GSC Advanced Engineering and Technology

Cross Ref DOI: 10.30574/gscaet

Journal homepage: https://gsconlinepress.com/journals/gscaet/

(Research Article)

Check for updates

Performance analysis and control of wastewater treatment plant using Adaptive Neuro-Fuzzy Inference System (ANFIS) and Multi-Linear Regression (MLR) techniques

Obasi Peter Okeke [1], Ismail I Aminu [2], Abdulazeez Rotimi [3], Bara'u Gafai Najashi [4], M.M Jibril [5,*] Awaisu Shafiu Ibrahim [6], Abba Bashir [6], Salim Idris Malami [5,7], Habibu M.A [8] and Mohammed Mukhtar Magaji [7]

[1] *Department of Electrical and Electronic Engineering, University of Abuja, Nigeria.*
[2] *Department of Civil Engineering Bayero University, Kano.*
[3] *Department of Civil Engineering, Baze University, Abuja, Nigeria.*
[4] *Department of Electrical and Electronic Engineering, Baze University, Abuja, Nigeria.*
[5] *Department of Civil Engineering, Kano University of science and technology, KUST, Wudil, Nigeria*
[6] *Department of Civil Engineering, Federal University, Dutsin-ma, Katsina, Nigeria.*
[7] *Department of Civil Engineering, Institute for Infrastructure & Environment, Heriot-Watt University, Edinburgh, UK.*
[8] *Department of Electrical Engineering, Federal University Gusau, Nigeria.*

GSC Advanced Engineering and Technology, 2022, 04(02), 001–016

Publication history: Received on 13 April 2022; revised on 06 October 2022; accepted on 09 October 2022

Article DOI: https://doi.org/10.30574/gscaet.2022.4.2.0033

Abstract

In order to provide an effective tool for the simulation of wastewater treatment plant performance and control, a reliable model is essential. In the present study, two different artificial intelligence (AI) models; Adaptive Neuro-Fuzzy inference system (ANFIS), and a classical multi-linear regression analysis (MLR) were applied for predicting the performance of Abuja wastewater treatment plant (AWWTP), in terms of Conductivity, pH, Iron content, BOD, COD, TSS and TDS. The daily data were obtained from Abuja Wastewater treatment plant, for this purpose, single and ensemble models were employed to compare and improve the prediction performance of the plant. The obtained result of single models proved that, MLR model provides an effective analysis in comparison to the other single model. The result showed that, conductivity influences the performance and efficiency of the water treatment plant by an increased efficiency performance of AI modelling up to 99.6% testing phase and 6.8% Error value of same phase. This shows that MLR model was more robust and reliable method for predicting the Abuja WWTP performance.

Keywords: Artificial Intelligence; Adaptive Neuro-Fuzzy Inference system; Multi-Linear Regression; Performance Indices; Training; Testing

1. Introduction

Water is essential to sustain life; therefore, affordable and adequate supply of water must be available [1]. Wastewater Treatment Plant (WWTP) are process that remove the contaminants from the untreated domestic wastewater with the goal of safeguarding the public health and natural environment [2]. 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 satisfy requirements demands [3]. WWTP is extremely complex and dynamic process due to its intricacy of the treatment method. Appropriate action, maintenance and control of WWTPs is very vital for monitoring the environmental and ecological.

* Corresponding author: MM Jibril
Department of Civil Engineering, Kano University of science and technology, KUST, Wudil.

A satisfactory treatment plant is quite vital in order to avoid the discharge of high pollutants and meet the required standards by law [4]. The parameters combination from physical, chemical and biological characteristics are often the major factors affecting the operation and control of WWTPs. Due to various composition and characteristic of wastewater treatment plant (WWTP) variables, it's performance can be assessed by considering certain sensitive variables such as Total Nitrogen (TN), Biological Oxygen Demand (BOD), Total Suspended Solids (TSS) and Chemical Oxygen Demand (COD), Total Dissolved Solids (TDS), Conductivity, pH, etc. Yet, the available literatures and published studies for predicting the WWTP used these parameters. The quality of untreated and treated sewage has a great effect on the operation and efficiency of any WWTP. However, WWTP comprises of large numbers of parameters and operations which are complex in terms of measurement and evaluation[5].Hence modeling this system is considered difficult due to the nature of the process and most of the available traditional models are based on the assumptions, estimation and requiring too much time and money, as such a reliable and appropriate tool are indispensable in predicting the performance of ANFIS [6].

Due to the importance of wastewater management, planning and control, the modeling field in this remain dynamic and active of study. Basically, the models applied in hydro-environmental studies can be categorized into two; namely, physical based and data-driven models. Physical models are based on the concept of distributed (white-box) models addressed the physical process and interaction for simulating hydro-environmental system. In contrast, data-driven models are based lumped (black-box) models that acquire the optimal links between inputs and outputs but neglect the physical process [7]. Various efforts have been presented to improve the accuracy and reliability of the effluent variables in the field of hydro-environmental studies, but no individual method has been proved applicable in modeling environmental process (Danandeh et al., 2018). With regards to this perspective, it could be pronounce that there is no acceptable single models which can perform better than the other in different hydro-environmental system, due to the dynamic and complex nature of the data [8]. This has neccessitate the development of the reliable and efficient models using the available data. In addition, the process of WWTP have both deterministic and stochastic system, stochastic time series model such as Multilinear regression analyses (MLR), Autoregressive (AR) models have been used in modeling and prediction of hydrological process especially time-series process. The AR is widely known by its moderation and simplicity among the linear models and is employed in several modeling studies. Owing to its linear nature, AR may not reliably and properly model the possibly intricate processes taking place in WWTP.

Based on the established WWTP, 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 AI methods which are considered as accurate and non-linear hydrologic tools. Meanwhile, several researchers have established different types of intelligence techniques such artificial intelligence (AI) 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. On the other side, the artificial intelligence (AI) models play a vital role and created great variations for forecasting several environmental and hydrological phenomena. Meanwhile, recent researches showed that the black-box models like Support Vector Machine (SVM), Artificial Neural Networks (ANN) and Adaptive Neuro Fuzzy Inference System (ANFIS) could be proper alternatives for any WWTPs performance analysis.

For example, Maleki et al. (2018) predicted the influent parameters in WTP using Auto Regressive Integrated Moving Average (ARIMA) and Neural Network Auto-regression (NNAR) models, despite an acceptable performance of ARIMA model, the results observed better prediction performance for NNAR with regard to ARIMA. Chen et al. (2001) developed ANN, genetic algorithm (GA) and fuzzy logic (FL) as new method for modeling the industrial WWTPs at Taiwan. The proposed new method served as the control strategies in successful management of the WWTP. Verma et al. (2013) demonstrated the ability of five different data mining approaches includes, MLP, K-nearest neighbor, SVM, random forest and multi-variate adaptive regression spline to estimates the total suspended solid (TSS) in a WWTP using different input parameters. The obtained results depicted that MLP outperformed all the models.

Memon et al., (2012) developed artificial neural network (ANN) with multi-layer perceptron (MLP) model to forecast the treated and untreated pH using the 17 measured input parameters in water treatment plant (WTP). The outcomes proved the suitability of MLR in modeling the drinking WTP parameters. Granata et al., (2017) made an attempt using several types of algorithms (i.e., support vector regression (SVR) and regression tree (RT)) to simulate wastewater quality indicator such as BOD, COD, TDS and TSS. From the outcomes, it was observed that both models showed the robustness and reliability in the prediction, however a significant performance of SVR was observed with regards to RT in modeling the effluent TSS, and COD.

Similarly, Gaya et al. (2017) developed the first implementation of ANN and Hammerstein-Wiene (H-W) models for forecasting the influent turbidity in Tamburawa WTP using different input parameters. The results indicated that ANN could outperform the H-W model and could serve as acceptable tool for modeling the turbidity of WTP. Guo et al. (2015)

used the influents of pH, Temperature, COD and SS to predict the concentration of TN effluent from the WWTP in Ulsan, Korea, by employing ANN and SVM models and concluded that, AI models could be reliable methods for prediction of the effluent conditions of the WWTPs. Civelekoglu et al. (2009) applied ANN and ANFIS methods to model the COD removal in biological WWTP, the overall results indicated that, ANFIS is a suitable model for prediction of the WWTP performance. Hamed et al. (2004) used the BOD and TSS values recorded at various positions as input parameters and outlet BOD and TSS as target variables to predict the performance of WWTP using ANN model. The results proved the ability of ANN model for predicting WWTP performance.

