Prediction of piezometric water level using artificial neural network optimized with particle swarm optimization

Prosper Basommi LAARI, Ghana

Keywords: Piezometric water level, particle swarm optimisation, artificial neural network, dam

SUMMARY

Recently, the government of Ghana re-organised some administrative regions to ensure Piezometric water level is a vital measure that contribute to the safety analysis of dam when performing seepage analysis. This study adopted hybrid machine learning models and performed statistical evaluation of their performance to predicting piezometric water level. These hybrid methods include Particle Swarm Optimisation and Backpropagation Neural Network (PSO-BPNN), Particle Swarm Optimisation and Radial Basis Function Neural Network (PSO-RBFNN) and Particle Swarm Optimisation and Generalized Regression Neural Network (PSO-GRNN). The standard Multiple Linear Regression (MLR) was also applied for comparison purposes. To assess the quality of the model and its efficiency, the study adopted Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and correlation coefficient (R). The results obtained showed that the applied hybrid machine learning models performed very well than the standalone methods of BPNN, RBFNN, GRNN and MLR. That is, PSO-BPNN, PSO- RBFNN and PSO-GRNN had the best RMSE values of 3.33E-07 m, 0.0001 m and 3.13E-04 m with corresponding MAE values of 3.47E-08 m, 0.0012 m and 3.86E-05 m. The PSO-BFNN, PSO-RBFNN and PSO-GRNN also recorded the largest R values of 0.8 m, 0.9 m and 0.8 m. The developed and tested hybrid models are a major contribution to industry players who are concerned about the safety of dams and the working environment.

Prediction of piezometric water level using artificial neural network optimized with particle swarm optimization

Prosper Basommi LAARI, Ghana

1. INTRODUCTION

Dams play a vital role in water resource management. They meet the needs of drinking, industrial water supply, flood control, and increase dry-weather flows. They can also maintain a wetland environment that is favourable to biodiversity. In some cases, they can provide indefinitely renewable hydroelectric power. Besides being a great source of wealth, dams can also become a source of accidents (Gourine and Khelifa, 2018). To prevent such accidents, dam safety must be considered when engineers design, construct, operate, and maintain the dams. The safety control of dams depends on the measurement of parameters such as seepage flows, deformations or movements, loading conditions, seepage water clarity, water levels, pressures, temperature variations, piezometric water levels among others.

A dam can also be exposed to significant water level variations and seasonal environmental temperature changes. All dams suffer to a certain extent from seepage, as water blocked by a dam will find ways to break through the dam and its surroundings. Such seepage can undermine the dam structure and cause the dam to fail. To monitor the seepage, piezometers are installed at certain sections of the dam. To accurately simulate and predict the dam safety monitoring and to monitor the running state of the dam and eliminate the danger in advance, an effective and efficient model of dam deformation is greatly significant.

In the past decades, several models such as the deterministic models, statistical models, Artificial Intelligence (AI) have been developed to predict piezometric water levels in dam deformation. Statistical regression models have been proposed to analyse and describe dam deformation data quantitively. To predict water level in piezometers, an Artificial Neural Network (ANN) model was developed by Tayfur et al. (2005). The authors used water levels on the upstream and the downstream sides of the dam as the input variables. It is important to note that a suitable choice of input variables for ANN is vital for accurate modelling of water levels in piezometers. Research conducted by Bonelli and Royet (2001) showed that piezometric measurements are generally affected by the water level effect and rainfall effect. As such, much research done in the prediction of piezometric water level considered rainfall as the input variable for the prediction (Ranković et al., 2014).

The studied dam is a 400-megawatt hydroelectric project in Ghana. As part of the temporary firmness of the dam, this study takes into consideration the piezometric water level which is mostly considered in seepage analysis of dam deformation evaluation. To achieve that, this study seeks to perform statistical evaluation of the performance of hybrid machine learning to predicting piezometric water level in dam deformation. The hybrid methods include Particle Swarm Optimisation and Backpropagation Neural Network (PSO-BPNN), Particle Swarm Optimisation and Radial Basis Function Neural Network (PSO-RBFNN) and Particle Swarm Optimisation and Generalized Regression Neural Network (PSO-GRNN). The hybrid methods were compared with standalone BPNN, RBFNN, GRNN and the conventional Multiple Linear Regression (MLR). The chosen methods have been extensively applied in different areas in dam deformation studies. Furthermore, there has been close to no application in literature on evaluating together the prediction competence of different AI methods (BPNN, RBFNN and GRNN) optimised with PSO in dam piezometric water level prediction. The achieved results in this study showed that the hybrid models enhanced the performance of the standalone methods can produce reasonable predictions.

2. STUDY AREA

The studied hydroelectric dam was built on the Black Volta River of Ghana with an energy generation capacity of 400-megawatt. The dam shares border with Northern Region and Brong-Ahafo Region (Anon., 2020). The dam is 108 m high above the foundation, 90 m above the riverbed and 492 m long crest with maximum and minimum operating level of 185 and 167 m. The dam consists of two saddle dams (saddle dams 1 and 2) and the main dam. A rock-fill embankment dam defines Saddle dam 1 which is located 500 m away from the southeastern gate of the main dam with a crest length of 300 m and a height of 37 m above ground level (Dietz et al. 2014). A zone earth fill dam defines Saddle dam 2 which can be found 1 km at the southwestern gate of the main dam with a crest length of 500m. Both Saddle dams 1 and 2 have a crest elevation of 187 m above mean sea level. The main and saddle dams have a combined reservoir capacity of 12,570 million cubic metres.

3. MATERIALS AND METHODS

3.1 Data used.

To evaluate the performance of machine learning to predicting piezometric water level in dam deformation, this study used dam deformation monitoring data from the Bui Power Authority which was collected weekly for a period of 2 years 4 months (January 2013 to April 2015). The data consist of observations of pressure, temperature, modulus (modulus of elasticity of the dam material in which piezometers are installed), and piezometric water level. Data from three blocks on the dam (Block 12, 17, and 24) alongside four different piezometers installed on each block were taken into consideration in this study. The study also considered rainfall as an extra predictor variable for the analysis.

The dam deformation monitoring data was obtained from the Bui Power Authority. As part of the safety control of the Bui dam, weekly observations are made on the dam to obtain geotechnical and geodetic parameters needed for the safety analysis of the dam. Given that, the data is valid and dependable for analytical research on the dam. Rainfall data for the Bui dam was obtained from the Ghana Meteorological Agency. Ghana Meteorological Agency is responsible for the analysis and forecast of meteorological conditions and they also provide efficient and reliable meteorological information by collecting, processing, analyzing among others. Therefore, there is belief and confidence in the data observation and its validity for any research analysis.

