



Sediment load prediction in Johor river: deep learning versus machine learning models

Sarmad Dashti Latif¹ · K. L. Chong² · Ali Najah Ahmed³ · Y. F. Huang² · Mohsen Sherif^{4,5} · Ahmed El-Shafie⁶

Received: 1 November 2022 / Accepted: 11 January 2023 / Published online: 13 February 2023
© The Author(s) 2023

Abstract

Sediment transport is a normal phenomenon in rivers and streams, contributing significantly to ecosystem production and preservation by replenishing vital nutrients and preserving aquatic life's natural habitats. Thus, sediment transport prediction through modeling is crucial for predicting flood events, tracking coastal erosion, planning for water supplies, and managing irrigation. The predictability of process-driven models may encounter various restrictions throughout the validation process. Given that data-driven models work on the assumption that the underlying physical process is not requisite, this opens up the avenue for AI-based model as alternative modeling. However, AI-based models, such as ANN and SVM, face problems, such as long-term dependency, which require alternative dynamic procedures. Since their performance as universal function approximation depends on their compatibility with the nature of the problem itself, this study investigated several distinct AI-based models, such as long short-term memory (LSTM), artificial neural network (ANN), and support vector machine (SVM), in predicting sediment transport in the Johor river. The collected historical daily sediment transport data from January 1, 2008, to December 01, 2018, through autocorrelation function, were used as input for the model. The statistical results showed that, despite their ability (deep learning and machine learning) to provide sediment predictions based on historical input datasets, machine learning, such as ANN, might be more prone to overfitting or being trapped in a local optimum than deep learning, evidenced by the worse in all metrics score. With RMSE = 11.395, MAE = 18.094, and $R^2 = 0.914$, LSTM outperformed other models in the comparison.

Keywords Long short-term memory (LSTM) · Artificial neural network (ANN) · Support vector machine (SVM) · Sediment transport prediction

✉ Y. F. Huang
huangyf@utar.edu.my

Sarmad Dashti Latif
Sarmad.latif@komar.edu.iq

K. L. Chong
chongkailun1993@gmail.com

Ali Najah Ahmed
mahfoodh@uniten.edu.my

Mohsen Sherif
MSherif@uaeu.ac.ae

Ahmed El-Shafie
elshafie@um.edu.my

Abdul Rahman, Jalan Sg. Long, Bandar Sg. Long, Cheras,
43000 Kajang, Selangor, Malaysia

³ Dam Safety Management & Engineering Group, Institute
of Energy Infrastructure and Department of Civil
Engineering, College of Engineering, Universiti Tenaga
Nasional (UNITEN), 43000 Kajang, Selangor, Malaysia

⁴ National Water and Energy Center, United Arab Emirates
University, P.O. Box 15551, Al Ain, United Arab Emirates

⁵ Civil and Environmental Engineering Department,
College of Engineering, United Arab Emirates University,
P.O. Box 15551, Al Ain, United Arab Emirates

⁶ Department of Civil Engineering, Faculty of Engineering,
University of Malaya, 50603 Kuala Lumpur, Malaysia

¹ Civil Engineering Department, College of Engineering,
Komar University of Science and Technology, Sulaimany,
Kurdistan Region 46001, Iraq

² Department of Civil Engineering, Lee Kong Chian
Faculty of Engineering and Science, Universiti Tunku

Abbreviations

LSTM	Long short-term memory
ANN	Artificial neural network
SVM	Support vector machine
ACF	Autocorrelation function
RMSE	Root mean square error
MAE	Mean absolute error
R^2	Coefficient of determination
AI	Artificial intelligence
MLP	Multilayer perceptron
FFNN	Feed-forward neural network
SSC	Suspended sediment concentration
Q	Discharge
SSL	Suspended sediment load
GPR	Gaussian process regression
Cv	Volumetric sediment concentration
HS-ANN	Harmony search–artificial neural network
PCC	Pearson correlation coefficient
RBF	Radial basis function
RNN	Recurrent neural network

