ISSN No:-2456-2165

Prediction of Performance Efficiency for Wastewater Treatment Plant's Effluent Biochemical Oxygen Demand Using Artificial Neural Network

Samson Balogun
Department of Civil Engineering, Faculty of Engineering,
University of Abuja P.M.B 117, Abuja Nigeria.

Toochukwu Chibueze Ogwueleka
Department of Civil Engineering, Faculty of Engineering,
University of Abuja P.M.B 117, Abuja Nigeria.

Abstract:- This study investigated the application of an artificial neural network (ANN) to predict the performance efficiency of the Abuja-based Wupa WWTP, Nigeria using effluent 5-day biochemical oxygen demand (BOD_5) as a performance indicator. Daily data of influent BOD_5, pH, total dissolved solids, total suspended solids, chemical oxygen demand, total coliform, *Escherichia coliform*, and fecal coliform; and effluent BOD_5 over a period of five years (2013 to 2017) for the Wupa WWTP was utilized for the plant's performance efficiency. The four most reliable multilayer perceptron ANN (MLP-ANN) algorithms namely, Levenberg-Marquardt (LM) backpropagation resilient backpropagation, Quasi-Newton backpropagation, and Fletcher-Reeves conjugate gradient backpropagation were adopted; and the most appropriate model was selected following training, validation and testing by altering the number of neurons and activation functions in both the hidden and output layers. The model efficiency was determined using mean square error (MSE) and correlation coefficient (R^2). The ML algorithm with Logsig-Tansig activation pairing and architecture [8-1270-1] performed the best in terms of convergence time and prediction error, with MSE and R^2 values of 1.522 and 0.922, respectively. Also, it revealed that the selected ANN model predicted the effluent BOD_5 with an overall correlation coefficient of 0.962; thus, demonstrating the efficacy of ANN models for accurate prediction of the Wupa WWTP performance. The novelty of this research is in evaluating the efficiency of the plant over the periods and determining the most precise ANN model for Wupa WWTP, Abuja, Nigerians a study which has never been carried out before now.

Keywords:- *Artificial Neural Network (ANN); Wastewater Treatment Plant (WWTP); 5-Day Biochemical Oxygen Demand (BOD_5); Wupa WWTP; Multilayer Perceptron (MLP)*

I. INTRODUCTION

The treatment and management of wastewater in our environment have increasingly gained attractive attention in the last decade, particularly in the face of the incessantly increasing volume of wastewater owing to population growth; rapid urbanization; increased agricultural; and industrial activities (Abba and Elkiran 2017; Arismendy et al., 2020; Alsulaili and Refaie, 2021).

Wastewater treatment plants (WWTPs) are built to clean wastewater and convert it into eco-friendlier water which is released into the environment (Varkeshi et al., 2019). However, due to the wide fluctuation in the quality and quantity of untreated wastewater transported to the treatment plant, the operation of WWTPs can be difficult and challenging (Szeląg et al. 2017).

Moreover, many treatment plants are constructed following the conventional activated sludge system which is allegedly riddled with inefficiencies associated with pollutant removal (Ogwueleka and Samson, 2020). In addressing the challenges of the conventional treatment systems, several alternative methods have been proposed, notable amongst which are the advanced oxidation processes (AOPs) (Deng and Zhao, 2015); nanomaterials (Adeleye et al., 2016); microalgae-activated sludge (MAAS) (Ogwueleka and Samson, 2020); microbial electrochemical system (Li et al., 2021). However, many of these emerging methods are still limited to pilot or laboratory scales and are yet to gain widespread practical applications due to a number of reasons such as the initial cost of installation, uncertainties with operations, adaptation and installation of new technologies, etc.; thus, there is still need to seek for means of attaining efficiency, even in the pre-existing installed treatment systems; which can be achieved by attaining and maintaining optimal conditions in WWTPs.

Attaining optimal operational conditions in WWTPs even with conventional systems is possible and can be achieved with the use of models to predict the WWTP performance based on previous measurements of major plant parameters (Jami et al., 2012).

An important parameter commonly utilized to examine the performance of WWTPs is the 5-day biochemical oxygen demand (BOD_5) (Dogan et al. 2008; Araromi et al., 2018). BOD_5 is an approximation of the quantity of biochemically degradable organic matter contained in a water sample, defined as the amount of oxygen necessary for the aerobic bacteria present in a sample to oxidize the organic matter to a stable organic form (Dogan et al., 2008). It is, however, difficult to measure, and requires five days for its determination (Dogan et al., 2008; Alsulaili and Refaie, 2021). Therefore, the determination of the output BOD_5 of a WWTP as a performance index using predictive tools could achieve safe and economic treatment process management (Araromi et al., 2018; Varkeshi et al., 2019).

Several models have been employed for the prediction of a variety of WWTP performance indicators; among them, artificial neural network (ANN) has risen in popularity and has proven to be excellent in terms of the following advantages

- Ability to model complex functions with high precision
- Modelling of multiple inputs and output concurrently
- It can accommodate workings with noisy and missing parametric
- It can train and update the model with dated data

In spite of the outlined advantages amongst others, there are also some pitfalls with the workings and the development of the ANN model

- There is no physical significance of model parameters
- Underfitting and overfitting may sometimes result due to a lack of standards in the determination of the network architecture (trial and error methods are used)
- ANN model is primarily computer dependent due to the enormous data needed to train the neurons.

In the last two decades, studies conducted on the modeling and prediction ability of the NN model. ANN was used to estimate wastewater treatment plant inlet biochemical oxygen demand. The results obtained show its flexibility, simplicity, the accuracy of prediction, and robust structure (Dogan et al., 2008; Banaei et al., 2013; Bekkari and Zeddouri, 2018; Abba et al., 2020; Alsulaili and Refaie 2021; Setshedi et al. 2021). ANN was used for the modeling of wastewater treatment and desalination using membrane processes, At the end of the study ANN shows high capability in terms of both accuracy and short time of computation compared to the conventional method (Jasir et al., 2021).

A variety of notable studies such as (Ogwueleka and Ogwueleka, 2009; Ogwueleka and Ogwueleka 2010; Vyas et al., 2011; Jami et al. ,2012; Banaei et al., 2013; Ahmadzadeh et al. 2015; Xue 2017; Bekkari and Zeddouri, 2018; Katip 2018; Arismendy et al., 2020; Gawdzik et al., 2020; Alsulaili and Refaie, 2021; Saleh, 2021; Saleh and Kayi, 2021) have documented the application of ANN to model various hydrology and environmental engineering issues. A remarkable application among these studies is the application of ANN to model and predict BOD in diverse WWTPs across the globe.