3.2 Methods

This study used four machine learning methods and one classical method for the piezometric water level prediction. These methods include Backpropagation Neural Network (BPNN), Radial Basis Function Neural network (RBFNN), Generalized Regression Neural Network (GRNN), Particle Swarm Optimization (PSO) and Multiple Linear Regression (MLR). The Particle Swarm Optimization, which is an optimization algorithm is then hybridized with the three Artificial Neural Network methods (BPNN, RBFNN and GRNN) for the piezometric water level prediction.

All the methods used pressure, temperature, modulus (modulus elasticity of the dam material in which the piezometers are installed), and rainfall as predictor (input) variables and the piezometric water level as the response (output) variable. The data were divided into a training set and testing set with the training set having a data span of 1 year 6 months (about 80 observations) and the testing set having a data span of 10 months (about 41 observations).

3.2.1. Particle Swarm Optimization and Backpropagation Neural Network (PSO-BPNN)

The PSO-BPNN is an optimised hybrid intelligent model formed by combining PSO with the BPNN. The PSO algorithm is a global algorithm that has a strong ability to find a global optimistic result. The backpropagation neural network algorithm, on the other hand, has a strong ability to find the local optimistic result but its ability to find global optimistic results is weak (Zhang *et al.,* 2007). So, this study combined the particle swarm optimization with the backpropagation neural network. The basic idea for this hybrid algorithm is that the particle swarm optimization is employed to accelerate the training speed. The hybrid algorithm uses the PSO algorithm to do a global search in the beginning and uses the backpropagation to do the local search. The particle swarm optimization is also hybridized with the backpropagation neural network to assist in finding optimal weight. The weight of the backpropagation neural network is the values that are used as particles in particle swarm optimization.

In the PSO-BPNN algorithm (Figure 1), first, the data for training and testing were divided randomly with the training set having a data span of 1 year 6 months (about 80 observations) and the testing set having a data span of 10 months (about 41 observations). Then, the network and parameters of the backpropagation and particle swarm optimization is initialized. The PSO parameters are also initialized. Then, each particle's weight is assigned into BPNN after which the network is trained and tested. The weights are then updated for every epoch and the personal best (pbest) position and the global best (gbest) position are also updated as well as the particle's velocity and position and finally set the stopping criteria.

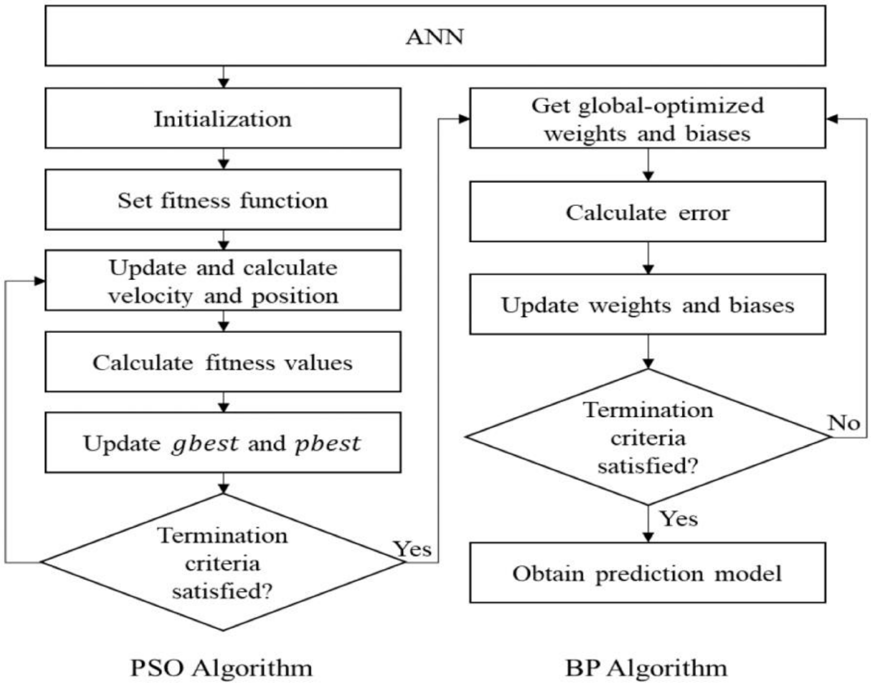

Figure 1 A Flow Chart of Hybrid PSO-BPNN

In setting of the PSO parameters, the various parameters are considered. These include inertia weight w, particle number m, accelerating constants c_1 and c_2, maximum limited velocity $v_{max.}$, the max iteration number T_{max}, and computed precision. Of the various parameters, w, m, c_1, c_2, and v_{max} are used as mainly controlling parameters, at the same time T_{max} is used as the conditions of stopping iteration. The inertia weight describes the previous velocity influence on current velocity. If $w = 0$, then the velocity of the particle depends on its current position pbest (the personal best position) and *gbest* (the global best position). The maximum and minimum inertia used were $w_{max} = 0.9$ ˈ $w_{min} = 0.4$.

Training of PSO-BPNN

In the training of the model, a data span of 1 year 6 months which covers about 66% of the entire dataset was used for the training to predict the piezometric water levels.

Testing of PSO-BPNN

The model is then tested to check the performance. A data span of 10 months which covers about 30% of the entire dataset was used to test the model. Here, the predictor variables are applied to the trained data without the response variable.

3.2.2. Particle Swarm Optimization and Radial Basis Function Neural Network (PSO-RBFNN)

Radial basis function neural network is an artificial neural network that uses radial basis functions as an activation function. The output of the network is a linear combination of radial basis functions of the inputs and neuron parameters and has an input layer, hidden layer, and an output layer.

The particle swarm optimization works through initializing a swarm randomly in the search space, attracting the particles to search for positions of high fitness. Each particle has an adaptive value determined by the optimized function.

Combining particle swarm optimization and radial basis function neural network, the global searching capability of the particle swarm optimization is used to optimize the topology of the radial basis function neural network, its connection weights, learning rules, improving the generalization capability, and learning efficiency which improve the performance of the radial basis function neural network.

