Received: 6 March 2023 | Revised: 11 May 2023 | Accepted: 26 May 2023 | Published online: 29 September 2023

RESEARCH ARTICLE

Artificial Intelligence and Applications
2024, Vol. 2(2) 141–154
DOI: 10.47852/bonviewAIA3202832

BON VIEW PUBLISHING

Comparative Study of Suspended Sediment Load Prediction Models Based on Artificial Intelligence Methods

Cynthia Borkai Boye[1], Paul Boye[2,*] and Yao Yevenyo Ziggah[1]

[1]*Department of Geomatic Engineering, University of Mines and Technology, Ghana*

[2]*Department of Mathematical Sciences, University of Mines and Technology, Ghana*

Abstract: Quantification of suspended load sediment is crucial for maintaining the ecosystem and quality of water/river bodies that serve as the habitat for many living organisms. Because the influencing factors are nonlinearly related to the suspended load sediment, it is a challenge to apply linear statistical models to predict accurately. To address such a problem, this study applied artificial intelligence (AI) methods to simulate and predict suspended load sediment. The AI methods are robust and can handle adequately issues related to nonlinearity in modelling. In the present study, four AI methods were developed to predict suspended sediment load (SSL) distribution. The methods include a backpropagation neural network, group method of data handling, least squares support vector machine, and generalised regression neural network (GRNN). In developing the respective models, drainage areas, river slopes, and length of rivers served as predictor variables while SSL was the response variable. The models were evaluated using the metrics of root mean square error (RMSE), percentage RMSE, uncertainty at 95%, RMSE observations standard deviation ratio, and Legates and McCabe index. According to the results, the GRNN model achieved higher prediction accuracy than the other competing methods. The performance of the GRNN model can be attributed to its ability to calibrate and generalise appropriately to the training and testing data set. Hence, in practice, the GRNN model is proposed for SSL prediction for the study area which can be useful to policymakers and managers of water resources.

Keywords: generalised regression neural network, suspended sediment load, water quality

1. Introduction

Globally, river quality and its management have become increasingly complex. With higher demands for clean water for domestic and industrial use, accurate prediction of suspended sediment load (SSL) could help to improve river quality, water monitoring design, river transportation research and management, water-related decision-making, and regulatory formulation.

Increasing soil erosion as a result of the indiscriminate clearing of vegetation around river catchments together with higher stormwater runoff has impacted greatly SSL due to climate change. Consequently, the majority of sediment load is carried in suspension or dissolved in solution, with the minority moved near the bed level (Alexandrov et al., 2009; Schenk & Bragg, 2014). Although sediment plays an important role in the physical and aquatic environments, it can carry microorganisms, pollutants, and nutrients along the stream, and when in excess may pose challenges to water resource management industries, especially for low-income countries. However, the processes involved in causing these changes and challenges are presently poorly understood. As a result, the prediction of SSL has become a great challenge, especially in developing countries in Africa where a limited number of monitoring stations are available (Chapman et al., 2016).

It is therefore a challenge to understand with certainty the spatial variability of the actual quantity of sediment load produced in Africa (Walling, 1977; Vanmaercke et al., 2014). In addition, the lack of continent-wide compilation of sediment load data has also resulted in little understanding of the spatial variability of sediment load in Africa (Milliman & Farnsworth, 2013; Walling & Webb, 1985; Walling & Webb, 1996). Moreover, a gathering of regional or country-wide sediment load data could serve as a substitute to supplement the existing small amount of data (Dunne, 1979).

To accurately predict SSL, various conventional methods (e.g. multiple linear regression and principal component regression) have been used in literature, but due to the nonlinear nature of SSL data coupled with model development computational complexities, model prediction accuracy has been poor (Alp & Cigizoglu, 2007; Lafdani et al., 2013). For decades, artificial intelligence (AI) methodologies have been utilised in the prediction of SSL as they have successfully shown their capability to find optimal mappings between input and output variables to give good accuracy in many prediction problems (Babanezhad et al., 2021; Lee et al., 2020; Nourani & Andalib, 2015).

For example, Melesse et al. (2011) used an AI modelling approach with an error backpropagation algorithm to develop models to predict the SSL of river systems. Results showed that backpropagation neural

*Corresponding author: Paul Boye, Department of Mathematical Sciences, University of Mines and Technology, Ghana. Email: pboye@umat.edu.gh

network (BPNN) is the best model to predict SSL as compared to the conventional methods. Rezaei et al. (2021) also developed AI models for SSL prediction. Comparative results revealed that least squares support vector machine (LSSVM) was superior to the other models in terms of prediction. Mehri et al. (2021) developed the group method of data handling (GMDH) model to predict SSL. Results showed that the GMDH model predicted the sediment load accurately. Wang et al. (2009) on the other hand developed feed-forward BP and generalised regression neural network (GRNN) models in comparison with the classical regression models to predict SSL. Statistical results showed that the GRNN model was superior to the other models in terms of prediction accuracy.

In literature, although these models (BPNN, GMDH, LSSVM, and GRNN) have been employed in the prediction of SSL, from the no-free-lunch theorem, a developed model cannot be used to solve all real-world problems. Thus, SSL prediction is a site-specific phenomenon (Mohamed & Shah, 2018). Moreover, accurate SSL prediction usually depends on some key factors such as data characteristics availability, geographical location of the catchment area, and data quality, but does not only depend on the method applied. Therefore, a developed model's prediction of a particular problem can vary from country to country (Adam et al., 2019; Wolpert, 2002). Consequently, this study aims to develop AI models to predict SSL and compare the models to determine the best one suitable for prediction in the study area. The models developed were the GMDH, LSSVM, BPNN, and GRNN. The reason is that these methods (BPNN, GMDH, LSSVM, and GRNN) have been applied successfully and evaluated in many fields, which also include SSL prediction with promising results (Cigizoglu & Alp, 2006; Hazarika et al., 2020; Kisi, 2012). Yet, the implementation of the applied AI methods (BPNN, GMDH, LSSVM, and GRNN) in the Ghanaian setup is to be thoroughly explored.

With the world gearing towards digital technology, AI has become one of the main focal technologies to achieve this. However, in most developing countries like Ghana, the consciousness of using such technologies to foster quick decision-making and increase productivity is still at the infant stage. Undoubtedly, the forthcoming AI has practically impacted substantially all aspects of our life, society, employment, and firms (Makridakis, 2017). Therefore, this study fills a research gap by way of creating awareness and communicating the significance of using AI as a computational tool to solve problems in SSL prediction. This study which is worth a scientific investigation in the Ghanaian setting can also provide reliable guidance to researchers and policymakers. Therefore, the main contributions of this study are to:

- Develop SSL AI prediction models: BPNN, GMDH, LSSVM, and GRNN and
- Compare and evaluate the developed models to determine the best model that suits SSL prediction for the study area.

Thus, the presented study has provided a comprehensive assessment of the developed AI models for improving SSL prediction accuracy.

2. Materials

Using the study area topographic map, the coastal river drainage areas were delineated and digitised in a geographical information system environment. A digital planimeter was then used to compute the coverage areas enclosed by the catchments of the rivers for validation purposes. Relying on the river flow existing data of the main rivers Butre, Ankobra, and Pra in the study area (southwestern Ghana), the discharge values were estimated for each river (Boye et al., 2019). Applying the logical method (Booth et al., 2002; Kim et al., 2003; Thompson, 2006), peak discharge was estimated from runoff rates of rainfall at maximum stormwater (Packman & Kidd, 1980). The logical method was employed due to its easiness in estimating the discharge values for small drainage basins, wide usage for computing peak discharge, and its efficiency in working with limited rainfall and drainage data (Cleveland et al., 2011; Montalto et al., 2007). Equation (1) shows the discharge of the peak flood.

$$Q_w = 277 \times 10^{-3} IAC \tag{1}$$

Where Q_w is the instantaneous water discharge (m^3/s), I is the rainfall intensity (mm/h), A is the catchment area (km^2), and C (constant) is the region's runoff coefficient. By considering the duration of a whole storm in the area of study, a calculation of volumes of runoff was made. The independent variables used to develop the model were obtained from a topographic map of the study area. Thus, the independent variables comprise catchment areas, length of longest rivers, and river slopes. From Equation (2), the coefficient of sediment rating (α) and index (m) were deduced from the analyses of data collected from 21 sediment loads monitored at various stations (Akrasi, 2011). These constants were substituted into the sediment rating regression relation (Equation (2)) (Nittrouer & Viparelli, 2014; Walling, 1977) to estimate the SSL (Q) in the catchment area. The SSL served as the dependent variable to develop the model.

$$Q = \alpha Q_w^m \tag{2}$$

2.1. Study area

The coast of the Western Region of Ghana was used for this study (Figure 1). The study area is located from longitudes 3°07′ to 1°40′ West and latitudes 4°40′ to 5°10′ North. It covers a land area of 23,921 km^2 constituting 10% of the size of the country (Boye, 2015).