Alsulaili and Refaie (2021) investigated the use of ANN in predicting BOD_5 in the WWTP, Kuwait, using seven years of data dating from 2013 to 2019; the outcome of the study indicated that the applied model for the BOD_5 prediction achieved a high level of precision with an R^2 value of 0.754, implying the viability of the model. Saleh (2021) applied ANN models to predict BOD_5, total suspended solids (TSS), and chemical oxygen demand (COD) in the effluent of the Muamirah WWTP, Al-Hillah using a two years dataset. The result of the study indicated the capability of ANN for COD, BOD_5, and TSS modeling.

II. ARTIFICIAL NEURAL NETWORK

An artificial neural network is an aspect of artificial intelligence that mimics the operation of the human central nervous system in receiving and computing information systems. it is based on a system of interconnected "neurons" forming the basis of neural network operation (Alsulaili and Refaie, 2021; Setshedi et al., 2021; Saleh, 2021). The model uses a network system similar to the human brain called neurons, to learn and predict various parameters. ANN is gaining quick popularity in the area of artificial intelligence and machine learning due to its robustness and precision accuracy. The significance of ANN is to determine the computational relationship between the dependent and the dependent variable especially due to the anthropogenic nature of the wastewater variable. The relationship is established by designing a network architecture in which previous information(data) is used to train the network during the learning phase to build models Figure l.

There are many types of ANN depending on their peculiarity in terms of architecture and parameters. the commonly used is backpropagation a feed-forward neural network that could be called multilayer perceptron (PLP). we have other types of ANN used for computational analysis such as; Radial Basial Function Neural Networks (RBFN) Networks current Neural Networks (RNN), Elman, Neural Network Networks Deep Neural Networks (DNN) are some of the typologies of MLP-ANN with different training algorithms and model architecture.

A. Multi-layer Perceptron

MLP type which this study is set-out to apply consist of the impute layer, one or more hidden layers, and an output. at each layer, the numbers of neurons at each layer are equal to the number of variables at both input and output in the architecture. do, at the hidden layer there could be a greater number of neurons irrespective of the input or the output layers. The quality of the computation ability at the hidden and the output layers is a function of the quantum of the calculated data assigned to the neurons called weights. A typical ANN MLP architecture is shown in the schematic diagram below:

ISSN No:-2456-2165

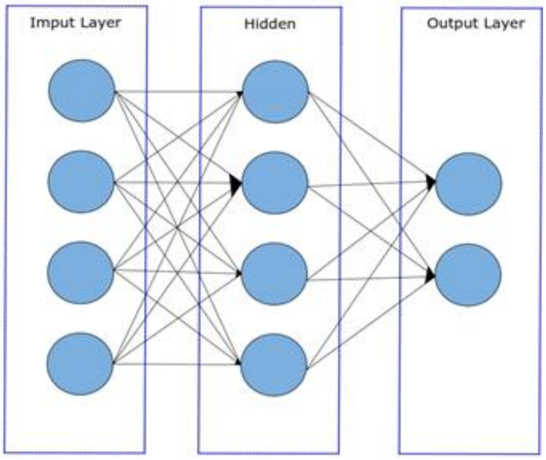

Fig.1. A typical schematic single hidden layer MLP architecture of (4-4-2)

III. PREPARE YOUR PAPER BEFORE STYLING

However, there are basically three phases of developing an ANN model: training, validation, and prediction. During this process, the data set is divided into three groups, usually for the purpose of this study and other development studies reported. the data set is divided into a ratio of 70:15:15 presents.70% of the data set is usually assigned to the training phase to train the neuron while about 15% is used for validation and the remaining 15% is used for the prediction of the ANN variables. At the training phase the weight attached to each neuron are updated after each epoch with the help of the training algorithm until the training is validated with high testing precision. Criterion for stoppage is usually defined as specified at the beginning of each training by using the numbers of iterations and the minimum mean square error as well as validation checks. The normalization checks are designed in such a way that each impute node contribute immensely to the prediction of the output to minimize local minimum convergence [29,30] The normalization equation can be expressed using Eq. (1) below

$$y = y_{min} + \frac{(x - x_{min})(y_{max} - y_{min})}{(x_{max} - x_{min})} \qquad 1$$

where and x_{min} and x_{min} are the maximum and minimum value of the data set, y_{max} and y_{min} are the range for normalization, and y the normalized value of x. basically, the range for normalization is either (0,1) or (-1,1). The output of each neutron is function of the training neurons and its eights assigned to it using Eq. (2)

$$a_{ij} = f_j \left(\sum_{k=1}^{(n(j-1))} a_{k(j-1)wki(j-1)+b_{ij}} \right) \qquad 2$$

where a_{ij} and b_{ij} are the output and bias of the i-th neutron in the j-th layer, $a_{k(j-1)}$ and $w_{ki(j-1)}$ are the output and the weight of neutron from the previous layer, respectively, $n_{(j-1)}$ is the number of neutrons in the (j-1) and f_j is the activation by introducing non-linearity to the network. The commonly used activation functions are logistic sigmoid (log-sigmoid), hyperbolic tangent sigmoid (tan-sigmoid), and linear transfer functions (purelin), whose output ranges and equations are given in Eq. (2).

B. *Significance of the study*

Vyas et al. (2011) studied the relevance of ANN techniques to predict influent and effluent BOD_5 for WWTP in Govindpura, Bhopal; using 3 years of data and two ANN models. The result for model 1 and model 2 showed R values of 0.9 and 0.73 respectively; which is an indication that ANN provides highly acceptable outcomes. Rene and Saidutta (2008) employed 12 AANN-based models to predict BOD_5 and COD levels in wastewater generated from the treatment plant of a petrochemical industry in Mangalore, India. The results revealed that; through its diverse training procedure, ANNs can accurately and effectively predict concentrations of water quality indicators.

Dogan (2008) established an ANN model to predict BOD in the inlet of WWTPs, and the results demonstrated that the ANN may be used to accurately estimate daily BOD at the input of wastewater treatment facilities. Hamed et al. (2004) developed two ANN models to forecast the effluent concentrations of BOD and suspended solids for a major WWTP in Cairo; using a 10 months dataset. The study reported that the prediction error fluctuated minimally and gradually throughout the range of data sizes utilized in training and testing, making ANN a reliable tool for prediction.

Despite these studies having established the viability of ANN for the prediction of BOD as a performance indicator in WWTP, it is important to note; no study has been conducted to predict the performance of BOD_5 as a performance indicator for the Wupa WWTP. Bearing in mind that each WWTP in the world is dynamic and unique with sometimes high variation in their contaminants as rightly noted by (Jami et al. 2012**).** It becomes expedient to apply ANN for the prediction of BOD_5 output in the Wupa WWTP. Therefore, this study applies ANN to predict the output BOD_5 for the Wupa WWTP using a 5 years' dataset.