In the PSO-RBFNN algorithm (Figure 2), first, the data for training and testing are divided randomly with the training set having a data span of 1 year 6 months (about 80 observations) and the testing set having a data span of 10 months (about 41 observations). Then the networking training specimen is collected after which the topology structure of the radial function network is built to determine the number of inputs, output, and hidden nodes. The population size is initialized. Then the fitness value of each particle is calculated and compared with the particle's best position. If the current value is better than the previous best solution, it is replaced, and the current solution is set as the local best position. The individual particles fitness is compared with the population best global position and if the current solution is better than the global best fitness, the current solution is set as the new global best solution. The particle's positions and velocities are updated and the personal best position (*pbest*) and global best position (*gbest*) of the particles are updated and the stopping criterion is finally set.

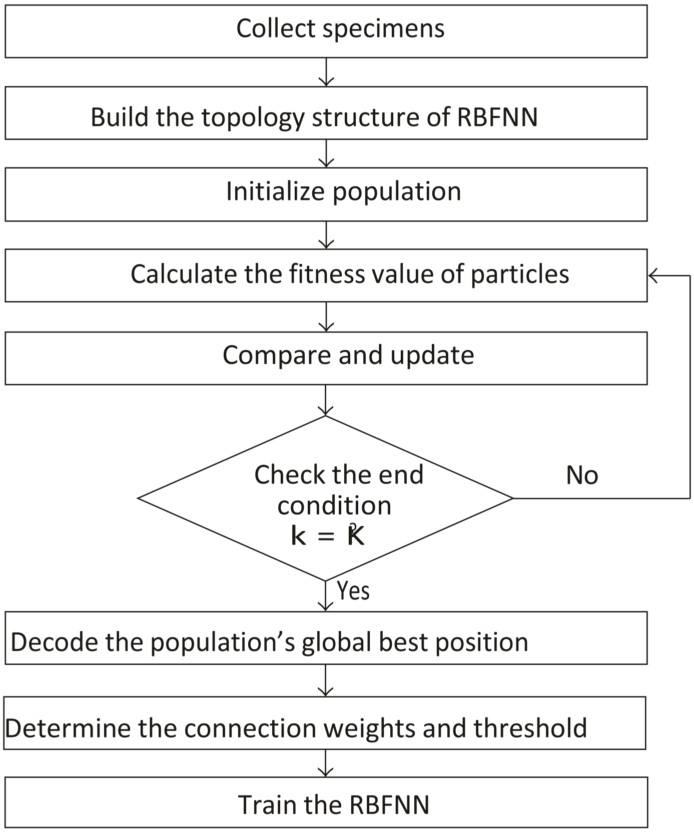

Figure 2 A Flow Chart of Hybrid PSO-RBFNN

Training of PSO-RBFNN

In training of the model, a data span of 1 year 6 months which covers about 66% of the entire dataset was used for the training to predict the piezometric water levels. Then the PSO parameters such as w, m, c_1, c_2, and v_{max} are determined. Then the connection weight is determined, and the radial basis function neural network is trained to predict the piezometric water level.

Testing of PSO-RBFNN

In testing the model, 34% of the entire data observation was used after the training to test how the model performs. The input variables were passed through the designed and trained network to predict the output variable which is the Piezometric water level.

3.2.3. Particle Swarm Optimization and Generalized Regression Neural Network (PSO-GRNN)

Generalized regression neural network is a kind of artificial neural network which uses a brain synapse-like structure to manage information (Tsuda, 1992). It is a variation to radial basis neural networks, and it represents an improved technique in the neural networks based on the nonparametric regression. The idea is that every training sample will represent a mean to a radial basis neuron.

In order to optimize the performance of the generalized regression neural network model and reduce the prediction error, we apply particle swarm optimization algorithm to search the optimal smoothing parameter to construct the generalized regression neural network model. The PSO algorithm simulates a simplified social model that is composed of a group of particles. The particle has two parameters, including position and velocity. In the solution process of an optimization problem, these particles will modify their positions according to their own learning experience and their neighbours and finally find their best positions.

In the optimization for GRNN model based on PSO algorithm (Figure 3), the smoothing parameter is set as the position of the particles and the searching space is one dimensional. The parameters in the tth iteration process of PSO algorithm were set as follows. The personal position of the kth particle is $p_k(t)$ and its velocity is $v_k(t)$. The best personal position of the particle is $pbest_k(t)$ and the global best position of the swarm is $gbest_k(t)$.

In PSO algorithm, Mean Square Error (MSE) between prediction values and actual values of the testing samples reflects the prediction accuracy of the model and is be used to judge the quality of the model. The value of the fitness function was calculated in every iteration. If the PSO algorithm doesn't reach the termination condition (usually the maximum iteration time), the velocity and position of the particle will then be modified.

In the setting of the PSO parameters, the various parameters are considered. These include C_1 and C_2 represent the self-cognitive and social-cognitive acceleration coefficient respectively, r_1 and r_2 which also represent the random variables in the range [0,1], and w represents the inertia weight. When completing the (t+1) th iteration, the best personal position of the particle and the global best position of the swarm will be updated according to the value of the fitness function.