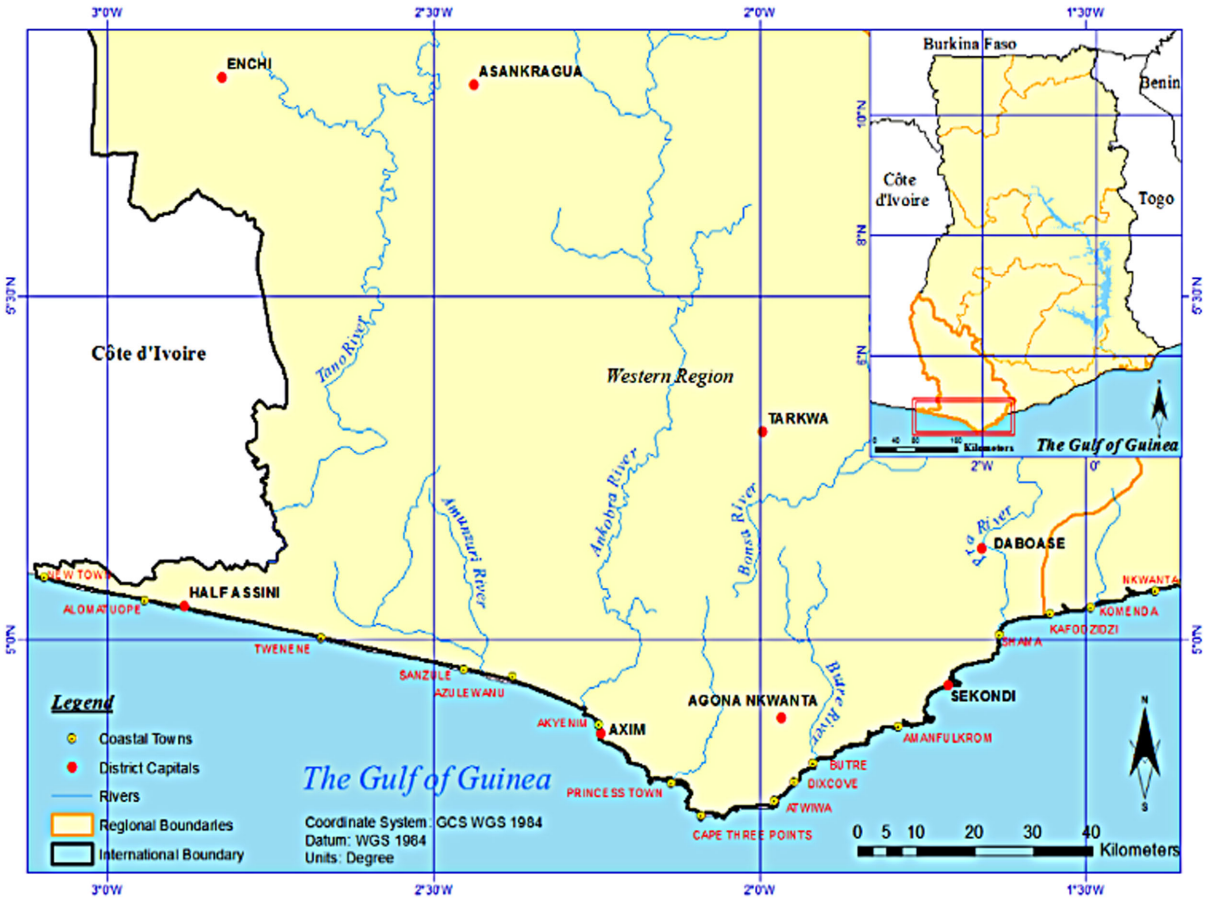
The region's coastline measures about 192 km out of the total of 540 km length of Ghana's coastline. It is characterised by a broad continental shelf with a maximum width of about 80 km around Cape Three Points (Boye et al., 2018). The eastern part of the area is bounded by rocky coast while the western section contains uninterrupted soft beaches that extend to about 100 km. The western section comprises sandy beaches which are seldomly crossed by lagoons and other wetlands. Pra, Tano, Bia, and Ankobra rivers are the four main rivers that drain through the region. Fluvial sediment from these rivers nourishes the shores, thus stabilising the beach from the sturdy west-east alongshore drift (Boateng, 2012). Most parts of the region are underlain by Pre-Cambrian rocks, that is, the Birimian and Tarkwaian series (Keates, 2021). These rocks contain precious minerals such as manganese, diamonds, and gold, which are mined in the country. In Ghana, the western region records the maximum rainfall with an average value ranging from 1250 to 2000 mm/year which nourishes the land to support the production of several crops. The western region is the largest producer of rubber, coconut, and cocoa.

3. Model Development

3.1. Backpropagation neural network

The BPNN is a proposed method of ANN by Rumelhart et al. (1986), and it has been used in many fields to solve problems. The BPNN architecture comprises input, hidden, and output layers arranged in a feed-forward manner (Figure 2). In Figure 2, consider the BPNN architecture to have m nodes in the input, g nodes in the hidden, and n nodes in the output layers. Let y_k and $x_i$denote

Figure 1
Ghana western coast



$$net_j = \sum_{i=1}^{m} \left(w_{ij}x_i + b_j\right) \text{ for } j = 1, 2 \ldots, g \tag{3}$$

the expected output and input data of units k and i, respectively. w_{ij} and b_j are the weights and thresholds from the input layer to the hidden layer, respectively, while w_{jk} and b_k are those from the hidden layer to the output layer, respectively. The learning rate and the incentive function are η and $f(x)$, respectively. The BPNN algorithm is described as follows. Firstly, the data received from the external environment by the input layer enter the network through the hidden layer nodes. The data are then multiplied by their respective weights and add them all together with a constant bias b. The computation process in the hidden layer node j is as shown in Equation (3) (Gupta et al., 2011; Zhang et al., 2021).

In this study, the Gaussian transfer function (Equation (4)) (Gundogdu et al., 2016) which was used in the hidden layer screens the added signals received from the neurons with a bias constant input value of one.

$$f(x) = \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{4}$$

Where μ and σ^2 are centre and dispersion parameters, respectively. The output results of the hidden layer unit j neurons are given as in Equation (5).

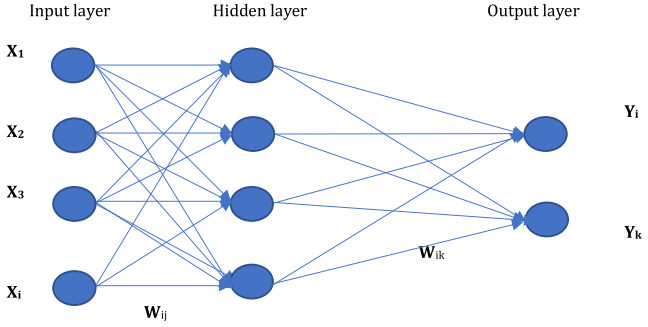
$$net_j = f\left(\sum_{i=1}^{m}\left(w_{ij}x_i + b_j\right)\right); \text{ for } j = 1, 2 \ldots, g \tag{5}$$

The output results of the output layer unit k neurons are given as in Equation (6).

$$out_k = \sum_{j=1}^{g}\left(w_{jk}net_j + b_k\right) \text{ for } k = 1, 2, \ldots, n \tag{6}$$

The mean square deviation (MSE) between the network and the expected outputs is known as the loss function (Equation (7)).

Figure 2
BPNN architecture



$$E = \frac{1}{2}\sum_{k=1}^{n}(y_k - out_k)^2 \tag{7}$$

The forecasted results from the network are compared with the true observed target value to estimate the error by using Equation (7). Errors that do not meet the minimum error threshold are reverted to the network for the connection weights and biases updation of the individual neurons using the generalised delta learning rule. The process is repeated till a minimum threshold of the error is obtained based on the loss function criterion (Toker & Kaçıranlar, 2013).

For each neuron in the output layer, Equations (8) and (9) are used to update the network weights based on the delta rule (Zurada, 1994).

$$w_{jk} = w_{jk} + \nabla w_{jk} \tag{8}$$

$$w_{ij} = w_{ij} + \nabla w_{ij} \tag{9}$$

The chain derivation rule (Equations (10)–(13)) is used to update the values of each parameter as shown as follows:

$$\nabla w_{jk} = -\eta\frac{\partial E}{\partial w_{jk}} = -\eta\left(\frac{\partial E}{\partial out_k}\right)\left(\frac{\partial out_k}{\partial w_{jk}}\right) = \eta e_k net_j \tag{10}$$

$$\nabla w_{ij} = -\eta\frac{\partial E}{\partial w_{ij}} = -\eta\left(\frac{\partial E}{\partial out_k}\right)\left(\frac{\partial out_k}{\partial w_j}\right)\left(\frac{\partial out_j}{\partial w_{ij}}\right)$$

$$= \eta\sum_{k=1}^{n} e_k w_{jk}\frac{\partial}{\partial x}\left(\sum_{i=1}^{m}(w_{ij}x_i + b_j)\right)x_i \tag{11}$$

$$w_{jk} = w_{jk} + \eta e_k net_j \tag{12}$$

$$w_{ij} = w_{ij} + \eta\sum_{k=1}^{n}(e_k w_{jk})\frac{\partial}{\partial x}\left(\sum_{i=1}^{m}(w_{ij}x_i + b_i)\right)x_i \tag{13}$$

Where $\frac{\partial}{\partial x}(.)$ is the derivative of the excitation function into the output layer which emanates from the input layer. Equations (14) and (15) are used to update the network thresholds in units k and j.

$$b_k = b_k + \nabla b_k \tag{14}$$

$$b_j = b_j + \nabla b_j \tag{15}$$

Where ∇ is the gradient vector. The chain derivative rule is used to compute the updated values of each threshold parameter as follows:

$$\nabla b_k = -\eta\frac{\partial E}{\partial b_k} = -\eta\left(\frac{\partial E}{\partial out_k}\right)\left(\frac{\partial out_k}{\partial b_k}\right) = \eta e_k \tag{16}$$

$$\nabla b_j = -\eta\frac{\partial E}{\partial b_j} = -\eta\left(\frac{\partial E}{\partial net_j}\right)\left(\frac{\partial b_k}{\partial b_j}\right)$$

$$= \eta\sum_{k=1}^{n}(e_k w_{jk})\frac{\partial}{\partial x}\left(\sum_{i=1}^{m}(w_{ij}x_i + b_j)\right) \tag{17}$$

$$b_k = b_k + \eta e_k \tag{18}$$

$$b_j = b_j + \eta\sum_{k=1}^{n}(e_k w_{jk})\frac{\partial}{\partial x}\left(\sum_{i=1}^{m}(w_{ij}x_i + b_j)\right) \tag{19}$$

Studies have shown that the scaled conjugate gradient algorithm was the preferred algorithm for training the BPNN. The reason being that the automatic algorithm does not require fine tuning of the parameters, and it is faster as well (Møller, 1993).

3.2. Group method of data handling

The GMDH is a feed-forward network formed to search for the optimum solution of a complex nonlinear problems based on the fundamentals of heuristic self-organisation systems. GMDH is an algorithm to find a linear parameter complex polynomial function. By external criteria, the algorithm performs selection of the best solution. GMDH can be described as subset components formulation of the function in Equation (20) (Nguyen et al., 2019; Stefenon et al., 2020).