As the literature review shows, there is no unique model to be superior to others in all cases and the performances of different models may be different according to condition of each WWTP. Therefore, it is tested and verified that the combination of outputs (from different models) through an ensemble method may lead to more accurate results. The idea of such an ensemble model has been already used at different fields of engineering, environmental and water quality modelling. However, since the pronouncement of ensemble methods in engineering, to the best of the authors' knowledge, there is no published study in the technical literature indicating the application of AI based ensemble approach in WWTP modeling.

1.1. Problem Statement

Essential for life, clean water is one of the most important natural resources on the planet. However, wastewater contains many harmful substances and cannot be released back into the environment until it is treated. Thus, the importance of wastewater treatment is twofold: to restore the water supply and to protect the planet from toxins. Look at a global drought map and you will see that many areas of the world simply do not have enough water just like is the case in FCT, Nigeria. All communities, especially areas with water scarcity, need to ensure they have good water treatment processes in place so that treated water can either be reused or returned to the water cycle, but never wasted.

Wastewater can include contaminants from both residential and commercial use. Untreated, the chemical compounds and pathogens in wastewater can harm the health of animals, plants and birds that live in or near the water. It can also contaminate crops and drinking water, affecting human health. Wastewater treatment is fundamental to protect the health of many different ecosystems Wastewater, properly treated, is a source of water for many purposes to Abuja residents. Good wastewater treatment allows the maximum amount of water to be reused instead of going to waste.

- Problems of water scarcity from all sectors
- Growing urban developments
- Operation and control of WWTPs is difficult and time consuming
- The general WWTP system is Complex
- Traditional linear model is based on the rough estimation, linear approximation and assumption.
- There is need for reliable and convenient modeling tool

The aims of this work are as found below:

- To perform the sensitivity analysis techniques to determine the most dominant parameters.
- To develop an independent model for the sensitive parameter
- To determine the performance of Abuja WTP using ANFIS
- To determine the performance of Abuja WTP using MLR
- To compare the performance accuracy between ANFIS and MLR

1.2. Significance of the Study

According to United Nations Educational, Scientific and Cultural Organization (UNESCO) 2015, WWTPs is paramount important for sustainable development and critical for human health ecosystems. This study will overcome the problems of water scarcity in the Nigeria particularly Northern Nigeria. This will serve as another important benchmark that will strengthen Water security. Another important significance of this study is the agricultural production, the treated effluents will substantially be used for irrigation, other farming activities and recharging the aquifers. Finally, the study may serve as the background for researchers carrying out further studies in Abuja WWTP.

1.3. Definition

Water treatment is any process that improves the quality of water to make it appropriate for a specific end-use. The end use may be drinking, industrial water supply, irrigation, river flow maintenance, water recreation or many other uses, including being safely returned to the environment. Water treatment removes contaminants and undesirable components, or reduces their concentration so that the water becomes fit for its desired end-use. This treatment is crucial to human health and allows humans to benefit from both drinking and irrigation use.

Treatment for drinking water production involves the removal of contaminants and/or inactivation of any potentially harmful microbes from raw water to produce water that is pure enough for human consumption without any short term or long term risk of any adverse health effect. In general terms, the greatest microbial risks are associated with ingestion of water that is contaminated with human or animal (including bird) feces. Feces can be a source of pathogenic bacteria, viruses, protozoa and helminths. The removal or destruction of microbial pathogens is essential, and commonly involves the use of reactive chemical agents such as suspended solids to remove bacteria, algae, viruses, fungi, and minerals including iron and manganese. These substances continue to cause great harm to several lower developed countries who do not have access to water purification.

Measures taken to ensure water quality not only relate to the treatment of the water, but to its conveyance and distribution after treatment. It is therefore common practice to keep residual disinfectants in the treated water to kill bacteriological contamination during distribution and to keep the pipes clean.

Water supplied to domestic properties, for tap water or other uses, may be further treated before use, often using an in-line treatment process. Such treatments can include water softening or ion exchange. Many proprietary systems also claim to remove residual disinfectants and heavy metal ions.

1.4. Work Structure

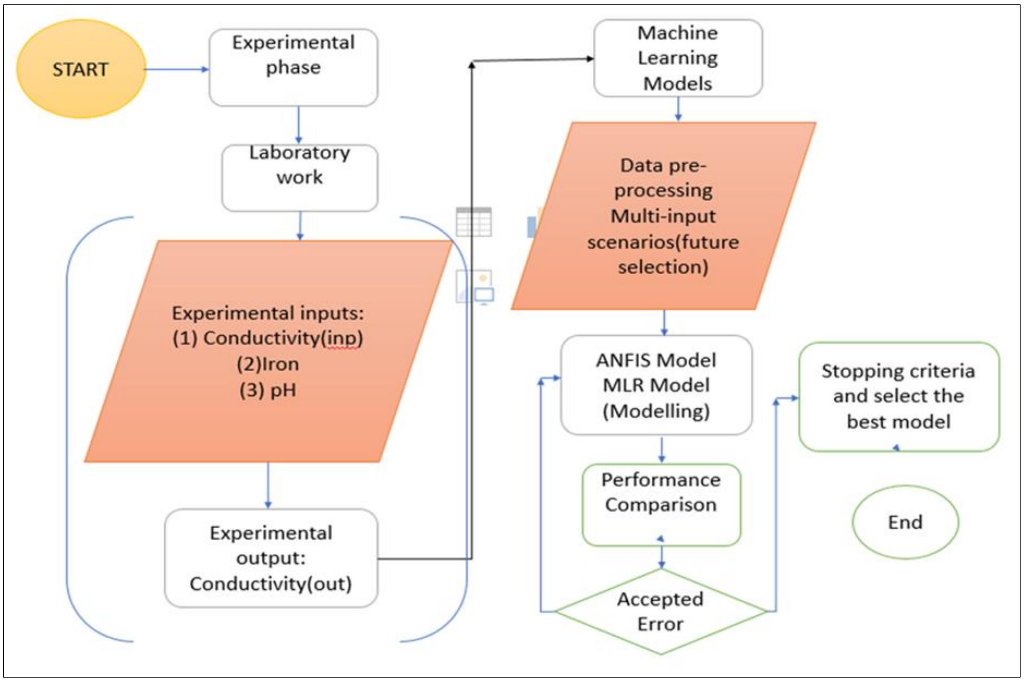

Figure 1 General proposed flowchart of the proposed methodology used in this study

2. Material and methods

2.1. Study area and location

2.1.1. Study area

Abuja, the Federal Capital Territory of Nigeria houses the Wupa Sewage Treatment Plant. The Territory is an 8000km2 area of land which became the new Federal Capital Territory (FCT) of Nigeria in 1976. It was not an existing city, hence a Master plan for its development was prepared to define the general structure and required infrastructure which includes buildings, road networks, and water and sewage system amongst others. Abuja is centrally located in Nigeria, bounded to the North by Kaduna State, to the South by Kogi State, to the West by Niger State and to the East by Nasarawa

State as shown in Figure 2 In line with the Master Plan, the Federal Capital City (FCC), is planned to be developed in Phases I, II, III, IV according to availability of resources or as requirements arise with centralized sewerage system made up sewage treatment plants connected with sewer lines (Oluwadamisi, 2013). Abuja is a tropical climatic area characterized by two annual weather conditions: the rainy season (April – October) and the dry season (November – March). The average temperature during rainy season ranges between 30.4 °C to 35.1 °C while during dry season, the range is between 34.5 °C – 39.2 °C. During the dry season, relative humidity falls in the afternoons, but the high altitudes and the undulating terrain of the FCT act as moderating influence on the weather of the Territory. The resultant effect is the frequent rainfall and a noticeable increase in the mean annual total rainfall from the South to the North within the range of 1100 mm to 1600 mm (Aondoakaa, 2012; Oluwadamisi, 2013). Also, the FCT is well drained by series of small and big rivers. Two main rivers flowing within the vicinity of Abuja: The Gurara River which flows in the northwest direction while the Usuma River flows in the northeast direction. There is also River Jere, Gerinya, Pegna, Izom, Wupa etc.