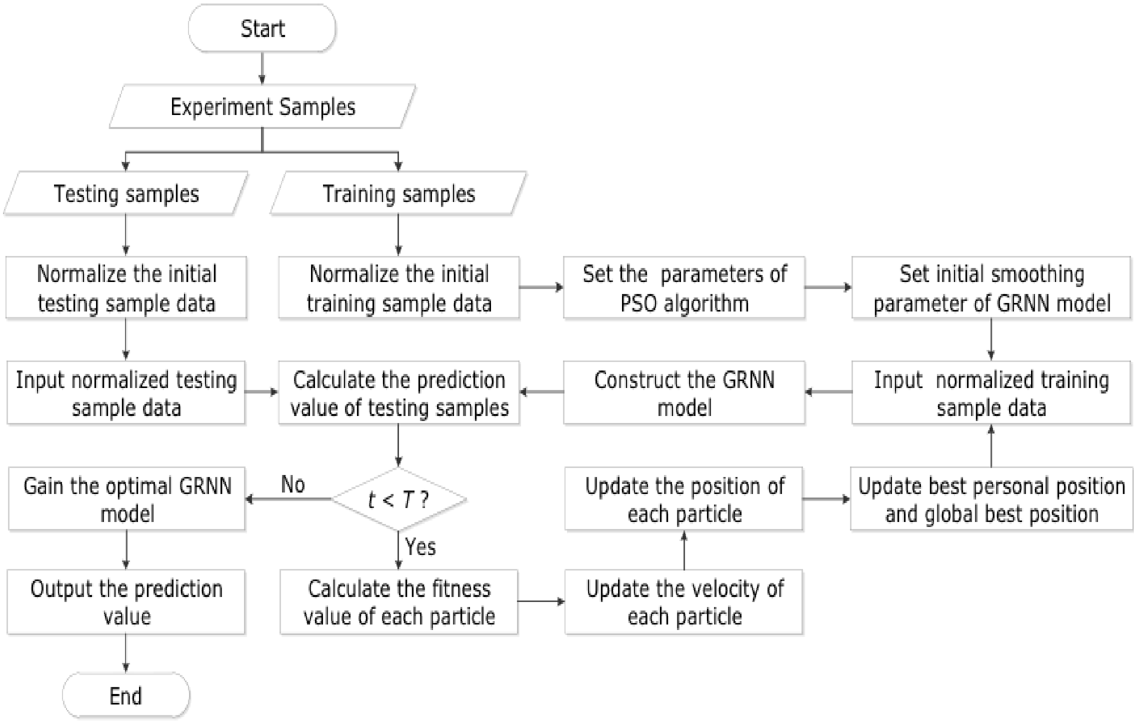

Figure 3 A Flow Chart of a Hybrid PSO-GRNN

3.2.4. Multiple Linear Regression

MLR is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. The multiple linear regression was used to model the linear relationship between the independent variables (modulus, pressure, temperature and rainfall) and the response or dependent variables (piezometric water level). The independent variables were used to calculate the dependent variable. In training of the model, a data span of 1 year 6 months which covers about 66% of the entire dataset was used for the training to predict the piezometric water levels. In testing the model, 34% of the entire data observation was used after the training to test how the model performs. The input variables (independent) were passed through the designed and trained network to predict the output variable (dependent) which is the Piezometric water level.

4. MODEL PERFORMANCE ASSESSMENT

Model evaluation is a vital part of the model development process. It helps to find the best model that represents our data and how well the chosen model will work in the future. Evaluating model performance is key to assessing the performance of the model. In assessing the performance of the models in this study, the data set were divided into training and testing data to perform a comparative analysis on the predicted Piezometric water levels and the observed Piezometric water levels. The distinctive differences between the observed values and the predicted Piezometric water level were calculated using the expression in Equation (1)

Error = (O-P_i) (1)

Where, O = the observed Piezometric water level
P = the predicted Piezometric water level from method *i*.

In this study, *i* can be any of the four methods used in the model (*i* = PSO-BPNN, PSO-RBFNN, and PSO-GRNN).

Moreover, the correlation coefficient (R), the root mean square error (RMSE), and the mean absolute error (MAE) were statistically used to evaluate the performance of the model.

The root mean square error (RMSE) is a measure of the variation of predicted Piezometric water levels to observed piezometric water levels. RMSE is estimated as the square root of the average of the squared residuals. This was achieved through the expression in Equation (2).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - y')^2} \tag{2}$$

Where, y_i = the observed piezometric water levels
y' = the piezometric water levels from the model's prediction

The RMSE indicates the absolute fit of the model to the data on how close the observed Piezometric water levels are to the model's predicted values. However, lower values of RMSE indicate better performance of the model.

The mean absolute error (MAE) is also another evaluation metric used in the assessment of the performance of the model. Mean absolute error of the model refers to the mean of the absolute values of each prediction error on all instances of the training and test data (Willmott and Matsuura, 2005). This was applied to the model to obtain the absolute prediction errors from all the four methods used using the expression in the Equation (3).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - y'| \tag{3}$$

Where, y_i = the observed piezometric water levels
$\grave{y}$ = the predicted piezometric water levels from the model.

4.3

The correlation coefficient (R) is the statistical measure which indicates the strength of the relationship between two variables. In the advent of this study, it was applied to calculate the relationship between the observed Piezometric water levels and the predicted Piezometric water levels (Meselhe and Rodrigue, 2013). This was done through the relationship in the Equation (4).

(4)

$$R = \frac{\sum_{i=1}^{N_0}(y_i - y')(m_i - m')}{\sqrt{\sum_{i=1}^{N_0}(y_i - y')^2 \sum_{i=1}^{N_0}(m_i - m')^2}}$$

Where, y_i and m_i = the prediction output and the measured value from the *i*th element
y' and m' = their average values, respectively
N_o = the number of observations

5. RESULTS AND DISCUSSION

The model was assessed by first estimating the results as predicted by the various methods of the observed piezometric water levels. The distinctive differences between the observed and the predicted water level values were calculated. Since evaluating predictive performance is the key to assessing the quality of the model and efficiency of the optimization. This study adopted various metrics to evaluate the performance of the models. The various metrics used include the correlation coefficient (R), the root mean square error (RMSE), and the mean absolute error (MAE) which were statistically used to evaluate the performance of the models.

The correlation coefficient was used to calculate the relationship between the observed and predicted piezometric water levels. The values of the correlation coefficient calculated in all the predictions performed by the various hybrid methods used were below 1. However, values for the correlation coefficient ranges from -1 to +1. Where negative values show a negative correlation between the two variables involved, a correlation coefficient value of 0 shows no correlation and positive values show a positive correlation. Regarding the combination of the optimization model, which is the particle swarm optimization with the various artificial neural network models, the performance of the model improved and there was a strong positive correlation between the predicted values and the observed values.

Another metric used for the evaluation of the models is the root mean square error (RMSE). The root mean square error is a standard way to measure the error of a model in predicting quantitative data. Root mean square error is a measure of the variation of predicted piezometric water level to the observed piezometric water level. It is estimated as the square root of the average of the squared residuals. As shown in Tables 1, 2, and 3, when the particle swarm optimization was combined with the various artificial neural network models, the results improved. The root mean square of the prediction performed on all the piezometers using the hybrid particle swarm optimization with the various neural network models were very minimal. When the particle swarm optimization was hybridized with the neural networks, the values of the root mean square error improved thereby improving the prediction accuracy as shown in Tables 1, 2 and 3.