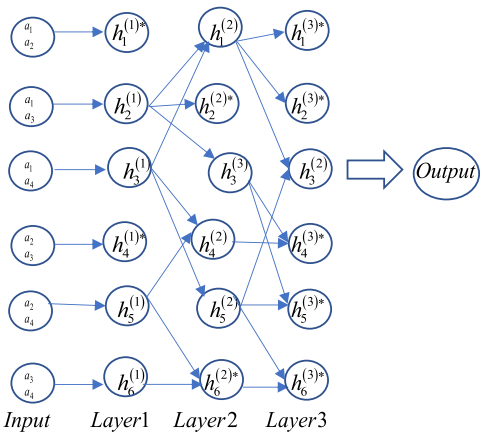
$$\hat{h}(\mathbf{A}) = b_0 + \sum_{u=1}^{m}(b_u f_u) \tag{20}$$

Where $\mathbf{A} = (a_1 a_2 \ldots a_n)$ is the input vector, $\mathbf{B} = (b_0 b_1 \ldots b_u)$ is the coefficients vector (weights), $\hat{h}$ is the each iteration prediction, f_u are the elementary functions that depend on diverse subsets as inputs, and m is the number of function components.

For the best solution to be obtained, GMDH uses many subsets of the base function (Equation (20)) components. The least squares method was used to estimate the coefficients of the models. As a regression technique, the use of the least squares was to approximate the system solution by minimising the residual sum of squares produced during the process. Gradually, the GMDH upsurges the partial components number and looks for a perfect complexity structure shown by a minimum value of an external criterion.

Prior to the development of the model, the external information was partitioned into training and testing sets according to a definite percentage. The data were trained to estimate the model coefficients, and the test set was used to check the model soundness. The neuron was assessed and verified by an external criterion and those with the worst prediction were rejected. In the next layer, reorganisation, training, testing, and selection processes were performed again until the prediction error of the neuron stops decreasing. Figure 3 shows the GMDH structure where the asterisk neurons were

Figure 3
GMDH network architecture



removed because of poor prediction (Armaghani et al., 2020; Rayegani & Onwubolu, 2014).

Consider the use of a sediment load data set. The relationships between the lags are learned by the algorithm and the path to follow is selected automatically by GMDH. In the GMDH neural network, the input and output variables mapping depict a nonlinear function as shown in Equation (21). Consider the input pair variables a_u and a_v, Equation (22) represents the regression method employed to solve for the vector of coefficients.

$$\hat{h}(\mathbf{A}) = b_0 + \sum_{u=1}^{n}(b_u a_u) + \sum_{u=1}^{n}\sum_{v=1}^{n}(b_{uv}a_u a_v) + \sum_{u=1}^{n}\sum_{v=1}^{n}\sum_{t=1}^{n}(b_{uvt}a_u a_v a_t) + .. \tag{21}$$

$$G(a_u, a_v) = b_0 + b_1 a_1 + b_2 a_v + b_3 a_u^2 + b_4 a_v^2 + b a_u a_v \tag{22}$$

Equation (23) represents the regularity external criterion, where h_u is the actual target values. The least w_t of the layer is noted, and when w_t is not reducing anymore as compared with the preceding layer's value, the network forecast error is said to be nondecreasing and the result of the preceding layer is produced.

$$w_t = \frac{\sum_{u=1}^{p}\left(\hat{h} - h_u\right)^2}{\sum_{u=1}^{p} h_u^2} \tag{23}$$

Where p is the number of testing data set. The vector of coefficients is estimated by the least squares error method as shown in Equation (24).

$$least\ square\ error = \begin{cases} \hat{h}(\mathbf{A}) = G(a_u, a_v) \\ error = \sum_{u=1}^{n}\left(h_u - \hat{h}\right)^2 \\ \frac{d}{d\hat{h}_t} error = 0, t = 1, 2, 3, 4, 5. \end{cases} \tag{24}$$

To facilitate the analysis, the computation of the vector of coefficients was done using matrix form as shown in Equation (25).

$$\hat{h} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{h} \tag{25}$$

Where Equation (26) depicts the input data set.

$$\mathbf{A} = \begin{bmatrix} 1 & a_{u1} & a_{v1} & a_{u1}a_{v1} & a_{u1}^2 & a_{v1}^2 \\ 1 & a_{u2} & a_{v2} & a_{u2}a_{v2} & a_{u2}^2 & a_{v2}^2 \\ 1 & a_{u3} & a_{v3} & a_{u3}a_{v3} & a_{u3}^2 & a_{v3}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_{un} & a_{vn} & a_{un} & a_{un}^2 & a_{vn}^2 \end{bmatrix} \tag{26}$$

3.3. Least squares support vector machine

LSSVM is an iteration of the improved standard SVM. It is one of the algorithms that gives important results of statistical learning theory. Using the least squares loss function (LSLF), LSSVM generates optimisation problem and it is based on equality constraints. The LSLF entails linear equation solution and it is considered to be simple compared to that in the ε-intensive loss function of the novel SVMs (Suykens et al., 2002). In general, LSSVM is employed for classification, regression problems, and optimal control (Kaytez et al., 2015; van Gestel et al., 2004).

This section is introducing the least squares support vector regression (LSSVR) briefly. The LSSVR method is used to approximate an incomprehensible function by relying on a training data set $[x_i, y_i]_{i=1}^{l}$. The regression is formulated as feature space representation shown in Equation (27).

$$y = f(\mathbf{x}) = w^T\phi(\mathbf{x}) + b \tag{27}$$

Where $\mathbf{x} \in R^n$ for a positive integer n, $y \in R$, w is the weight vector of the same dimension as the feature space, b is a bias, and $\phi(.) : R^n \rightarrow R^{nh}$ is a nonlinear mapping to the high dimensional feature space. The minimisation of the error together with the regularisation is given as in Equations (28) and (29).

$$\text{Minimise } J(wbe) = \frac{1}{2}w^T w + \frac{\delta}{2}\sum_{i=1}^{l} e_i^2 \tag{28}$$

$$\text{Subject to } y_i = w^T\phi(x_i) + b + e_i; i = 1, 2, \ldots, l \tag{29}$$

Where $e = (e_1 e_2 \ldots e_l) \in R^l$, δ is the regularised parameter balancing the trade-off between the margin and the error.

The Lagrangian function (L_g) of the optimisation problem for Equations (28) and (29) is given as in Equation (30) (de Kruif & de Vries, 2003; Yang et al., 2014).

$$Lg(w\ b\ e; \beta) = J(w\ b\ e) + \sum_{i=1}^{l}\beta_i(y_i - w^T\phi(x_i) - b - e_i) \tag{30}$$

Where the Lagrange multipliers which can either be negative or positive in the formulation of LSSVM are $\beta = (\beta_1\beta_2 \ldots \beta_l) \in R$. The Lagrangian optimality conditions are shown in Equations (31)–(34).

$$\frac{\partial Lg}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^{l}\beta_i\phi(x_i) \tag{31}$$

$$\frac{\partial Lg}{\partial b} = 0 \Rightarrow -\sum_{i=1}^{l}\beta_i = 0 \tag{32}$$

$$\frac{\partial Lg}{\partial e_i} = 0 \Rightarrow \beta_i = \delta e_i \tag{33}$$

$$\frac{\partial Lg}{\partial \beta_i} = 0 \Rightarrow w^T\phi(x_i) + b + e_i - y_i = 0 \text{ for } i = 1, 2, \ldots, l \tag{34}$$

The linear system (Equation (35)) represents the Lagrangian optimal conditions.

$$\begin{pmatrix} 0 & e^T \\ e & \Omega + \frac{I}{\delta} \end{pmatrix}\begin{pmatrix} b \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ Y \end{pmatrix} \tag{35}$$

Where $I \in R^{l\times l}$ is an identity matrix. $Y = (y_1\ \ y_2\ \ \ldots\ \ y_l)^T$, $\beta = (\beta_1\ \beta_2 \ldots \beta_l)^T$, $\Omega = (\Omega_{ij}) = k(x_i, x_j)k(x_i, x_j) = \langle\phi(x_i), \phi(x_j)\rangle$.

The output (Equations (36) and (37)) of the approximator is computed for new input values of x with β and b as

$$\hat{y}(x) = \langle w, \phi(x)\rangle + b \tag{36}$$

$$= \left\langle \sum_{i=1}^{l} \beta_i \phi(x_i), \phi(x) \right\rangle + b$$

$$= \sum_{i=1}^{l} \beta_i \langle \phi(x_i), \phi(x) \rangle + b$$

$$= \sum_{i=1}^{l} \beta_i K(x_i, x) + b \tag{37}$$

3.4. Generalised regression neural network

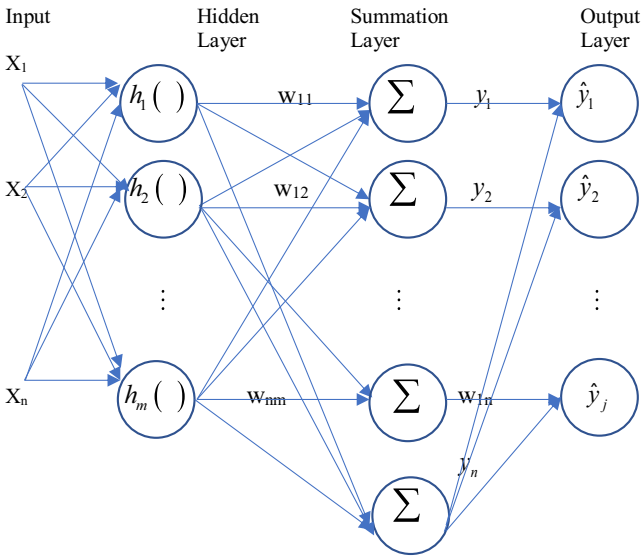
GRNN is a variant of radial basis function network which uses a kernel regression network and it is used to solve linear or nonlinear approximation problems. The network computes the most probable output that minimises the mean squared error (MSE) value. The GRNN does not require an iterative training procedure as in backpropagation method, but each layer is passed through forward computation. The method has excellent performance in learning speed and robust function approximation ability. The reason is the method provides rapid convergence to the optimum regression surface by using a probability distribution. (He et al., 2021; Kisi et al., 2006; Wang & Peng, 2018). GRNN comprises input, hidden, summation, and output layers (Figure 4). External information is passed directly to the hidden layer without weighting through the input layer. The output of the *i*th neuron in the hidden layer is computed as in Equation (38) (Hou et al., 2022).

$$h_i(\mathbf{X}) = \exp\left(-\frac{\|\mathbf{X} - c_i\|}{2r^2}\right) \text{ for } 1 \le i \le p \tag{38}$$

Where $\mathbf{X} = (x_1 \quad x_2 \quad \ldots \quad x_n)$ is the independent variable of the actual sediment load data set, r is the radius of the RBF and it determines the generalisation capability of the GRNN by controlling the degree of smoothness, c_i is the training input vector, and the number of training set is p.