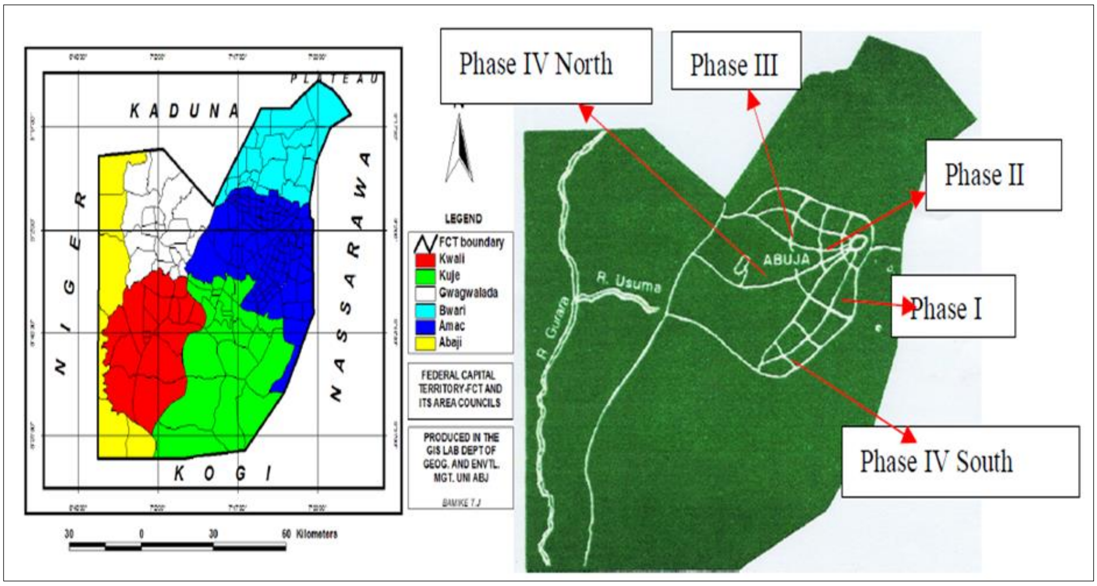

Figure 2 Map of Federal Capital Territory showing Abuja

2.1.2. Plant's Description

Wupa is a small community in Idu District of the FCC and an Industrial Zone. This is where the Wupa Sewage treatment plant is located and close to the Wupa River as showed in Figure 1a. The Wupa Sewage Treatment Plant (WSTP) covers an area of 297,900 square meters. The Plant has an average dry weather inflow of 131,250 cubic meters per day and designed to meet the requirement of 700,000 PE (Population Equivalence). There are five drainage basins in the FCC, one of which is the Wupa Basin with the WSTP. Light industries, institutions, commercial and public buildings located within Idu Industrial zone and Districts of the FCC have their wastewater channeled through network of sewer lines to the Wupa Basin Sewage Treatment Plant. The basic wastewater treatment process in Wupa Basin Sewage treatment plant is activated sludge processes with its flow diagram shown as Figure 1.2. WSTP was designed with a peak inflow of 9,000 m^3/d, COD load of 84,000 kg/d, BOD and TSS load of 42,000 kg/d and the plant's final effluent expected to fall within the nationally acceptable discharge limits into waterways. Sewage treatment by the WSTP is divided into mechanical and biological processes. From the sewer lines to the treatment plant's inlet, sewage is received and passed through the coarse screen which removes debris and wastes larger than 5 cm. Three sets of alternating screw-pumps homogenize and lifts the sewage to about 8 m to the fine screen for removal and dewatering of debris and wastes larger than 4 mm. The grit chamber, made up of two compartments removes grease, grit, scum and colloidal particles while the sewage is passed into six distribution wells which channels sewage into each Bioreactors for the biological processes. The bioreactor is the heart of the aerobic treatment where biological processes take place with a volume of 22,700 m^3. It consists of twelve (12) mammoth rotors and eight (8) submersible mixers. Aeration occurs with the aid of the rotors introducing atmospheric oxygen into the sewage while the submersible mixers keep the biomass in suspension. In 30.4 hours, Hydraulic Retention Time (HRT), nitrification, denitrification and BOD removal takes place in the bioreactor. Sedimentation Tanks separates the clear water from sludge while the sludge is removed by the return sludge pump. The disinfection Tanks consist of three (3) channels of 144 UV lamps each where the treated wastewater passes through for disinfection before the effluent is discharged into Wupa River. The returned sludge is passed through the gravity thickener, dewatering system and the final thickened sludge is sent to the drying bed effluent before it gets discharge into the river (UNDP, 2014).

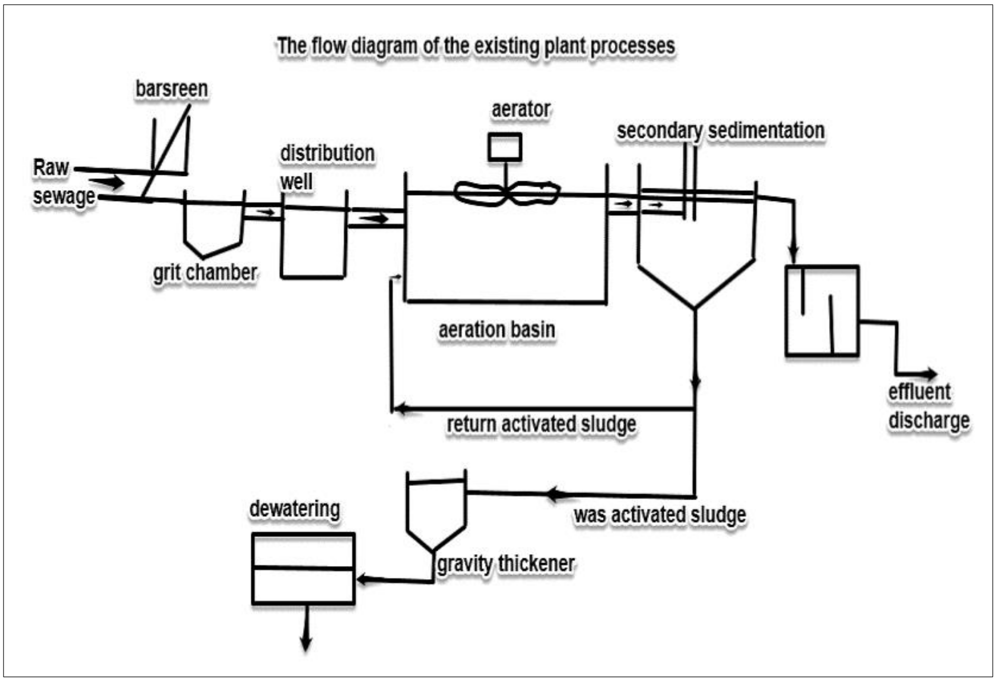

Figure 3 Flow chart of Wua Sewage Treatment Plant, Abuja

2.2. Plant and Operational line of Treatment

In Wupa WWTP the line of treatment comprises of 11 stages from the raw to treated sewage effluents as shown in Fig. 3 First of all, the sewage is separated into liquid and solid waste and goes through the first chamber called *screening chamber* (1) where the solids larger than 6 mm will be removed. The inflow then flows down slowly so that the heavy solids (grit, sand) can fall to the bottom and oil and grease float to the surface at *grit and grease chamber* (2). The pump station (3) was used to pump up the water to the next unit called *fine sieve* (4), this unit removes solids larger than 2mm. The next step is the biological treatment of waste which is the stage that creates the condition to encourage bacteria to consume the waste comprising of three units (5, 6 and 7). Stage (8) is where the separation and treatment of the by-products of the hall process into clean water, fertilizer and biogas is taking place named membrane treatment. After that, the water is disinfected at *chlorine contact tank* (10). Tank (11) is the treated sewage Fig. 3 shows the schematic of the Wupa WWTP line of treatments.