Table 1 Performance Assessment of the RBFNN

RBFNN Piezometers	RMSE (m)	R	MAE (m)

		Training	Testing	Training	Testing	Training	Testing
BLOCK 12	PO1_12	3.18E-04	0.7601	1	1	1.87E-03	0.3682
	PO2_12	4.54E-06	0.0773	1	1	3.29E-05	0.0221
	PO3_12	4.80E-04	0.2454	1	1	3.70E-04	0.1310
	PO4_12	2.02E-05	0.1626	1	1	1.43E-05	0.0414
BLOCK 17	PO1_17	1.95E-06	0.7210	1	1	1.43E-06	0.2039
	PO2_17	1.62E-04	0.4948	1	1	1.13E-05	0.1266
	PO3_17	2.74E-05	0.9048	1	1	1.10E-06	0.2111
	PO4_17	5.95E-06	0.0786	1	1	1.36E-06	0.0198
BLOCK 24	PO1_24	3.43E-06	0.9216	1	1	2.48E-07	0.9227
	PO2_24	1.87E-06	0.8083	1	1	8.79E-03	0.2140
	PO3_24	1.13E-07	0.0671	1	1	8.08E-06	0.0340
	PO4_24	1.46E-05	0.6748	1	1	1.18E-07	0.4123

Table 2 Performance Assessment of PSO-RBFNN

PSO-RBFNN Piezometers		**RMSE (m)**		**R**		**MAE (m)**	
		Training	**Testing**	**Training**	**Testing**	**Training**	**Testing**
BLOCK 12	PO1_12	6.47E-05	0.0001	0.9	0.9	4.83E-05	0.0873
	PO2_12	6.35E-07	0.0013	0.9	0.9	4.81E-06	0.0012
	PO3_12	4.91E-05	0.0056	0.9	0.9	3.82E-06	0.0048
	PO4_12	2.11E-06	0.0084	0.9	0.9	1.52E-06	0.0036
BLOCK 17	PO1_17	2.34E-07	0.0417	0.9	0.9	2.64E-07	0.0015
	PO2_17	4.82E-06	0.0650	0.9	0.9	3.52E-07	0.0195
	PO3_17	8.22E-07	0.0321	0.9	0.9	2.72E-07	0.0025
	PO4_17	2.17E-07	0.0038	0.9	0.9	1.58E-08	0.0029
BLOCK 24	PO1_24	1.24E-08	0.0094	0.9	0.9	1.24E-08	0.0087
	PO2_24	8.25E-04	0.0073	0.9	0.9	8.25E-04	0.0030
	PO3_24	1.79E-08	0.0013	0.9	0.9	1.79E-08	0.0035
	PO4_24	3.24E-08	0.0528	0.9	0.9	3.24E-08	0.0197

Table 3 Performance Assessment of GRNN

GRNN Piezometers		**RMSE (m)**		**R**		**MAE (m)**	
		Training	**Testing**	**Training**	**Testing**	**Training**	**Testing**
BLOCK	PO1_12	0.8421	0.9732	1	1	0.7065	0.9621

12	PO2_12	0.0677	0.0950	1	1	0.0513	0.0819
	PO3_12	0.0936	0.2182	1	1	0.0712	0.1826
	PO4_12	0.1569	0.3144	1	1	0.1143	0.2573
BLOCK 17	PO1_17	0.0124	0.4532	1	1	0.6134	0.8231
	PO2_17	0.3425	0.4031	1	1	0.2460	0.3068
	PO3_17	0.5423	0.8242	1	1	0.6451	0.8023
	PO4_17	0.1529	0.1304	1	1	0.1179	0.0995
BLOCK 24	PO1_24	0.3452	0.6521	1	1	0.3471	0.5243
	PO2_24	0.6246	0.9273	1	1	0.5290	0.8293
	PO3_24	0.1719	0.5094	1	1	0.1477	0.4056
	PO4_24	0.1127	0.1333	1	1	0.0920	0.1099

Comparing the neural networks and hybrid models, the values of the root mean square for the hybrid model were far better than the single neural networks and the multiple linear regression which indicate that the hybrid models performed well than the neural networks and the multiple linear regression as shown in Tables 1, 2 and 3. The lower values of the root mean square error of the hybrid models which are approximately approaching zero in Tables 1, 2 and 3 show how well the models performed and correspond to a better fit between the predicted and observed water levels hence the lower values of the root mean square of the hybrid model which are approximately approaching zero indicates better performance of the model.

Mean square error was also applied to obtain the absolute prediction errors from all the hybrid methods used. Mean square error measures the average magnitude of the errors in a set of predictions, without considering their direction. Analyzing the calculated mean absolute errors, it was seen from all the hybrid methods to have very lower values which are approximately approaching zero than the various neural networks. So, the combination of the particle swarm optimization with the neural networks method improved the results. The values of the mean absolute error which are approximately approaching zero as shown in the Tables 1, 2 and 3 give an indication of the high performance of the hybrid model for the prediction of the piezometric water levels in the dam.

The lower values of the root mean square error (RMSE), the mean absolute error (MAE) and the correlation coefficient (R) of the hybrid models in the Tables 2, 4 and 6 indicates that the hybrid models performed better than the single neural network models and the classical method (multiple linear regression) in the Tables 1, 3, 5 and 7 respectively. Therefore, hybridizing the particle swarm optimization (PSO) with the various neural network models improved the prediction accuracy to predicting piezometric water level in the Bui dam and overcomes their shortcomings and achieves better prediction and optimization results. Therefore, the hybrid machine learning models has a better prediction accuracy and performs much better than the classical method (multiple linear regression). The hybrid machine learning model has a better prediction accuracy to predicting piezometric water level.

Table 4 Performance Assessment of PSO-GRNN

PSO-GRNN Piezometers		RMSE (m)		R		MAE (m)	
		Training	Testing	Training	Testing	Training	Testing
BLOCK 12	PO1_12	0.0552	3.13E-04	0.8	0.9	0.0046	0.0023
	PO2_12	2.57E-05	4.20E-04	0.9	0.8	2.08E-05	3.59E-04
	PO3_12	1.89E-04	0.0635	0.9	0.9	4.97E-04	0.0430
	PO4_12	3.91E-04	0.0025	0.9	0.8	2.12E-07	3.86E-05
BLOCK 17	PO1_17	5.31E-04	0.0018	0.8	0.9	3.91E-06	4.12E-04
	PO2_17	0.0004	0.0028	0.8	0.9	3.21E-04	0.0024
	PO3_17	0.0003	0.0242	0.9	0.9	0.0006	0.0171
	PO4_17	0.0046	0.0654	0.9	0.9	5.24E-04	0.0002
BLOCK 24	PO1_24	0.0003	0.0321	0.9	0.9	1.22E-04	0.0042
	PO2_24	1.24E-05	0.0216	0.9	0.9	0.0007	0.0463
	PO3_24	2.19E-04	0.0631	0.9	0.9	6.11E-04	0.0434
	PO4_24	0.0006	0.0351	0.9	0.8	0.0002	0.0258