The summation layer comprises simple summation, S_s, and weighted summation, S_w. The S_s computes the arithmetic sum of the hidden layer output as given in Equation (39).

Figure 4
GRNN architecture



$$S_s = \sum_{i=1}^{p} h_i(\mathbf{x}) \tag{39}$$

The *j*th weighted summation (Equation (40)) calculates the hidden layer outputs weighted sum.

$$S_{wj} = \sum_{i=1}^{p} y_i h_i(\mathbf{x}) \tag{40}$$

Where y_i is the interconnection weight in the *i*th desired response. Finally, the output of the *j*th output neuron for GRNN model (Equation (41)) is computed as a weighted average of the desired response.

$$\hat{y}_j = \frac{S_{wj}}{S_s} \tag{41}$$

3.5. Model building framework

This section presents a flowchart showing the model development, evaluation, and implementation to predict the SSL. The input and output variables were first selected and data partitioning was performed. To have homogeneity, the data were scaled into a specific interval using the normalisation process. The network was then trained and the optimum model was obtained based on a termination criterion. The optimum trained model was however tested and the respective performance metrics were computed based on the observed and predicted SSL. A detailed presentation of the computational process presented in Figure 5 is given in the subsequent sections.

3.5.1. Data specification and variables selection

In developing the various AI models, 47-sample data set was used. These data were obtained with the help of a sediment rating curve by extracting drainage areas, river slopes, and length of rivers from the topographic maps provided by the Survey and Mapping Division of Lands Commission, Ghana. The extracted features served as the independent variables while SSL served as the dependent variable. From the 47-sample data set, 33 (approximately 70%) of the training set were used to develop the AI prediction models and the remaining 14 data points (approximately 30%) served as the unseen data for the validation of the trained models. The partitioning of data was carried out using the widely used hold-out cross-validation approach.

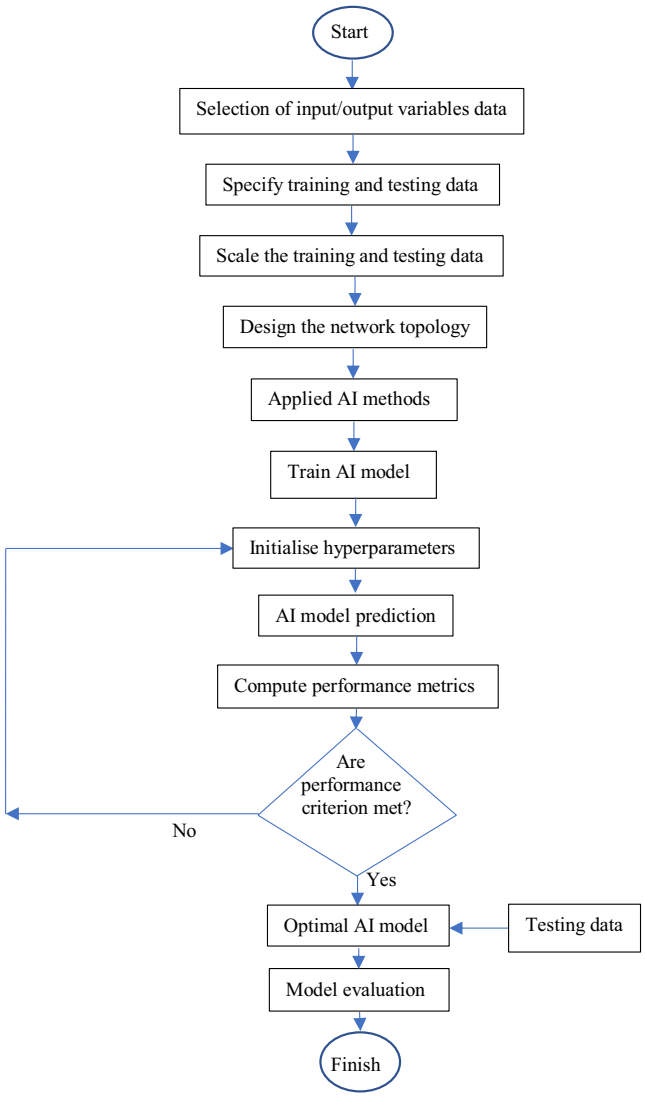
3.5.2. Normalisation of data

Data normalisation was performed to minimise the impact of larger input values on the smaller ones during the model process. This helps to put the training set into a common range which improves the convergence speed of the AI during training. Equation (42) was used to carry out the normalisation (Mueller & Hemond, 2013).

$$d_N = d_{mn} + \frac{(d_{mx} - d_{mn})(u_o - u_{mn})}{(u_{mx} - u_{mn})} \tag{42}$$

Where d_N denotes the normalised data, $u_o$is the observed SSL, u_{mn} and u_{mx} represent minimum and maximum values, respectively, of the observed SSL with d_{mx} and d_{mn} ranged from -1 to 1, respectively.

Figure 5
Flowchart of AI methods for SSL prediction



3.5.3. Network training

In modelling using ANN, data sets are trained to produce the desired output for a given input. Similarly, in this study, ANN was trained to find the functional relationship between the independent variables (drainage areas, river slopes, and length of rivers) serving as the input and the dependent variable (SSL) as the output. To perform the network training, the 47 data sets were put into a specific range and partitioned thereafter into 33 training and 14 testing subsets. The training set was used as parameterisation (adjustment of weight) to minimise the error function, while the testing set was used for model validation. In the network training, the Levenberg–Marquardt backpropagation algorithm was employed to train the BPNN. During this phase, the network was allowed to train until no additional effective enhancement happened. After training the network, the unseen testing data were used to give a general independent assessment of the performance of the BPNN. In determining the optimum BPNN model, the MSE of the model was monitored at each stage of training and testing. In addition, the correlation coefficient (R) and coefficient of determination (R^2) were used to judge the performance of the ANN model. After several trials, the model with the highest R and R^2 values with the lowest MSE was selected as the best model (Nandy et al., 2012).

The GMDH is a feed-forward multilayer network of quadratic neurons that are used to map the input–output variables' functional relationship. Thus, the key idea of the GMDH is to find a mapping $\hat{g}$ as an approximation of the actual function g for the difference between the actual output y and the predicted output $\hat{y}$ to be small as possible. The technique has a faster learning speed, converges to the optimal nonlinear or linear regression surface, and has good approximation capability. This phenomenon is achieved because the GMDH relies on optimisation technique that determines the optimal structure automatically by a layer-by-layer pruning process based on the mean square error criterion (MSE). The GMDH network automatically stops adding layers the moment the MSE of the proceeding layer exceeds the preceding layer. In this case, the network selects the lowest MSE component in the highest layer as its final model outcome (Srinivasan, 2008).

The LSSVM which is a variant of SVM adopts equality constraints. The LSSVM technique is to fit a functional model $y(x)$ on the training data sets such that the function could be used to infer the target y for a new input data point x later. The training process involves two linear systems with an identical positive definite coefficient matrix, followed by the application of the conjugate gradient method. Thus, the underlying optimisation problem follows a system of a linear equation and this improves the training efficiency for large-scale learning tasks (Liu et al., 2013; Xia, 2018).

The GRNN is based on a standard statistical technique known as kernel regression. Considering the training, the output for the input is computed in two steps. Firstly, the hidden layer produces a set of weights associated with the closeness of the input vector to the training patterns. Here, the weighted sum is one and it represents the contribution of every training pattern to the final result. Secondly, the output layer computes the output as the sum of the product of the weights and the targets. The GRNN technique approximates any arbitrary function between the input and output vectors, drawing the function estimate directly from the training data. That is, as the training set increases in number, the estimated error approaches zero with only mild restrictions on the function (Cigizoglu & Alp, 2006).

3.5.4. Model performance evaluation

This study used root mean square error (RMSE), percentage RMSE (PRMSE), uncertainty at 95%, RMSE observations standard deviation ratio (RSR), and Legates and McCabe's (E_{LM}) (Legates & McCabe, 2013; Tian et al., 2016; Willmott & Matsuura, 2005) statistical indices to determine the efficiency of the developed models for sediments load prediction.

i. Root mean square error

The RMSE (Equation (43)) is a dimensioned measure of average model metric. The metric expresses average model prediction error in the units of the variable of interest.

$$RMSE = \sqrt{\frac{\sum_{j=1}^{\tau} (O_j - P_j)^2}{\tau}} \tag{43}$$

Where τ is the test observations size, and O and P are the test observation and prediction values, respectively.

ii. Percentage root mean square error

The PRMSE (Equation (44)) which is capable of evaluating the precision of a models' predictive performance is a scale-independent measure.

$$\text{PRMSE} = \frac{RMSE}{x_m} \times 100 \tag{44}$$

iii. Uncertainty at 95%

U_{95} (Equation (45)), the uncertainty at 95% confidence level, is a statistical analysis indicator that reveals more information about the developed model prediction deviations from the actual observations.

$$U_{95} = 1.96\sqrt{S_d^2 + (RMSE)^2} \tag{45}$$

Where S_d is the standard deviation between the prediction and the observation values, 1.96 is the coverage factor corresponding to 95% confidence level, and RMSE is the root mean square error.

iv. RMSE observations standard deviation ratio

The RSR (Equation (46)) which depends largely on the RMSE varies from an optimal value of 0 to a large positive value. The smaller the RMSE value, the least RSR value becomes. Hence, the better the predictive power of the developed model.

$$RSR = \frac{RMSE}{S_d} = \frac{\sum_{i=j}^{\tau}(o_j - p_j)^2}{\sqrt{\sum_{i=j}^{\tau}(o_j - o_m)^2}} \tag{46}$$

Where S_d is the standard deviation of the test observations, τ is the test observations size, and O and P are test observation and prediction values, respectively. O_m is the mean of the observation values.

v. Legates and McCabe

The E_{LM} (Equation (47)) reveals the correlation between the predictions and the observed values. Closer the E_{LM} value to 1 is an indication of the developed model having strong predictive power.

$$E_{LM} = 1 - \frac{\sum_{i=j}^{\tau} Abs(p_j - o_j)}{\sum_{i=j}^{\tau} Abs(o_j - o_m)} \tag{47}$$

Where O and P are test observation and prediction values, respectively, O_m is the mean of the observation values, and τ is the test observations size.