Introduction

Sediment is a naturally occurring substance that decomposes during weathering and erosion processes and is subsequently transported by the action of natural forces such as water, wind, or ice or by the gravitation force on the particles (Fassoni-Andrade and de Paiva 2019; Liu et al. 2019). There are three basic types of sediments: clastic, biogenic, and chemical sediments. Firstly, clastic sediments are inorganic sediments formed due to rock erosion and weathering; they are classified based on their dominant grain size. The fragment's primary grain size is stone, followed by a cobble, pebble, sand, silt, and clay as the smallest grain. Secondly, biogenic sediments are the organic sediments produced by organisms through their life activities. They may compose of shell remnants, skeletal remains, or fragments of coral reefs. Finally, chemical sediments are a mixture of all evaporite minerals, including halite, sylvite, bassanite, and gypsum. In addition, it contains deposits of iron phosphates, minerals, manganese, silica, and metal sulfide. These mixtures are the products of inorganic processes in the sedimentary environment. Their common characteristics are typically dissolved, entrained, and transported in surface water, such as rivers, lakes, and oceans, through sediment transport (Matos et al. 2019; Żarczyński et al. 2019).

When sediments flow through a fluid medium, often rivers and streams, due to a combination of forces acting on them, such as gravity and fluid movement, they are transported along the streams (Addo-Bediako et al. 2021; Kuriqi et al. 2020). The three commonly known modes of sediment movement are bed load, suspended load, and dissolved load.

Bed load means the load that goes upstream to downstream on the riverbed (Goldstein et al. 2019; Ma et al. 2020). The bed loads contain clastic sediments such as rocks, pebbles, and sand, as well as biogenic sediments, which are heavier and denser than water, and these sediments can only roll, slip, and salt down the river bed. A river also holds suspended loads of mud, silt, and fine sand. While these materials are denser than water, they remain suspended due to the instability of the streamflow, forming swirls and eddies caused by friction between the stream and its canal. Dissolved loads come mostly from groundwater when bedrock and aquifers are chemically degraded as surface water passes through the rocks and aquifers into streams. Suspended load is the most typically borne by streams and rivers (Mao et al. 2019; Wilkes et al. 2019). Furthermore, the rate of being transported rises in relation to human land use and practices such as agriculture and urbanization, which decreases vegetation coverage and makes rainwater quickly wash sediments into streams (Harada et al. 2019; Wesselman et al. 2019).

Extreme suspended loads in streams may endanger aquatic life and have a negative impact on the ecosystem (Sathya et al. 2022). After all, the suspended solids can clog the fish gills and impede light penetration into the stream bed, limiting the rate of photosynthesis of aquatic plants and resulting in a decrease in the dissolved oxygen level of the streams, which can be lethal to aquatic life. Besides, siltation happens when suspended solids such as silt settle in the stream bed; it can alter the bottom of the stream and impact the organisms that live on the surface. A high concentration of suspended solids in a water body also might make drinking water treatment difficult. The increased coagulation and sediment filtering operations will necessitate additional funding, including extra chlorine for turbid water disinfection. However, the suspended solids have a positive side that reacts with the toxic chemicals that dissolve in the stream and make it less harmful to aquatic life. Thus, it is essential to precisely forecast the flow of sediments in order to analyze the conditions of the stream, reduce their impacts and enable the planning of a better solution or mitigation method (Hapsari et al. 2019; Zhang and Yang 2020).

Background

Over the last few decades, hydrologists have simulated sediment transport using linear and nonlinear models to determine the optimum one for accurately predicting sediment transport (Lu and Chiang 2019). Nevertheless, the river sediment is not only time-varying and spatially distributed, but it also inhibits nonlinear behavior. Simple models can not effectively characterize and simulate sediment transport, including conventional linear and nonlinear statistical models, such as regression analysis. The effort to innovate and build a detailed, balanced, and accurate sediment

transport model has thus never been halted (Mèžinè et al. 2019). Recently, many researchers have adopted artificial intelligence (AI) based model to develop robust models for solving engineering problems (Kumar et al. 2022) (Rahman et al. 2022). In forecasting sediment transport, a variety of AI approaches, spanning from machine learning to deep learning, have been used:

Gaussian process regression (GPR)

Roushangar and Shahnaz (2020) used Gaussian process regression (GPR) to estimate the sediment transport rate of 19 gravel-bed rivers in the USA. They compared their developed SVM to GPR, which is a specific type of kernel-based model. Furthermore, this study investigated two scenarios: hydraulic characteristics only in the first scenario, and hydraulic and sediment properties in the second. Findings showed that the GPR model performed better. They also noticed that input based on the second scenario contributed to more accurate forecasts.