Table 5 Performance Assessment of BPNN

BPNN Piezometers		RMSE (m)		R		MAE (m)	
		Training	Testing	Training	Testing	Training	Testing
BLOCK 12	PO1_12	1.99E-04	3.95E-04	1	1	1.59E-04	3.19E-04
	PO2_12	0.0021	0.0026	1	1	0.0014	0.0019
	PO3_12	3.85E-04	0.0065	1	1	2.87E-04	0.0037
	PO4_12	2.49E-04	1.91E-04	1	1	1.98E-05	1.24E-04
BLOCK 17	PO1_17	3.88E-05	1.93E-04	1	1	4.86E-05	1.19E-04
	PO2_17	3.36E-04	3.93E-05	1	1	2.53E-05	3.02E-05
	PO3_17	6.77E-05	3.97E-04	1	1	4.31E-05	1.03E-04
	PO4_17	1.93E-05	1.57E-05	1	1	1.18E-05	1.19E-05
BLOCK 24	PO1_24	1.99E-05	1.01E-04	1	1	1.59E-05	6.34E-05
	PO2_24	1.15E-04	4.77E-04	1	1	6.07E-05	3.48E-04
	PO3_24	2.68E-04	0.0012	1	1	1.94E-05	7.63E-04
	PO4_24	1.86E-05	7.35E-05	1	1	1.53E-05	6.38E-05

Table 6 Performance Assessment of PSO-BPNN

PSO-BPNN Piezometers		RMSE (m)		R		MAE (m)	
		Training	Testing	Training	Testing	Training	Testing
BLOCK 12	PO1_12	1.46E-06	2.16E-06	0.9	0.9	6.75E-08	3.47E-08
	PO2_12	1.04E-05	2.20E-04	0.9	0.8	1.22E-04	2.39E-04
	PO3_12	1.77E-06	4.81E-04	0.9	0.9	1.95E-06	5.22E-04
	PO4_12	2.37E-06	4.46E-06	0.8	0.9	2.09E-06	4.58E-06
BLOCK 17	PO1_17	1.74E-06	2.56E-05	0.9	0.9	4.33E-06	2.38E-05
	PO2_17	4.23E-08	3.33E-07	0.9	0.9	9.45E-07	6.34E-07
	PO3_17	6.78E-08	1.27E-06	0.8	0.9	4.44E-07	5.16E-06
	PO4_17	2.71E-08	7.89E-07	0.9	0.8	3.56E-08	5.66E-07
BLOCK 24	PO1_24	5.12E-07	1.23E-05	0.9	0.9	8.95E-06	9.34E-06
	PO2_24	5.65E-06	4.66E-05	0.9	0.9	2.12E-06	1.15E-05
	PO3_24	1.87E-06	8.91E-04	0.9	0.8	4.67E-06	5.67E-05
	PO4_24	1.39E-07	6.88E-07	0.9	0.9	4.45E-06	6.77E-07

Table 7 Performance Assessment of MLR

MLR Piezometers		RMSE (m)	R	MAE (m)
		Testing	Testing	Testing
BLOCK 12	PO1_12	2.8898135	0.20343	2.582305
	PO2_12	0.1556161	-0.0943	0.135728
	PO3_12	0.3959383	0.59862	0.290842
	PO4_12	0.4136606	0.49088	0.310539
BLOCK 17	PO1_17	3.6164465	0.22154	2.931165
	PO2_17	0.7415096	0.04619	0.522044
	PO3_17	11.522543	0.55991	5.857131
	PO4_17	0.3261856	0.06569	0.275713
BLOCK	PO1_24	5.278815	-0.1789	4.471037

24	PO2_24	2.5290819	0.79957	2.455271
	PO3_24	2.8898135	0.20343	2.582305
	PO4_24	0.1556161	-0.0943	0.135728

Comparing the multiple linear regression and the hybrid machine learning model, the lower values of the root mean square error and the mean absolute errors which are approaching zero as shown in the Tables 2, 4 and 6 indicates that the hybrid machine learning models performs better than the multiple linear regression as shown in Table 7. The correlation coefficient of the hybrid machine model as shown in Tables 2, 4 and 6 has a strong positive correlation than the multiple linear regression in Table 7.

Therefore, the hybrid machine learning model has a good prediction accuracy and performs better as shown in the Tables 2, 4 and 6 than the multiple linear regression in Table 7. The lower values of the root mean square error of the hybrid models which are approaching zero indicates a better performance of the hybrid model which correspond to a better fit between the predicted and observed piezometric water level than the multiple linear regression. The lower values of the mean absolute error of the hybrid models also indicates a better performance of the hybrid models than the multiple linear regression.

Figure 4 shows the root mean errors of the hybrid models, artificial neural network models and the classical method (multiple linear regression). The lower values of the root mean error of the hybrid models which are approximately approaching zero indicates a better performance of the hybrid models than the multiple linear regression which correspond to a better fit between the predicted and observed piezometric water level.