Figure 6
Topographic map of the study area

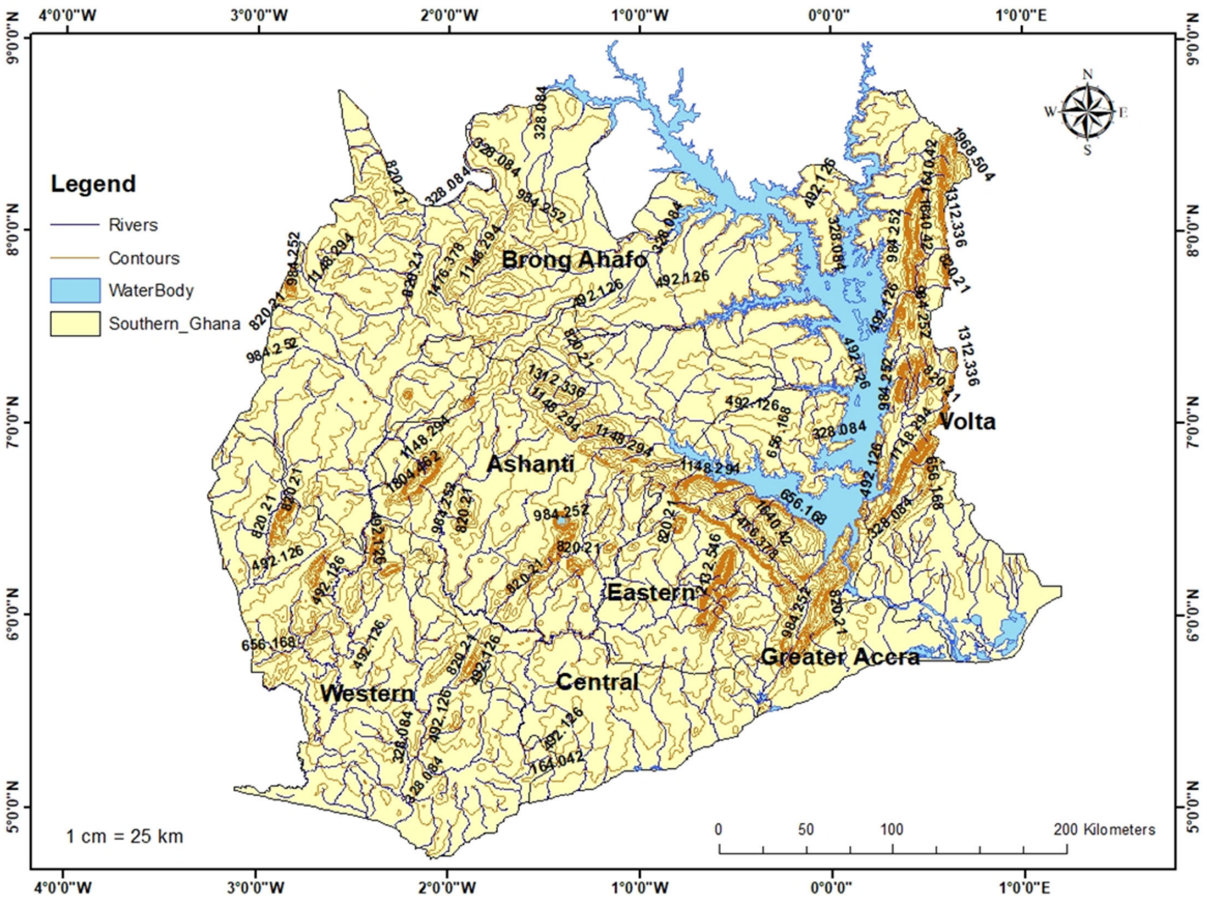


Table 1
Summary of data set

Statistics	Min	Max	Mean	SD
Drainage areas (km^2)	0.8860	873.0910	56.1632	151.8754
River slopes (m/km)	4.0105	129.3718	29.5497	26.5039
Length of rivers (m)	124.0000	3808.7000	929.4156	994.2973
Suspended sediment load (mg/L)	0.4368	12665.8140	327.5531	1885.3659

4. Numerical Application

4.1. Data used

Topographic maps from Survey and Mapping Division of Lands Commission, Ghana, West Africa and sediment rating curve were used to extract the data for the analysis (Figure 6). Drainage areas, river slopes, and length of rivers served as independent variables while SSL served as the dependent variable. Table 1 shows the statistical summary results of the data used to develop the SSL prediction models.

4.2. Optimum model developed

The BPNN optimum trained model had a model structure comprising three inputs, eight hidden neurons, and one output, that is, [3-8-1]. For the GRNN, the optimum model that produced the best performance had a smoothing parameter of 2.6 with three inputs and one output. In training the LSSVM, the best values for regularisation parameter (gamma) and the width of the kernel function (sigma) were 34177922.666 and 30.2519838598, respectively. The optimum GMDH model consisted of two layers and single neuron. In the GMDH training process, two input variables (drainage areas and river slopes) were automatically selected and seen to be very relevant in SSL prediction. This showcases the feature extraction capability of the GMDH model. The final GMDH model for SSL prediction is shown in Equation (48).

$$\begin{aligned} X_4 &= -1.73313143729 - 0.00170794657055X_2 + 1.1237213778X_1 \\ &\quad - 3.49331727402 \times 10^{-05}X_1X_2 + 1.76142505781 \times 10^{-07}{X_2}^2 \\ &\quad + 0.00102950304895{X_1}^2 \\ \text{Final GMDH Model} &= 1.02012760202 + 1.00280017581X_4 \\ &\quad - 0.154325304254X_1 - 0.00558111412937X_1X_4 \\ &\quad + 0.00102456552567{X_4}^2 + 0.00644401995933{X_1}^2 \end{aligned} \tag{48}$$

4.3. Developed model efficiency assessment

In developing the models (BPNN, GMDH, LSSVM, and GRNN), the data set acquired was partitioned into 70% training set (32 data points) and the remaining 30% as testing set (13 data points). The training set was used to fit the models and the testing set served as an independent data to authenticate the forecasting strength of the developed models. To predict the SSL, model assessment was performed for the competing methods BPNN, GMDH, LSSVM, and GRNN to find not only how the best method fits the data set well, but how it will work in future applications. This was accomplished by using the five test-case performance indices RMSE, PRMSE, U_{95}, RSR, and E_{LM} as shown in Table 2.

Table 2
Summary test results of statistical indicators

Model	RMSE	PRMSE	U_{95}	RSR	E_{LM}
BPNN	4.5127	15.7721	12.3461	0.0551	0.9485
GMDH	7.2736	25.4217	20.2807	0.0888	0.9418
LSSVM	3.1522	11.0171	8.8422	0.0385	0.9476
GRNN	2.6779	9.3594	7.1446	0.0327	0.9637

The E_{LM} is a model efficiency criterion used to determine the correlation between the prediction and the observed values. The E_{LM} ideal value is 1 and it is an indication that the model is having a strong predictive strength. From Table 2, the GRNN model had the best E_{LM} value of 0.9637. This means that 96.37% of the total variability in the SSL predicted was explained by the independent variables (drainage areas, length of rivers, river slopes) used to develop the model. In other words, the GRNN model E_{LM} value of 96.37% shows clearly that the GRNN model predictions are closely related to the observed SSL. This close association can additionally be viewed in Figures 7 and 8. Consequently, the GRNN model does not only fit the observed data very well, but it has a strong predictive power as well.

The RMSE, PRMSE, U_{95}, and RSR indices show the models' bias in the prediction of the SSL. The smaller these indices values are, the better the acceptable accuracy of the developed model. From Table 2, RMSE, PRMSE, U_{95}, and RSR values for the GRNN method are 2.6779, 9.3594, 7.1446, and 0.0327, respectively. This is an indication that the GRNN method fitted the data very well than the other competing methods based on their index's values.

From Table 2, the intercomparison among the methods employed was also perceptible. The GRNN method had the best RMSE value of 2.6779 followed by LSSVM, BPNN, and GMDH with 3.1522, 4.5127, and 7.2736 values, respectively. The performances of the methods in descending order for PRMSE are GRNN, LSSVM, BPNN, and GMDH with values 9.3594, 11.0171, 15.7721, and 25.4217 respectively. In the case of U_{95}, GRNN method had the least value of 7.1446 followed by LSSVM, BPNN, and GMDH with values 8.8422, 12.3461, and 20.2807, respectively. For RSR, the best method in a descending order is GRNN, LSSVM, BPNN, and GMDH with values 0.0327, 0.0385, 0.0551, and 0.0888, respectively.

The reason for the accurate forecasting ability of the GRNN model is based on the reliability of its feed-forward additional summation layer in selecting the best output which is used for the weighted average computation for the desired response (Chen et al., 2019; Jian et al., 2019).

To determine further the accuracy of the four ANN methods (BPNN, GMDH, LSSVM, and GRNN), the line graph (Figure 9) was employed. The line graph is very useful when assessing the performance of many methods as it graphically summarises the methods into a single plot and allowing methods comparison with the observed data set. The methods performances are expressed in

Figure 7
AI developed models prediction

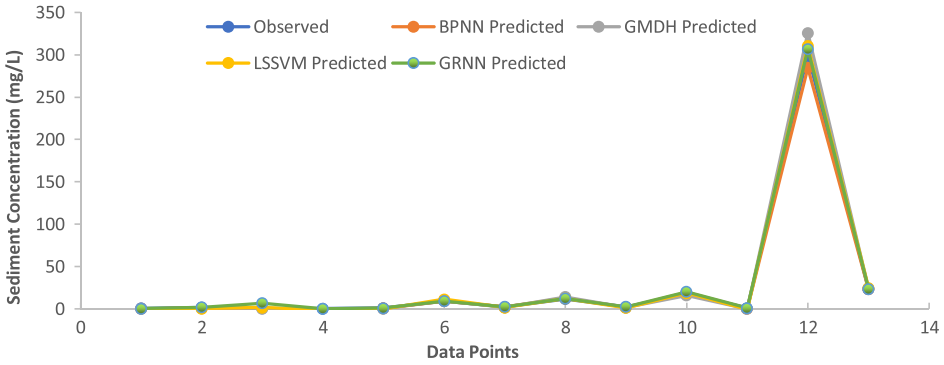
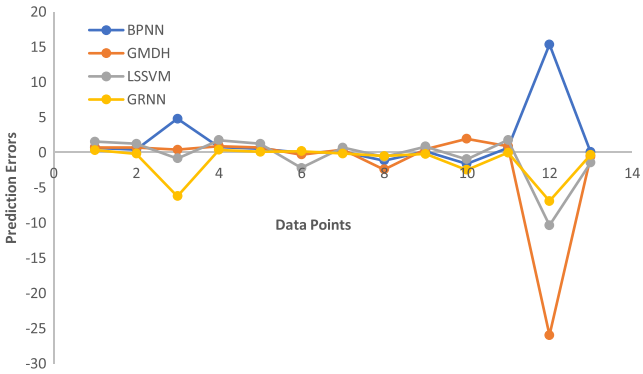


Figure 8
AI models prediction errors



terms of their standard deviations from the 45° diagonal line. Thus, the standard deviation is proportional to the method's distance from the 45° diagonal line.