Artificial neural network (ANN)

Afan et al. (2015) used ANN to estimate the average load of sediments. For this reason, two distinct ANN algorithms, FFNN and RBF, were used. The neural networks are fed and validated using sediment flow data from the Johor river station in Rantau Panjang on a regular basis. The results show that combining sediment load data with flow data provides a detailed model for predicting sediment load. Comparing the findings reveals that the FFNN model has better efficiency in estimating the daily sediment load than the RBF model.

Support vector machine (SVM)

Shafaghat and Dezvareh (2020) utilized the support vector machine (SVM) in estimating the sediment transport forecast. Due to the diversity of SVM, they approached the forecasting of sediment transport in two different machine learning tasks, namely the classification and regression. The findings showed that classifying the sediment transport despite without inputting information such as bed and beach profiles, it could achieve a satisfactory performance. Also, the other forecast form of regression further substantial the capability of SVM in predicting sediment transport rate.

Long short-term memory (LSTM)

Kaveh et al. (2021) considered LSTM to predict suspended sediment concentration in the Schuylkill River, USA. In their study, the LSTM expanded the recurring neural network with memory cells to store and output information instead of recurring modules, enabling the learning of long-term

relationships. Observed time series of river discharge in the daily interval and sediment concentrations were used as input to develop their proposed model. Their results were tested and compared with the feed-forward neural network and the adaptive neuro-fuzzy inference models used algorithms in predicting sediment concentration in literature. The model prediction accuracy comparison showed that the LSTM model could estimate not only sediment concentration but also cumulative suspended sediment load to a high degree of accuracy.

Hybrid ML with other techniques

Zounemat-Kermani et al. (2020) proposed a hybrid model that comprises a metaheuristic harmony search algorithm (HS) and artificial neural network (ANN) for forecasting sediment transport in sewer pipe systems in terms of volumetric sediment concentration. They used the factor analysis methodology to solve the difficulties of selecting the optimal input variables number and considering the model's effective parameters. A multiple linear regression (MLR), a comparison model to the developed model, is also used to tackle the same formulated problem. Based on the comparative analysis, the HS-ANN model surpassed the existing MLR model in successfully predicting sediment transport in sewage networks. Apart from metaheuristic, another hybridization involve the mechanism of transforming the raw dataset into simpler and much meaningful input sub-series. For instance, Ebtehaj et al. (2016) adopted the wavelet transform in order to decompose the observed sediment transport into smaller components, known as decomposed wavelets. They demonstrated that a smoother and decomposed sub-series was relatively easier to predict.

Problem statement

The problems with sediment transport prediction have garnered significant attention because of how crucial sediment movement is in shaping the Earth's surface. The adoption of a well-grounded tool to estimate the suspended sediment load is often the point of interest. Although contemporary numerical models have improved, river sediment load movement remains challenging and has yet to be fully elucidated. A direct technique, for example, that requires the deployment of a hydrometric station for monitoring and sample purposes can be costly and time-consuming, especially in distant areas (Bandini et al. 2020). While indirect procedures are less expensive, reconciling theoretical results to observations is challenging due to sediment particle sensitivity to many environmental conditions (Asadi et al. 2021). Moreover, the bulk of experimentally validated equations is limited to a narrow range of environment situations, restricting their application even further (Ebtehaj et al. 2014).

Alternatively, to avoid the computation of the complex sediment transport rate, the rising of artificially intelligent

algorithms, such as ANN and SVM, has revolutionized various time series forecasting, including sediment transport prediction. One of the main advantages of such approach is that the underlying physical process of the complex sediment transport process is not requisite (Goldstein et al. 2019). While they excel in automatic adaptivity and self-learning, they nonetheless have significant drawbacks. One of them is inadequate to ensure that the search technique will result in an ideal solution only with their competent exploitation search (Chong et al. 2020). Furthermore, given the spatial and temporal aspects, multiple studies have raised concerns that the black box model fails to capture long-term reliance when provided with a substantial historical datasets (Samal et al. 2021; Xing et al. 2020).

