International Journal of Earth Sciences Knowledge and Applications (2021) 3 (2) 134-150

International Journal of Earth Sciences Knowledge and Applications
www.ijeska.com

Research Article

e-ISSN: 2687-5993

Lake Water Level Prediction Model Based on Artificial Intelligence and Classical Techniques – An Empirical Study on Lake Volta Basin, Ghana

Michael Stanley Peprah[1]*, Edwin Kojo Larbi [1]

[1]*Department of Cartography, Ghana Cocoa Board, Ghana West Africa*

INFORMATION

Article history
Received 17 October 2020
Revised 04 January 2021
Accepted 13 January 2021
Available 15 March 2021

Keywords
Artificial Intelligence
Lake Volta Basin
Stochastic Models
Time Series Analysis
Water Level Modelling

Contact
* Michael Stanley Peprah
E-mail: mspeprah91@gmail.com

ABSTRACT

Several studies in the past and recent years have suggested numerous mathematical models for Lake Water Level (LWL) modelling to a good precision. This study considered an empirical evaluation of Artificial Intelligence and Classical Techniques such as Wavelet Transform (WT), Bayesian Regularization Backpropagation Artificial Neural Network (BRBPANN), Levenberg-Marquardt Backpropagation Artificial Neural Network (LMBPANN), Scaled Conjugate-Gradient Backpropagation Artificial Neural Network (SCGBPANN), Radial Basis Functions Artificial Neural Network (RBFANN), Generalized Regression Artificial Neural Network (GRANN), Multiple Linear Regression (MLR), and Autoregressive Integrated Moving Average (ARIMA) for LWL modelling. The motive is to apply and assess for the first time in our study area, the working efficiency of the aforementioned techniques. Satellite altimetry data provided by the United States Department of Agriculture was used in this study. The input and output variables used in this study were the decomposed LWL by the WT. Each model technique was assessed based on statistical measures such as Arithmetic Mean Error (*AME*), Arithmetic Mean Square Error (*AMSE*), arithmetic mean absolute percentage deviation (*AMAPD*), minimum error value (r_{min}), maximum error value (r_{max}), and arithmetic standard deviation (*ASD*). The statistical analysis of the results revealed that, all the hybridized models successfully estimate the LWL heights at a good precision for the study area. However, Discrete Wavelet Transform (DWT)-MLR model outperforms DWT-BRBPANN, DWT-LMBPANN, DWT-SCGBPANN, DWT-RBFANN, DWT-GRANN, and DWT-ARIMA techniques in estimating the LWL heights for the study area. In terms of *AME*, *AMSE* and *ASD*, DWT-MLR achieved 0.1988 m, 0.0024 m, and 0.0017 m respectively. The main conclusion drawn from this study is that, the method of using novel ensemble models is promising and can be adopted for LWL modelling in the study area. This study seeks to contribute to the existing knowledge on understanding the hydrodynamic processes in Lake Volta Basin and support water resource management.

1. Introduction

LWL predictions using time series analysis has become one of the major researches focuses among geophysicists, geodesists and oceanographers. As a significant characteristic of the Lake ecosystem, precise LWL forecasting is important for lake water resources management and proper planning purposes (Barzeger et al., 2020; Aytek et al., 2014). Additionally, LWL studies help geospatial professionals and

governmental authorities in decision making on climate adaptation and mitigating strategies to reduce risk in spatial contest (Grgic et al., 2017). Moreover, water level (WL) studies contribute massively to the understanding nature of processes, patterns and interactions of the changing earth (Pashova and Popova, 2011). Precise determination of the WL can be done using available monthly and the associated meteorological data (El-Shazly, 2005). However, this is a difficult task in many scientific and practical fields of applications (Sithara et al., 2020; Ebtehaj et al., 2019; Al-Krargy et al., 2017). This is due to the variation of the WL known as sea surface topography (SST) and meteorological conditions (El-Shazly, 2005). Conversely, this can be as a result of non-climatic factors such as tectonic activity, changes in local catchment morphology and anthropogenic activities (Srivasta et al., 2020). These factors by climate and human activities have affected the accessibility of ecological LW and has decreased tremendously especially in arid and semi-arid areas (Liu et al., 2019). LWL fluctuation is a complex and dynamic process, characterized by highly stochasticity, non-linearity problems, and difficult to model and forecast (Zhu et al., 2020a). However, this has drawn the attention of many researchers to develop models for simulating of the extreme or abnormal LWL variations in order to control future LWL changes (Aytek et al., 2014), In view of that, several scholars have proposed different methodologies to aid in estimating the LWL to a good precision.

The WL could be determined either by a relative technique (tide gauges) (Tolkatchev, 1996) with reference to a local geodetic datum (geoid) on land or an absolute technique (satellite altimetry) (Mitchum, 2000) with reference to a global reference datum (ellipsoid). Additionally, the classical levelling techniques can be used to determine the WL (El-Shazly, 2005). The presented study considered the latter approach (absolute technique) in determining the LWL due to the availability of satellite altimeter data for the study area. Notably among some of the applications of absolute techniques that have been applied in the past and recent years include tides measurement (Tziavos et al., 2005), precise geoid determination (Andersen and Knudsen, 1998; Cazenave et al., 1996), global mean sea surface models (Lemoine et al., 1998; Wenzel, 1998) and regional mean sea surface models (Arabelos and Tziavos, 1996; Vergos et al., 2003) as well as the recovery of gravity anomalies from the altimetry measurements.

Recent studies reveal that, there are limitations in the relative techniques for WL determination such as inadequate and inhomogeneous distribution of the used geodetic data (Amin, 2003) and geophysical process within the earth's system that cause changes in global WL (Srivasta et al., 2020; Makarynskyy et al., 2004; Tziavos et al., 2005). In addition, the classical levelling techniques does not truly represent WL and it varies from place to place (El-Shazly, 2005). The inconsistency of the relative techniques led to the investigation of how modern techniques and abundant data source may be utilized for geodetic purpose (Turner et al., 2013). Precise prediction of WL can be done using sophisticated mathematical models which include time series and climate data (Pashova and Popova, 2011; Ledolter, 2008). The traditional method for WL prediction includes the classical least square (El-Shazly, 2005; Jian-Jun, 2003), Kalman Filter (Okwuashi and Olayinka, 2017; Adnan et al., 2012; Peprah and Larbi, 2021), Hydro-Balance models (Fischer et al., 2020; Lin et al., 2015), Soil and Water Assessment Technique (SWAT) model (Muthuwatta, 2004), and ARIMA (Ebtehaj et al., 2019; Farajzadeh et al., 2014; Fernandez et al., 2018; Srivastava et al., 2016; Fernandez et al., 2017; Makwinga et al., 2017; Ledolter, 2008). However, in recent times, the level of uncertainty in the existing approaches for LWL forecast methods have increased significantly due to climate change (Ebtehaj et al., 2019). Hence, there is the need to develop more sophisticated accurate models for precise estimation of LWL. These aforementioned techniques are very viable but have limitations due to their inabilities to model noisy data, nonlinearity between the dataset and data availability (Yakubu et al., 2018) (such as lake discharge and meteorological data which were not available) and for that matter not considered in this study. For that reason, soft computing methods which have the advantage of using available data directly for forecasting without any simplification and requirements was adopted (Adnan et al., 2012). Notably, the Artificial Neural Network (ANN) is one of the commonly used soft computing methods (Yaseen et al., 2020; Zhu et al., 2020a; Zhu et al., 2020b; Demir and Ulke Keskin, 2020; Deo and Chaudhari, 1998).

Soft computing is a machining learning technique which is built on how the neurons in the human brain operates (Piasecki et al., 2015). ANN is one of the most widely used soft computing method which is capable of directly correlating the multiple input variables with output variables through iteratively learning (Piri and Kahkha, 2016; Pashova and Popova, 2011). Some examples of ANN techniques that have been applied in the recent decades in WL studies include Least Squares Support Vector Machine (LSSVM) (Kaloop et al., 2020; Sithara et al., 2020; Kaloop et al., 2017; Okwuashi and Ndehedehe, 2017; Shafaei and Kisi, 2015), Backpropagation Artificial Neural Networks (Ebtehaj et al., 2019; Yaseen et al., 2020; Piasecki et al., 2015; Adnan et al., 2012; Pashova and Popova, 2011; Ghorbani et al., 2010; Pozzi et al., 2000), Genetic Programming (Aytek et al., 2014), Radial Basis Function Neural Network (Nikentari, 2017), Generalized Regression Neural Network (GRNN) (Pashova and Popova, 2011), Wavelet Transform analysis (Wang et al., 2020; Sithara et al., 2020; Zhu et al., 2020c; Zhu et al., 2019; Jie-Xing, 2012; Sehgal et al., 2014; Sahay and Sehgal, 2014), Convolutional Neural Networks (CNN) (Barzegar et al., 2020) and Multivariate Adaptive Regression Splines (MARS) (Zahra and Xiaoli, 2015) are some of the methods been applied. The researchers concluded from their research findings that, the ANN techniques are promising and can be adopted to model and predict the WL to a good precision. The authors were motivated to embark on this study due to the results found in the works of (Zhu et al., 2020a; Piasecki et al., 2018; Young et al., 2015; Kisi et al., 2012).

In Africa, of which the Republic of Ghana is an example, the WL and WL variations have affected related applications in coastal engineering, geophysics studies and space research

(El-Shazly, 2005). The LWL was established at Lake Volta, Ghana by using the available altimetry data between the years (October 1992 to September 2020). The study for the first time in Ghana, applied and assessed the performance of DWT, soft computing and classical techniques such as BRBPANN, LMBPANN, SCGBPANN, RBFANN, GRANN, MLR and ARIMA as an effective tool for modelling and predicting LWL in the study area. In order to resolve the problems associated with nonlinearity problems with the ANN techniques and limitations of the classical techniques, wavelet transform (WT) based on time series for LWL forecasting was applied. WT is nonlinear function that is defined for a range of finite and infinite integrals (Veitch, 2005; Kumi-Boateng and Peprah, 2020).

In this study, DWT which is an example of finite integral was applied. Predicting LWL with noise in a high precision is difficult, hence there is the need to remove noise from the original data before prediction (Chao-Long et al., 2011). WT has the capabilities to denoise noisy dataset. In this present study, WT was applied to remove noise from LWL original data at first, then predict it at high precision using soft computing and classical techniques. This was done because ANN techniques have problem with nonlinearity between the input and output variables (Adnan et al., 2012) and in order to improve its performance, there was the need to applied WT. Upon carefully reviewing of existing literature, these aforementioned techniques have not been applied within Ghana. Therefore, this study will help scientist in Ghana to know the efficacy of using soft computing and classical techniques for future forecasting of the LWL. This is because, the knowledge of LWL is very vital for protection of catchment areas and monitoring changes in lake ecosystem.

2 Study Area

The study area (Fig. 1a to 1d) is one of the highly esteemed projects from the period of Africa's decolonization (Lawson, 1970). Lake Volta is the largest artificial reservoir by surface area in the world, covering an approximated total area of 8500 km^2 which represent 3.2 % of Ghana's total land surface (Ndehedehe et al., 2017; Ni et al., 2017). The lake which lies entirely in Ghana is shared by six West African countries namely, Benin, Burkina Faso, Ivory Coast, Ghana, Mali and Togo. The aim of its construction was to produce hydroelectric power (Owusu et al., 2008; Béné, 2007; Gyau-Boakye, 2001), but the reservoirs fisheries have been a significant socio-economic importance to Ghana (Béné and Obirih-Opareh, 2009; Béné and Russell, 2007; Braimah, 2003; Abban, 1999).

The Volta basin is primarily underlain by a Voltarian formation consisting of Sandstones, Shales and Mudstones. Another formation is Precambrian, classified into Birimian, Buem and Tarkwaian rocks (Dickson and Benneh, 1977). The basin stretches over four climate regions. From lowland rainforest in the South with approximated geographical location of 006° 17′ 20.96" North (N) and 000° 0′ 50.92" West (W) to the Sahel-Sudan desert in the North with geographical location 009° 3′ 8.26" N and 001° 7′ 33.26" W (Rodger et al., 2007).

The average topographic elevation of the study is 400 m and the average water level is 4 m. Much of the Volta basin lies in a rainfall region which starts from July to September. The North has only one wet season in September. In the South, there are two rainy seasons which starts from June to July and September to October (Ni et al., 2017).