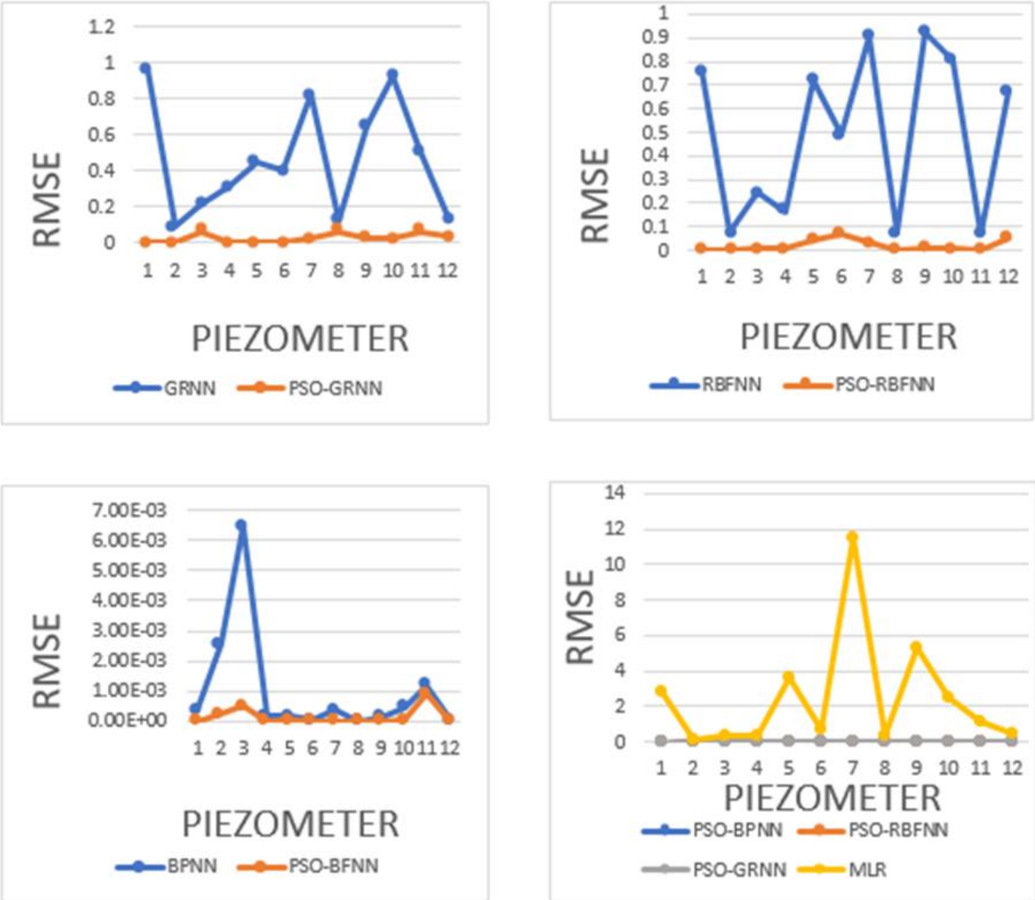

Figure 4 Plots of Root Mean Square Error of the Hybrid Models, ANN and MLR.

Figure 5 shows the mean absolute errors of the hybrid models, artificial neural network models and the classical method (multiple linear regression). The values of the mean absolute errors of the hybrid model which are approaching zero indicates a better performance of the hybrid models than the multiple linear regression which gives an indication of a high performance of the hybrid model for the prediction of piezometric water level in the dam.

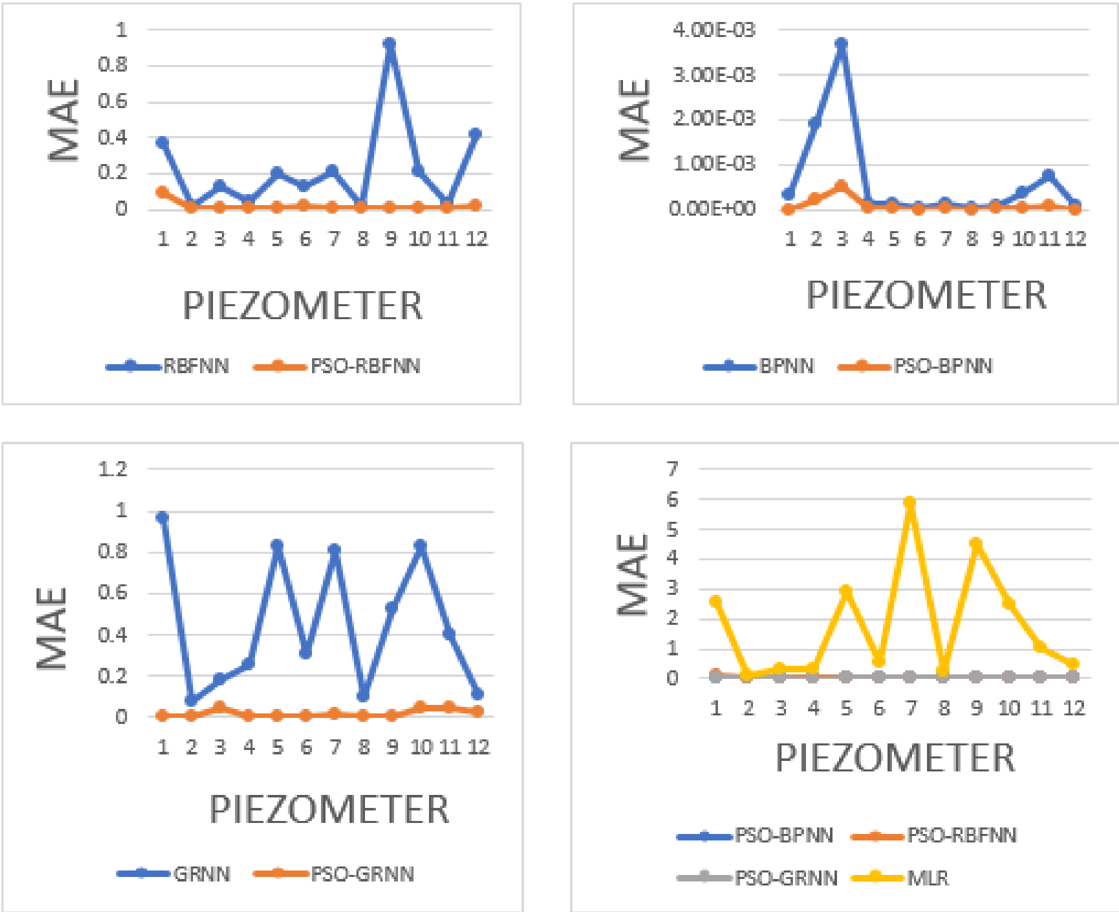

Figure 5 Plots of Mean Absolute Error of the Hybrid Models, ANN and MLR

6. CONCLUSIONS

This study developed and evaluated the performance of three hybrid machine learning models of PSO-BPNN, PSO-RBFNN and PSO-GRNN for predicting dam piezometric water level. These hybrid models were compared with their respective standalone methods (BPNN, RBFNN and GRNN) and the conventional MLR model. The statistical performance analysis used for the evaluations include RMSE, R and MAPE. From the results of the study, it is concluded that:

i. All the hybrid machine learning models (PSO-BPNN, PSO-RBFNN and PSO-GRNN) developed to predicting piezometric water level in dam deformation have shown that they are good potential in producing good prediction accuracy.
ii. The hybrid models enhanced their respective standalone models (BPNN, RBFNN and GRNN) and was better than the conventional MLR approach.

REFERENCES

Alhassan, E. H. (2013), “Hydro-Biology and Fish Production of the Black Volta near the Bui Dam during the Pre and Post Impoundment Periods”, *PhD thesis*, University of Ghana, 190pp.