A recurrent neural network (RNN), a variant of ANN, was initially adopted due to its compatibility to sequential data as RNN processes memory which allows it to learn long-term dependency. However, as indicated in the study by Vaswani et al. (2017), they showed that when RNNs are trying to learn long-term dependency information, they suffer from vanishing or exploding gradient, leading to deteriorate in model performance. Consequently, addressing the spatial and temporal aspects has led the researchers to the field of deep learning, such as the application of LSTM or CNN. The current research indicates that applying CNN may solve the spatial dependency of a time series, but the temporal domain aspect is still challenging (Vo et al. 2017). This research was therefore motivated by the lack of a reliable sediment transport model that could account for temporal inequality. While there might be a number of AI-based applications in forecasting sediment transport, such as the study by Kaveh et al. (2021), a thorough investigation is necessary due to the characteristics of sediment load transport being influenced by various factors that individually different from one another.

Contribution of the study

- In view of the mentioned issues, the following are the study's contributions:
- In tackling a machine learning regression task in the sediment transport application for the temporal inequality, the proposed deep learning LSTM model is compared to traditional machine learning.
- The research of input lag selection for developed auto-correlation function (ACF) based sediment transport prediction models is investigated and addressed.
- An investigation is carried out into how differentiated SVM and ANN are in sediment transport prediction, which typically has identical competence in many areas.

Methodology

Study area

The Johor River is located south of the Malaysian Peninsular (Fig. 1). The basin surface elevation ranges from 20 to 540 m (Katimon et al. 2018). It has five main branches (Gemuroh, Linggui, Lemekik, Lebak, and Semanggar). The channel flows north to south, culminating in the Johor Strait. Johor river water is used for water supply in the Malaysian state of Johor, containing the city of Johor Bahru. Hence, the Johor river's water quality is critical for the state. The collected data between 01/01/2008 to 10/12/2018 in Sungai Rantau, Johor river basin, was recorded daily for ton per day. The reason for the selected station is that, compared to other stations in the Sungai Rantau area of the Johor River Basin, the data collected throughout the research period is to be the most trustworthy. Furthermore, several stations have incomplete data, rendering the sediment movement forecasting technique difficult. Figure 2 shows the daily sediment of the Johor river from 01/01/2008 to 10/12/2018. Table 1 presents the statistical properties of the daily sediment dataset of the Johor river.

Long short-term memory (LSTM)

The LSTM approach was developed in 1997. Schmidhuber and Hochreiter (1997) proposed LSTM, which proved to be much faster in terms of convergence and more accurate and reliable than other models. LSTM is a particular kind of RNN (called recurrent neural networks, also known as RNN) used in deep learning applications. However, unlike a neural network, apart from processing the input via multiple hidden layers, they also stored the information in them and iterated to offer new information.

$$g_t = \sigma(U_g x_t + W_g h_{t-1} + b_g) \quad (1)$$

$$i_t = \sigma(U_i x_t + W_i h_{t-1} + b_i) \quad (2)$$

$$\tilde{c}_t = \tanh(U_c x_t + W_c h_{t-1} + b_c) \quad (3)$$

$$c_t = g_t * c_{t-1} + i_t * \tilde{c}_t \quad (4)$$

$$o_t = \sigma(U_o x_t + W_o h_{t-1} + b_o) \quad (5)$$

$$h_t = o_t * \tanh(c_t) \quad (6)$$

U and W represent the weights in different gates: input gate (i_t), gate modulates input (\tilde{c}_t), forget gate (g_t), and gate output (o_t). b represents the bias term, c_t represents a cell state, and h_t represents the hidden state. These gates handle the