The novelty of this research lies in the application of four reliable multilayer perceptron ANN algorithms including the Levenberg-Marquardt (LM) backpropagation, resilient backpropagation (RP), Quasi-Newton backpropagation (BFG) and Fletcher-Reeves conjugate gradient backpropagation (CGF); for performance prediction of the Abuja-based Wupa WWTP using effluent 5-day biochemical oxygen demand (BOD_5) as a performance indicator.

III. MATERIALS AND METHODS

A. *Study Area*

The location for this study is Abuja, the Federal Capital City of Nigeria; using the Wupa WWTP Abuja, Nigeria. Wupa WWTP occupies an area of 297,900 square meters and lies between UTM coordinate N998183.603, E321889.651, N998183.603, E322283.340 and 321889.65, E321889.651, N997495.399. E322283.340. The Wupa WWTP is an oxidation ditch plant; a type of activated sludge system developed to a capacity of 131.3 million liters per day with

700,000 population equivalent (P.E.) which is expandable to 1,000,000 P.E. to treat Abuja wastewater. It has three operational units, one of which is currently in service; the other two are standbys in case one unit fails. Wupa WWTP's maximum operating capacity of waste per day is 131,250 m^3. Notably, treated effluents from the plant is released into River Wupa. An area map of the Wupa WWTP is presented in Figure 2 and a schematic diagram showing the plant's flow process is presented in Figure 3.

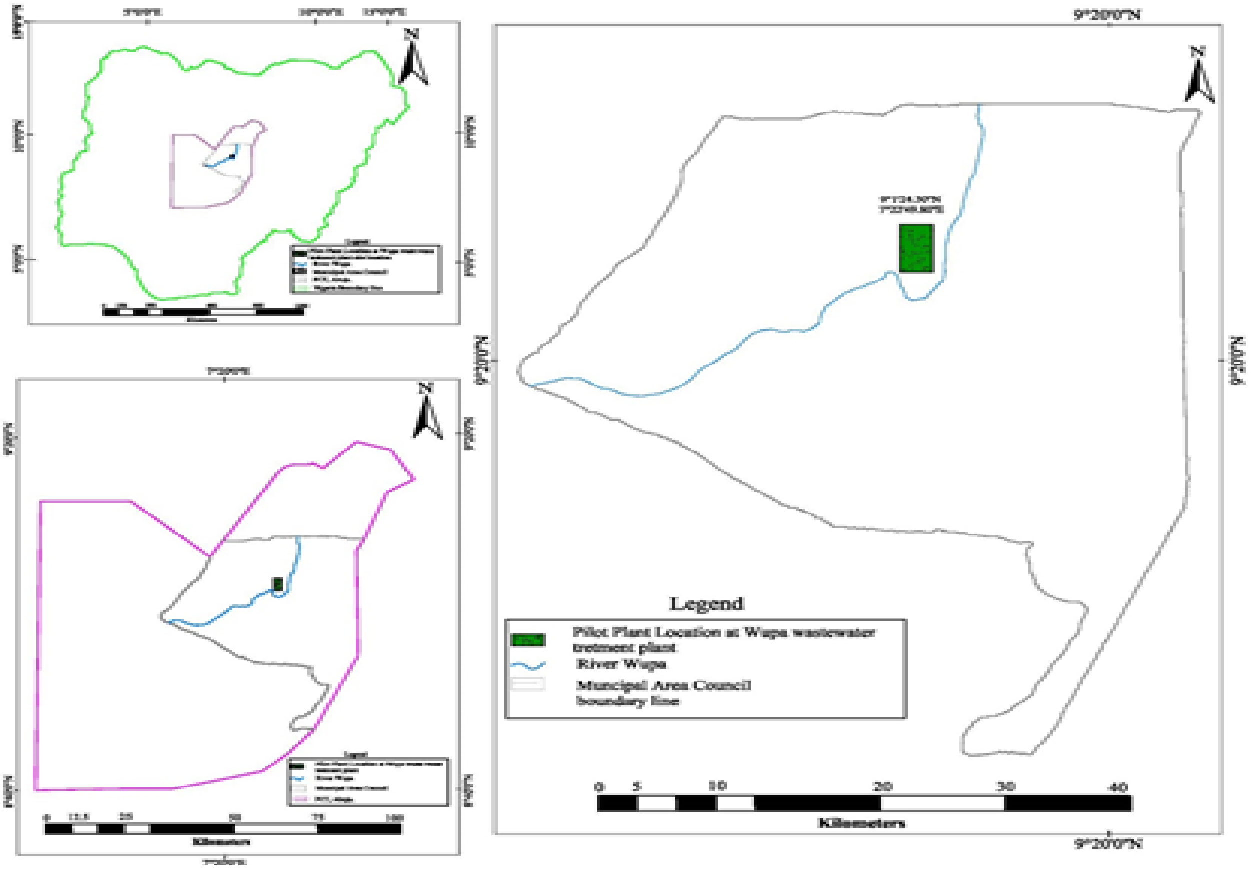

Fig 2: Area map of Wupa WWTP

B. Equations

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Number equations consecutively. Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in.