From Figure 9, clearly the simulation point of the GRNN model is closer to the observed 45° diagonal line than any of the other methods used. This implies that there are similarities in the GRNN model predictions and the observed data set in terms of its obtained lowest standard deviation (least distance from the 45° diagonal line). This shows that the GRNN method had quality in simulating the test data well than any of the other methods studied.

4.4. Validation of results with literature

It is of great importance to validate the present study's results with the established research results in the literature. In that regard, a review of literature from 2019 to 2023 (5 years) has been made, with two papers reviewed each year (Table 3). As per the data set range, the absolute error metrics might vary from one case study to another. Consequently, the coefficient of determination value (R^2) was selected for validation. For example, Latif et al. (2023) developed an AI model to predict sediment load. The maximum R^2 value realised from the developed models ranges from 0.79 to 0.91. In Keshtegar et al. (2023), the prediction of sediment yields using a data-driven model achieved an $R^2 = 0.72$, RMSE = 0.51, and MAPE = 11.99%. Furthermore, several AI models which include ANN-GA, SEA/Balance, ITD-EPR, WATEM/SEDEM, ElasticNet LR, MLP, EGB, LSTM, RS, SVM-RBF, SVM-NPK, RF, MM-ANNs, FNN-PSOGSA, FNN-PSO, FNN, ANFIS, WM5, WANN, LSTM, and M5T have also been developed. Based on the revealed prediction accuracy from the literature and in comparison, with this study, the GRNN was able to obtain comparable predictability performance as evident in Table 2.

Figure 9
Models performance assessment

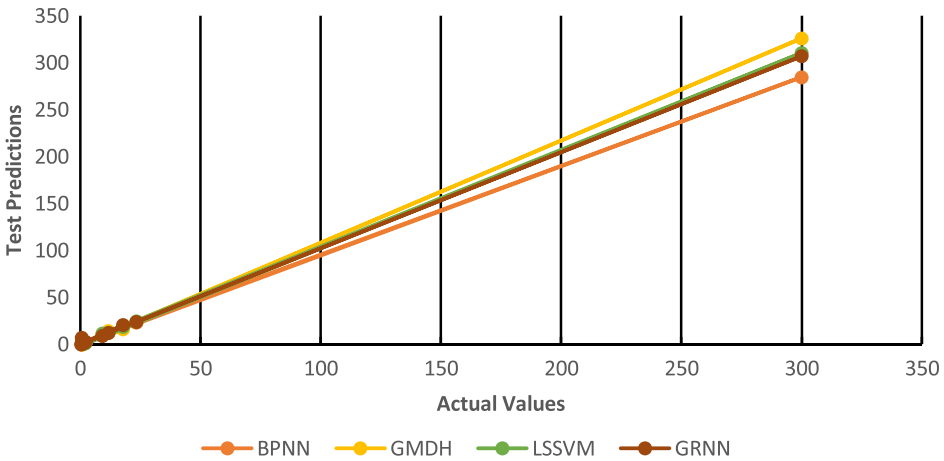


Table 3
Validation of AI methods and results used over time

Authors	Paper title	Methods applied	Validation results
Latif et al. (2023)	Sediment load prediction in Johor River: Deep learning versus machine learning models	LSTM, ANN, SVM	R^2 values range from 0.79 to 0.91
Keshtegar et al. (2023)	Prediction of sediment yields using a data-driven	RM5T	RM5T ($R^2 = 0.72$), RM5T (RMSE = 0.51), RM5T (MAPE = 11.99%)
Yadav et al. (2022)	Optimised scenario for estimating suspended sediment yield using an artificial neural network coupled with a genetic algorithm	ANN-GA	ANN-GA ($R^2 = 0.8710$), ANN-GA (RMSE = 0.0088)
Maltsev et al. (2022)	Assessment of net erosion and suspended sediments yield within river basins of the agricultural belt of Russia	SEA/Balance, WATEM/SEDEM	SEA/Balance ($R^2 = 0.78$), WATEM/SEDEM ($R^2 = 0.79$)
Zhao et al. (2021)	A decomposition and multi-objective evolutionary optimisation model for suspended sediment load prediction in rivers	ITD-EPR	ITD-EPR ($R^2 = 0.92$), ITD-EPR (WI = 0.93)
AlDahoul et al. (2021)	Suspended sediment load prediction using long short-term memory neural network	ElasticNet LR, MLP, EGB, LSTM	ElasticNet LR ($R^2 = 92.01$), MLP ($R^2 = 96.56$), EGB ($R^2 = 96.71$), LSTM ($R^2 = 0.9945$)
Nhu et al. (2020)	Monthly suspended sediment load prediction using artificial intelligence: Testing of a new random subspace method	RS, SVM-RBF, SVM-NPK, RF	RS (NSE = 0.83), SVM-RBF (NSE = 0.80), SVM-NPK (NSE = 0.78), RF (NSE = 0.68)
Meshram et al. (2020)	Application of artificial neural networks, support vector machine, and multiple model-ANN to sediment yield prediction	MM-ANNs	MM-ANNs ($R^2 = 0.921$), MM-ANNs (NSE = 0.74), MM-ANNs (RAE = 0.360)
Meshram et al. (2019)	New approach for sediment yield forecasting with a two-phase feed-forward neuron network-particle swarm optimisation model integrated with the gravitational search algorithm	FNN-PSOGSA, FNN-PSO, FNN, ANFIS	FNN-PSOGSA (NSE = 0.612), FNN-PSO (NSE = 0.500), FNN (NSE = 0.331), ANFIS (NSE = 0.244), FNN-PSOGSA (WI = 0.832), FNN-PSO (WI = 0.771), FNN (WI = 0.692), ANFIS (WI = 0.726)
Nourani et al. (2019)	A wavelet-based data mining technique for suspended sediment load modelling	WM5, WANN, M5T	WM5 (NSE = 0.94), WANN (NSE = 0.89), M5T (NSE = 0.77)

Note: RM5T, Radial M5 tree; ElasticNet LR, ElasticNet Linear Regression; MLP, multi-layer perceptron neural network; EGB, extreme gradient boosting; LSTM, long short-term memory; ANN-GA, artificial neural network genetic algorithm; SEA/Balance, soil erosion-accumulation balance; WATEM/SEDEM, water and tillage erosion model/sediment delivery model; ITD-EPR, intrinsic time-scale decomposition evolutionary polynomial regression; RS, random subspace; SVM-NPK, support vector machine normalised polynomial kernel; SVM-RBF, support vector machine radial basis function; MM-ANNs, multiple model artificial neural networks; FNN-PSOGSA, feed-forward neural network particle swarm optimisation gravitational; FNN-PSO, feed-forward neural network particle swarm optimisation; FNN, feed-forward neural network, search algorithm; ANFIS, adaptive neuro-fuzzy inference system; WM5, wavelet M5; WANN, wavelet artificial neural network; M5T, M5 tree; RF, random forest.

4.5. Research implications

Although rivers have natural sediment transports that vary with time, many significant threats to these rivers arise primarily from human activities such as climate change, pollution, landscape changes, and urbanisation; of which the excess of such phenomenon causes river channels to become unstable and flood capacity decreased due to infilling. Therefore, it is imperative to know the level of SSL because a higher level can cause flash floods in the advent of rains causing extensive damage to property, the social well-being of people, and human lives. As a result, developing a model to predetermine the level of SSL is vital. The GRNN model approach of predicting SSL can help governments and policymakers to formulate appropriate measures to prevent water/river pollution and this would reduce SSL levels. This would lead to substantial improvements in water quality and maintain the ecosystem that serves as the habitat for many living organisms. This is achievable because the accurate SSL prediction is very useful since it brings solutions to the devastating natural events (flood hazards) occurring around the world which are partly caused by SSL accumulation on river banks and cause these flash floods.

5. Conclusions

In this study, four AI models BPNN, GMDH, LSSVM, and GRNN have been developed for accurate SSL prediction based on case study data obtained from Ghana. The statistical analysis results showed that the developed AI models are good and can be used to predict SSL based on their performance indicators. The BPNN method had RMSE (4.5127), PRMSE (15.7721), U_{95} (12.3461), RSR (0.0551), and E_{LM} (0.9485). For the GMDH, 7.2736, 25.4217, 20.2807, 0.0888, and 0.9418 were produced correspondingly. On the contrary, the LSSVM had 3.1522, 11.0171, 8.8422, 0.0385, and 0.9476 for RMSE, PRMSE, U_{95}, RSR, and E_{LM}. The respective values obtained by the GRNN were 2.6779, 9.3594, 7.1446, 0.0327, and 0.9637. However, the GRNN has proven to be the best method suitable for predicting SSL based on its low RMSE, PRMSE and

U_{95} achieved values, and high E_{LM} value when compared with the other contending methods.

The predictive power of the BPNN, GMDH, LSSVM, and GRNN models was also presented visually using a line graph. This revealed how closely the GRNN model prediction results were to the observed SSL and how its performance is comparable to the other models. Based on all the statistical results, it is clear that the GRNN method has confirmed good learning and generalisation power as compared with the other methods. Therefore, the developed GRNN method can provide reliable guidance to researchers and policymakers in Africa.