The summary of the line of treatment are as follows:

- Screening Chamber
- Grit and grease chamber
- Pump Station
- Fine Sieve
- Biological treatment
- Membrane treatment
- Chlorine contact

2.3. Data used and Pre-processing

For this project, the available daily data were obtained from Abuja WWTPs. The measurement of the selected parameters covers all the seasonal variations and consists of various sets of inputs and outputs parameters. The daily measured data obtained from the WWTP which includes (pH, Conductivity, TDS, TSS, BOD, COD, Total Nitrogen, Total Phosphor, Oxygen etc.) as the corresponding output respectively. Table 1.0 show the input and output parameters used in this project. Table 1 shows the general parameters used for the study. Note the selection of the parameters are quite line with the previous research done in the field of WWTP performance analysis. However, other studied engaged some of these parameters in the control and management of the WWTP from the angle of AI- based models.

Table 1 Parameter used (Influents and Effluents)

Parameters	Influents Parameter	Effluents Parameters
pH	pHinf	pHeff
Conductivity	Condinf	Condeff
Biological Oxygen Demand	BODinf	BODeff
Chemical Oxygen Demand	CODinf	CODeff
Total Nitrogen	T-Ninf	T-Neff
Iron	Iron (Inf)	Iron (eff)
Total Phosphorous	T-Pinf	T-P eff
Suspended Solid	TDSinf	TDSeff
Total Suspended Solid	TSSinf	TSSeff

2.4. Data Processing and Statistical analysis

Data processing is the process of turning the raw data into appropriate and meaningful information, prior to the model training, the data must be scale between the internal of 0 and 1 this process is called normalization as seen in equation (1) below. The process was applied in order to deduce the data redundancy and increased data integrity (Abba and Elkiran, 2017). The normalized data were divided in to 75 % and 25 % for both calibration and verification, respectively. The validation methods are implemented using different approach. Here, Normalization (Xi) approach was used to analyze the acquired data.

$$\boldsymbol{X_i} = \frac{x_u - x_{min}}{x_{max} - x_{min}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots . (1)$$

Where "X_i" is the normalized quantity, "x_u" is un- normalized quantity, "x_{min}" is the minimum and "x_{max}" is the maximum quantity of the data set (Nourani et al., 2012).

Statistical analysis used to explain the data trend series are smoothing and normalization, the former was carried out by fitting the data in to regression function to eliminate the noise from the data and latter was to ensure the uniformity of the input-output value (scaling to fall within a small, specified range). The descriptive statistic of the selected parameter can be presented in Table 2 Every data analysis concerning AI base models relies normally on the historical data [10], [11], [12], [13], [14], [15], [16]. Therefore, the data and statistical analysis of the input-output is essentials because it identifies the type and strength of the relations between inputs and outputs. In order to efficiently train AI base model, these data need to be clean and filtered properly, because the raw data often comprised of missing records, outliers, noise, discrepancies of codes and names or was infected by all kind of error including human and instrumental. More information on normalization and statistical analysis can be found in [17], [18], [19], [16], [20].

2.5. Proposed Methodology

In this project work, three different scenarios were proposed separately for modeling and prediction the performance of new Nicosia WWTP. The first scenario I, explored the application of data-driven algorithms (i) to develop and compare the potential of some AI based models ANFIS and conventional multi-linear model (MLR) for prediction of the Wupa WWTP performance considering four different combinations of input parameters. Other feasible alternatives models may also be used, but they were adopted here due to their outstanding performances in various literature in hydro-environmental studies. Some alternatives are Genetic programing, ARIMA models, machine learning models as presented in [21]–[28].

2.6. Reasons for combining linear and non-linear models

It is difficult to determine in practice whether one model in particular is better than others. Thus, selecting the proper method for a particular case is a difficult task for the predictors. The complexity of selecting the appropriate models can be resolved by choosing assembly of various models. The traditional linear models are still used despite the inability to provide the accurate outputs due to their various limitations and inconsistences to handle non-stationary and non-linearity data. Such linear models are still applicable because, a) traditional linear models are economical, uncomplicated, and the natural phenomenon can be employed in a functional linear system, b) non-linear models magnify the noise for additional time steps while the linear models increase the noise included in the data linearly. Therefore, applying the traditional linear model for linear portions of the process is recommended. The natural and real-world processes may contain both the linear and non-linear characteristic. As such, ARIMA, AR and MLR models are not capable of handling non-linear system solely. On the other hand, an AI model may expand the noise of the linear pattern, and therefore both models cannot adequately estimate the time series of the process individually. Hence, by combining results of traditional model and AI models, the magnify non-linear behaviors of the noise and complex architecture can be addressed in a simple approach. More reason on combining both traditional models and AI models can equally be justified from [29], [30], [31], [31], [32], [33], [34], [35] and [36]

2.7. Methodology

2.7.1. Adaptive Neuro-Fuzzy Inference System (ANFIS)

The integration of ANN and fuzzy logic creates a robust hybrid system that is used to solve nonlinear complex functions. ANFIS is a kind of ANN that is based on Takagi-Sugeno fuzzy interference system. It is a Multi-Layer Feed-Forward (MLFF) neural network, used to map inputs with outputs. The output of each fuzzy rule could be a linear combination of input variables plus a constant term. Generally, two types of learning algorithms are employed in ANFIS; the BP and hybrid learning. The BP learning in ANFIS is similar to that of ANN. The hybrid learning consists of the combination of BP and least squares methods. ANFIS hybrid learning algorithm is faster to converge than the conventional back propagation method. The ANFIS model has the advantages of both ANN and fuzzy logics [37]. Since ANFIS combine the topology of ANN and Fuzzy logic, it covers both their methodology and limitations. Therefore, the ANFIS model provides the optimum desired outcomes quickly with less error and without any uncertainty or vagueness. Moreover, in terms of learning duration, ANFIS model learns within a very short time compared to ANN model. The general structure of ANFIS is shown in figure 4.

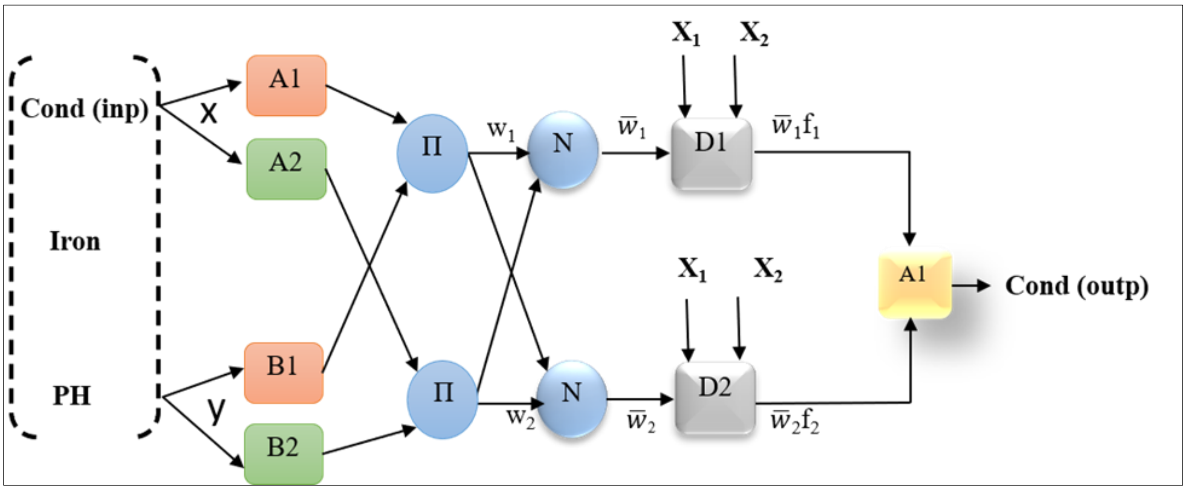

Figure 4 Proposed structure for ANFIS model used in this study

Assume that the ANFIS have two inputs 'x' and 'y' and one output 'f'. A first- order Sugeno fuzzy has the following rules.

$$\text{Rule 1: if } \mu(x) \text{ is } A_1 \text{ and } \mu(y) \text{ is } B_1 \text{ then } f_1 = p_1x + q_1y + r_1 \quad \ldots\ldots\ldots\ldots\ldots (2)$$

$$\text{Rule 2: if } \mu(x) \text{ is } A_2 \text{ and } \mu(y) \text{ is } B_2 \text{ then } f_2 = p_2x + q_2y + r_2 \quad \ldots\ldots\ldots\ldots\ldots (3)$$

Where A_1, B_1, A_2, B_2 are membership function parameters for x and y inputs, and $p_1, q_1, r_1, p_2, q_2, r_2,$ are the outlet function parameters. The structure and formulation of ANFIS follow a five-layer neural network arrangement.