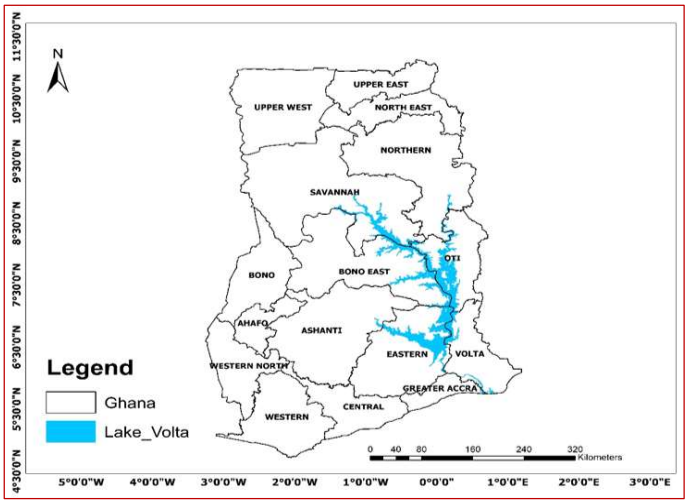

Fig. 1a. Regional map of Ghana showing the study area

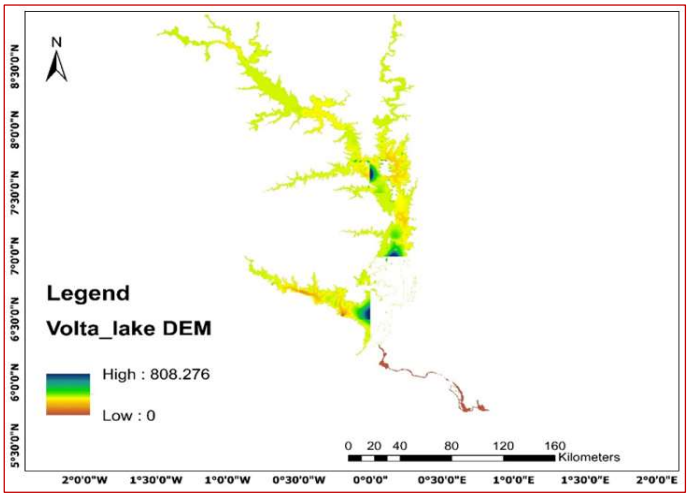

Fig. 1b. Digital Elevation Model of the Study area

Fig. 1c Satellite Image showing the study area

3. Material and Methods Used

3.1. Data used

The LWL heights for Lake Volta Basin were obtained from Satellite altimetry data of observed lake height variations computed from (Topex/Poseidon, Jason-1, Jason-2. Jason-3, TP/J1, TPJOS and Jason-2/OSTM) provided by the United States Department of Agriculture (USDA). The geodetic datum for these heights was the world ellipsoid. The use of altimetry-based measurements for a data deficient region is beneficial since they are continuous and potentially available few days after measurement (Ndehedehe et al., 2017; Coe and Birkitt, 2004). Thus, time series data of Lake Volta level heights covering the period of 28 years from (October 1992 to September 2020) were downloaded from (https://www.pecad.fas.usda.gov/cropexplorer/global_reservoir/gr_regional_chart_jason1.aspx?regionid=wafrica&reservoir_name=Volta) database and used in this study to model and predict Lake Volta WL. As presented in the works of (Ndehedehe et al., 2017; Ni et al., 2017), satellite altimetry data has been successfully validated by comparing altimetric time series and in situ observations. Additionally, all classical corrections (polar and solid earth tides, ionospheric and tropospheric delay, and altimeter biases) have been applied to the altimetry data. This was done by using a median type filter to eliminate outliers and reduce high frequency noise.

3.2. Computations of monthly MLWL

The available time series of Lake Volta WL data for a period of 336 months were used. This period started from October 1992 to September 2020. The monthly MLWL from the given altimetry data were computed according to Eq. 1. However, the monthly meteorological and lake discharge data were not available, hence not included.

$$MLWL_{i,j} = \sum_{j=1}^{i} \frac{x_i}{n} \quad (1)$$

where x_i is the time series data, i is an integer ranging from 1 to n, and n is the number of observations. Table 1 is the statistical analysis of the satellite altimeter data, thus, maximum (max), minimum (min), mean value and standard deviation (SD) of the computed monthly MLWL. Figs. 2a and 2b are the monthly distribution and histogram graph analysis of the computed monthly MLWL. Fig. 3 is the flowchart for the present study methodology.

Table 1. Statistical analysis of computed monthly MLWL (units in meter)

PCI	min	max	mean	SD
MLWL	-0.0400	7.9967	0.2272	0.0114

3.3. DWT

WT analysis was applied since noise is an important factor that influences the precision and accuracy of predictions (Chao-Long et al., 2011). WT is one of the time frequency analysis. The classical signal analysis is built on the basis of Fourier transform (Jie-Xing et al., 2012). WT has been widely used to analyze and denoise signals, images and its applications in hydrological studies and earth sciences is well acknowledged (Sithara et al., 2020; Zhu et al., 2020c). A detailed theoretical literature of WT can be found in the works of (Xue et al., 2017; Shafaei and Kisi, 2015; Jie-Xing et al., 2012; Veitch 2005; De-Bao et al., 2012). WT decompose a given data into different resolution levels and analyses each level with a resolution mapped to its scale (Huang et al., 2016; Zhou and Yin, 2014).

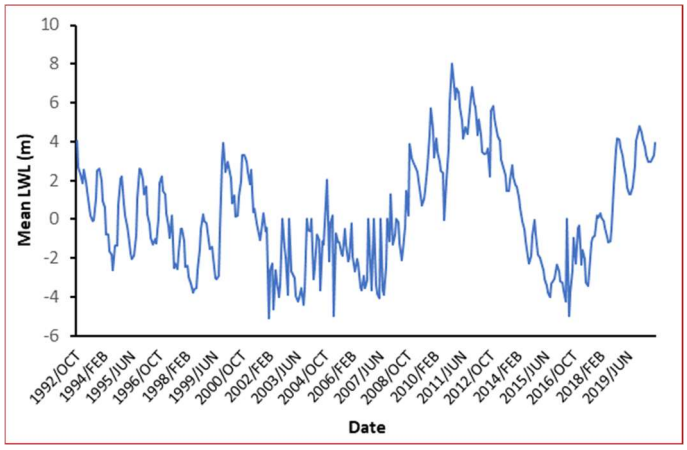

Fig. 2a. Monthly mean distributions of Lake Volta WL

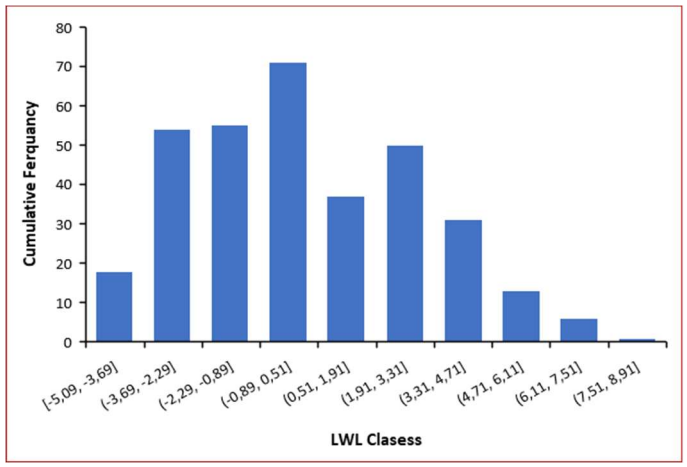

Fig. 2b. Histogram graph analysis of the computed data MLWL

For LWL modelling, the discrete DWT was applied to decompose the computed MLWL heights $X(t)$ decomposition into low frequency component and high frequency component. DWT helps the user to achieved one or more detail series and approximation at different scales (Zhu et al., 2019; Ismail et al., 2018; Shafaei and Kisi, 2015). DWT was adopted in this study to decompose the raw time series data of Lake Volta Basin since existing studies have shown that, DWT has a higher efficiency and is simpler to use which requires less computation time (Zhu et al., 2019; De-Bao et al., 2012). DWT is calculated by successive low pass and high pass filtering of the discrete time domain signal (Ismail et al., 2018). This is known as the Mallat algorithm or Mallat-tree decomposition. The Mallat algorithm was implemented because is a fast wavelet algorithm based on multi-resolution analysis including two parts of decomposition and reconstruction (Shafaei and Kisi, 2015; Zhou and Yin, 2014).

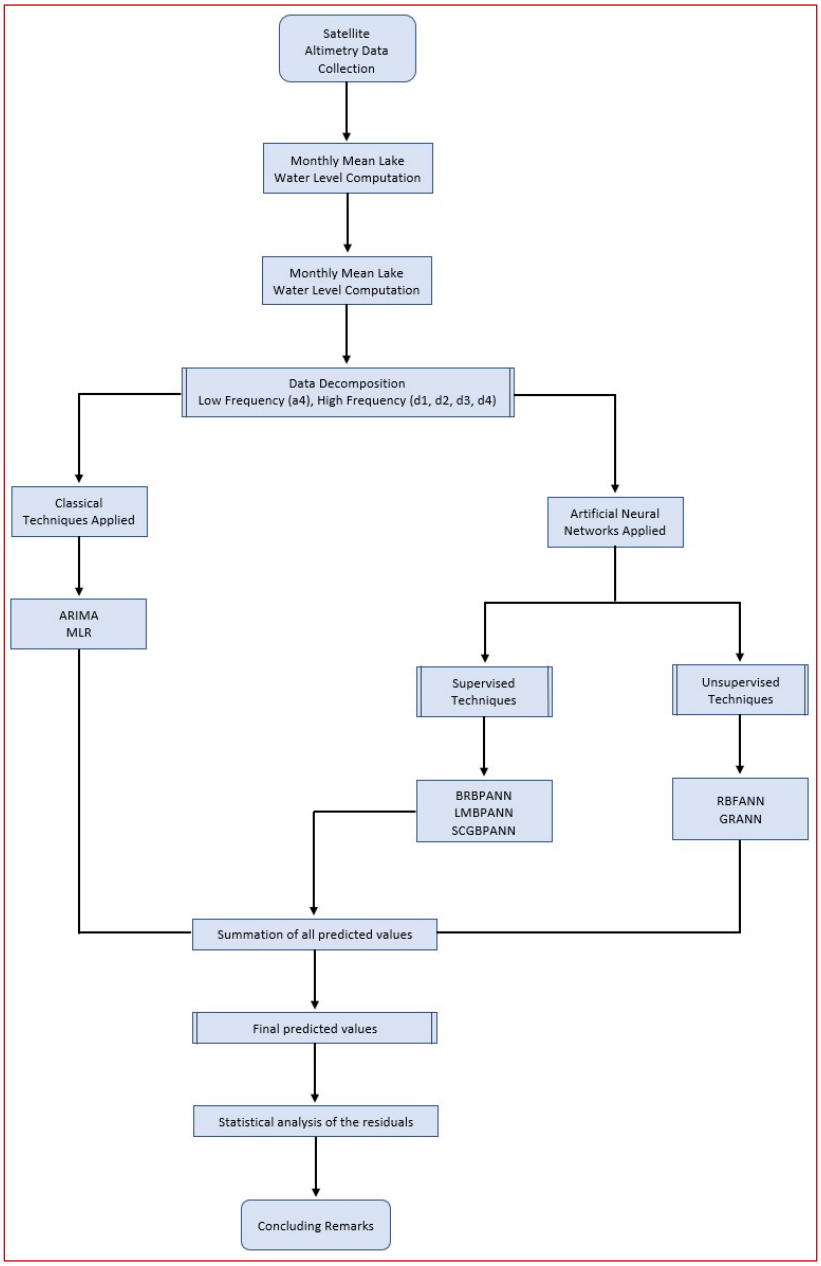

Fig. 3. Flowchart of the project methodology

However, its significance is in the manner it connects the continuous time multiresolution to discrete time filters (Zhu et al., 2019; Ismail et al., 2018). One of the most important factors when implementing the DWT algorithm is the selection of an appropriate wavelet function (Zhu et al., 2019). The *dbN* wavelet function is effective for geodetic purpose because of its localization capability in both time and frequency domains (Huang et al., 2016). Previous studies suggest that, the best results can be achieved from a decomposition of level three, four or five (Zhu et al., 2019; Huang et al., 2016; Zhou and Yin, 2014). In this study, the best decomposition level for MLWL modelling is four. Wavelet function $\psi(t)$ which is called the mother wavelet define as $\int_{-\infty}^{+\infty}\psi(x)dt=0$ can be achieved through compressing and expanding $\psi(t)$ given by Eq. 2:

$$\psi_{a,b}(t)=\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (2)$$

where a is positive which define the scale, b is any real number which define the shift or time factor, $\psi_{a,b}(t)$ is the successive wavelet.