The study findings offer several avenues for future research to extend and strengthen prediction-oriented model assessment and comparison in SSL prediction. The possible challenges or restrictions that could impede the model's performance are data quality and its associated characteristics. Hence, fine-tuning the hyperparameters of the AI model using metaheuristic optimisation algorithms must be considered for future analysis. This will help overcome the manual setting of the hyperparameters in the model development phase. In addition to that, environmental and climatic factors should also be considered in the future when developing SSL prediction models for different jurisdictions.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data available on request from the corresponding author upon reasonable request.

References

Adam, S. P., Alexandropoulos, S. A. N., Pardalos, P. M., & Vrahatis, M. N. (2019). No free lunch theorem: A review. In I. C. Demetriou & P. M. Pardalos (Eds.), *Approximation and Optimization* (pp.57–82). Springer. https://doi.org/10.1007/978-3-030-12767-1_5

Akrasi, S. A. (2011). Sediment discharges from Ghanaian rivers into the sea. *West African Journal of Applied Ecology*, *18*, 1–13. https://doi.org/10.4314/wajae.v18i1.70320

AlDahoul, N., Essam, Y., Kumar, P., Ahmed, A. N., Sherif, M., Sefelnasr, A., & Elshafie, A. (2021). Suspended sediment load prediction using long short-term memory neural network. *Scientific Reports*, *11*(1), 7826. https://doi.org/10.1038/s41598-021-87415-4

Alexandrov, Y., Cohen, H., Laronne, J. B., & Reid, I. (2009). Suspended sediment load, bed load, and dissolved load yields from a semiarid drainage basin: A 15-year study. *Water Resources Research*, *45*(8), W08408. https://doi.org/10.1029/2008WR007314

Alp, M., & Cigizoglu, H. K. (2007). Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environmental Modelling & Software*, *22*(1), 2–13. https://doi.org/10.1016/j.envsoft.2005.09.009

Armaghani, D. J., Momeni, E., & Asteris, P. G. (2020). Application of group method of data handling technique in assessing deformation of rock mass. *Metaheuristic Computing and Applications*, *1*(1), 1–18.

Babanezhad, M., Behroyan, I., Marjani, A., & Shirazian, S. (2021). Artificial intelligence simulation of suspended sediment load with different membership functions of ANFIS. *Neural Computing and Applications*, *33*, 6819–6833. https://doi.org/10.1007/s00521-020-05458-6

Boateng, I. (2012). An application of GIS and coastal geomorphology for large scale assessment of coastal erosion and management: A case study of Ghana. *Journal of Coastal Conservation*, *16*, 383–397. https://doi.org/10.1007/s11852-012-0209-0

Booth, D. B., Hartley, D., & Jackson, R. (2002). Forest cover, impervious-surface area, and the mitigation of stormwater impacts. *Journal of the American Water Resources Association*, *38*(3), 835–845. https://doi.org/10.1111/j.1752-1688.2002.tb01000.x

Boye, C. B., Boateng, I., Appeaning, A. K., & Wiafe, G. (2019). An assessment of the contribution of fluvial sediment discharge to coastal stability: A case study of western region of Ghana. *African Journal of Environmental Science and Technology*, *13*(5), 191–200. https://doi.org/10.5897/AJEST2017.2387

Boye, C. B. (2015). *Causes and trends in shoreline change in the western region of Ghana*. PhD Thesis, University of Ghana.

Boye, C. B., Appeaning Addo, K., Wiafe, G., & Dzigbodi-Adjimah, K. (2018). Spatio-temporal analyses of shoreline change in the western region of Ghana. *Journal of Coastal Conservation*, *22*, 769–776. https://doi.org/10.1007/s11852-018-0607-z

Chapman, A. D., Darby, S. E., Hồng, H. M., Tompkins, E. L., & Van, T. P. (2016). Adaptation and development trade-offs: Fluvial sediment deposition and the sustainability of rice-cropping in Giang Province, Mekong Delta. *Climatic Change*, *137*, 593–608. https://doi.org/10.1007/s10584-016-1684-3

Chen, R. P., Zhang, P., Kang, X., Zhong, Z. Q., Liu, Y., & Wu, H. N. (2019). Prediction of maximum surface settlement caused by earth pressure balance (EPB) shield tunneling with ANN methods. *Soils and Foundations*, *59*(2), 284–295. https://doi.org/10.1016/j.sandf.2018.11.005

Cigizoglu, H. K., & Alp, M. (2006). Generalized regression neural network in modelling river sediment yield. *Advances in Engineering Software*, *37*(2), 63–68. https://doi.org/10.1016/j.advengsoft.2005.05.002

Cleveland, T. G., Thompson, D. B., & Fang, X. (2011). *Use of the rational and modified rational method for hydraulic design*. Retrieved from: https://trid.trb.org/View/1128254

de Kruif, B. J., & de Vries, T. J. (2003). Pruning error minimization in least squares support vector machines. *IEEE Transactions on Neural Networks*, *14*(3), 696–702. https://doi.org/10.1109/TNN.2003.810597

Dunne, T. (1979). Sediment yield and land use in tropical catchments. *Journal of Hydrology*, *42*(3–4), 281–300. https://doi.org/10.1016/0022-1694(79)90052-0

Gundogdu, O., Egrioglu, E., Aladag, C. H., & Yolcu, U. (2016). Multiplicative neuron model artificial neural network based on Gaussian activation function. *Neural Computing and Applications*, *27*, 927–935. https://doi.org/10.1007/s00521-015-1908-x

Gupta, V. K., Khani, H., Ahmadi-Roudi, B., Mirakhorli, S., Fereyduni, E., & Agarwal, S. (2011). Prediction of capillary gas chromatographic retention times of fatty acid methyl esters in human blood using MLR, PLS and back-propagation artificial neural networks. *Talanta*, *83*(3), 1014–1022. https://doi.org/10.1016/j.talanta.2010.11.017

Hazarika, B. B., Gupta, D., & Berlin, M. (2020). A comparative analysis of artificial neural network and support vector regression for river suspended sediment load prediction. In

First International Conference on Sustainable Technologies for Computational Intelligence, 339–349. https://doi.org/10.1007/978-981-15-0029-9_27

He, H., Kojima, S., Omura, T., Maruta, K. & Ahn, C. J. (2021). Generalized regression neural network-based channel identification and compensation using scattered pilot. *Radioengineering*, *30*(4), 695–703. https://doi.org/10.13164/re.2021.0695

Hou, J., Lu, X., Zhang, K., Jing, Y., Zhang, Z., You, J., & Li, Q. (2022). Parameters identification of rubber-like hyper elastic material based on general regression neural network. *Materials*, *15*(11), 3776. https://doi.org/10.3390/ma15113776

Jian, J., Guo, Y., Jiang, L., An, Y., & Su, J. (2019). A multi-objective optimization model for green supply chain considering environmental benefits. *Sustainability*, *11*(21), 5911. https://doi.org/10.3390/su11215911

Kaytez, F., Taplamacioglu, M. C., Cam, E., & Hardalac, F. (2015). Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. *International Journal of Electrical Power & Energy Systems*, *67*, 431–438. https://doi.org/10.1016/j.ijepes.2014.12.036

Keates, S. G. (2021). Notes on the Palaeolithic finds from the Walanae valley, southwest Sulawesi, in the context of the Late Pleistocene of Island Southeast Asia. In S. G. Keates & J. M. Pasveer (Eds.), *Quaternary Research in Indonesia*(pp. 95–109). CRC Press.

Keshtegar, B., Piri, J., Hussan, W. U., Ikram, K., Yaseen, M., Kisi, O., . . . , & Waseem, M. (2023). Prediction of sediment yields using a data-driven radial M5 tree model. *Water*, *15*(7), 1437. https://doi.org/10.3390/w15071437

Kim, D. K., Mikhaylova, M., Zhang, Y., & Muhammed, M. (2003). Protective coating of superparamagnetic iron oxide nanoparticles. *Chemistry of Materials*, *15*(8), 1617–1627. https://doi.org/10.1021/cm021349j

Kisi, O. (2012). Modeling discharge-suspended sediment relationship using least square support vector machine. *Journal of Hydrology*, *456–457*, 110–120. https://doi.org/10.1016/j.jhydrol.2012.06.019

Kisi, O., Karahan, M. E., & Şen, Z. (2006). River suspended sediment modelling using a fuzzy logic approach. *Hydrological Processes: An International Journal*, *20*(20), 4351–4362. https://doi.org/10.1002/hyp.6166

Lafdani, E. K., Nia, A. M., & Ahmadi, A. (2013). Daily suspended sediment load prediction using artificial neural networks and support vector machines. *Journal of Hydrology*, *478*, 50–62. https://doi.org/10.1016/j.jhydrol.2012.11.048

Latif, S. D., Chong, K. L., Ahmed, A. N., Huang, Y. F., Sherif, M., & El-Shafie, A. (2023). Sediment load prediction in Johor River: Deep learning versus machine learning models. *Applied Water Science*, *13*(3), 79. https://doi.org/10.1007/s13201-023-01874-w

Lee, K., Hong, C., Lee, E. H., & Yang, W. (2020). Comparison of artificial intelligence methods for prediction of mechanical properties. In *IOP Conference Series: Materials Science and Engineering*, *967*(1), 012031. https://doi.org/10.1088/1757-899X/967/1/012031

Legates, D. R., & McCabe, G. J. (2013). A refined index of model performance: A rejoinder. *International Journal of Climatology*, *33*(4), 1053–1056. https://doi.org/10.1002/joc.3487

Liu, F., Wang, J., & Qin, S. (2013). Training seats-square support vector machine by a recurrent neural network based on fuzzy c-mean approach. In *Proceedings of Advances in Swarm Intelligence: 4th International Conference*, 106–113.