Layer 1: In this layer, every node i is an adaptive node having a node function as equation 4.

$$Q^1_{\ i} = \mu_{Ai}(x) \text{ for } i = 1,2 \text{ or } Q^1_{\ i} = \mu_{Bi}(x) \text{ for } i = 3,4 \ldots\ldots\ldots\ldots\ldots (4)$$

Where $Q^1_{\ i}$ is the membership grade for input x or y. The membership function chosen was Gaussian because it has the lowest prediction error.

Layer 2: In this layer, every rule between inputs is connected by T-norm operator that perform as an 'AND' operator.

$$Q^2_{\ i} = w_i = \mu_{Ai}(x).\,\mu_{Bi}(y)\ for\ i = 1,2 \qquad \ldots\ldots\ldots\ldots\ldots (5)$$

Layer 3: In this layer, every neuron is labelled Norm, and the output is called 'Normalized firing strength'.

$$Q^3_{\ i} = \overline{w_i} = \frac{w_1}{w_1+w_2}\ , 1,2\ldots\ldots\ldots\ldots\ldots (6)$$

Layer 4: In this layer, every node i is an adaptive node having a node function as in equation 7.

$$Q^4_{\ i} = \overline{w_i}(p_1 x + q_1 y + r_1) = \overline{w_i} f_i \ldots\ldots\ldots\ldots\ldots (7)$$

Where $p_1, q_1, r_{1,}$ are irregular parameters referred to as 'consequent parameters?

Layer 5: In this layer, the overall output is computed as the summation of all incoming signals.

$$Q^5_{\ i} = \sum \overline{w_i}\, f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \qquad \ldots\ldots\ldots\ldots\ldots (8)$$

2.7.2. Multi-Linear Regression (MLR)

Multi-linear regression (MLR) is a famous method of modeling mathematically, the linear relationship between one or more independent variables and dependent variable. In general, the dependent variable y, and n regressor variables may be related. The model is defined with n regressor variables and is called MLR model. The equation is given by Ladlani et al. (2014).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \ldots\ldots\ldots + \beta_n x_n \ldots\ldots\ldots\ldots\ldots (9)$$

Where β_i are n regressor coefficients with$i = 1 \ldots\ldots\ldots.n$. The model explained a hyperplane in $n -$ dimensional space of regressor variablesxj. The change in y response expected per unit change in xj is represented by parameter β_j when the entire independent variables remaining x_i $(i \neq j)$ are kept constant [31].

3. Results and discussion

Table 2 Performance Criteria of the Models

Models	Calibration				Verification			
	R^2	R	MSE	RMSE	R^2	R	MSE	RMSE
M1-ANFIS	0.008737	0.093473	0.03526	0.187777	0.973304	0.986562	0.004817	0.069405
M2-ANFIS	0.017124	0.130858	0.034962	0.186981	0.972583	0.986196	0.004947	0.070336
M3-ANFIS	0.263842	0.513655	0.026186	0.161821	0.981342	0.990627	0.003367	0.058023
M1-MLR	0.006887	0.082986	0.035326	0.187952	0.996128	0.998062	0.004722	0.068718
M2-MLR	0.007701	0.087753	0.035297	0.187875	0.99482	0.997407	0.004793	0.069232
M3-MLR	0.008176	0.090422	0.03528	0.18783	0.995071	0.997533	0.00477	0.069064

Performance of Wastewater treatment plant has been determined using AI models of ANFIS and MLR. In this study, various input variables (pH, Iron, Conductivity, TDS, TSS, BOD, COD) were put together to determine the performance of Abuja Wastewater treatment plant using its output variable. Sensitivity analysis was carried to determine the contribution of each of the input parameters to the overall performance of the treatment plant. The predictive accuracy of the models (MLR & ANFIS) was evaluated using R^2, MSE, R, and RMSE as seen in table 2 below. Both models were implemented using MATLAB software.

3.1. Simulation Results

This study presented the comparative analysis of both ANFIS model and MLR for the estimation of the performance and control of wastewater treatment plant. The modeling results were evaluated using R^2, R, RMSE, and MSE in both training and testing phases.

The most dominant and suitable input combinations with the targeted variables were investigated using traditional sensitivity analysis and a correlation matrix. The type of linear relationship between the variables is represented as seen in **Error! Reference source not found.** It can also be used as a basic indicator for the correlation of variable sets. As shown in **Error! Reference source not found.**, the stationary and significant variables with probability less than 0.05 ($P<0.05$) indicates the high strength of the linear correlations. Also, the negative correlation values show an inverse relationship between two variables. As a result, the correlation value's weakness indicates that traditional methods are ineffective in modeling such complex interactions, and that there is a significant need to introduce more robust tools. The AI models used, and their corresponding combined variables are seen in table 3 below.

Table 3 Correlation between the experimental variables

	Conductivity (input)	**PH**	**IRON**	**Conductivity (output)**
Conductivity(input)	1			
pH	-0.23174	1		
Iron	-0.3775	0.235057	1	
Conductivity (outp)	-0.09404	0.010185	0.057194	1

Table 4 AI Models and their corresponding Variables

M1	Cond (inp)+Cond (outp)
M2	Cond (inp) +pH+ cond (outp)
M3	Cond (inp)+pH+ Iron+ cond (outp)

From the results in table 3, the model combinations were generated, from M1, M2, and M3 using both ANFIS and MLR models. Table 4 shows the variables that were used as inputs for each model to forecast observed parameters with Adaptive Neuro-Fuzzy inference system (ANFIS) Designer tool of MATLAB was used. The input and output parameters of the membership function (MF) was tuned in order to generate a Sugeno-type fuzzy inference system. For the input parameter a triangular MF type was selected and for the output parameter, a constant MF type was selected. The FIS was trained with an error tolerance of 0.005 for 50 iterations (epochs).

The normalized parameters produced by both models were partitioned into training (75 %) and testing (25 %) in order to properly evaluate the performance of ANFIS and MLR in determining the performance of the WWTP. The results of the performance criteria are displayed in table 2. It can be seen that the MLR models, MLR-M1 produced the best results with values of R^2 = 0.996128 R= 0.998062 MSE=0.004722 and RMSE=0.068718 while ANFIS-M3 produced the best results of the ANFIS models with values of R^2=0.981342 R=0.990627 MSE=0.003367 and RMSE=0.058023. Hence, MLR model is more accurate than ANFIS for this study.

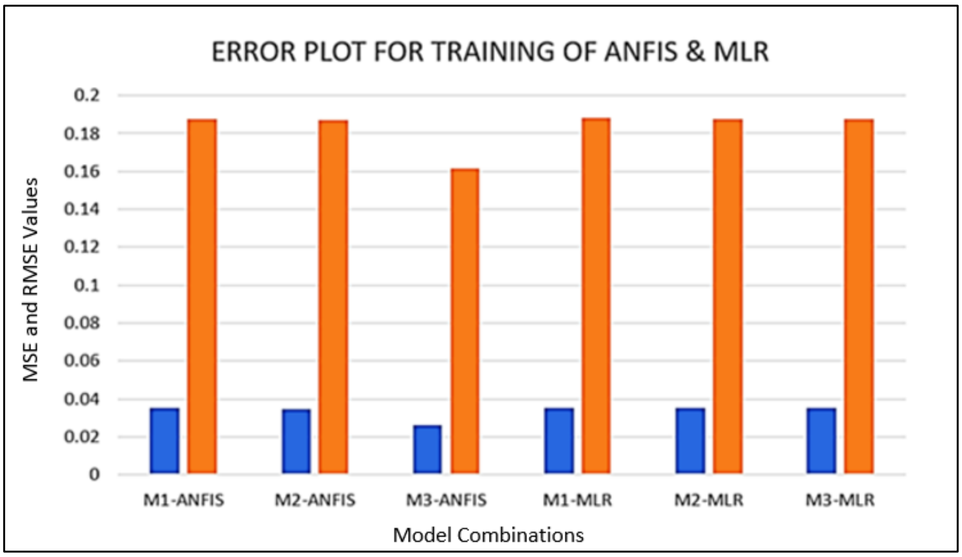

Figure 5 Error plot for Training of ANFIS & MLR

From Fig 5 above, the MES and RMSE for M2-ANFIS is higher than those of the M3-ANFIS which implies that M3-ANFIS gives a better result relative to M2-ANFIS.

