrics and Intelligent Laboratory Systems*. Elsevier B.V., p. 104007. doi: 10.1016/j.chemolab.2020.104007.

Abdullahi, H. U., Usman, A. G. and Abba, S. I. (2020) 'Modelling the Absorbance of a Bioactive Compound in HPLC Method using Artificial Neural Network and

Multilinear Regression Methods', 6(2), pp. 362–371.

Alizadeh, M. J. and Kavianpour, M. R. (2015) 'Development of wavelet-ANN models to predict water quality parameters in Hilo Bay, Pacific Ocean', *Marine Pollution Bulletin*, 98(1–2). doi: 10.1016/j.marpolbul.2015.06.052.

Elkiran, G. *et al.* (2018) 'Artificial intelligence-based approaches for multi-station modelling of dissolve oxygen in river', *Global Journal of Environmental Science and Management*, 4(4), pp. 439–450. doi: 10.22034/gjesm.2018.04.005.

Elkiran, G., Nourani, V. and Abba, S. I. (2019) 'Multi-step ahead modelling of river water quality parameters using ensemble artificial intelligence-based approach', *Journal of Hydrology*. Elsevier, 577(April), p. 123962. doi: 10.1016/j.jhydrol.2019.123962.

Gaya, M. S. *et al.* (2014) 'ANFIS Modelling of Carbon and Nitrogen Removal in Domestic Wastewater Treatment Plant', *Jurnal Teknologi*, 67(5). doi: 10.11113/jt.v67.2839.

Gaya, M. S. *et al.* (2017) 'Estimation of turbidity in water treatment plant using hammerstein-wiener and neural network technique', *Indonesian Journal of Electrical Engineering and Computer Science*, 5(3), pp. 666–672. doi: 10.11591/ijeecs.v5.i3.pp666-672.

Ghali, U. M. *et al.* (2020) 'Applications of Artificial Intelligence-Based Models and Multi-Linear Regression for the Prediction of Thyroid Stimulating Hormone Level in the Human Body', 29(4), pp. 3690–3699.

Gómez, T. *et al.* (2017) 'Assessing the efficiency of wastewater treatment plants: A double-bootstrap approach', *Journal of Cleaner Production*, 164, pp. 315–324. doi: 10.1016/j.jclepro.2017.06.198.

Nasr, M. S. *et al.* (2012) 'Application of Artificial Neural Network (ANN) for the prediction of EL-AGAMY wastewater treatment', *Alexandria Engineering Journal*. Faculty of Engineering, Alexandria University, 51(1), pp. 37–43. doi: 10.1016/j.aej.2012.07.005.

Nourani, V. (2017) 'An Emotional ANN (EANN) approach to modeling rainfall-runoff process', *Journal of Hydrology*. Elsevier B.V., 544, pp. 267–277. doi: 10.1016/j.jhydrol.2016.11.033.

Parmar, K. S. and Bhardwaj, R. (2015) 'River Water Prediction Modeling Using Neural Networks, Fuzzy and Wavelet Coupled Model', *Water Resources Management*, 29(1), pp. 17–33. doi: 10.1007/s11269-014-0824-7.

Pham, Q. B. *et al.* (2019) 'Potential of Hybrid Data-Intelligence Algorithms for Multi-Station Modelling of Rainfall', *Water Resources Management*, 33(15). doi: 10.1007/s11269-019-02408-3.

S. I. Abba, V. N. and G. E. (2019) 'Multi-parametric modeling of water treatment plant using AI-based non-linear ensemble', 2, pp. 1–15. doi: 10.2166/wst.2011.079.

Selin, A. G. U. and Abba, I. S. I. (2020) 'A Novel Multi ‑ model Data ‑ Driven Ensemble Technique for the Prediction of Retention Factor in HPLC Method Development', *Chromatographia*. Springer Berlin Heidelberg, (0123456789). doi: 10.1007/s10337-020-03912-0.

Yaseen, Z. M. *et al.* (2018) 'Hybrid Adaptive Neuro-Fuzzy Models for Water Quality Index Estimation', *Water Resources Management*. Water Resources Management, 32(7), pp. 2227–2245. doi: 10.1007/s11269-018-1915-7.