Alhassan, H. S. (2009), “Viewpoint-butteries vs Hydropower: Reflections on large dams In Contemporary Africa”, *Water Alternatives*, Vol. 2, No.1, pp. 148-160.

Anon. (2002), “Structural deformation surveying”, *Engineer Manual*, 237pp.

Anon. (2020a), “Bui Dam”, *https://en.wikipedia.org/wiki/Bui_Dam*. Accessed: March 20, 2020.

Anon. (2020b), “Dam”, *https://en.wikipedia.org/wiki/Dam*. Accessed: February 17, 2020.

Anon. (2020c), “Support Vector Machine”, *https://en.wikipedia.org/wiki/Support_Vector Machine.* Accessed: February 10, 2020.

Anon. (2020d), “Neural Networks”, *https://en.wikipedia.org/wiki/Neural Networks.* Accessed: February 20, 2020.

Anon. (2020e), “Machine Learning”, *https://en.wikipedia.org/wiki/Machine_Learning.* Accessed: January 25, 2020.

Bonelli S. and Royet P. (2001), “Delayed Response Analysis of Dam Monitoring Data” *Proceedings of the ICOLD European symposium on Dams in a European Context, GEIRANGER*: Nor-ve`ge, pp. 91–99.

Buscema, M. (1998), “Back Propagation Neural Networks”. *Subst. Use Misuse 1998*, 33, pp. 233-270.

Calgary A. (2012), “Seepage Modeling With SEEP/W: GEO-SLOPE International Ltd.

Dietz AJ, Hees S, Seuren G, Veldkamp F (2014) Water dynamics in the seven African countries of Dutch policy focus: Benin, Ghana, Kenya, Mali, Mozambique, Rwanda, South Sudan. Report on Ghana: the African Studies Centre Leiden and commissioned by VIA Water, Programme on water innovation in Africa. https://aquaf orall. org/ viawa ter/ files/ asc_ water_ ghana_3. pdf (accessed on 5th May 2021)

Fell, R., Wan, C.F., Cyganiewicz, J. and Foster, M. (2003), “Time for Development of Internal Erosion and Piping in Embankment dams”, *Jr. Geotech. Geoenviron. Eng.*, pp. 307-314.

Gourine, B and Khelifa, S. (2018), “Analysis of Dam Deformation Using Artificial Neural Networks Methods and Singular Spectrum Analysis”, *Advances in Science, Technology and Innovation*, pp. 817-874.

Hassan, R., Cohanim, B., De Weck, O. and Venter, G. (2005), "A Comparison of Particle Swarm Optimization and the genetic algorithm", *In Proceedings of the 46th AIAA/ASME/ASCE/AHS/ASC structures, structural dynam*ics *and materials conference, Austin,* TX, USA, 1897pp.

Khan, M.S and Coulibaly, P. (2006), "Application of Support Vector Machine in Lake Water Level Prediction", *Journal of Hydrologic Engineering,* pp. 119-205.

Lach, S. (2018), "An analysis of the dynamics of changes to water levels in the open piezometers of the Pieczyska dam in the study period between January 2016 and April 2017", E3S *Web of Conferences,* Vol. 45, p. 00044.

Li, G.C. and Desai, C.S. (1983), "Stress and Seepage Analysis of Earthen Dams", *Journal of Geotechnical Engineering,* ASCE 109, pp. 946-960.

Meselhe, E. A. and Rodrigue, M. D. (2013), "Models Performance Assessment Metrics and Uncertainty Analysis", *Baton Rouge,* LA: *Louisiana Coastal Area Program,* 27pp.

Nengcheng, C., Xiong, C., Du, W., Wang., Lin, X. and Chen, Z. (2019), "An Improved Genetic Algorithm Coupling a Back Propagation Neural Network Model for Water-Level Predictions", *Innovation Center of Geospatial Technology,* Wuhan 430079, China.

Nwobi-Okoye, C.C and Igboanugo, A.C. (2013), "Predicting Water Level at Kainji Dam Using Artificial Neural Networks", *International Journal*.

Ranković, V., Novaković, A., Grujović, N., Divac, D. and Milivojević, N. (2014),

"Predicting piezometric water level in dams via artificial neural networks", *Neural* Computing & Applications, Vol. 24, No. 5, pp. 1115-1121.

Raschid-Sally, L., Twum-Korangteng, R. and Akoto-Danso, E. K. (2008), "Research, development and capacity building for the sustainability of dam development with special reference to the bui dam project", Bui Power Project, Ghana, 47pp.

Sivanandam, S and Deepa, S. (2008), "Genetic algorithm optimization problems", *In Introduction to Genetic Algorithms;* Springer: Berlin/Heidelberg, Germany, pp. 165-209.

Tayfur, G., Swiatek D., Wita A. and Singh, V.P. (2005), Case study: "finite element method and artificial neural network models for flow through Jeziorsko Earthfill Dam in Poland", J Hydraul Eng, pp. 431-440.

Tsuda, I. (1992), "Dynamic link of memory-chaotic memory map in nonequilibrium neural networks. *Neural Networks,* pp. 313-326.

Willmott, C. J. and Matsuura, K. (2005), "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance", *Climate research,* Vol. 30, No.1, pp. 79-82.

Zhang, J.R., Zhang, J., Lok, T.M, and Lyu, M.R. (2007), "A hybrid particle swarm optimization-back-propagation algorithm for feedforward neural network training. *Applied Mathematics and Computation*, pp. 1026-1037.

BIOGRAPHICAL NOTES

Prof. Prosper Basommi Laari is an associate Professor of Geospatial and Environmental Science at the SD-Dombo University for Business and Integrated Development Studies. He has over thirty-four refereed publications hinging on spatial analysis, land use, environment, and geomatics. Prof. Laari is a member of both the Ghana Institution of Surveyors and the Licensed Surveyors of Ghana. He is also a member of the Local Organizing Committee of FIG 2024 working week in Ghana. He has worked as a lead consultant on the World Bank-funded Ghana National Household Registry in the Ministry of Gender, Children and Social Protection and now serves as a consultant to the Ghana Productive Safety Net Project 2. He is the chairman of the Lands Commission in the North East region of Ghana and a member of the National Lands Commission.

CONTACTS

Prof. Prosper Basommi Laari
Department of Environment and Resource Studies,
SDD-University for Business and Development Studies,
P.O. Box 520,
Wa
GHANA
Tel. +233 244853165

Email: pbasomini@uds.edu.gh