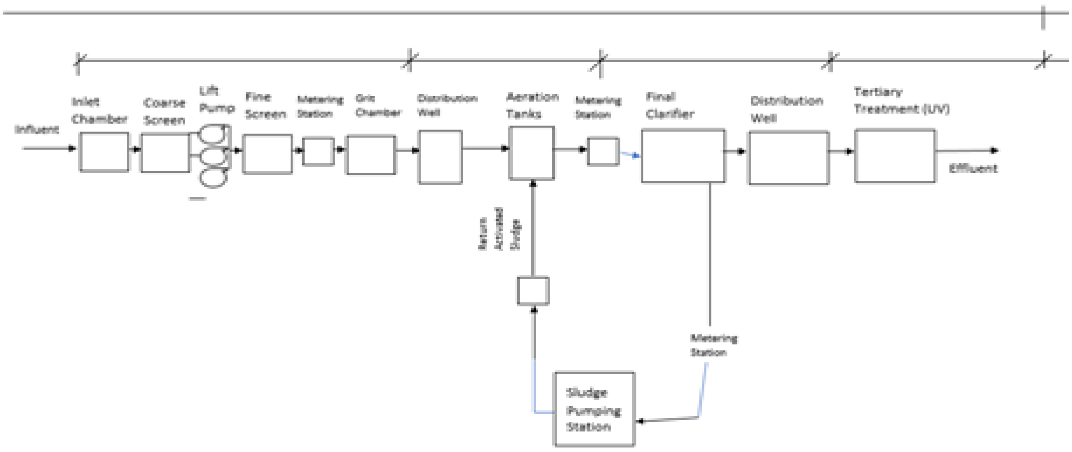

Fig 3: Schematic diagram showing the Wupa WWTP flow process

B. Data Collection

This study utilized data obtained from Wupa WWTP, Abuja. The data contained a total of 14,600 dataset collected daily for a period of five years from 2013 to 2017. Notably, the obtained data was consistent without any missing record. The large volume of data used for this study is preferred because it encompasses all seasonal fluctuations in the parameter capable of affecting the data pattern. Also, ANN modeling is heavily reliant on the quality of historical data, therefore the larger the data collection, the more dependable the developed model (Jami et al. 2012; Alsulaili and Refaie, 2021). The data comprised notable influent($_{inf.}$) Authors and Affiliations and effluent ($_{eff.}$) parameters commonly used for performance evaluation of aerobic wastewater treatment process; which include potential hydrogen (pH), total suspended solid (TSS), BOD_5, total dissolved solid (TDS), COD, Escherichia coliform (EC), total coliform (TC) and facal coliform (FC). Table 1 shows the descriptive statistics for the Wupa WWTP *applied variables. The mean, minimum, maximum, standard deviation, and variance are denoted by Xmean, Xmin, Xmax, S, and V. Prior to the ANN modelling process, a correlation matrix of the obtained data was developed to evaluate the dependence between the variables, using Pearson's correlation* as shown in table.2

Table 1: Descriptive Statistics of the applied parameters

Parameter	Unit	X_{mean}	X_{min}	X_{max}	S	V
Input parameter						
$PH_{inf.}$		**7.31**	**1.781**	**12.513**	**1.539**	**2.367**
$TDS_{inf.}$	**Mg/L**	**151.856**	**125.835**	**183.317**	**15.954**	**254.518**
$TSS_{inf.}$	**Mg/L**	**197.623**	**155.498**	**259.43**	**35.485**	**1259.167**
$B0D_{inf.}$	**Mg/L**	**130.983**	**104.775**	**185.663**	**26.162**	**684.446**
$COD_{inf.}$	**Mg/L**	**262.028**	**136.186**	**403.842**	**83.533**	**6977.752**
$T.C_{inf.}$	**Mg/L**	**259999.99**	**199995.72**	**300005.43**	**49003.3**	**2401324017**
$E,C_{inf.}$	**Mg/L**	**499.99**	**196.109**	**1003.432**	**309.927**	**96054.474**
$F.C_{inf.}$	**Mg/L**	**5732.137**	**3159.007**	**8917.195**	**2304.64**	**5311370.1**
Output parameter						
$B0D_{eff.}$	**Mg/L**	**9.739**	**0.417**	**24.991**	**5.564**	**30.96**

C. Analytical Procedure

➢ *ANN Model*

The multilayer perceptron (MLP) is the applied ANN architecture for this study; it is the most extensively used ANN and is known to exceed others in precision (Alsulaili and Refaie, 2021; Setshedi et.al., 2021). MPL is a feed-forward ANN model that uses a supervised learning technique involving the backpropagation (BP) algorithm to map sets of input data into appropriate outputs (Setshedi et.al., 2021). Notably, forward feed ANN is one that creates connections in a single direction, from input to output, without causing cycles (Saleh, 2021). The MLP-ANN is made up of numerous basic neurons that operate simultaneously in three layers namely; input, hidden, and output layers as shown in Figure 4. The network function can be determined by connecting the neurons and using operators to link the signal phases of one neuron to the other. This can be explained by Equation 1 (Bekkari and Zeddouri 2018).

$$y_i = \mathrm{f}\left(\sum_{j=1}^{n} w_{ij}\, x_j + b_i\right) \qquad (3)$$

where y_i denotes the ith nodal value in the current layer, f denotes the activation function; w_{ij} denotes the weight allocated to each input. $x_j$is the previous layer's jth nodal value; $b_i$represent the bias for each output; and N is the total amount of inputs.

Three layers of architecture (input, hidden, and output) for a feedforward ANN were created to predict the Wupa WWTP. Because poor node fitting could affect the network training and validation phase (Bekkari and Zeddouri 2018), the model employs fewer hidden nodes to avoid the over-fitting problem that may emerge from generalization.

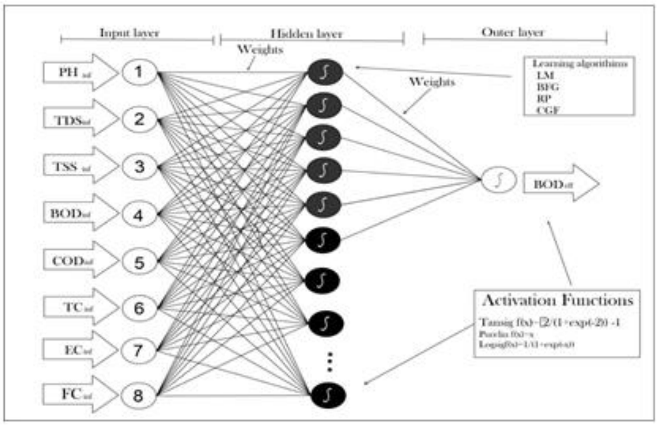

Fig 4: Architecture of the applied ANN

D. Training, Validation and testing

The appropriate ANN model development in this study depended on three major steps including; training, validation, and testing. Training was used to build the model by altering weights; whilst the validation dataset was used to identify the BP algorithm's stopping point as well as the ANN architecture determination; and the testing dataset was used to determine performance parameters like accurateness and model generalization testing.

The efficiency of the ANN model was determined using mean square error (MSE) and correlation coefficient (R^2) which were estimated between actual and predicted data, as stated in Equations 2 and 3 below.

$$R^2 = \left(\frac{\left(\sum(x_i - \bar{x})\,(y_i - \bar{y})\right)}{\sqrt{(x_i - \bar{y})^2 (y_i - \bar{y})^2}}\right)^2 \qquad (2)$$

$$MNSE = \frac{\sum_{t=1}^{n}(x_i - y_i)^2}{N} \qquad (3)$$

where $\boldsymbol{x_i}$ represent measured value; $\boldsymbol{y}$ indicates predicted value; $\overline{\boldsymbol{x}}$ denotes the mean of the measured values; $\bar{y}$ is the mean of the predicted values; while N denotes the total amount of model outputs. When the MSE was the lowest and the R was the highest (near to 1), the best model was created.

➢ *ANN Model Simulation and Architecture*

The MLP-ANN architecture and simulation was performed using MATLAB RB 2021. Data splitting, in which data is partitioned into training, validation and testing subsets to guarantee strong model generalizability, is a critical phase in the building of an ANN and has a major impact on the model performance (May et al., 2010). The MATLAB's default splitting ratio of 70:15:15 was used for the training, validation, and testing sets respectively. The experimental data was randomly imputed into the work area, with $COD_{eff.}$ as the dependent variable and the independent variables being $pH_{inf.}$, $TDS_{inf.}$, $TSS_{inf.}$, $BOD_{inf.}$, $COD_{inf.}$, $TC_{inf.}$, $FC_{inf.}$, $EC_{inf.}$

This research work adopted four most reliable MLP-ANN algorithms including the Levenberg-Marquardt (LM) backpropagation, resilient backpropagation (RP), Quasi-Newton backpropagation (BFG) and Fletcher-Reeves conjugate gradient backpropagation (CGF). A single hidden layer and several neurons was used in the model.