The pair (a, b) defines in the right half plane $R\times R$. If $\psi_{a,b}(t)$ satisfy Eq. 2 for the time series $f(t)\in L^2(R)$, successive wavelet transforms of $f(t)$ is defined by Eq. 3 as:

$$\omega_\psi f(a,b)=|a|^{-1/2}\int_R f(t)\psi-\left(\frac{t-b}{a}\right)dt \quad (3)$$

where; $\omega_\psi f(a,b)$ is complex conjugate functions of $\psi(t)$.

For a discrete time, series $f(t)$ which occurs at a different time t can be defined as Eq. 4 given as:

$$\omega_{\psi} f(j,k) = 2^{-1/2} \sum_{t=0}^{N-1} f(t) - \left(2^{-j} t - k\right) \quad (4)$$

where $\omega_{\psi} f(j,k)$ is the wavelet coefficient for discrete wavelet of scale $a = 2j, b = 2jk$.

3.4. RBFANN

RBFANN model is an unsupervised learning algorithm which is constructed based on functional approximation. It consists of three functionally distinct layers namely; an input layer, a hidden layer and an output layer. The input layer is made up of sensory units that connect the network to its environment. In the second layer, the only hidden layer in the network applied a nonlinear transformation from the input space to the hidden space. The output layer is linear, supplying the response of the network to the activation pattern applied to the output layer. In this study, the input and output variables were frequency component values denoted as $[a4_{i-3}, a4_{i-2}, a4_{i-1}]$, $[d1_{i-2}, d1_{i-1}]$, $[d2_{i-2}, d2_{i-1}]$, $[d3_{i-2}, d3_{i-1}]$, $[d4_{i-3}, d4_{i-2}, d4_{i-1}]$ and $[a4_i]$, $[d1_i]$, $[d2_i]$, $[d3_i]$, $[d4_i]$, respectively. The dataset used for the formulation of the model were divided as training data which consist of 70 % of the total LWL dataset and testing data which consists of 30 %. RBFNN is an exact interpolator (Erdogan, 2009), hence a linear function is used in the input neurons and the connection between the input and hidden layers are not weighted (Yakubu and Dadzie, 2019; Kaloop et al., 2017). In this presented study, the Gaussian function is applied, and the output neuron is a summation of the weighted hidden output layer given by Eq. 5 (Erdogan, 2009):

$$y(x) = \sum_{j=1}^{n} \kappa_j \chi_j(x) \quad (5)$$

where n is the number of hidden neurons $x \in R^M$ is the input, K_j are the output layer weights of the radial basis function network, $\chi_j(x)$ is Gaussian radial basis function given by Eq. 6 (Srichandan, 2012; Idri et al., 2010):

$$y(x) = \sum_{j=1}^{n} \kappa_j \chi_j(x) \quad (6)$$

where $c_j \in R^M$ and σ are the centre and width of j_{th} hidden neurons respectively, $\|\|$ denotes the Euclidean distance.

3.5. BPANN

BPANN is a supervised kind of multilayer feedforward neural network which consist of three layers namely, the input layer, hidden layer and output layer. In this study, the input and output variables were the frequency component values defined in section 3.4. In BPANN, the dataset has to be normalized to eliminate the impact on the model due to different dimensions and units of variables (Yakubu et al., 2018a; Huang et al., 2016). The original data to be used for the BPANN iterating and its model formulation are expressed in different units with different physical meanings. Hence, for constant variation in the BPANN model, datasets are frequently normalized to a certain interval such as [-1, 1], [0, 1] or other scaled criteria (Ziggah et al., 2016). The selected input and output variables were normalized between the intervals [−1, 1] according to Eq. 7 (Mueller and Hemond, 2013):

$$y_i = y_{\min} + \frac{(y_{\max} - y_{\min}) \times (x_i - x_{\min})}{(x_{\max} - x_{\min})} \quad (7)$$

where y_i represents the normalized frequency component values, x_i is the frequency component values while $x_{\min}$ and $x_{\max}$ represent the minimum and maximum values of the frequency component values with $y_{\max}$ and $y_{\min}$ values set at 1 and -1, respectively. To find the optimum weight combination and best learning training algorithm for the study area, the network was trained by using Bayesian Regularization, Levenberg-Marquardt, and Scaled Conjugate Gradient learning algorithms. The datasets were divided into training (70 %) and testing (30 %). The objective of training is to find the set of weights between the neurons that determine the global minimum of error function. The main function of the testing set is to evaluate the generalization ability of a trained network. Training is stopped when the error of the testing dataset starts to increase (Chakraborty and Goswani, 2017). The tansig and purelin activation functions were used for the hidden and output layer respectively in the network training. BPANN is an iterative training procedure, therefore the network was trained varying the number of hidden neurons ranging from 1 to 50 until the optimal model was achieved.

3.5.1. SCGBPANN

SCGBPANN training algorithm denoted as (trainscg) belongs to a class of conjugate gradient methods (Arthur *et al.*, 2020) developed by Moller (1993). It is a basic BPANN algorithm which adjusts the weights in the steepest descent direction, thus, the most negative of the gradient (Baghirli, 2015). This happens to be the direction in which the performance function is decreasing rapidly. It was revealed in the work of Hagan et al. (1996) that, the function decreases most rapidly along the negative of the gradient which does not necessarily produce the fastest convergence. In SCGBPANN model building, a search is performed along such a direction, which yields generally faster convergence than the steepest descent direction while preserving the error minimization obtained in all previous steps (Baghirli, 2015; Kisi and Uncuoglu, 2005). This kind of direction is called the conjugate direction (Baghirli, 2015). Additionally, the step-size is adjusted at each iteration and a search is made along the conjugate gradient direction to determine the step size (Arthur et al., 2020; Baghirli, 2015). This helps minimizing the performance function along that line. SCGBPANN commence by searching in the steepest descent direction at first iteration given by Eq. 8 (Baghirli, 2015). Moreover, SCGBPANN models are used with line search. This implies that, the step size is approximated with a line search technique which ignore the computation of the Hessian matrix to determine the optimal distance to move along the

current search direction given by Eq. 9. The next search direction is obtained to conjugate to previous search direction as according to Eq. 10. Hence, the general procedure for obtaining the new search direction is achieved by combining the new steepest descent direction with the previous direction (Arthur et al., 2020; Baghirli, 2015; Sandhu and Chhabra, 2011; Hagan et al., 1996).

$$p_o = -g_o \tag{8}$$

$$x_{k+1} = x_k + a_k g_k \tag{9}$$

$$\rho_k = -g_k + \beta_k \rho_{k-1} \tag{10}$$

The various types of conjugate algorithms are distinguished by the manner in which the factor β_k is calculated (Baghirli, 2015; Kisi and Uncuoglu, 2005). Conversely, there is a possibility to use another approach in estimating the step size than the line search technique. The notion behind is to combine the model trust region approach (Bahirli, 2015), known from the LMBPANN algorithm with the SCGBPANN approach. This approach is known as SCG and detailed in the work of Moller (1993). In this model technique as represented by Eq. (11), *is* the Hessian matrix approximation, E is the total error function and $E^{'}$ is the gradient of E, scaling factors λ_k and δ_k are introduced to approximate the Hessian matrix and initialized by the user at the beginning of the algorithm iterating such that $0 < \lambda_k < 10^{-b}$ and $0 < \delta_k < 10^{-4}$ (Baghirli, 2015). In SCGBPANN model building, β_k factor calculation and direction are done according to Eqs. 12 and 13 (Moller 1993) given as:

$$S_k = \frac{E\left(w_k + \delta_k \rho_k\right) - E^{'}\left(w_k\right)^{'}}{\delta_k} + \lambda_k \rho_k \tag{11}$$

$$\beta_k = \frac{\left(\left|g_{k+1}\right|^2 - g_{k+1} g_k^T\right)}{g_k^T g_k} \tag{12}$$

$$\rho_{k+1} = -g_{k+1} + \beta_k \rho_k \tag{13}$$

In addition, design parameters are updated at each iteration stage independently, which is an important for the success of the algorithm. This is a major advantage compared to the line search-based algorithms (Baghiri, 2015). Moreover, the trainscg algorithm denotes a quadratic approximation to the E in a neighborhood of a point w by $E_{qw}(y)$ given by Eq. 14 (Arthur et al., 2020):

$$E_{qw}(y) = E(w) + E^{'}(w)_y^r + \frac{1}{2} y^r E^{''}(w)_y \tag{14}$$

Hence, to determine the minimum of $E_{qw}(y)$, the critical points for $E_{qw}(y)$must be found (Arthur et al., 2020). The critical points are the solution to the linear system defined by Moller (1993). The SCG algorithm can train any networks as long as its weight, net input, and transfer functions having derivatives functions (Arthuer et al., 2020; Sanhu and Chhabra, 2011).

3.5.2. LMBPANN

LMBPANN training algorithm denoted as (trainlm) is an iterative model for finding the minimum of a multivariate error function (Z) according to Eq. 15 and is expressed as the sum of squares of the difference between the actual output (y_i) and target output (t_i) (Arthur et al., 2020; Adeoti and Osanaiye, 2013).

$$Z_{i,j} = \frac{1}{2}\Sigma\left(y_i - t_i\right) \tag{15}$$

The trainlm algorithm was designed to approach second-order training speed without having to compute the Hessian Matrix (Baghirli, 2015). The Hessian matrix (H) and the gradient (g) can be computed using Eqs. 16 and 17 respectively, when the performance function has a form of sum of squares (Kisi and Uncuoglu, 2005; Baghirli, 2015; Arthur et al., 2020).

$$H = J^T J \tag{16}$$

$$g = J^T e \tag{17}$$

where; J is the Jacobian matrix containing the first derivatives of the network errors with respect to the biases and weights, and e is the network error vector. The Jacobian matrix can be computed through a standard backpropagation technique that is much less complex than computing the Hessian matrix (Arthur et al., 2020; Baghirli, 2015). The trainlm algorithm uses this approximation to the Hessian matrix in following Newton-like update according to Eq. 18:

$$w_{i+1} = w_i - \left|J^T J + \delta I\right|^{-1} J^T e \tag{18}$$

where; w is the connection weights, δ is the damping term and I is the identity matrix. The trainlm uses the combination of Gauss-Newton method and gradient descent in its iterative process (Arthur et al., 2020). When the δ is zero, it becomes a Gauss-Newton method, using the approximate Hessian matrix. When δ is large, it becomes a gradient descent method having a small step size. Newton's method is faster and more accurate near an error minimum, so the aim is to shift towards Newton's method as quickly as possible (Baghirli, 2015). Thus, δ is decreased after each successful step (reduction in performance function) and is increased only when a tentative step would increase the performance function. In this way, the performance function will always be reduced at each iteration of the algorithm (Arthur et al., 2020; Baghirli, 2015). The trainlm optimization technique is more powerful than the conventional gradient descent techniques (Wilamowski, 2009).