Makridakis, S. (2017). The forthcoming artificial intelligence revolution: Its impact on society and firms. *Futures*, *90*, 46–60. https://doi.org/10.1016/j.futures.2017.03.006

Maltsev, K., Golosov, V., Yermolaev, O., Ivanov, M., & Chizhikova, N. (2022). Assessment of net erosion and suspended sediments yield within river basins of the agricultural belt of Russia. *Water*, *14*(18), 2781. https://doi.org/10.3390/w14182781

Mehri, Y., Nasrabadi, M., & Omid, M. H. (2021). Prediction of suspended sediment distributions using data mining algorithms. *Ain Shams Engineering Journal*, *12*(4), 3439–3450. https://doi.org/10.1016/j.asej.2021.02.034

Melesse, A. M., Ahmad, S., McClain, M. E., Wang, X., & Lim, Y. H. (2011). Suspended sediment load prediction of river systems: An artificial neural network approach. *Agricultural Water Management*, *98*(5), 855–866. https://doi.org/10.1016/j.agwat.2010.12.012

Meshram, S. G., Ghorbani, M. A., Deo, R. C., Kashani, M. H., Meshram, C., & Karimi, V. (2019). New approach for sediment yield forecasting with a two-phase feedforward neuron network-particle swarm optimisation model integrated with the gravitational search algorithm. *Water Resources Management*, *33*, 2335–2356. https://doi.org/10.1007/s11269-019-02265-0

Meshram, S. G., Singh, V. P., Kisi, O., Karimi, V., & Meshram, C. (2020). Application of artificial neural networks, support vector machine and multiple model-ANN to sediment yield prediction. *Water Resources Management*, *34*(15), 4561–4575. https://doi.org/10.1007/s11269-020-02672-8

Milliman, J. D., & Farnsworth, K. L. (2013). *River discharge to the coastal ocean: A global synthesis*. UK: Cambridge University Press.

Mohamed, I., & Shah, I. (2018). Suspended sediment concentration modeling using conventional and machine learning approaches in the Thames River, London Ontario. *Journal of Water Management Modeling, 26*, 1–12. https://doi.org/10.14796/JWMM.C453

Møller, M. F. (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, *6*(4), 525–533. https://doi.org/10.1016/S0893-6080(05)80056-5

Montalto, F., Behr, C., Alfredo, K., Wolf, M., Arye, M., & Walsh, M. (2007). Rapid assessment of the cost-effectiveness of low impact development for CSO control. *Landscape and Urban Planning*, *82*(3), 117–131. https://doi.org/10.1016/j.landurbplan.2007.02.004

Mueller, A. V., & Hemond, H. F. (2013). Extended artificial neural networks: Incorporation of a priori chemical knowledge enables the use of ion-selective electrodes for in-situ measurement of ions at environmentally relevant levels. *Talanta*, *117*, 112–118. https://doi.org/10.1016/j.talanta.2013.08.045

Nandy, S., Sarkar, P. P., & Das, A. (2012). Analysis of a nature-inspired firefly algorithm based back-propagation neural network training. *arXiv Preprint:1206.5360*.

Nguyen, T. N., Lee, S., Nguyen-Xuan, H., & Lee, J. (2019). A novel analysis-prediction approach for geometrically nonlinear problems using group method of data handling. *Computer Methods in Applied Mechanics and Engineering*, *354*, 506–526. https://doi.org/10.1016/j.cma.2019.05.052

Nhu, V. H., Khosravi, K., Cooper, J. R., Karimi, M., Kisi, O., Pham, B. T., & Lyu, Z. (2020). Monthly suspended sediment load prediction using artificial intelligence: Testing of a new random subspace method. *Hydrological Sciences Journal*, *65*(12), 2116–2127. https://doi.org/10.1080/02626667.2020.1754419

Nittrouer, J. A., & Viparelli, E. (2014). Sand as a stable and sustainable resource for nourishing the Mississippi River delta. *Nature Geoscience*, *7*(5), 350–354. https://doi.org/10.1038/ngeo2142

Nourani, V., & Andalib, G. (2015). Daily and monthly suspended sediment load predictions using wavelet based artificial intelligence approaches. *Journal of Mountain Science*, *12*, 85–100. https://doi.org/10.1007/s11629-014-3121-2

Nourani, V., Molajou, A., Tajbakhsh, A. D., & Najafi, H. (2019). A wavelet-based data mining technique for suspended sediment load modeling. *Water Resources Management*, *33*, 1769–1784. https://doi.org/10.1007/s11269-019-02216-9

Packman, J. C., & Kidd, C. H. R. (1980). A logical approach to the design storm concept. *Water Resources Research*, *16*(6), 994–1000. https://doi.org/10.1029/WR016i006p00994

Rayegani, F., & Onwubolu, G. C. (2014). Fused deposition modeling process parameter prediction and optimization using group method for data handling and differential evolution. *The International Journal of Advanced Manufacturing Technology*, *73*(1), 509–519. https://doi.org/10.1007/s00170-014-5835-2

Rezaei, K., Pradhan, B., Vadiati, M., & Nadiri, A. A. (2021). Suspended sediment load prediction using artificial intelligence techniques: Comparison between four state-of-the-art artificial neural network techniques. *Arabian Journal of Geosciences*, *14*, 215. https://doi.org/10.1007/s12517-020-06408-1

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0

Schenk, L. N., & Bragg, H. M. (2014). *Assessment of suspended-sediment transport, bedload, and dissolved oxygen during a short-term drawdown of Fall Creek Lake, Oregon, winter 2012-13*. Retrieved from: https://doi.org/10.3133/ofr20141114

Srinivasan, D. (2008). Energy demand prediction using GMDH networks. *Neurocomputing*, *72*(1–3), 625–629. https://doi.org/10.1016/j.neucom.2008.08.006

Stefenon, S. F., Ribeiro, M. H. D. M., Nied, A. C., Mariani, V., dos Santos Coelho, L., da Rocha, D. F. M., . . . , & de Barros Ruano, A. E. (2020). Wavelet group method of data handling for fault prediction in electrical power insulators. *International Journal of Electrical Power & Energy Systems*, *123*, 106269. https://doi.org/10.1016/j.ijepes.2020.106269

Suykens, J. A. K., van Gestel, T., de Brabanter, J., de Moor, B., & Vandawalle, J. (2002). *Least squares support vector machines*. Singapore: World Scientific Publishing.

Thompson, S. (2006). *The political theory of recognition: A critical introduction*. UK: Polity.

Tian, Y., Nearing, G. S., Peters-Lidard, C. D., Harrison, K. W., & Tang, L. (2016). Performance metrics error modeling, and uncertainty quantification. *Monthly Weather Review*, *144*(2), 607–613. https://doi.org/10.1175/MWR-D-15-0087.1

Toker, S., & Kaçıranlar, S. (2013). On the performance of two parameter ridge estimator under the mean square error criterion. *Applied Mathematics and Computation*, *219*(9), 4718–4728. https://doi.org/10.1016/j.amc.2012.10.088

van Gestel, T., Suykens, J. A., Baesens, B., Viaene, S., Vanthienen, J., Dedene, G., . . . , & Vandewalle, J. (2004). Benchmarking least squares support vector machine classifiers. *Machine Learning*, *54*, 5–32. https://doi.org/10.1023/B:MACH.0000008082.80494.e0

Vanmaercke, M., Poesen, J., Broeckx, J., & Nyssen, J. (2014). Sediment yield in Africa. *Earth-Science Reviews*, *136*, 350–368. https://doi.org/10.1016/j.earscirev.2014.06.004

Walling, D. E. (1977). Assessing the accuracy of suspended sediment rating curves for a small basin. *Water Resources Research*, *13*(3), 531–538. https://doi.org/10.1029/WR013i003p00531

Walling, D. E., & Webb, B. W. (1985). Estimating the discharge of contaminants to coastal waters by rivers: Some cautionary comments. *Marine Pollution Bulletin*, *16*(12), 488–492. https://doi.org/10.1016/0025-326X(85)90382-0

Walling, D. E., & Webb, B. W. (1996). Erosion and sediment yield: A global overview. In *Proceedings of the Exeter Symposium*, 3–19.

Wang, Y. M., Kerh, T., & Traore, S. (2009). Neural networks approach for modeling river suspended sediment concentration due to tropical storms. *Global NEST Journal*, *11*(4), 457–466.

Wang, Y., & Peng, H. (2018). Underwater acoustic source localization using generalized regression neural network. *The Journal of the Acoustical Society of America*, *143*(4), 2321–2331. https://doi.org/10.1121/1.5032311

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error over the root mean square error in assessing average model performance. *Climate Research*, *30*(1), 79–82. http://doi/org/10.3354/cr030079

Wolpert, D. H. (2002). The supervised learning no-free-lunch theorems. In R. Roy, M. Köppen, S. Ovaska, T. Furuhashi & F. Hoffmann (Eds.), *Soft Computing and Industry* (pp. 25–42). Springer.

Xia, X. L. (2018). Training sparse least squares support vector machines by the QR decomposition. *Neural Networks*, *106*, 175–184. https://doi.org/10.1016/j.neunet.2018.07.008

Yadav, A., Hasan, M. K., Joshi, D., Kumar, V., Aman, A. H. M., Alhumyani, H., . . . , & Mishra, H. (2022). Optimized scenario for estimating suspended sediment yield using an artificial neural network coupled with a genetic algorithm. *Water*, *14*(18), 2815. https://doi.org/10.3390/w14182815

Yang, X., Tan, L., & He, L. (2014). A robust least squares support vector machine for regression and classification with noise. *Neurocomputing*, *140*, 41–52. https://doi.org/10.1016/j.neucom.2014.03.037

Zhang, G., Xia, B., & Wang, J. (2021). Intelligent state of charge estimation of lithium-ion batteries based on LM optimized back-propagation neural network. *Journal of Energy Storage*, *44*, 103442. https://doi.org/10.1016/j.est.2021.103442

Zhao, N., Ghaemi, A., Wu, C., Band, S. S., Chau, K. W., Zaguia, A., . . . , & Mosavi, A. H. (2021). A decomposition and multi-objective evolutionary optimisation model for suspended sediment load prediction in rivers. *Engineering Applications of Computational Fluid Mechanics*, *15*(1), 1811–1829. https://doi.org/10.1080/19942060.2021.1990133

Zurada, J. (1994). End effector target position learning using feedforward with error back-propagation and recurrent neural networks. In *Proceedings of 1994 IEEE International Conference on Neural Networks*, *4*, 2633–2638. https://doi.org/10.1109/ICNN.1994.374637

How to Cite: Boye, C. B., Boye, P., & Ziggah, Y. Y. (2024). Comparative Study of Suspended Sediment Load Prediction Models Based on Artificial Intelligence Methods. *Artificial Intelligence and Applications*, *2*(2), 141–154. https://doi.org/10.47852/bonviewAIA3202832