For each algorithm, the number of neurons in the hidden layer as well as the activation functions pairing were altered, and three (3) activation functions namely; linear activation function (purelin), hyperbolic tangent sigmoid activation functions (tansig), and logistic sigmoid activation function (logsig) were applied. Equations 4-6 represent the mathematical expressions of activation functions utilized. The trial-and-error method was utilized to tune the hyperparameters because it is one of the most practical and widely used approaches for choosing the ideal number of activation functions, batch sizes, neurons, epochs and learning rates among other things (Bashiri and Geranmayeh, 2011).

Learning rate parameter (LRP), which is used to maximize the likelihood of the training process not becoming trapped in a local minimum rather than the global minimum, maybe critical in network convergence, based on the application and network configuration (Hamed et al., 2004; Bekkari and Zeddouri, 2018). LRP of 0.01 was applied in this study. To achieve suitable timely convergence, the input and output data were normalized to a range of 0–1 using Equation 7, and the real values were calculated once the ANN was completed by modifying the output n_i data with Equation 8.

$$f(n) = n \qquad (4)$$

a). Purelin function
b). Tangent sigmoid function

$$f(n) = 2/((1 + e\hat{}(-2n))) - 1 \qquad (5)$$

c). Sigmoid function

$$f(n) = 1/((1 + e\hat{}(-n))) \qquad (6)$$

The outer boundaries of the activation functions are [0.1, [-∞, +∞] and [-1, 1].

$$x_{ni} = (x_{ni} - x_{min})/(x_{max} - x_{min}) \qquad (7)$$

$$x_i = x_{ni}\,(x_{max} - x_{min}) + x_{min} \qquad (8)$$

IV. RESULTS AND DISCUSSION

Table 2: Correlation Matrix for the of the plant data variables

	$pH_{inf.}$	$TDS_{inf.}$	$TSS_{inf.}$	$BOD_{inf.}$	$COD_{inf.}$	$T.C_{inf.}$	$E.C_{inf.}$	$F.C_{inf.}$	$BOD_{eff.}$
$pH_{inf.}$	1								
$TDS_{inf.}$	-0.024	1							
$TSS_{inf.}$	0.016	0.621	1						
$BOD_{inf.}$	-0.021	0.709	0.842	1					
$COD_{inf.}$	-0.013	0.608	0.862	0.925	1				
$T.C_{inf.}$	0.019	0.606	0.369	0.174	-0.029	1			
$E.C_{inf.}$	-0.013	0.783	0.538	0.634	0.347	0.791	1		
$F.C_{inf.}$	-0.014	0.488	0.407	0.58	0.242	0.583	0.905	1	
BOD_{eff}	0.034	-0.802	-0.123	-0.235	-0.1	-0.638	-0.629	-0.32	1

A. *ANN Model Performance for $BOD_{eff.}$ Output*

Tables 3 to 6 present the best results for each of the applied MLP-ANN algorithms (LM, BFG, RP and CGF) using various combinations of activation pairs; while Figure 4 shows the best ANN model's regression plot between actual and predicted data on the BOD_{eff}; and Figure 5 shows the BODeff output actual values against prediction values for the best ANN model.

Tables 3 to 6 indicate the following for each of the applied algorithm:

- **LM Algorithm:** MSE and R^2 values for the LM algorithm during training varied between 1.49 to 5.307 and 0.872 to 0.929 respectively. MSE and R^2 values for LM during validation varied between 1.404 to 5.483 and 0.845 to 0.929 respectively. MSE and R^2 values for LM during testing varied between 1.522 to 5.576 and 0.845 to 0.925 respectively. Table 3 indicates that the LM algorithm with the Purelin-Tansig activation pairing and architecture [8-1265-1] performed best.
- **RP Algorithm:** MSE and R^2 values for RP during training ranged between 1.519 to 5.322 and 0.053 to 0.949 respectively. The MSE and R^2 values for the RP during validation ranged between 1.493 to 5.499 and 0.077 to 0.918 respectively. The MSE and R^2 values for RP during testing ranged between 1.556 to 5.576 and 0.085 to 0.929 respectively. Table 4 indicates that the RP algorithm with the Logsig-Tansig activation pairing and architecture [8-1170-1] performed best.
- **BGF Algorithm:** The MSE and R^2 values for BFG during training ranged between 1.505 to 10.347 and 0 to 0.945 respectively. The MSE and R^2 values for BFG during validation varied between 1.417 to 10.591 and 0 to 0.929 respectively. The MSE and R^2 values for BFG during testing ranged between 1.424 to 10.978 and 0 to 0.939 respectively. Table 5 indicates that the BGH algorithm with the Logsig-Tansig activation pairing and architecture [8-1270-1] performed best.
- **CGF Algorithm:** MSE and R^2 values for CGF during training ranged between 1.523 to 5.308 and 0.677 to 0.966 respectively. MSE and R^2 values for CGF during validation ranged between 1.457 to 5.484 and 0.714 to 0.953 respectively. MSE and R^2 values for CGF during testing ranged between 1.545 to 5.573 and 0.676 to 0.935 respectively. Table 6 indicates that the CGF algorithm with the Logsig-Purelin activation pairing and architecture [[8-1024-1] performed best.

Notably, with the exception of the Tansig-Logsig and Logsig-Logsig functions, where the BFG had lower iteration counts, LM converged with the fewest iteration number for most activation functions pairs. Additionally, LM presented the best level of MSE indicators (i.e., the least MSE values) in all of the training, validation and testing stages amongst the four applied algorithms. As a result, the LM algorithm is regarded as the most efficient in function approximation. Consequently, the best ANN model for the prediction of output BOD_5 for this study is the MLP that applied Purelin-Tansig activation pairing and architecture [8-1265-1]. The findings in this study agrees with studies such as Hamed et al. (2004); Bekkari and Zeddouri (2018); and Banaei et al., (2013) that have demonstrated that the LM algorithm outperforms other algorithms in terms of convergence time and prediction error.

The result of the best ANN model's regression plot between actual and predicted data on the $BOD_{eff.}$ in Figure 4 indicate a high level of compatibility between the actual data and predicted ANN values with the noted high correlation coefficients of 0.965, 0.950, and 0.956 for the training, validation and testing phases respectively; and an overall correlation coefficient of 0.962. The comparison between the measured data and predicted data for the BOD_{eff} presented in Figure 5 demonstrates high accuracy of the applied ANN model for the prediction of the BOD_{eff} for Wupa WWTP. Although some deviations appear in some data points; these deviations can be attributed to factors such as, noise in the training; the wight of the dataset and input parameters etc.