3.5.3. BRBPANN

BRBPANN denoted as (trainbr) sequentially updates the weights and bias values according to Levenberg-Marquardt optimization (Foresee and Hagan, 1997; Kaur and Salaria, 2013). It minimizes a combination of squared errors and

weights, and then determines the correct combination to produce a network that generalizes well (Kaur and Salaria, 2013; Pan et al., 2013). The method of improving generalization is referred to as regularization (Arthur et al., 2020; Foresee and Hagan, 1997). The aim of training is to reduce the sum of squared error (ε_D). This implies that, the training objective function is $F_\omega = \varepsilon_D$. However, regularization adds an additional term, ε_ω. The objective function is then expressed as shown in Eq. 19 (Foresee and Hagan, 1997; Yue et al., 2011):

$$F_\omega = \tau\varepsilon_D + \vartheta\varepsilon_\omega \tag{19}$$

where; ε_ω is the sum of the squares of the network weights, ε_D is the sum of network errors, τ and ϑ are the objective function parameters. In the trainlm framework, the weights of the network are viewed as random variables, and the distribution of the network weights and training set are considered as gaussian distribution (Baghirli, 2015). According to Foresee and Hagan (1997), the relative size of the objective function parameters dictates the emphasis for training. If $\tau \ll \vartheta$, then the training algorithm will drive the errors smaller and if $\tau \gg \vartheta$, training will emphasize weight size reduction at the expense of network errors. Thus, producing a smoother network problem (Arthur et al., 2020). Moreover, the main problem with implementing regularization is setting the correct values for the objective function parameters. The τ and ϑ factors are defined using the Bayes' theorem. A detailed procedure for calculating the correct values of τ and ϑ is reported in Foresee and Hagan (1997). The Bayes' then relates two variables (or events), α and β, based on their prior (or marginal) probabilities and posterior (or conditional) probabilities according to Eq. 20 (Li and Shi, 2012; Baghirli, 2015):

$$P(\alpha|\beta) = \frac{p(\beta|\alpha)P(\alpha)}{P(\beta)} \tag{20}$$

where $P(\alpha|\beta)$ is the posterior probability of α conditional on β, $P(\alpha|\beta)$ is the prior of β conditional on α, and $P(\beta)$ is the non-zero prior probability of event β, which functions as a normalizing constant. In order to find the optimal weight space, the objective function (Eq. 19) needs to be minimized. This is equivalent to maximizing the posterior function given by Eq. 21:

$$P(\alpha,\beta|D,M) = \frac{P(D|\alpha,B,M)p(\alpha,\beta|M)}{P(D|M)} \tag{21}$$

where α and β are the factors needed to be optimized, D is the weight distribution, M is the particular neural network architecture, $P(D|M)$ is the normalization factor, $P(\alpha,\beta|M)$ is the uniform prior density for the regularization parameters and $P(D|\alpha,\beta,M)$ is the likelihood function of D given α,β,M. Maximizing the posterior function $P(\alpha,\beta|D,M)$ is equivalent to maximizing the likelihood function $P(D|\alpha,\beta,M)$ In this process, optimum values for α and β for a given space are found. Hence, trainbr model moves into trainlm phase, where Hessian matrix computations take place and updates the weight space in order to minimize the objective function (Baghirli, 2015). If the convergence is not met, algorithm estimates new values for α and β and the entire procedure repeats itself until convergence is reached (Yue et al., 2011).

3.6. GRANN

GRANN was first introduced by Specht (1991) and is a different kind of RBFANN which is based on Kernel regression networks (Yakubu and Dadzie, 2019; Hannan et al., 2010) with one pass learning algorithm and highly parallel structure (Dudek, 2011). GRANN consist of four layers namely; input layer, pattern layer (radial basis layer), summation layer, and output layer. In this study, the input and output variables were the frequency component values defined in section 3.4. The number of input units in the first layer depends on the total number of the observational parameters. The first layer is connected to the pattern layer and in this layer, each neuron is being presented by a training pattern and its output. The pattern layer is connected to the summation layer. The summation layer consists of two different types of summation namely, single division unit and summation unit (Hannan et al., 2010).

The summation with output layer combined perform a normalization of output datasets. In training of the network, radial basis and linear activation functions are used in hidden and output layers. Each pattern layer unit is connected to two neurons in the summation layer. One neuron unit computes the sum of the weighted response of the pattern, and the other neuron unit computes unweighted outputs of pattern neurons. The output layer divides the output of each neuron unit by each other yielding the estimated output variables as according to Eq. 22 as:

$$y_i = \frac{\sum_{i=1}^{n} y_i . \exp - G(x, x_1)}{\sum_{i=1}^{n} \exp - G(x, x_1)} \tag{22}$$

where; y_i is the weighted connection between the i_{th} neuron in the pattern layer and the summation neuron, n is the number of training patterns, G is the Gaussian function given by Eq. 23.

$$G(x, x_1) = \sum_{k=1}^{m} \frac{(x_1 - x_{1k})^2}{\sigma} \tag{23}$$

where, m is the number of elements of an input vector, x_1 and x_{1k} are the kth element of x and x_i respectively. During the network training, the spread parameter was varied between 0 and 1 until the output with minimal residuals in terms of statistical analysis were achieved. This same procedure was also done when training the RBFNN.

3.7. ARIMA

The ARIMA model introduced by Box and Jenkins (1976) is a widely used technique for time series forecasting (Boye and Ziggah, 2020; Akyen et al., 2016). This type of model is a hybridized model which consists of autoregressive (AR) and moving average (MA) respectively (Boye and Ziggah, 2020;

Yakubu et al., 2018b). In ARIMA (p,d,q) modelling, the first step is to check the stationarity of the time series data. When the used time series data is not stationary, it is transformed into a stationary time series by applying the appropriate order of differencing (d) (Yakubu et al., 2018; Akyen et al., 2016). The desired values of autoregressive order (p) and moving average (q) is acquired by checking the autocorrelation function and partial autocorrelation function of the time series data (Makwinja et al., 2017; Yusof et al., 2013). The $AR(p)$ model is a discrete time linear equation with noise as expressed by Eq. 24 (Yakubu et al., 2018b):

$$\chi_t = \alpha_1 \chi_{t-1} + \cdots + \alpha_p \chi_{t-p} + \xi_t \quad (24)$$

where χ_t is the current forecasted model, p is the order, $\alpha_1, \cdots, \alpha_p$ are the parameters of coefficients of the model formed, χ_{t-1}, χ_{t-p} are the previous observations, and ξ_t is the error of the forecast. The $MA(q)$ model is an explicit formula for χ_t in terms of noise as given by Eq. 25:

$$\chi_t = \xi_t - \beta_1 \xi_{t-1} - \cdots - \beta_p \xi_{t-p} \quad (25)$$

The difference operator Δ is given by Eq.26:

$$\Delta\chi_t = \chi_t - \chi_{t-1} = (1-L)\chi_t \quad (26)$$

The ARIMA model with orders (p,d,q) is given by Eq. 27 as:

$$\left(1 - \sum_{j=1}^{p} \alpha_j L^j\right)(1-L)^d \chi_t = \left(1 + \sum_{j=1}^{q} \beta_j L^j\right)\xi_t \quad (27)$$

where; L^j is the time lag operator, ξ_t is an error term, and d_i is the order of integration. In this present study, the first order $AR(1)$, and $MA(1)$ was adopted due to its simplest non-degenerated time-series process (Boye and Ziggah, 2020) as recommended by Makwinja et al. (2017) for LWL forecasting. The basic systematic approach utilizing Box-Jenkins methodology was applied in this present study to build the first order ARIMA model giving by Eq. 24. This includes stationarity checks, model identification, parameter estimation, model selection, and diagnostic checking. A detailed literature review about these aforementioned techniques can be found in (Makwinja et al., 2017; Akyen et al., 2016). ARIMA model was implemented and coded in MATLAB environment.

3.8. MLR

It has been observed that LWL varies from time to time and is highly correlated with climatic conditions. Therefore, the MLR model can generate internal dynamics between inputs and outputs. MLR is a nonlinear regression model in which observational data are modelled by a function $Z(f(x))$ and depends on one or more independent variables (Tiryaki, 2008). The MLR was adopted in this study to estimate the decompose LWL heights data defined in section 3.4. MLR fits a linear combination of the components of multiple input parameters to a single output parameter (Ziggah et al., 2016; Sheta et al., 2015). A ρ^{th} order MLR, thus, $MLR(\rho)$ refers to the correlation size between values in a time series that are ρ periods apart given by Eq. 28 (Sehgal *et al.*, 2014):

$$Z(x) = \sum_{i=1}^{\rho} \alpha_i x_{i-1} + \varepsilon_t \quad (28)$$

where, α_i are the MLR coefficients, $Z(x)$ is the time series under investigation, ρ is the order (length of the MLR model), and ε_t is the residual term which is assumed to the Gaussian white noise (Sehgal et al., 2014). ρ is generally much less than the data length of the series. Hence, in MLR model building, the LWL can be estimated by a linear weighted sum of previous decomposed LWL. He weights are the MLR coefficients which are normally estimated using the Least Squares method (Sehgal et al., 2014). The Least Squares approach has been successfully and frequently applied in geodetic sciences. Therefore, the mathematical backgrounds and theories of the method will not be repeated here. A more comprehensive detail on them can be found in the works of (Yakubu et al., 2018a; Peprah and Kumi, 2017; Peprah and Mensah, 2017; Ghilani 2010).

3.9. Model Performance Assessment

In order to determine the accuracies of the models being used, statistical error analysis was carried out. The statistical measures applied were the *AME*, *AMSE*, *AMAPD*, r_{min}, r_{max} and *ASD*. Their mathematical expressions are given by Eqs. 29 to 34 respectively.

$$AMAPD = \frac{1}{n}\sum_{i=1}^{n}\left\|\frac{a_i - \beta_i}{a_i}\right\| \quad (29)$$

$$AME = \frac{1}{n}\sum_{i=1}^{n}(\alpha_i - \beta_i) \quad (30)$$

$$AMSE = \frac{1}{n}\sum_{i=1}^{m}(\alpha_i - \beta_i)^2 \quad (31)$$

$$r_{\max} = \alpha_i - \beta_i \quad (32)$$

$$r_{\min} = \alpha_i - \beta_i \quad (33)$$

$$ASD = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\mu - \bar{\mu})^2} \quad (34)$$

where; n is the total number of the observations, α_i and β_i are the measured and predicted decomposed LWL heights from the various techniques, μ denote the residual between the measured and estimated decomposed LWL height, $\bar{\mu}$ is the mean of the residual and i is an integer varying from 1 to n.

4. Results and Discussions

4.1. Developing of ANN models

A single layer BPANN model was trained using Bayesian Regularization, Levenberg-Marquardt and Scaled Conjugate Gradient learning algorithm respectively. Tansig and Purelin functions were both used for the hidden and output layer when training BPANN model with trainbr, trainlm, trainscg respectively. The optimal model structure, which is highly dependent on the number of hidden neurons was achieved

through a sequential trial and error approach based on the lowest *AME*, *AMSE*, *AMAPD*, r_{mim}, r_{max} and *ASD*.
In this present study, the model was trained varying the number of hidden neurons from 1 to 50. The network was allowed to train for 5000 epochs with a learning rate of 0.03, minimum performance gradient of 0.0000001, a goal of 0, maximum validation failures of 6, and momentum coefficient of 0.9 for each iterative training process. In training the neural network, validation stops when the minimum gradient and maximum epoch is reached.

In the case of GRANN and RBFANN model training, the model's output is highly based on the value of the width parameter (spread constant). Therefore, the optimal width parameter value for GRANN and RBFANN was also achieved based on a sequential trial and error approach for each iterative training process. Moreover, a gradient descent rule was implemented to train the GRANN and RBFANN model respectively. The ANN models (BRBPANN, LMBPANN, SCGBPANN, GRANN and RBFANN) were coded and implemented in MATLAB (R2018a) software.

After several trial-and-error methods, the optimal model achieved by the BRBPANN model after successive iterative training in estimating the variables defined in section 3.4 were [3 35 1], [2 33 1], [2 17 1], [2 31 1], and [2 39 1] respectively. That is, (3 and 2) input variables of the decomposed LWL by the WT (independent dataset), (35, 33, 17, 31, and 39) hidden neurons respectively in estimating the low and high frequencies, and 1 output variable (dependent dataset). The optimal LMBPANN model after successive iterative training in predicting the frequency variables were [3 7 1], [2 21 1], [2 6 1], [2 7 1], and [2 8 1] respectively. Also, the optimal SCGBPANN model in estimating the frequency variables were [3 9 1], [2 18 1], [2 13 1], [2 16 1], and [2 11 1] respectively. These optimal BRBPANN, LMBPANN, and SCGBPANN models' structures gave the lowest minimum value in terms of their statistical analysis (*AME*, *AMSE*, *AMAPD*, r_{mim}, r_{max} and *ASD*).