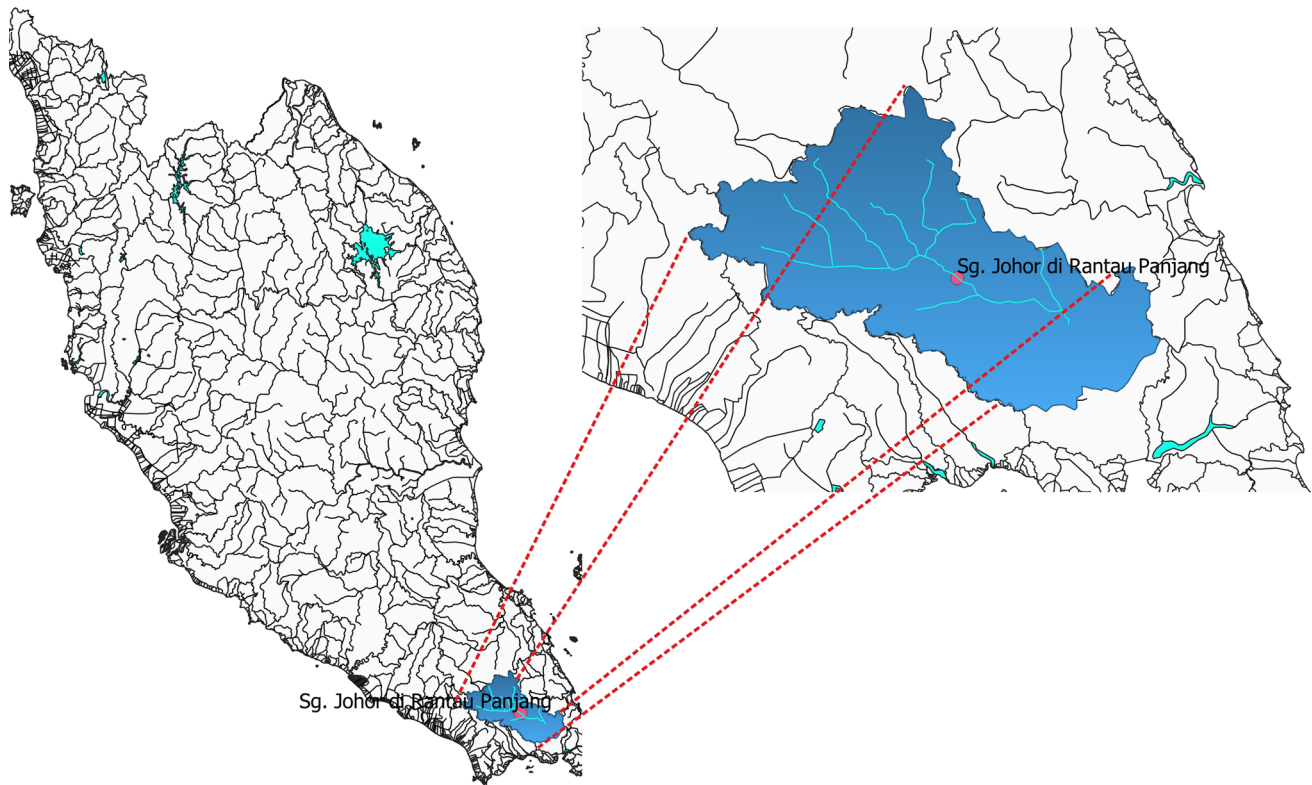


Fig. 1 Johor river and its sampling station

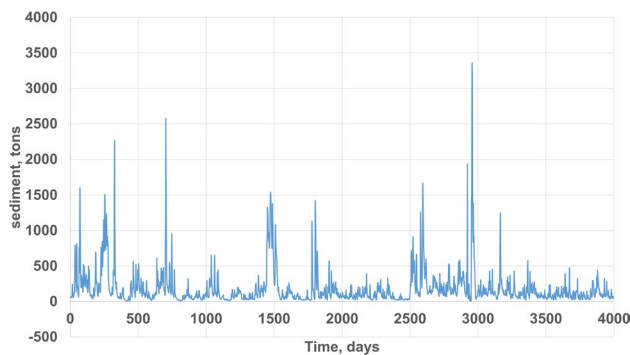


Fig. 2 Daily sediment at Johor river from 01/01/2008 to 10/12/2018

information obtained from the previous cycle and decide how much of the recent information to forward to the subsequent phase. LSTM can learn the long-term dependencies resulting from its unique architectural nature, as shown in Fig. 3 (Li et al. 2017). It is one of the main advantages of LSTM compared to RNN, which previously lost the information reported in dealing with long-term dependence problems.

Time series forecasting using deep learning code has been applied to predict daily sediment at the Johor river. The training progress of the developed LSTM model is shown in Fig. 4.