Table 3: MLP-ANN model performance statistics for training; validation; and testing of the LM method for $BOD_{eff.}$

		LM								
			Training		Validation		Testing			
HLAF	**OLAF**	**Designation**	**MSE**	**R^2**	**MSE**	**R^2**	**MSE**	**R^2**	**IN**	**Architecture**
Logsig	Purelin	LP	1.512	0.929	1.520	0.904	1.523	0.925	11	[8-1003-1]
Logsig	Tansig	LT	1.492	0.929	1.576	0.912	1.534	0.918	10	[8-1270-1]
Tansig	Purelin	TP	1.49	0.929	1.605	0.906	1.527	0.925	16	[8-1200-1]
Tansig	Logsig	TL	5.307	0.872	5.432	0.861	5.576	0.845	15	[8-1275-1]
Purelin	Logsig	PL	5.306	0.874	5.483	0.845	5.513	0.857	15	[8-1201-1]
Logsig	Logsig	LL	5.305	0.872	5.432	0.861	5.571	0.846	2	[8-899-1]
Purelin	Purelin	PP	1.526	0.927	1.404	0.929	1.617	0.906	4	[8-1098-1]
Tansig	Tansig	TT	1.504	0.929	1.578	0.912	1.528	0.920	14	[8-1044-1]
Purelin	Tansig	PT	1.531	0.927	1.480	0.916	1.522	0.922	16	[8-1265-1]

Note: OLAF = output layer activation function; HLAF = hidden layer activation function; IN, iteration number

ISSN No:-2456-2165

Table 4: MLP-ANN model performance statistics for training; validation; and testing of the RP method for $BOD_{eff.}$

			RP							
			Training		Validation		Testing			
HLAF	**OLAF**	**Designation**	**MSE**	**R^2**	**MSE**	**R^2**	**MSE**	**R^2**	**IN**	**Architecture**
Logsig	Purelin	LP	1.557	0.927	1.641	0.918	1.588	0.924	27	[8-1045-1]
Logsig	Tansig	LT	1.519	0.949	1.621	0.903	1.556	0.925	113	[8-1170-1]
Tansig	Purelin	TP	1.527	0.933	1.620	0.912	1.580	0.929	61	[8-1200-1]
Tansig	Logsig	TL	5.309	0.523	5.434	0.388	5.576	0.415	48	[8-1275-1]
Purelin	Logsig	PL	5.322	0.053	5.499	0.077	5.527	0.085	38	[8-1201-1]
Logsig	Logsig	LL	5.306	0.599	5.431	0.634	5.573	0.619	29	[8-1211-1]
Purelin	Purelin	PP	1.615	0.869	1.493	0.897	1.724	0.893	24	[8-1034-1]
Tansig	Tansig	TT	1.553	0.933	1.653	0.916	1.584	0.927	39	[8-1052-1]
Purelin	Tansig	PT	2.397	0.867	2.478	0.839	2.603	0.867	27	[8-1033-1]

Note: OLAF = output layer activation function; HLAF = hidden layer activation function; IN, iteration number

Table 5: MLP-ANN model performance statistics for training; validation; and testing of the BFG method for $BOD_{eff.}$

		BFG								
			Training		Validation		Testing			
HLAF	**OLAF**	**Designation**	**MSE**	**R^2**	**MSE**	**R^2**	**MSE**	**R^2**	**IN**	**Architecture**
Logsig	Purelin	LP	1.506	0.945	1.607	0.908	1.530	0.937	34	[8-1203-1]
Logsig	Tansig	LT	1.521	0.931	1.566	0.906	1.524	0.939	31	[8-1270-1]
Tansig	Purelin	TP	1.505	0.929	1.608	0.904	1.538	0.925	31	[8-1001-1]
Tansig	Logsig	TL	10.347	0.213	10.591	0.191	10.978	0.244	1	[8-1245-1]
Purelin	Logsig	PL	6.321	0.089	6.179	0.108	6.346	0.084	4	[8-1121-1]
Logsig	Logsig	LL	6.299	0.000	6.282	0.000	6.346	0.000	1	[8-689-1]
Purelin	Purelin	PP	1.543	0.925	1.417	0.929	1.645	0.904	17	[8-1048-1]
Tansig	Tansig	TT	1.570	0.924	1.588	0.910	1.619	0.908	46	[8-1044-1]
Purelin	Tansig	PT	1.531	0.927	1.480	0.916	1.527	0.920	46	[8-1065-1]

Note: OLAF = output layer activation function; HLAF = hidden layer activation function; IN, iteration number

Table 6: MLP-ANN model performance statistics for training; validation; and testing of the CGF method for $BOD_{eff.}$

		CGF								
			Training		Validation		Testing			
HLAF	**OLAF**	**Designation**	**MSE**	**R^2**	**MSE**	**R^2**	**MSE**	**R^2**	**IN**	**Architecture**
Logsig	Purelin	LP	1.523	0.937	1.624	0.814	1.545	0.935	52	[8-1024-1]
Logsig	Tansig	LT	1.576	0.910	1.649	0.867	1.546	0.927	57	[8-1220-1]
Tansig	Purelin	TP	1.546	0.924	1.633	0.819	1.583	0.850	37	[8-1230-1]
Tansig	Logsig	TL	5.306	0.767	5.432	0.714	5.573	0.676	41	[8-1275-1]
Purelin	Logsig	PL	5.308	0.696	5.484	0.778	5.515	0.677	46	[8-1005-1]
Logsig	Logsig	LL	5.306	0.794	5.432	0.745	5.573	0.785	27	[8-989-1]
Purelin	Purelin	PP	1.567	0.887	1.457	0.953	1.678	0.872	16	[8-1031-1]
Tansig	Tansig	TT	1.530	0.966	1.613	0.947	1.552	0.929	28	[8-1052-1]
Purelin	Tansig	PT	4.679	0.677	4.617	0.796	4.262	0.750	64	[8-1033-1]

Note: OLAF = output layer activation function; HLAF = hidden layer activation function; IN, iteration number