Moreover, the optimal RBFANN predictive model in estimating the frequency variables with the least statistical assessment values were [3 20 1 0.1], [2 5 1 0.8], [2 30 1 0.1], [2 40 1 0.2], and [2 25 1 0.1] respectively. That is, (3 and 2) input variables (independent dataset), (20, 5, 30, 40, and 25) hidden neurons, 1 output (dependent dataset), and a width parameter (between 0 to 1). The optimal constant was achieved by varying the spread parameter (from 0 to 1) in each iterating training until the best results was achieved. The optimal GRANN model for predicting the frequency variables were [3 50 1 0.1], [2 50 1 0.1], [2 50 1 0.1], [2 50 1 0.1], and [2 50 1 0.1] respectively. That is, (3 and 2) inputs (independent datasets), a hidden layer with a maximum of 50 hidden neurons, 1 output (dependent dataset), and a width parameter of 0.1. The summarized results of the training and testing of the predicted frequency components $[a4_i]$, $[d1_i]$, $[d2_i]$, $[d3_i]$, $[d4_i]$ by the various soft computing techniques are tabulated in Tables 2a to 2e.

Based on the statistical results tabulated in Tables 2a to 2e, it is observed that soft computing techniques provide satisfactory results in successfully predicting the frequency variables with much better accuracy for the study area. The minimum and maximum residuals are quiet encouraging. The *AME*, *AMSE*, and *ASD* of both training and testing are quite good and equally encouraging. However, ANN proved to be a powerful realistic alternative tool for LWL modelling for the study area with much better accuracy.

4.2. Developing the Classical Regression Models

In formulating the MLR model for predicting the frequency variables discussed in section 3.4, a statistical description of the data was performed by using Sigma Plot Version 12.5 to find the correlation between the independent variables (input datasets) and the dependent dataset (output datasets). The optimal MLR equation generated by the Sigma Plot software for estimating the low and high frequency variables is given by Eqs. 35 to 39:

$$a_{4i} = 0.00371 - (1.928 \times a_{41}) + (2.915 \times a_{42}) \quad (35)$$

$$d_{1i} = -0.000102 - (0.579 \times d_{11}) - (0.847 \times d_{12}) \quad (36)$$

$$d_{2i} = -0.000177 - (0.784 \times d_{21}) + (0.016 \times d_{22}) \quad (37)$$

$$d_{3i} = -0.000180 - (0.868 \times d_{31}) + (1.566 \times d_{32}) \quad (38)$$

$$d_{4i} = 0.000138 - (0.920 \times d_{41}) + (1.805 \times d_{42}) \quad (39)$$

where; a_{4i}, d_{1i}, d_{2i}, and d_{4i} are the dependent variables (frequency variables), (0.00371, 1.928, 2.915, -0.000102, 0.579, 0.847, and so on) are the generated unknown parameters generated by the least square's method. The final estimated a_{4i}, d_{1i}, d_{2i}, and d_{4i} values with the given equation and parameters were coded and implemented in MATLAB environment. Moreover, in developing an optimal ARIMA model using the MATLAB software, non-stationarity which existed in the decomposed data which will results in wrong statistical inferences was resolved by differencing the data to ensure that the data is stationary. In this study, ARIMA (0, 1, 1) model which is also known as first degree integrated moving average was the optimal model been developed from the decomposed data to predict the low and high frequency variables for the study area. According to Makwinja et al. (2017), the best model for forecasting Lake WL should have adequate accurate statistical performance as low as possible for it to have accurate forecasts.

Therefore ARIMA (0, 1, 1) model was selected as the optimal model for WL forecasting for the study area based on the lowest statistical performance achieved in this study and recommendations by researchers. It was further observed that the coefficients of the parameters of ARIMA (0, 1, 1) model was significant ($p<0.05$). According to the works of (Boye and Ziggah, 2020; Makwinja et al., 2017; Akyen et al., 2016), the ARIMA model which indicate lowest normalized Bayesian Information Criterion (NBIC) is significant ($p<0.05$), hence, a better model in terms of forecasting performance than with large NBIC. Based on the results achieved in the present study findings, the most suitable model for forecasting Lake Volta WL for the study area was confirmed to be ARIMA (0,

1, 1). Table 3 shows the model performance results for the classical techniques.

Table 2a. Models result for $a4_i$ (units in meters)

Training						
PCI	**AME**	**AMSE**	**AMAPD**	$r_{\min}$	$r_{\max}$	**ASD**
BRBPANN	-0.0014	5.3919E-05	0.0121	4.8856E-05	0.0963	0.0043
LMBPANN	0.0014	6.2808E-05	0.0093	-2.2595E-05	-0.0893	6.9332E-05
SCGBPANN	0.0013	0.0001	0.0205	-6.1036E-05	-0.2828	0.0001
RBFANN	-6.3960E-05	5.9891E-05	0.0121	-0.0001	-0.0995	5.9873E-05
GRANN	-0.0009	2.6722E-05	-0.0036	-1.1000E-05	-0.0961	2.2860E-05
Testing						
PCI	**AME**	**AMSE**	**AMAPD**	$r_{\min}$	$r_{\max}$	**ASD**
BRBPANN	-0.0027	6.5311E-05	0.0050	8.9126E-05	-0.1611	9.2353E-05
LMBPANN	0.0040	0.0002	0.0011	-0.0003	0.0543	0.0002
SCGBPANN	0.0103	0.0004	0.0171	-0.0014	0.0794	0.0003
RBFANN	-0.0018	4.7458E-05	0.0031	0.0002	0.0431	6.5621E-05
GRANN	-0.1645	0.0012	0.0494	-0.0014	-0.8542	0.0004

Table 2b. Models result for $d1_i$ (units in meters)

Training						
PCI	**AME**	**AMSE**	**AMAPD**	$r_{\min}$	$r_{\max}$	**ASD**
BRBPANN	0.0012	0.0006	1.8957	0.0022	-2.1847	0.0006
LMBPANN	0.0041	0.1656	1.3527	-0.0001	-2.1460	0.0006
SCGBPANN	-0.0110	7.4018E-05	2.2698	0.0017	-2.1681	0.0001
RBFANN	-1.1375E-12	0.0007	1.3121	0.0011	-2.2336	0.0007
GRANN	-0.0071	0.0006	0.2939	-4.4400E-16	-1.9621	0.0006
Testing						
PCI	**AME**	**AMSE**	**AMAPD**	$r_{\min}$	$r_{\max}$	**ASD**
BRBPANN	-0.0086	0.0019	0.2247	0.0010	-1.8006	0.0018
LMBPANN	-0.0092	0.0021	0.0346	-0.0009	-1.8528	0.0020
SCGBPANN	-0.0724	0.0010	2.5509	-0.0018	-1.7377	0.0003
RBFANN	0.0148	0.0024	0.8620	0.0007	-1.7472	0.0026
GRANN	-0.0268	0.0025	0.3741	-0.0074	-1.9304	0.0023

Table 2c. Models result for $d2_i$ (units in meters)

Training						
PCI	**AME**	**AMSE**	**AMAPD**	$r_{\min}$	$r_{\max}$	**ASD**
BRBPANN	-0.0166	0.0001	0.2402	0.0004	-0.9282	7.3363E-05
LMBPANN	0.0169	0.0002	0.2613	0.0002	0.7155	0.0002
SCGBPANN	-0.0022	0.0009	-0.1842	0.0024	-0.7994	0.0009
RBFANN	1.6723E-09	8.2598E-05	0.2027	0.0007	-0.8467	8.2958E-05
GRANN	0.0022	0.0002	0.0690	-3.1500E-09	-0.6034	0.0002
Testing						
PCI	**AME**	**AMSE**	**AMAPD**	$r_{\min}$	$r_{\max}$	**ASD**
BRBPANN	0.0141	0.0009	0.0119	0.0076	0.4237	0.0008
LMBPANN	0.0466	0.0010	0.0240	-0.0032	0.4124	0.0006
SCGBPANN	0.0209	0.0015	0.0705	6.7430E-05	0.4355	0.0014
RBFANN	0.0339	0.0015	0.0239	0.0024	-0.4632	0.0012
GRANN	0.0270	0.0004	0.0345	-0.0071	-1.9304	0.0007

Table 2d. Models result for $d3_i$ (units in meters)

Training						
PCI	**AME**	**AMSE**	**AMAPD**	$r_{\min}$	$r_{\max}$	**ASD**
BRBPANN	-0.0033	0.0002	0.1351	0.0013	0.5274	0.0002
LMBPANN	-0.0021	0.0003	0.1195	-0.0021	-0.5024	0.0003
SCGBPANN	0.0038	0.0002	0.0674	-0.0015	-0.9021	0.0002
RBFANN	0.0006	8.2401E-05	0.1023	-0.0016	0.8286	5.5251E-05
GRANN	0.0027	0.0003	0.0830	1.4000E-14	-0.4600	0.0002
Testing						
PCI	**AME**	**AMSE**	**AMAPD**	$r_{\min}$	$r_{\max}$	**ASD**
BRBPANN	-0.0138	0.0009	0.0287	-0.0036	-0.4268	0.0011
LMBPANN	-0.0074	0.0009	0.0969	0.0044	-0.4627	0.0009
SCGBPANN	0.0203	0.0007	0.2362	-0.0003	-0.4657	0.0005

RBFANN	0.0023	0.0015	0.0494	0.0048	-0.4589	0.0015
GRANN	0.0028	0.0013	0.5394	-0.0060	-0.6093	0.0013

Table 2e. Models results for $d4_i$ (units in meters)

Training						
PCI	**AME**	**AMSE**	**AMAPD**	r_{min}	r_{max}	**ASD**
BRBPANN	-0.0034	0.0002	0.1951	0.0001	-0.3434	0.0002
LMBPANN	-0.0013	0.0001	0.1585	-0.0004	-0.2160	0.0001
SCGBPANN	-0.0018	0.0002	0.2449	0.0001	0.1839	0.0002
RBFANN	-4.0936E-10	0.0002	0.2047	-0.0005	0.1722	0.0002
GRANN	-0.0014	0.0001	0.4272	1.1700E-05	-0.3009	0.0001
Testing						
PCI	**AME**	**AMSE**	**AMAPD**	r_{min}	r_{max}	**ASD**
BRBPANN	-0.0012	0.0001	0.0234	0.0005	-0.0908	0.0002
LMBPANN	-0.0014	7.0606E-05	0.0423	0.0017	-0.0607	8.5040E-05
SCGBPANN	-0.0059	6.2526E-05	0.0830	0.0004	-0.0964	0.0001
RBFANN	-0.0019	4.2465E-05	0.0594	-0.0001	-0.0808	2.7507E-05
GRANN	-0.0079	0.0007	0.2629	0.0023	0.2178	0.0007

4.3. Comparing the predictive performance results of the ANN models with the classical models

The best optimal ensemble soft computing techniques (DWT-BRBPANN, DWT-LMBPANN, DWT-SCGBPANN, DWT-RBFANN, and DWT-GRANN) discussed in Section 4.1 have been compared to the conventional techniques (DWT-ARIMA and DWT-MLR) using the summation of all the predicted values. The statistical analysis is represented by Table 4. From Table 4, it is seen that the proposed hybrid soft computing techniques produce similar satisfactory results as compared to the classical techniques. The reason is related to the reported statistical assessment. Statistical analysis of Table 4 indicates that the proposed ANN models' predictions were closely related to the computed MLWL height with a higher prediction accuracy. The same was observed for the classical model with a high prediction accuracy in estimating the MLWL heights at a good precision. ARIMA and MLR models were observed to predict the LWL heights at a good accuracy. The maximum and minimum residual values of the conventional techniques and the soft computing techniques were -1.9967 m and -0.0003 m respectively. When comparing their statistical analysis, the MLR and ARIMA model had r_{min} values of -0.0016 m, -0.0007 m and *ASD* of 0.0017 m, 0.0049 m respectively.