Also, in Fig 6 below MSE and RMSE for M1-MLR are both lower than those of M2-MLR and M3- MLR which indicates that the result of M1-MLR is better than those of M2 and M3-MRL models. Same performance is also evident in figure 6 as seen below.

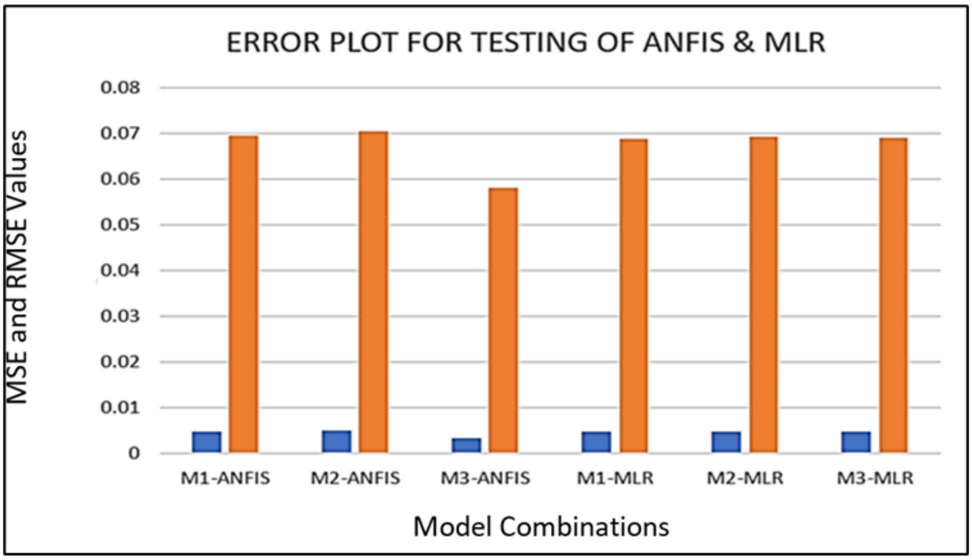

Figure 6 Error plot for Testing of ANFIS and MLR

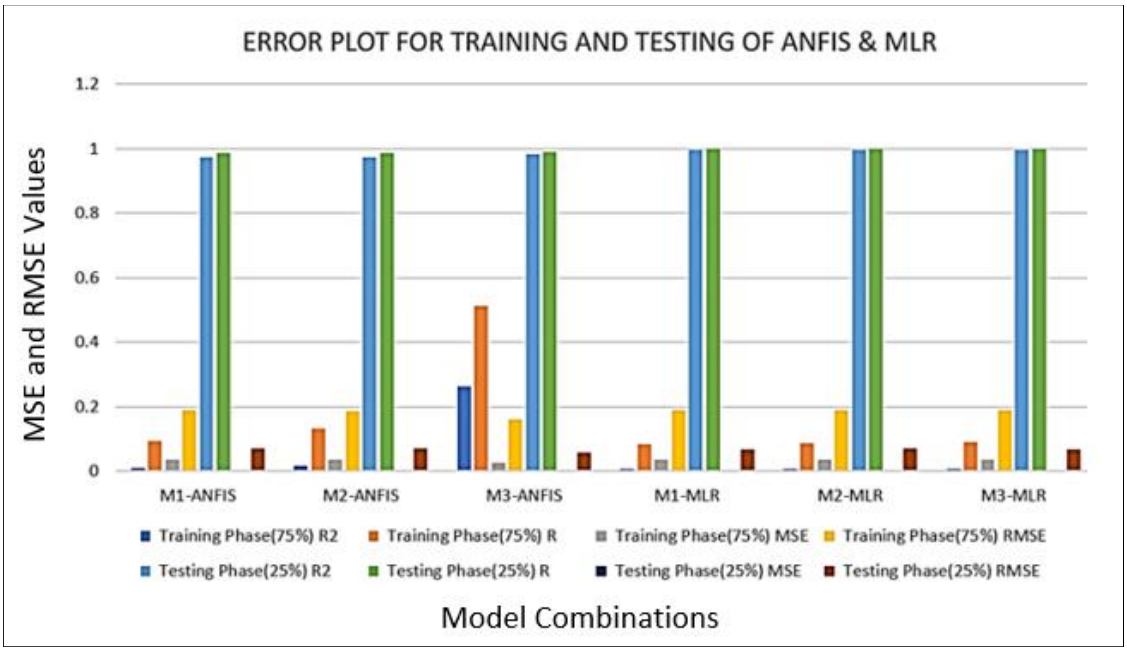

Figure 7 Error plot for Training and Testing of ANFIS & MLR

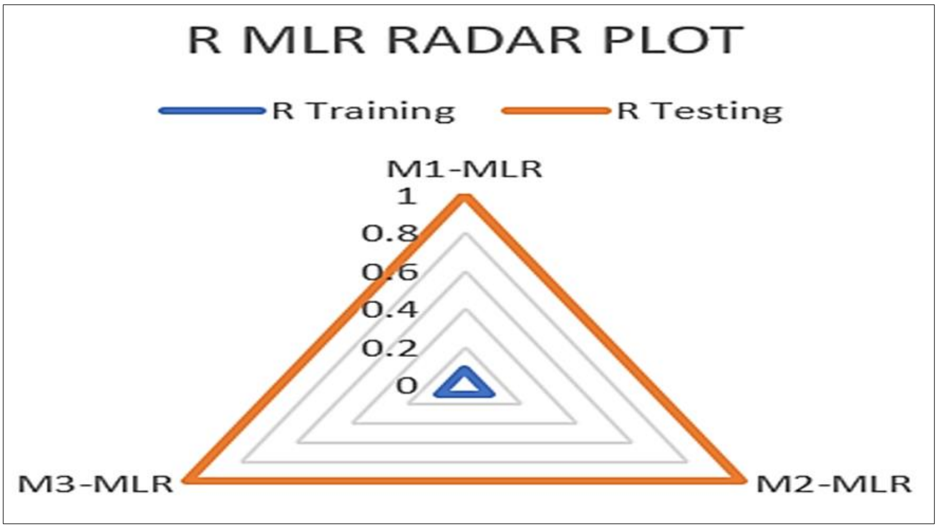

Figure 8 MLR Radar plot for R for both Training & Testing

The R value for M1-MLR is higher compared to the R value of M3-ANFIS as seen in fig. 8 above and Fig 9 below. This is an indication that the M1-MLR gives a better result compared to M3-ANFIS.

From the time series plot for M1-MLR (fig 10 above), the prediced performance is higher than the of M3-ANFIS as seen in fig 11 below. Also, the perdicted and observed performance are closer for M1-MRL relative to M3-ANFIS which reflects that MLR model provides a better performance for this study compared to ANFIS model as analysed.