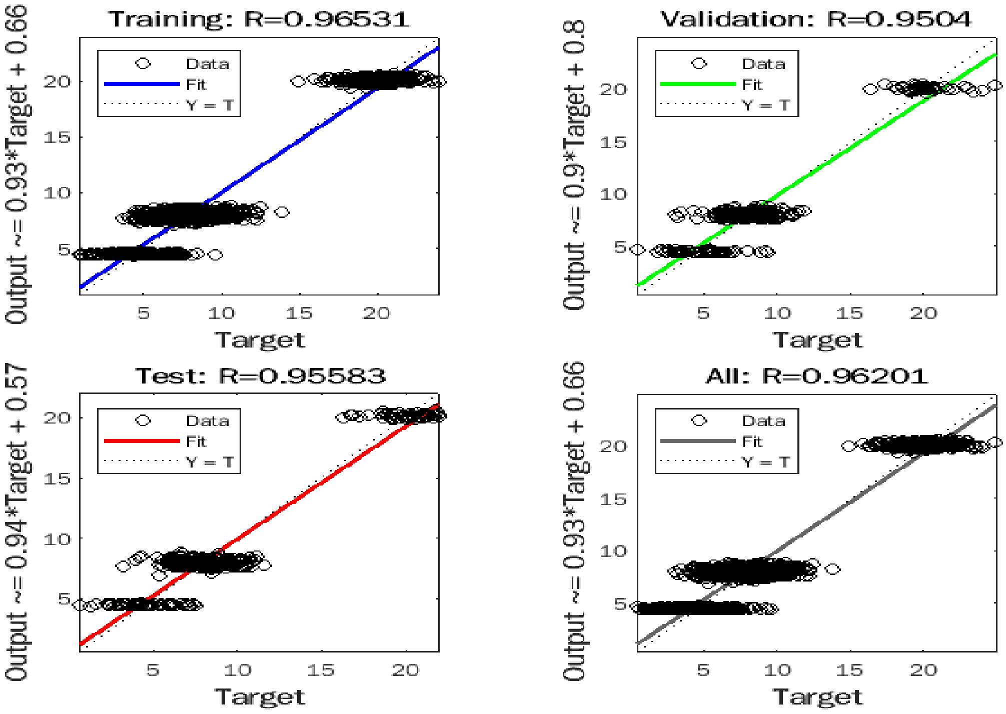

Fig 5: The best ANN model's regression plot between actual and predicted data on the BOD_{eff}

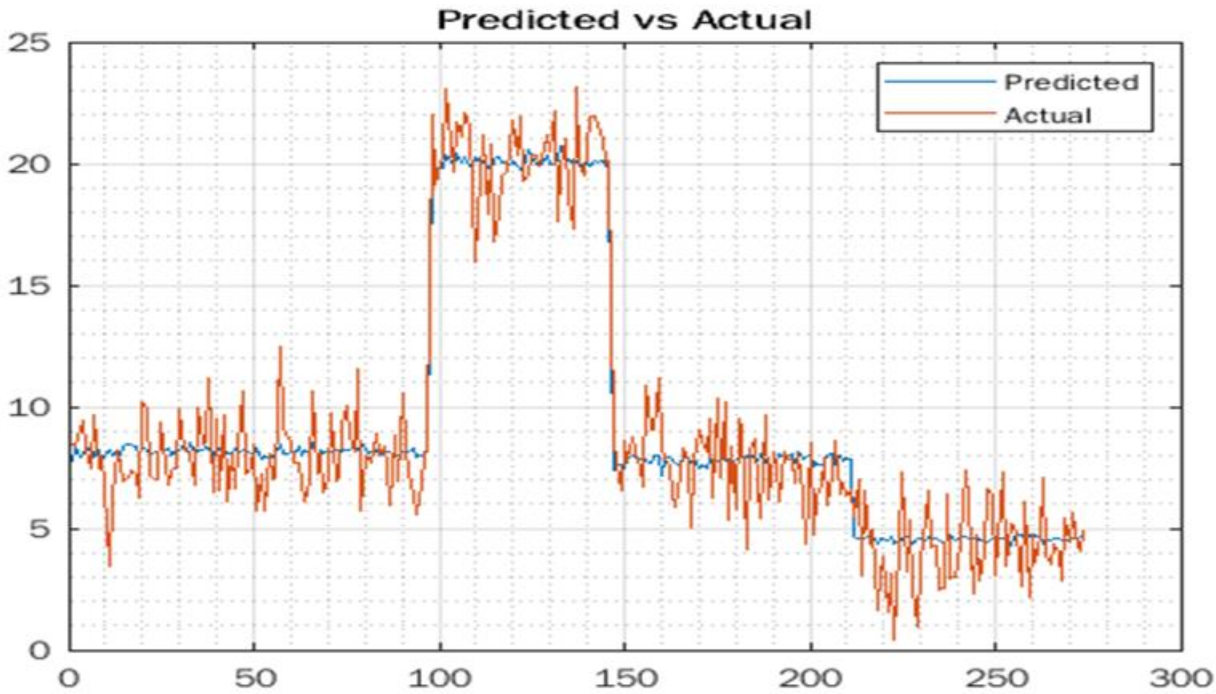

Fig 6*:* BOD_{eff} output actual values against predicted values for the best ANN model

V. CONCLUSION

Models that estimate WWTP performance based on historical data of key plant characteristics can help achieve optimal operational conditions in WWTPs, which are crucial for sustainable water resource management. The use of MLP-ANN, a renowned prediction model recognized for its high accuracy, to estimate the performance of the Abuja-based Wupa WWTP utilizing effluent BOD_5 as a performance indicator was investigated in this study. Over a five-year period (2013–2017), daily data for influent BOD_5, pH, TDS, TSS, COD, TC, EC, and FC; and effluent BOD_5 for the Wupa WWTP were used.

Four reliable multilayer perceptron ANN algorithms namely, LM, RP, BFG and CGF were adopted; and the most appropriate model was selected following training, validation and testing of the models by changing the number of neurons and activation functions (tansig, purelin and logsig) in the hidden and output layers. The model efficiency was determined using mean square error (MSE) and the correlation coefficient (R^2).

The study revealed that the LM algorithm with the Logsig-Tansig activation pairing and architecture [8-1270-1] exhibited the best results in terms of convergence time and prediction error with MSE and R^2 values of 1.522 and 0.922 respectively. Also, it showed that the selected ANN adequately predicted the effluent BOD_5 with an overall correlation coefficient of 0.962; thus, demonstrating the efficacy of ANN models for accurate prediction of the Wupa WWTP performance.

FUNDING

This study received no particular support from governmental, commercial, or not-for-profit funding entities.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Ethical approval

There are no studies involving human participants or animals done by any of the authors in this article.

REFERENCES

[1]. Abba SI, Elkiran G (2017) Effluent prediction of chemical oxygen demand from the wastewater treatment plant using artificial neural network application. Procedia Computer Science 120: 156-163. https://doi.org/10.1016/j.procs.2017.11.223

[2]. Abba SI, Usman AG, Danmaraya YA, Usman AG, Abdullahi HU (2020) Modeling of water treatment plant performance using artificial neural network: case study Tamburawa Kano-Nigeria. Dutse Journal of Pure and Applied Sciences 6(3): 135-144.