Table 1 Statistical properties of the daily sediment at Johor river

Daily sediment for the selected period (tons)	
Mean	185.404
Standard error	4.101
Median	108.000
Mode	23
Standard deviation	259.262
Sample variance	67,200.040
Kurtosis	31.218
Skewness	4.521
Range	3347.400
Minimum	5.600
Maximum	3353.000
Sum	740,690.700
Count	3995

Artificial neural network (ANN)

MLP and FFNN are the simplest ANN in which the node’s connections do not form a loop; in other words, the flow of information is only in a single forward direction. FFNN can be a single-layer perceptron or an MLP, depending on the number of hidden layers available. The single-layer perceptron consists of a single-layer output node connected directly

Fig. 3 Basic structure of LSTM

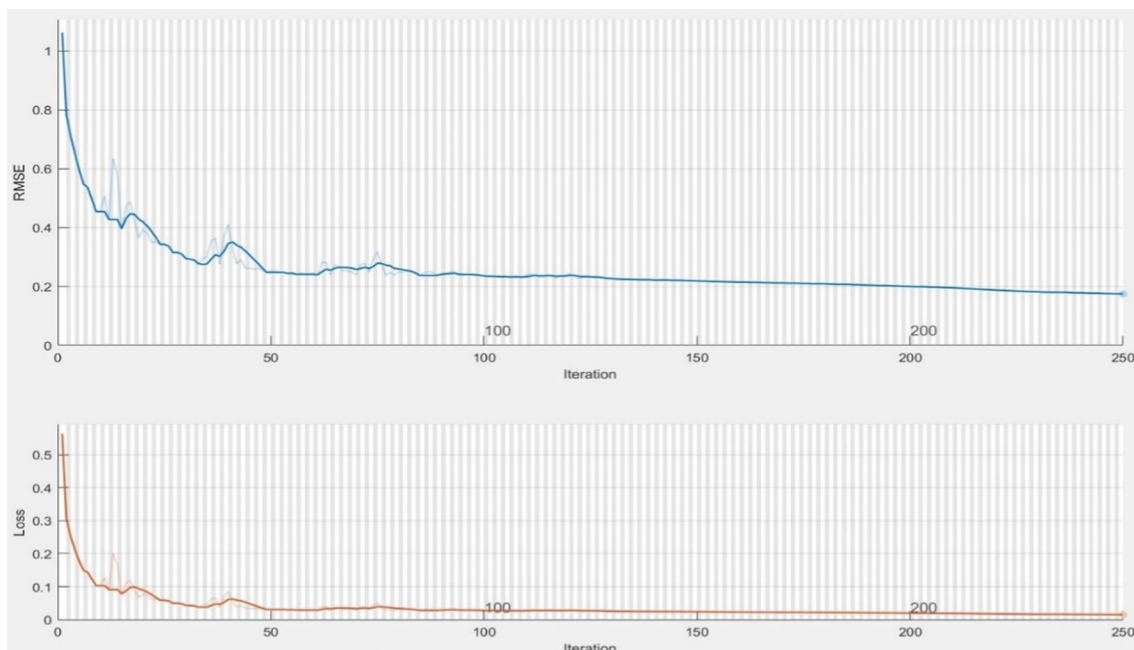
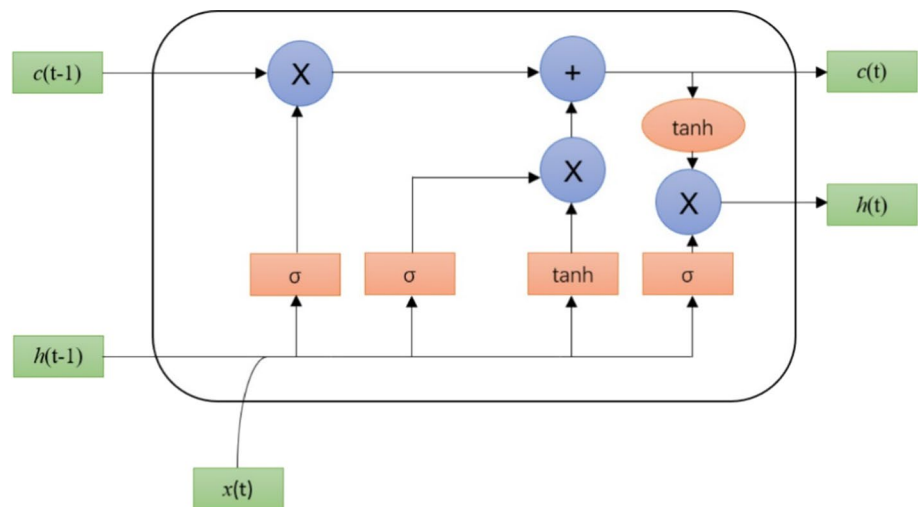