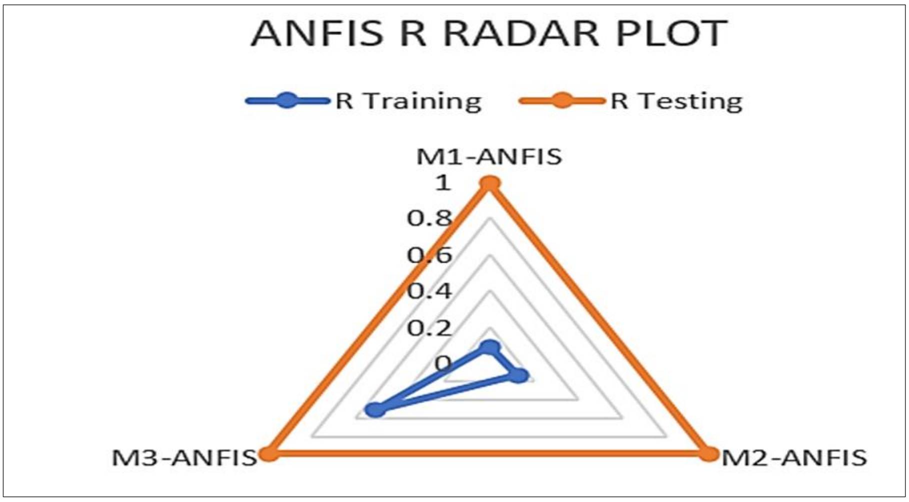

Figure 9 ANFIS Radar plot for R for both Training & Testing

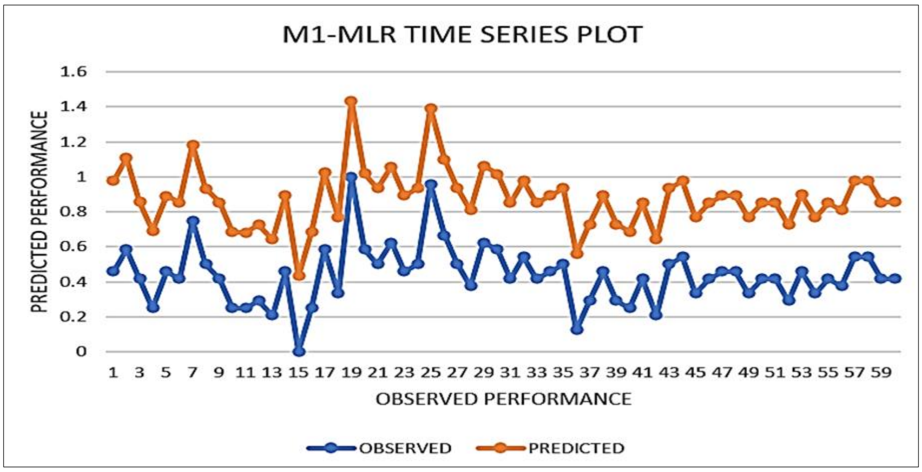

Figure 10 M1-MLR TIME SERIES PLOT

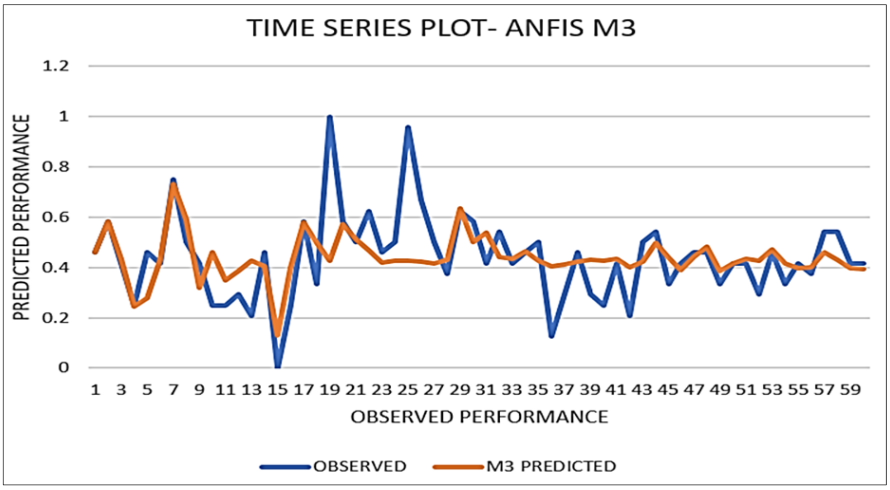

Figure 11 M3-ANFIS TIME SERIES PLOT

Observation

It can be clearly observed from Table 2, Fig 10 and 11 that multi-linear regression (MLR) model gives out a better result as compared to Adaptive neuro fuzzy inference system (ANFIS). The performances index (R2) for MLR gives a higher value compared to that of ANFIS. Also, from the scatter plot, the predicted performance for M1-MLR is higher than that of M3-ANFIS, hence, MLR model gives a better result/performance for the WWTP compared to ANFIS for this project work.

4. Conclusion

The analytical technique for determining the performance and control of a Wastewater treatment plant(WWTP) have several computations of performance indices, therefore, analyzing WWTP performance and control using traditional sensitivity method may be tedious, require more time with reduced accuracy of performance indices, hence, with ANFIS & MLR approaches this technique can be more suitable which consist of training and testing phases for different variables that influences the performance of a water treatment plant. Accurate result was obtained from the testing phases with the performance indices (conductivity, iron and pH). MLR and ANFIS models were used to analyze performance of the WWTP and MLR model proved to be more accurate in determining the WWTP performance and error indices.

Recommendation

Due to the complexity surrounding the analysis of wastewater treatment plant performance and control, traditional sensitivity method cannot give the expected high accuracy of result, hence, the need for AI models of ANFIS and MLR.

Future project work of similar AI model analysis is advised to be carried out with data acquired over a period of years in order to achieve high accuracy of performance indices.

Compliance with ethical standards

Acknowledgments

The authors acknowledged this research to Humans and Machines LTD, and Civil Lake LTD for their funding assistance.

Disclosure of conflict of interest

No conflict of interest.

References

[1] WHO, "WHO | Guidelines for safe recreational water environments," vol. 1, pp. 118–127, 2013, [Online]. Available: http://www.who.int/water_sanitation_health/bathing/srwe1/en/.

[2] Y. Hong, L. Fei, X. Yuge, and L. Jin, "GA based LS-SVM classifier for waste water treatment process," 2008 27th Chinese Control Conf., pp. 436–439, 2008, doi: 10.1109/CHICC.2008.4605860.

[3] G. Elkiran and A. Turkman, "Water Scarcity Impacts On Northern Cyprus And Alternative Mitigation Strategies Provided for non-commercial research and education use . Not for reproduction , distribution or commercial use .," no. January 2016, 2008, doi: 10.1007/978-1-4020-8960-2.

[4] S. I. Abba et al., "Emerging evolutionary algorithm integrated with kernel principal component analysis for modeling the performance of a water treatment plant," J. Water Process Eng., vol. 33, no. December 2019, p. 101081, 2020, doi: 10.1016/j.jwpe.2019.101081.

[5] S. I. Abba and G. Elkiran, "Effluent prediction of chemical oxygen demand from the astewater treatment plant using artificial neural network application," in Procedia Computer Science, 2017, vol. 120, doi: 10.1016/j.procs.2017.11.223.

[6] Q. B. Pham et al., "Potential of Hybrid Data-Intelligence Algorithms for Multi-Station Modelling of Rainfall," Water Resour. Manag., vol. 33, no. 15, pp. 5067–5087, 2019, doi: 10.1007/s11269-019-02408-3.

[7] S. J. Hadi, A. J. Hadi, K. S. Ismail, and M. Tombul, "Homogeneity and Trend Analysis of Rainfall and Streamflow of Seyhan Basin (Turkey)," Advances in Sustainable and Environmental Hydrology, Hydrogeology, Hydrochemistry and Water Resources. Springer International Publishing, pp. 17–19, 2019, doi: 10.1007/978-3-030-01572-5_4.

[8] S. I. Abba and A. G. Usman, "Modeling of Water Treatment Plant Performance using Artificial Neural Network: Case Study Tamburawa Kano-Nigeria," Dutse J. Pure Appl. Sci. (DUJOPAS), Vol., vol. 6, no. 3, pp. 135–14, 2020.

[9] V. Nourani, H. Hakimzadeh, and A. B. Amini, “Implementation of artificial neural network technique in the simulation of dam breach hydrograph,” J. Hydroinformatics, vol. 14, no. 2, p. 478, 2012, doi: 10.2166/hydro.2011.114.