[3]. Ahmadzadeh S, Rezayi M, Karimi-Maleh H, Alias Y (2015) Conductometric measurements of complexation study between 4-Isopropylcalix [4] arene and Cr3+ cation in THF–DMSO binary solvents Measurement70:214–224. http://dx.doi.org/10.1016/j.measurement.2015.04.005

[4]. Alsulaili and A.Refaie. (2021). Artificial Neural Network Modeling Approach for the Prediction of Five-Day Biological Oxygen Demand and Wastewater Treatment Plant Performance. *Water Supply , 21*(5), 1861-1877.

[5]. Gawdzik, J. Gawdzik, B. Gawdzik, A. Gawdzik, and M. Rybotycki . (2020). Prediction of BOD5 Content of the Inflow to the Treatment Plant using Different Methods of Black Box – the Case Study . *Desalination and Water Treatment, 196*, 58-66.

[6]. Katip. (2018). The usage of Artificial Neural Networks in Microbial Water Quality Modeling: a Case Study from the Lake Iznik. *Applied Ecology and Environmental Research, 16*(4), 3897-3917.

[7]. A. Saleh and H. Kayi. (2021). Prediction of Chemical Oxygen Demand from the Chemical Composition of Wastewater by Artificial Neural Networks. *Journal of Physics: Conference Series, 1818*(012035), 1-11.

[8]. Szeląg, K. Barbusiński, J. Studziński and L. Bartkiewicz . (2017). Prediction of Wastewater Quality Indicators at the Inflow to the Wastewater Treatment Plant using Data Mining Methods. *E3S Web of Conferences, 22*(00174), 1-8.

[9]. O. Araromi, O. T. Majekodunmi, J. A. Adeniran and T. O. Salawudeen. (2018). Modeling of an Activated Sludge Process for Effluent Prediction—a Comparative Study using ANFIS and GLM Regression. *Environmental Monitoring and Assessment, 190*(495), 1-17.

[10]. Dogan, A. Ates, E. C. Yilmaz and B. Eren. (2008). Application of Artificial Neural Networks to Estimate Wastewater Treatment Plant Inlet Biochemical Oxygen Demand. *Environmental Progress, 27*(4), 439-446.

[11]. R. Rene and M. B. Saidutta . (2008). Prediction of BOD and COD of A Refinery Wastewater using Multilayer Artificial Neural Networks. *Journal of Urban and Environmental Engineering, 2*(1), 1-7.

[12]. K. Banaei, A. A. L. Zinatizadeh, M. Mesgar, and Z. Salari. (2013). Dynamic Performance Analysis and Simulation of a Full Scale Activated Sludge System Treating an Industrial Wastewater Using Artificial Neural Network. *International Journal of Engineering, 26*(5), 465-472.

[13]. A. A. Saleh. (2021). Wastewater Pollutants Modeling Using Artificial Neural Networks. *Journal of Ecological Engineering, 22*(7), 35-45.

[14]. Xue. (2017). Prediction of Chemical Oxygen Demand Emissions in Wastewater Treatment Plant Based on Improved Artificial Neural Network Model. *Chemical Engineering Transactions, 62*, 1453-1458 .

[15]. Jasir Javad, Alaa H.Hawari, Sayed Javaid Zaidi. (2021). *Artificial neural network modelling of wastewater tretment and dessalination using membrane processes.* Centre for Advanced Materials, Qatar University, P.O. Box 2713, Doha, Qatar: Chemical Engineering Journal 419 (2021) 129540.

[16]. J. Setshedi, N. Mutingwende and N. P. Ngqwala. (2021). The use of Artificial Neural Networks to Predict the Physicochemical Characteristics of Water Quality in Three District Municipalities, Eastern Cape Province, South Africa. *International Journal of Environmental Research and Public Health, 8*(10), 5248.

[17]. Arismendy, D. Cárdenas, D. Gómez, A. Maturana, R. Mejía and C. G. M. Quintero. (2020). Intelligent System for the Predictive Analysis of an Industrial Wastewater Treatment Process. *Sustainability, 12*(6348), 1-19.

[18]. B. Varkeshi, K. Godini, M. ParsiMehr and M. Vafaee. (2019). Predicting the Performance of Gorgan Wastewater Treatment Plant Using ANN-GA, CANFIS, and ANN Models. *Avicenna Journal of Environmental Health Engineering, 6*(2), 1-8.

[19]. M. Hamed, M. G. Khalafallah and E. A. Hassanien. (2004). Prediction of Wastewater Treatment Plant Performance using Artificial Neural Networks. *Environmental Modelling and Software*, *19*, 919–928.

[20]. S. Jami, I. A. F. Husain, N. A. Kabashi and N. Abdullah. (2012). Multiple Inputs Artificial Neural Network Model for the Prediction of Wastewater Treatment Plant Performance. *Australian Journal of Basic and Applied Sciences*, *6*(1), 62-69.

[21]. M. Vyas, B. Modhera, V. Vyas and A. K. Sharma. (2011). Performance Forecasting of Common Effluent Treatment Plant Parameters by Artificial Neural Network. *ARPN Journal of Engineering and Applied Sciences, 6*(1), 38-42.

[22]. Bekkari and A. Zeddouri. (2018). Using Artificial Neural Network for Predicting and Controlling the Effluent Chemical Oxygen Demand in Wastewater Treatment Plant. *Management of Environmental Quality*, *30*(3), 593-608.

[23]. R. J. May, H. R. Maier and G. C. Dandy. (2010). Data Splitting for Artificial Neural Networks using SOM-Based Stratified Sampling. *Neural Networks*, *23*(2), 283-294.

[24]. S. I. Abba and G. Elkiran. (2017). Effluent Prediction of Chemical Oxygen Demand from the Wastewater Treatment Plant using Artificial Neural Network Application. *Procedia Computer Science*, *120*, 156-163 .

[25]. S. I. Abba, A. G. Usman, Y. A. Danmaraya, A. G. Usman and H. U. Abdullahi. (2020). Modeling of Water Treatment Plant Performance using Artificial Neural Network: Case Study Tamburawa Kano-Nigeria. *Dutse Journal of Pure and Applied Sciences, 6*(3), 135-144.

[26]. T. C. Ogwueleka and N. F. Ogwueleka. (2009). Application of Artificial Neural Networks in Estimating Wastewater Flows. *The IUP Journal of Science and Technology, 5*(3), 20-30.

[27]. T. C. Ogwueleka and N. F. Ogwueleka. (2010). Data Mining Application in Predicting Cryptosporidium Spp. Oocysts and Giardia Spp. Cysts Concentrations in Rivers. *Journal of Engineering Science and Technology, 5*(3), 342-349.

[28]. W. Wu , R. May, G. C. Dandy and H. R. Maier. (2012). A Method for Comparing Data Splitting Approaches for Developing Hydrological ANN Models. *2012 International Congress on Environmental Modelling and Software.* Leipzig, Germany.