Table **3** Model Performance assessment of the classical techniques (units in meters)

a_4						
PCI	AME	AMSE	AMAPD	r_{min}	r_{max}	ASD
MLR	3.0920E-05	0.0001	0.0217	0.0005	-0.3548	0.0001
ARIMA	-0.0119	0.0056	0.0191	0.0004	-1.8513	0.0055
d_1						
PCI	AME	AMSE	AMAPD	r_{min}	r_{max}	ASD
MLR	-4.6671E-07	0.0004	0.8248	5.0770E-05	-2.2625	0.0004
ARIMA	-0.0885	0.0008	1.6332	-0.0019	2.4628	0.0005
d_2						
PCI	AME	AMSE	AMAPD	r_{min}	r_{max}	ASD
MLR	-1.5776E-06	0.0012	0.1091	-0.0025	0.9872	0.0012
ARIMA	0.0007	0.0028	0.4610	0.0024	1.6662	0.0028
d_3						
PCI	AME	AMSE	AMAPD	r_{min}	r_{max}	ASD
MLR	3.7776E-07	0.0002	0.2056	-0.0009	0.9216	0.0002
ARIMA	-0.0015	0.0015	0.1641	0.0001	0.7095	0.0015
d_4						
PCI	AME	AMSE	AMAPD	r_{min}	r_{max}	ASD
MLR	-6.6421E-07	0.0001	0.2137	4.8525E-05	-0.2178	0.0001
ARIMA	0.0039	0.0016	0.1386	-7.9500E-05	-0.6966	0.0016

Table 4 Statistical analysis of all the models (units in meters)

PCI	r_{min}	r_{max}	**AME**	**AMSE**	**AMAPD**	**ASD**
DWT-BRBPANN	-0.0327	1.9793	0.2768	0.0036	0.2008	0.0026
DWT-LMBPANN	-0.0021	-1.9835	0.2776	0.0039	0.2405	0.0029
DWT-SCGBPANN	-0.0003	-1.9951	0.2500	0.0029	0.1450	0.0020
DWT-RBFANN	-0.0094	-1.9967	0.2820	0.0034	0.1585	0.0024
DWT-GRANN	0.0007	-1.9936	0.1905	0.0033	0.3151	0.0026
DWT-MLR	-0.0016	1.9937	0.1988	0.0024	0.3824	0.0017

DWT-ARIMA	-0.0007	-1.9661	0.2152	0.0041	0.0313	0.0049

According to the works of (Boye and Ziggah, 2020; Akyen et al., 2016; Makwinja et al., 2017; Poku-Gyamfi, 2009; Chen and Hill, 2005; Peprah et al., 2017), it was observed that, the more complicated stochastic models, there is a likelihood the achieved results may deviate from their true mean. Hence, there was the need to keep the order as low as possible. The performance of the hybridized WT with the classical techniques outperforms the ANN techniques in estimating LWL heights.

After comparing the soft computing techniques to the conventional techniques in terms of their statistical analysis, the classical techniques were much better as compared to the soft computing methods in estimating LWL heights for study area. However, DWT-MLR model compared to DWT-BRBPANN, DWT-LMBPANN, DWT-SCGBPANN, DWT-RBFANN, DWT-GRANN and DWT-ARIMA showed better performance in predicting the LWL heights for the study area. Figures 5a to 5c show the *AME*, *AMSE* and *AMAPD* model graphs of all the applied techniques.

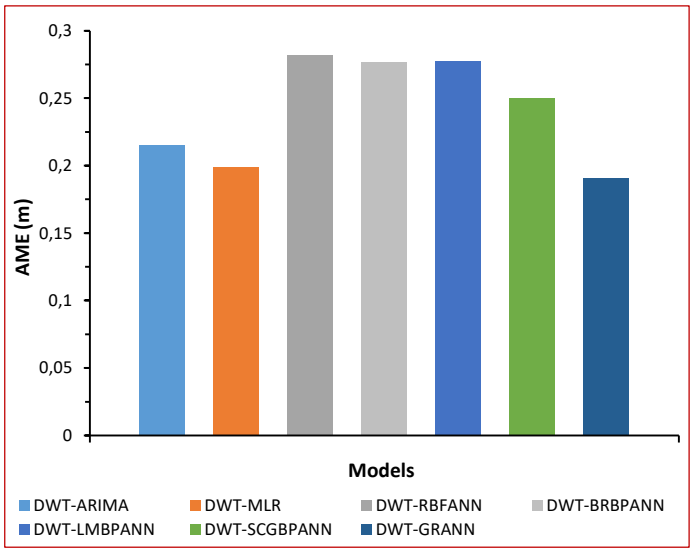

Fig. 5a. AME graph of the models

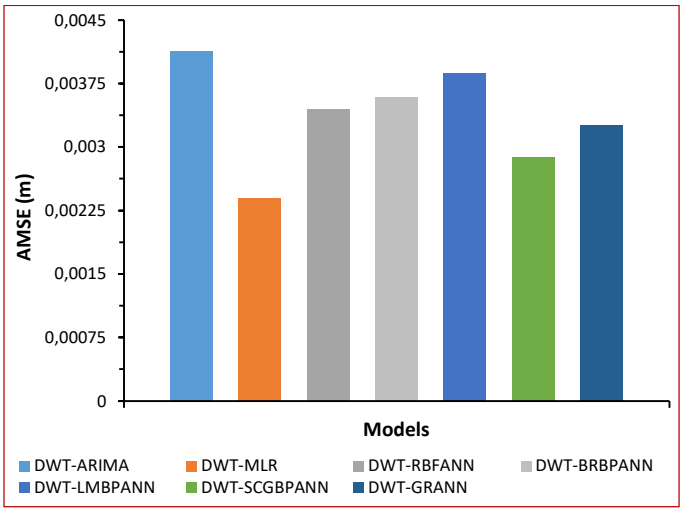

Fig. 5b. AMSE graph of the models

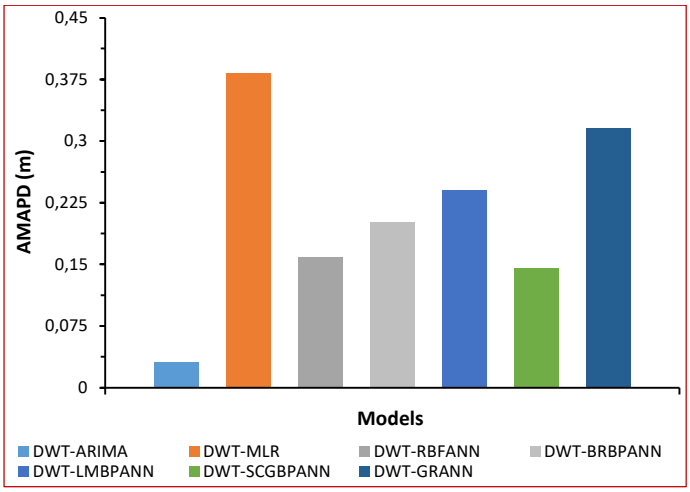

Fig. 5c AMAPD graph of the models

5. Conclusions and Recommendations

LWL study is an important characteristic of the lake ecosystem. Precise predicting of LWL using time series analysis is a crucial step to understand the hydrodynamics of the lake basin. Hence, this is problematic with major economic, social and environmental implications. This study for the first time in Ghana investigates the theoretical and practical analysis of Artificial Intelligence and Classical Regression models in modelling and predicting the monthly LWL of Lake Volta Basin. An ensemble model approach of DWT-BRBPANN, DWT-LMBPANN, DWT-SCGBPANN, DWT-RBFANN, DWT-GRANN, DWT-MLR, and DWT-ARIMA were adopted in predicting the monthly LWL for Lake Volta Basin. The available satellite altimetry data for twenty-eight years, thus, (October 1992 – September 2020) was used in the model building process. The statistical measures that were applied in assessing the performance of the proposed models include *AME*, *AMSE*, *AMAPD*, r_{min}, r_{max}, and *ASD*. The present study selected DWT-MLR model as the best model for modelling the monthly LWL of the Volta Basin based on its statistical performance. DWT-MLR achieved *AME*, *AMSE*, *AMPAD*, and *ASD* of 0.1988 m, 0.0024 m, 0.3824 m, and 0.0017 m respectively which indicated a good forecast of the model as compared to the DWT-ARIMA, DWT-GRANN, DWT-RBFANN, DWT-BRBPANN, DWT-LMBPANN, and DWT-SCGBPANN which achieved 0.2152 m, 0.0041 m, 0.0313 m, 0.0049 m, 0.1905 m, 0.0033 m, 0.3151 m, 0.0026 m, 0.2820 m, 0.0034 m, 0.1585 m, 0.0024 m, 0.2768 m, 0.0036 m, 0.2008 m, 0.0026 m, 0.2776 m, 0.0039 m, 0.2405 m, 0.0029 m, 0.2500 m, 0.0029 m, 0.1450 m, and 0.0020 m respectively. Based on the results achieved in this present study, it is concluded that utilizing hybridized models for LWL modelling have proven to be purposeful and can be adopted for management of lake ecosystem. However, it is recommended that, more works should be done in Lake Volta Basin of Ghana utilizing other soft computing techniques which were not considered in this study. Notably among them are deep learning Convolutional Neural Networks (CNN), Least Squares Support Vector

Machine (LSSVM), Extreme Learning Techniques (ELM), Group Method of Handling Data (GMHD), Genetic Programming and many others to classical improved regression techniques such as Gaussian Regression model, Kernel Ridge Regression, Autoregressive Fractionally Integrated Moving Average (ARFIMA) to evaluate its effectiveness for future further analysis of the Lake Volta Basin.

Acknowledgement
We are much grateful to Dr Christopher E. Ndehedehe, Department of Spatial Sciences, Curtin University, Perth, Australia and Mr. Abubakar Saani Mohammed, Department of Surveying and Mapping Science and Engineering, Hohai University, Nanjing, China for assisting in acquisition of data used for the study area.

Conflict of Interest
The Authors declare that, there is no any potential conflict of interest in this present study. The current study does not infringe any copyright, proprietary right, or any other right of a third party. The Authors are the allodial owners of the work.

References

Abban, E.K., 1999. Integrated Development of Artisanal Fishes. Integrated Development of Artisanal Fisheries Project, GHA/93/008, 1-43.

Adeoti, O.A., Osanaiye, P.A., 2013. Effect of Training Algorithms on The Performance of ANN for Pattern Recognition of Bivariate Process. International Journal of Computer Applications 69 (20), 8-12.

Adnan, R., Ruslan, F.A., Samad, A.M., Zain, Z.M., 2012. Artificial Neural Network Modelling and Flood Water Level Prediction Using Extended Kalman Filter. 2012 IEEE International Conference on Control System, Computing and Engineering, 23-25 November 2012, Penang, Malaysia, 535-538.

Akyen, T., Boye, C. B., Ziggah, Y.Y., 2016. Municipal Solid Waste Estimation and Landfill Lifespan Prediction. 4th UMaT Biennial International Mining and Mineral Conference, Tarkwa, Ghana, GG 127-133.

Al-Krargy, E.M., Mohamed, H.F., Hosney, M.M., Dawod, G.M. 2017. A High-precison Geoid for Water Resources Management: A Case Study in Menofia Governorate, Egypt. National Water Research Center (NWRC) Conference on: Research and Technology Development for Sustainable Water Resources Management, Cairo, Egypt, 1-13.

Amin, M.M., 2003. An up to Date Precise 5′ * 5′ Geoid Grid for Egypt by Collocation Technique. Port-Said Engineering Research Journal (PSERT), Published by faculty of Engineering, Suez Canal University, Port-Said, Egypt, 1-22.

Andersen, O.B., Knudsen, P., 1998. Global marine gravity field from the ERS1 and Geosat geodetic mission. J Geophys Res, 103, 8129–8137.

Arabelos, D., Tziavos, I.N., 1996. Combination of ERS1 and TOPEX altimetry for precise geoid and gravity recovery in the Mediterranean Sea. Geophysical Journal International 125, 285-302.

Arthur, C.K., Temeng, V.A., Ziggah, Y.Y., 2020. Performance Evaluation of Training Algorithms in Backpropagation Neural Network Approach to Blast-Induced Ground Vibration Prediction. Ghana Mining Journal 20 (1), 20-33.

Aytek, A., Kisi, O., Guven, A., 2014. A Genetic Programming Technique for Lake Level Modelling. Hydrology Research 45 (4-5), 529-539.

Baghirli, O., 2015. Comparison of Lavenberg Marquardt, Scaled Conjugate Gradient and Bayesian Regularization Backpropagation Algorithms for Multistep Ahead Wind Speed Forecasting Using Multilayer Perceptron Feedforward Neural Network", Published MSc Thesis Report, Uppsala University, Gotland, 1-35.