[10] U. M. Ghali et al., “Applications of Artificial Intelligence-Based Models and Multi- Linear Regression for the Prediction of Thyroid Stimulating Hormone Level in the Human Body,” vol. 29, no. 4, pp. 3690–3699, 2020.

[11] G. M. Khalid and A. G. Usman, “Application of data-intelligence algorithms for modeling the compaction performance of new pharmaceutical excipients,” 2021.

[12] A. N. Alsharksi, Y. A. Danmaraya, H. U. Abdullahi, U. M. Ghali, and A. G. Usman, “Potential of Hybrid Adaptive Neuro Fuzzy Model in Simulating Clostridium Difficile Infection Status,” no. 1, pp. 1–6, 2020, doi: 10.35940/ijbsac.A0191.073120.

[13] S. I. Abba et al., “Comparative implementation between neuro-emotional genetic algorithm and novel ensemble computing techniques for modelling dissolved oxygen concentration,” Hydrol. Sci. J., vol. 0, no. 0, 2021, doi: 10.1080/02626667.2021.1937179.

[14] S. I. Abba, A. G. Usman, and S. Is, “Simulation for response surface in the HPLC optimization method development using arti fi cial intelligence models : A data-driven approach,” vol. 201, no. March, 2020, doi: 10.1016/j.chemolab.2020.104007.

[15] A. G. Usman, S. Işik, S. I. Abba, and F. Meriçli, “Chemometrics-based models hyphenated with ensemble machine learning for retention time simulation of isoquercitrin in Coriander sativum L. using high-performance liquid chromatography,” J. Sep. Sci., vol. 44, no. 4, pp. 843–849, 2021, doi: 10.1002/jssc.202000890.

[16] A. G. Usman, S. Işik, and S. I. Abba, “Hybrid data-intelligence algorithms for the simulation of thymoquinone in HPLC method development,” J. Iran. Chem. Soc., vol. 18, no. 7, pp. 1537–1549, 2021, doi: 10.1007/s13738-020-02124-5.

[17] M. H. Ahmad, A. G. Usman, and S. I. Abba, “Comparative performance of extreme learning machine and Hammerstein–Weiner models for modelling the intestinal hyper-motility and secretory inhibitory effects of methanolic leaf extract of Combretumhypopilinum Diels (Combretaceae),” Silico Pharmacol., vol. 9, no. 1, 2021, doi: 10.1007/s40203-021-00090-1.

[18] U. M. Ghali et al., “Advanced chromatographic technique for performance simulation of anti-Alzheimer agent: an ensemble machine learning approach,” SN Appl. Sci., vol. 2, no. 11, 2020, doi: 10.1007/s42452-020-03690-2.

[19] A. G. U. Selin and I. S. I. Abba, “A Novel Multi - model Data - Driven Ensemble Technique for the Prediction of Retention Factor in HPLC Method Development,” Chromatographia, no. 0123456789, 2020, doi: 10.1007/s10337-020-03912-0.

[20] W. Asnake Metekia, A. Garba Usman, B. Hatice Ulusoy, S. Isah Abba, and K. Chirkena Bali, “Artificial intelligence-based approaches for modeling the effects of spirulina growth mediums on total phenolic compounds,” Saudi J. Biol. Sci., no. xxxx, 2021, doi: 10.1016/j.sjbs.2021.09.055.

[21] S. Ismail, R. A. A. A. G. Usman, and S. I. Abba, “Development of chemometrics - based neurocomputing paradigm for simulation of manganese extraction using solid - phase tea waste,” Model. Earth Syst. Environ., 2022, doi: 10.1007/s40808-022-01369-8.

[22] R. Costache et al., “Flash-flood hazard using Deep Learning based on H2O R package and fuzzy-multicriteria decision-making analysis,” J. Hydrol., vol. 609, no. March, p. 127747, 2022, doi: 10.1016/j.jhydrol.2022.127747.

[23] A. G. Usman et al., “Simulation of liver function enzymes as determinants of thyroidism: a novel ensemble machine learning approach,” Bull. Natl. Res. Cent., vol. 46, no. 1, 2022, doi: 10.1186/s42269-022-00756-6.

[24] S. S. Sammen et al., “Assessment of climate change impact on probable maximum floods in a tropical catchment,” Theor. Appl. Climatol., pp. 15–31, 2022, doi: 10.1007/s00704-022-03925-9.

[25] B. Tawabini, M. A. Yassin, M. Benaafi, J. A. Adetoro, A. Al-Shaibani, and S. I. Abba, “Spatiotemporal Variability Assessment of Trace Metals Based on Subsurface Water Quality Impact Integrated with Artificial Intelligence-Based Modeling,” Sustain., vol. 14, no. 4, 2022, doi: 10.3390/su14042192.

[26] S. I. Abba et al., “Integrating feature extraction approaches with hybrid emotional neural networks for water quality index modeling,” Appl. Soft Comput., vol. 114, p. 108036, 2022, doi: 10.1016/j.asoc.2021.108036.

[27] M. Jamei et al., "Earth Skin Temperature Long-Term Prediction using Novel Extended Kalman Filter Integrated with Artificial Intelligence Models and information gain feature selection," Sustain. Comput. Informatics Syst., vol. 35, no. March, p. 100721, 2022, doi: 10.1016/j.suscom.2022.100721.

[28] H. Tao et al., "Groundwater Level Prediction using Machine Learning Models: A Comprehensive Review," Neurocomputing, vol. 489, pp. 271–308, 2022, doi: 10.1016/j.neucom.2022.03.014.

[29] M. S. Gaya et al., "Estimation of turbidity in water treatment plant using Hammerstein-Wiener and neural network technique," Indones. J. Electr. Eng. Comput. Sci., vol. 5, no. 3, pp. 666–672, 2017.

[30] V. Nourani, G. Elkiran, and S. I. Abba, "Wastewater treatment plant performance analysis using artificial intelligence – an ensemble approach," Water Sci. Technol., 2018, doi: 10.2166/wst.2018.477.

[31] S. I. Abba and G. Elkiran, "Effluent prediction of chemical oxygen demand from the astewater treatment plant using artificial neural network application," Procedia Comput. Sci., vol. 120, pp. 156–163, 2017, doi: 10.1016/j.procs.2017.11.223.

[32] S. I. Abba et al., "Evolutionary computational intelligence algorithm coupled with self-tuning predictive model for water quality index determination," J. Hydrol., vol. 587, no. April, p. 124974, 2020, doi: 10.1016/j.jhydrol.2020.124974.

[33] G. Elkiran, V. Nourani, and S. I. Abba, "Multi-step ahead modelling of river water quality parameters using ensemble artificial intelligence-based approach," J. Hydrol., p. 123962, Jul. 2019, doi: 10.1016/J.JHYDROL.2019.123962.

[34] Z. M. Yaseen et al., "Predicting compressive strength of lightweight foamed concrete using extreme learning machine model," Adv. Eng. Softw., vol. 115, no. April 2017, pp. 112–125, 2018, doi: 10.1016/j.advengsoft.2017.09.004.

[35] S. I. Abba, Y. S. Said, and A. Bashir, "Assessment of Water Quality Changes at Two Location of Yamuna River Using the National Sanitation Foundation of Water Quality," J. Civ. Eng. Environ. Technol., vol. 2, no. 8, pp. 730–733, 2015.

[36] S. I. Haruna et al., "Compressive Strength of Self-Compacting Concrete Modified with Rice Husk Ash and Calcium Carbide Waste Modeling: A Feasibility of Emerging Emotional Intelligent Model (EANN) Versus Traditional FFNN," Arab. J. Sci. Eng., vol. 46, no. 11, pp. 11207–11222, 2021, doi: 10.1007/s13369-021-05715-3.

[37] R. A. Abdulkadir, S. I. A. Ali, S. I. Abba, and P. Esmaili, "Forecasting of daily rainfall at Ercan Airport Northern Cyprus: a comparison of linear and non-linear models," Desalin. Water Treat., vol. 177, no. May 2019, pp. 297–305, 2020, doi: 10.5004/dwt.2020.25321