Barzegar, R., Adamowski, J., Quilty, J., Aalami, M.T., 2020. Using a Boundary-Corrected Wavelet Transform Coupled with Machine Learning and Hybrid Deep Learning Approaches for Multi-Step Water Level Forecasting in Lakes Michigan and Ontario. EGU General Assembly 2020.

Béné, C., 2007. Diagnostic study of the Volta Basin fisheries Part 1 - Overview of the fisheries resources. Volta Basin Focal Project Report No 6. World Fish Center Regional Offices for Africa and West Asia, Cairo Egypt, and CPWF, Colombo, Sri Lanka, 1-31.

Béné, C., Russell, A.J.M., 2007. Diagnostic study of the Volta Basin fisheries. Part 1 - Livelihoods and poverty analysis, current trends and projections. Volta Basin Focal Project Report No 7. World Fish Center Regional Offices for Africa and West Asia, Cairo Egypt, and CPWF, Colombo, Sri Lanka, 1-67.

Béné, C., Obirih-Opareh, N., 2009. Social and Economic Impacts of Agricultural Productivity Intensification: The Case of Brush Park Fisheries in Lake Volta. Agricultural Systems, 102, 1-10.

Box, G.E.P., Jenkins, G.M., 1976. Time Series Analysis: Forecasting and Control, Holden-Day, Boca Raton, Fla, USA.

Boye, P., Ziggah, Y.Y., 2020. A Short-Term Stock Exchange Prediction Model Using Box-Jenkins Approach. Journal of Applied Mathematics and Physics 8, 766-779.

Braimah, L.I., 2003. Fisheries Management Plan for the Volta Lake, Ministry of Food and Agriculture, Directorate of Fisheries, Accra, Ghana, 1-77.

Cazenave, A., Schaeffer, P., Berge, M., Brosier, C., Dominh, K., Genero, M.C., 1996. High-resolution mean sea surface computed with altimeter data of ERS1 (Geodetic Mission) and TOPEX/POSEION. Geophysical Journal International 125, 696-704.

Chakraborty, A., Goswami, D., 2017. Slope Stability Prediction using Artificial Neural Network (ANN). International Journal of Engineering and Computer Science 6 (6), 21845-21848.

Chao-Long, Y., Li-Long, L., Si, X., 2011. Wavelet Denoising and Dynamic Fuzzy Neural Network in the Application of Deformation Analysis. 2011 International Conference on Instrumentation, Measurement, Computer, Communication and Control, 270-273.

Chen, W., Hill, C., 2005. Evaluation Procedure for Coordinate Transformation. Journal of Surveying Engineering 131 (2), 43-49.

Coe, M.T., Birkett, C.M., 2004. Calculation of River Discharge and Prediction of Lake Height from Satellite Radar Altimetry: Example for the lake Chad Basin. Water Resources Research 40 (W10205), 1-11.

De-Bao, Y., Ximin, C., Jinging, J., Guo, W., Wanyang, X., 2012. Application Research of Wavelet Neural Network in Prediction of Deformation. 2012 Asia Pacific Conference on Environmental Science and Technology, Advances in Biomedical Engineering, 202-209.

Deo, M.C., Chaudhari, G., 1998. Tide Prediction Using Neural Networks. Computer-Aided Civil and Infrastructure Engineering 13, 113-120.

Demir, V., Ulke Keskin, A., 2020. Height Modelling with Artificial Neural Networks (Samsun_Mert River Basin). Gazi Journal of Engineering Sciences 6 (1), 54-61.

Dickson, K.B., Benneh, G., 1977. A New Geography of Ghana. Longman Group, London, UK, 1-173.

Dudek, G., 2011. Generalized Regression Neural Network for Forecasting Time Series with Multiple Seasonal Cycles.

Springer-Verlag Berlin Heidelberg 1, 1-8.

Ebtehaj, I., Bonakdari, H., Gharabaghi, B., 2019. A Reliable Linear Method for Modelling Lake Level Fluctuations. Journal of Hydrology 570, 236-250.

El-Shazly, A. H., 2005. The Efficiency of Neural Networks to Model and Predict Monthly Mean Sea Level from Short Spans Applied to Alexandria Tide Gauge. From Pharaohs to Geoinformatics, FIG Working Week 2005 and GSDI-8, Cairo, Egypt, April 16-21, 2005, 1-13.

Erdogan, S., 2009. A Comparison of Interpolation Methods for Producing Digital Elevation Models at the Field Scale. Earth Surface Processes and Landforms, 34, 366-376.

Farajzadeh, J., Fard, A. F., Lotfi, S., 2014. Modelling of Monthly Rainfall and Runoff of Urmia Lake Basin Using Feed-Forward Neural Network and Time Series Analysis Model. Water Resources and Industry 7 (8), 38-48.

Fernandez, F.R.Q., Montero, N.B., III, R B.P., Addawe, R.C., Diza, H.M.R., 2018. Forecasting Manila South Harbour Mean Sea Level Using Seasonal ARIMA Models. Journal of Technology Management and Business 5 (1), 1-7.

Fernandez, F.R., III, R.P., Montero, N., Addawe, R., 2017. Prediction of South China Sea Level Using Seasonal ARIMA Models. Proceedings of the 13th IMT-GT International Conference on Mathematics, Statistics and their Applications (ICMSA 2017), 050018-1-050018-6.

Fischer, M.L., Markowska, M., Bachofer, F., Foerster, V.E., Asrat, A., Zielhofer, C., Trauth, M.H., Junginger, A., 2020. Determining the Pace and Magnitude of Lake Level Changes in Southern Ethiopia Over the Last 20,000 Years Using Lake Balance Modelling and SEBAL. Frontiers in Earth Science 8 (197), 1-21.

Foresee, F.D., Hagan, M.T. 1997. Gauss-Newton approximation to Bayesian learning. Proceedings of the International Joint Conference on Neural Networks, 3, 1930-1935.

Ghilani, D.C., 2010. Adjustment Computations, Spatial Data Analysis. Fifth Edition, Wiley & Sons, INC. Hoboken, New Jersey, USA, 1-674.

Ghorbani, M.A., Khatibi, R., Aytek, A., Makarynskyy, O., Shiri, J., 2010. Sea Water Level Forecasting Using Genetic Programming and Comparing the Performance with Artificial Neural Networks. Computers & Geosciences 36, 620-627.

Grgic, S.M., Jukic, S., Nerem, R.S., Basic, S.T., 2017. The Assessment of an Impact of Mean Sea Level change in the Mid-Adriatic Region Based on Satellite Altimeter Records. 17th International Multidisciplinary Scientific Geoconference SGEM 2017, Section Photogrammetry and Remote Sensing, 263-270.

Gyau-Boakye, P., 2001. Environmental Impacts of the Akosombo Dam and Effects of Climate Change on the Lake Levels. Environ. Dev. Substain, 3(1), 17-29.

Hagan, M.T., Demuth, H.B., Beale, M.H., 1996. Neural Network Design, Boston, MA: PWS Publishing.

Hannan, S.A., Manza, R.R., Ramteke, R.J., 2010. Generalized Regression Neural Network and Radial Basis Function for Heart Disease Diagnosis. International Journal of Computer Applications 7 (13), 7-13.

Huang, F.M., Wu, P., Ziggah, Y.Y., 2016. GPS Monitoring Landslide Deformation Signal Processing Using Time Series Model. International Journal of Signal Processing, Image Processing and Pattern Recognition 9 (3), 321-332.

Idri, A., Zakrani, A., Zahi, A., 2010. Design of Radial Basis Function Neural Networks for Software Effort Estimation. International Journal of Computer Science Issue, 7 (4), 11-17.

Ismail, S., Pandiali, S.M., Shabri, A., Mustapha, A., 2018. Comparative Analysis of River Flow Modelling by Using Supervised Learning Technique. IOP Conf. Series, Journal of Physics: Conference Series 995 (2018) 012045, 1-9.

Jie-Xing, Z., Jing, W., Peng-Fei, W., Yang, H., Bin, L., Jing, L., 2012. Wavelet Analysis of Water Quality Changes in Dianchi Lake during the Past 7a. International Conference on Structural Computation and Geotechnical Mechanics, Procedia Earth and Planetary Sciences 5, 280-288.

Jian-Jun, S., 2003. Prediction and Analysis of Tides and Tidal Currents. International Hydrographic Review 4 (2), 24-29.

Kaloop, M.R., Beshr, A.A.A., Zarzoura, F., Ban, W.H., Hu, J.W. 2020. Predicting Lake Wave Height Based on Regression Classification and Multi Input-Single Output Soft Computing Models. Arabian Journal of Geosciences 13 (591), 1-14.

Kaloop, M.R., Rabah, M., Hu, J.W., Zaki, A., 2017. Using Advanced Soft Computing Techniques for Regional Shoreline Geoid Model Estimation and Evaluation. Marine Georesources & Geotechnology 36 (6) 1-11.

Kaur, H., Salaria, D.S., 2013. Bayesian Regularization Based Neural Network Tool for Software Effort Estimation, Global Journal of Computer Science and Technology 13 (2), 44-50.

Kisi, O., Shiri, J., Nikoofar, B., 2012. Forecasting Daily Lake Levels Using Artificial Intelligence Approaches. Computers and Geoscience 41, 169-180.

Kişi, Ö., Uncuoğlu, E., 2005. Comparison of three back-propagation training algorithm for two case studies. Indian Journal of Engineering and Materials Science 12, 434-442.

Kumi-Boateng, B., Peprah, M.S., 2020. Modelling Local Geometric Geoid using Soft Computing and Classical Techniques: A Case Study of the University of Mines and Technology (UMaT) Local Geodetic Reference Network. International Journal of Earth Sciences Knowledge and Applications 2 (3), 166-177.

Lawson, G.W., 1970. Lessons of the Volta – A New Man-made Lake in Tropical Africa. Africa Biological Conservation 2 (2), 91-96.

Ledolter, J., 2008. A Statistical Analysis of the Lake Levels at Lake Neusiedl. Austrian Journal of Statistics 37 (2), 147-160.

Lemoine, F.G., Kenyon, S.C., Factor, J.K., Trimmer, R.G., Pavlis, N.K., Chinn, D.S., Cox, C., Klosko, S.M., Luthcke, S.B., Torrence, M.H., Wang, Y.M., Williamson, R.G., Pavlis, E.C., Rapp, R.H., Olson, T.R., 1998. The development of the join NASA GSFC and NIMA geopotential model EGM96. NASA Technical Paper, 1-4.

Li, G., Shi, J., 2012. Applications of Bayesian methods in wind energy conversion systems. Renewable Energy 43, 1-8.

Lin, P., Yang, Z.L., Cai, X., David, C.H., 2015. Development and Evaluation of a Physically-based Lake Level Model for Water Resource Management: A Case Study for lake Buchanan, Texas. Journal of Hydrology: Regional Studies 4, 661-674.

Liu, D., Wang, X., Zhang, Y.L., Yan, S.J., Cui, B.S., Yang, Z.F., 2019. A Landscape Connectivity Approach for Determining Minimum Ecological Lake Level: Implications for Lake Restoration. Water 11 (2237), 1-14.

Makwinja, R., Phiri, T., Kosamu, I.B.M., Kaonga, C.C., 2017. Application of Stochastic Models in Predicting Lake Malawi Water Levels. International Journal of Water Resources and Environmental Engineering 9 (9), 191-200.

Makarynskyy, O., Kuhn, M., Makarynska, D., Featherstone, W.E., 2004. The Use of Artificial Neural Networks to Retrieve Sea Level Information from Remote Data Sources. IAG International Symposium: Gravity, Geoid and Space Missions, Porto, Portugal, August 30-September 3, 2004, Springer Verlag, Berlin, Germany, 1-4.

Mitchum, G.T., 2000. An Improved Calibration of Satellite Altimetric Heights using Tide Gauge Sea Levels with Adjustment for land Motion. Marine Geodesy 23, 145-166.

Møller, M.F., 1991. A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning. Neural Networks 6, 525-533.

Mueller, V.A., Hemond, F.H., 2013. Extended artificial neural networks: in-corporation of a priori chemical knowledge enables use

of ion selective electrodes for in-situ measurement of ions at environmental relevant levels. Talanta 117, 112-118.

Muthuwatta, L.P., 2004. Long Term Rainfall-Runoff-Lake Level Modelling of the Lake Naivasha Basin, Kenya. Published MSc Dissertation, Water Resources Survey and Environmental Systems Analysis and Management, International Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands, 1-89.

Ndehedehe, C.E., Awange, J.L., Kuhn, M., Agutu, N.O., Fukuda, Y., 2017. Analysis of Hydrological Variability over the Volta River Basin using in-situ Data and Satellite Observations. Journal of Hydrology: Regional Studies 12, 88-110.

Ni, S., Cehn, J., Wilson, C. R., Hu, X. (2017). Long-Term Water Storage Changes of Lake Volta from GRACE and Satellite Altimetry and Connections with Regional Climate, Remote Sensing, 9 (842).

Nikentari, N., 2017. Tide Forecast Using Radial Basis Function Neural Network. ADRI International Journal of Semantic Technology 1, 38-40.

Okwuashi, O., Olayinka, D.N., 2017. Tide Modelling Using the Kalman Filter. Journal of Spatial Science 62 (2), 353-365.

Okwuashi, O., Ndehedehe, C.E., 2017. Tide Modelling Using Support Vector Machine Regression. Journal of Spatial Science 62 (1), 29-46.

Owusu, K., Waylen, P., Qiu, Y., 2008. Changing Rainfall Inputs in the Volta Basin: Implications for Water Sharing in Ghana. GeoJournal 71 (4), 201-210.

Pan, X., Lee, B., Zhang, C., 2013. A comparison of neural network backpropagation algorithms for electricity load forecasting. Intelligent Energy Systems (IWIES), 2013 IEEE International Workshop, 22-27.

Pashova, L., Popova, S., 2011. Daily Sea Level Forecast at Tide Gauge Burgas, Bulgaria Using Artificial Neural Networks. Journal of Sea Research 66, 154-161.

Peprah, S.M., Yevenyo, Y.Y., Issaka, I., 2017. Performance Evaluation of the Earth Gravitational Model (EGM2008) – A Case Study. South African Journal of Geomatics 6 (1), 47-72.

Peprah, M.S., Mensah, I.O., 2017. Performance Evaluation of the Ordinary Least Square (OLS) and Total Least Square (TLS) in Adjusting Field Data: An Empirical Study on a DGPS Data. South African Journal of Geomatics 6 (1), 73-89.

Peprah, M.S., Kumi, S.A., 2017. Appraisal of Methods for Estimating Orthometric Heights – A Case Study in a Mine. Journal of Geoscience and Geomatics 5 (3), 96-108.

Peprah, M.S., Larbi, E.K., 2021. Lake Water Level Prediction Model Based on Autocorrelation Regressive Integrated Moving Average and Kalman Filtering Techniques – An Empirical Study on Lake Volta Basin, Ghana. International Journal of Earth Sciences Knowledge and Applications 3 (1), 1-11.

Poku-Gyamfi, Y., 2009. Establishment of GPS Reference Network in Ghana. Published MPhil Dissertation, Universitat Der Bundeswehr Munchen Werner Heisenberg-Weg 39, 85577, Germany, 1-218.

Piri, J., Kahkha, M.R.R., 2016. Prediction of Water Level Fluctuations of Chahnimch Reservoirs in Zabol Using ANN, ANFIS, and Cuckoo Optimization Algorithm. Iranian Journal of Health, Safety and Environment 4 (2), 706-715.

Piasecki, A., Jurasz, J., Adamowski, J.F., 2018. Forecasting Surface Water-Level Fluctuations of a Small Glacial Lake in Poland Using a Wavelet-based Artificial Intelligence Method. Acta Geophysica 66 (5), 1093-1107.

Piasecki, A., Jurasz, J., Skowron, R., 2015. Application of Artificial Neural Networks (ANN) in Lake Drweckie Water Level Modelling. Limnological Review 15 (1), 21-29.

Pozzi, M., Malmgren, B.A., Monechi, S., 2000. Sea Surface-Water Temperature and Isotopic Reconstructions from Nannoplankton Data Using Artificial Neural Networks. Palaeontologia Electronica 3 (2), 1-14.

Rodgers, C., Van de Giesen, N., Laube, W., Vick, P.L.G., Youkhana, E., 2007. The GLOWA Volta Project: A Framework for Water Resources Decision-making and Scientific Capacity Building in a Transnational West African Basin. In Integrated Assessment of Water Resources and Global Change: A North-South Analysis, Springer: Dordrecht, The Netherlands, 295-313.

Sahay, R.S., Sehgal, V., 2014. Wavelet-ANFIS Models for Forecasting Monsoon Flows: Case Study for the Gandak River (India). Water Resources 41 (5), 574-582.

Sandhu, P.S., Chhabra, S., 2011. A Comparative Analysis of Conjugate Gradient Algorithms and PSO Based Neural Network Approaches for Reusability Evaluation of Procedure Based Software Systems. Chiang Mai Journal of Science 38 (2), 123-135.

Sehgal, V., Tiwari, M.K., Chatterjee, C., 2014. Wavelet Bootstrap Multiple Linear Regression based Hybrid Modelling for Daily River Discharge Forecasting. Water Resources Management 28, 2793-2811.

Shafaei, M., Kisi, O., 2015. Lake Level Forecasting Using Wavelet-SVR, Wavelet-ANFIS, and Wavelet-ARMA Conjunction Models. Water Resources Management 29, 1-21.

Sheta, A.F., Ahmed, S.E.M., Faris, H., 2015. A Comparison between Regression, Artificial Neural Networks and Support Vector Machines for Predicting Stock Market Index. International Journal of Advanced Research in Artificial Intelligence 4 (7), 55-63.

Sithara, S., Pramada, S. K., Thampi, S.G., 2020. Sea Level Prediction Using Climatic Variables: A Comparative Study of SVM and Hybrid Wavelet SVM Approaches. Journal Acta Geophysica 68 (1779-1790).

Specht, D., 1991. A General Regression Neural Network. IEEE Transactions on Neural Networks 2 (6), 568-576.

Srichandan, S., 2012. A New Approach of Software Effort Estimation using Radial Basis Function Neural Networks. ISSN (Print), 1(1), 2319-2526.

Srivastava, P., Kumar, A., Singh, R., Deepak, O., Kumar, A.M., Ray, Y., Jayangondaperumal, R., Phartiyal, B., Chahal, P., Sharma, P., Ghosh, R., Kumar, N., Rajesh, A., 2020. Rapid Lake Level Fall in Pangong Tso (Lake) in Ladakh, NW Himalaya: A Response of Late Holocene Aridity. Current Science 119 (2), 219-231.

Srivastava, P.K., Islam, T., Singh, S.K., Petropoulos, G.P., Gupta, M., Dai, Q., 2016. Forecasting Arabian Sea Level Rise Using Exponential Smoothing State Space Models and ARIMA from TOPEX and Jason Satellite Radar Altimeter Data. Meteorological Applications 23, 633-639.

Tiryaki, B., 2008. Predicting intact rock strength for mechanical excavation using multivariate statistics, Artificial Neural Networks and Regression Trees. Engineering Geology 99, 51-60.

Tolkatchev, A., 1996. Global Sea Level Observing System. Marine Geodesy 19, 21-62.

Turner, J.F., Iliffe, J.C., Ziebart, M.K. Jones, C., 2013. Global Ocean Tide Models: Assessment and use within a surface Model of Lowest Astronomical Tide. Marine Geodesy 36, 123-137.

Tziavos, I.N., Vergos, G.S., Kotzev, V., Pashova, L., 2005. Mean Sea Level and Sea Level Variation Studies in the Black Sea and the Aegean. Gravity, Geoid and Space Missions, 254-259.

Vergos, G.S., Tziavos, I.N., Andritsanos, V.D., 2003. On the Determination of Marine Geoid Models by Least Squares Collocation and Spectral Methods using Heterogeneous Data. Presented at Session G03 of the 2004 IUGG General Assembly, Sapporo, Japan, July 2–8, 2003, 1-5.

Veitch, D., 2005. Wavelet Neural Networks and their Applications in the Study of Dynamical Systems. Published MSc Thesis in

Data Analysis, Networks and Nonlinear Dynamics, Department of Mathematics, University of York, UK, 1-90.

Wenzel, H.G., 1998. Ultra-high degree geopotential model GPM3E97 to degree and order 1800 tailored to Europe. Presented at the 2nd Continental Workshop on the geoid in Europe, Budapest, Hungary, 1-4.

Wilamowski, B.M., 2009. Neural network architectures and learning algorithms. Industrial Electronics Magazine, IEEE, 3 (4), 56-63.

Xue, Y.J., Cao, J.X., Zhang, G.I., Du, H.K., Wen, Z., Zeng, X.H., Zou, F., 2017. Application of Local Wave Decomposition in Seismic Signal Processing. Intech, Open Science, Open Minds, 1-25.

Yakubu, I., Dadzie, I., 2019. Modelling Uncertainties in Differential Global Positioning System Dataset. Journal of Geomatics 13 (1), 16-23.

Yakubu, I., Ziggah, Y.Y., Peprah, M.S., 2018a. Adjustment of DGPS Data using artificial intelligence and classical least square techniques. Journal of Geomatics 12 (1), 13-20.

Yakubu, I., Ziggah, Y.Y., Baafi K.A., 2018b. Prediction of Tidal Effect on the Earth Crust for Geodetic Deformation Monitoring. Ghana Journal of Technology 2 (2): 63 - 69.

Yaseen, Z.M., Naghshara, S., Salih, S.Q., Kim, S., Malik, A., Ghorbani, M.A. 2020. Lake Water Level Modelling Using Newly Developed Hybrid Data Intelligence Model. Theoretical and Applied Climatology.

Young, C.C., Liu, W.C., Hsieh, W.L., 2015. Predicting the Water Level Fluctuation in an Alpine Lake Using Physically Based, Artificial Neural Network, and Time Series Forecasting Models. Mathematical Problems in Engineering, 1-5.

Yusof, F., Kane, I.L., Yusof, Z., 2013. Hybrid of ARIMA-GARCH Modelling in Rainfall Time Series. Jurnal Teknologi (Sciences and Engineering) 63 (2), 27-34.

Wang, Y., Yuan, Y., Pan, Y., Fan, Z., 2020. Modelling Daily and Monthly Water Quality Indicators in a Canal Using a Hybrid Wavelet-based Support Vector Regression Structure. Water 12 (1476), 1-21.

Zahra, G., Xiaoli, D., 2015. Application of the Multi-Adaptive Regression Splines to Integrate Sea Level Data from Altimetry and Tide Gauges for Monitoring Extreme Sea Level Events. Marine Geodesy 38 (3), 261-276.

Zhou, C., Yin, K., 2014. Landslide Displacement Prediction of WA-SVM Coupling Model Based on Chaotic Sequence. Electronic Journal of Geotechnical Engineering 19, 2973-2987.

Zhu, S., Lu, H., Ptak, M., Dai, J., Ji, Q., 2020a. Lake Water-Level Fluctuation Forecasting Using Machine Learning Models: A Systematic Review. Environmental Science and Pollution Research 27, 44807-44819.

Zhu, S., Hrnjica, B., Ptak, M., Choinski, A., Sivakumar, B., 2020b. Forecasting of Water Level in Multiple Temperature Lakes Using Machine Learning Models. Journal of Hydrology 585 (124819), 1-13.

Zhu, S., Ptak, M., Yaseen, Z.M., Dai, J., Sivakumar, B., 2020c. Forecasting Surface Water Temperatures in Lakes: A Comparison of Approaches. Journal of Hydrology 585 (124809), 1-10.

Zhu, S., Hadzima-Nyarko, M., Gao, A., Wang, F., Wu, J., Wu, S., 2019. Two Hybrid Data-Driven Models for Modelling Water-Air Temperature Relationship in Rivers. Environmental Science and Pollution Research, 1-10

Ziggah, Y.Y., Youjian, H., Yu, X., Basommi, L.P., 2016. Capability of Artificial Neural Network for forward Conversion of Geodetic Coordinates (Φ, λ, h) to Cartesian Coordinates (X, Y, Z). Mathematical Geosciences 48, 687-721.