Fig. 4 Training progress of the developed LSTM model

to the input by a series of weights. On the contrary, a multilayer perceptron is a network interconnected with multiple layers of hidden or computational units. The sigmoid function is a commonly used activation function in a multilayer perceptron (Fiyadh et al. 2019). Figure 5 depicts the general architecture of the MLP model.

Support vector machine (SVM)

This section provides a basic overview of SVM. The support vector machine (SVM) is an efficient data categorization and regression approach, which was initially proposed

by Cortes and Vapnik (1995). There are accountable decision functions, for example, and hyperplanes capable of delineating positive and negative data that defined the maximum margins. It displays the variance from the nearest positive to a hyperplane sample and maximizes the variance between the nearest negative sample and the hyperplane. Data points closest to the separating hyperplane are support vectors (Fig. 6).

The main principle behind SVM in regression is to use a nonlinear kernel function to translate the input data into a higher-dimensional space where a linear estimate function is defined as follows:

Fig. 5 General structure of MLP

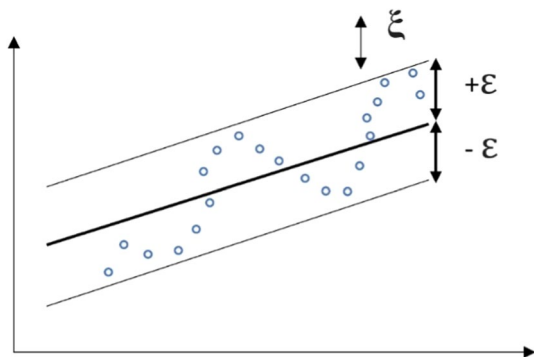
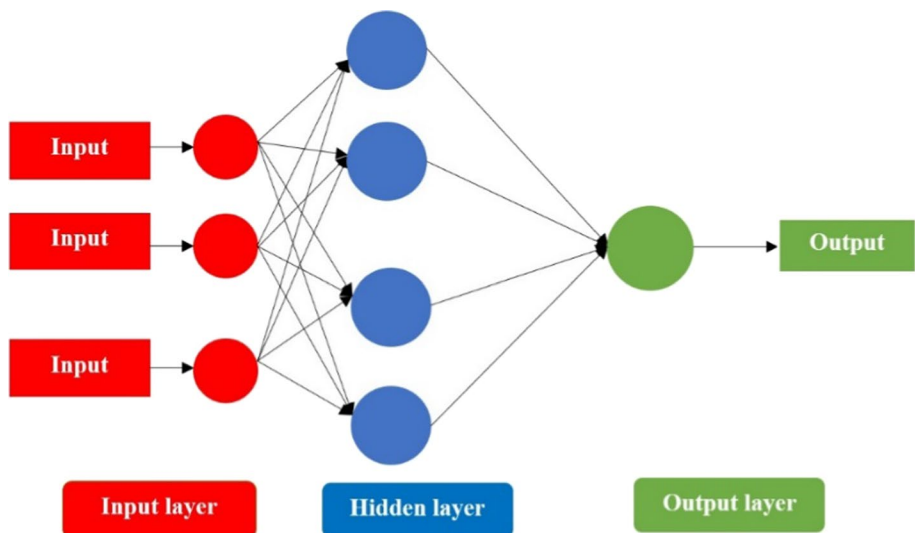


Fig. 6 Basic concept of SVM

$$y(x) = w^T \phi(x) + b \tag{7}$$

where $\phi(x)$ represents the high-dimensional function spaces, mapped nonlinearly from the x input space along with the weight vector W and bias b . By minimizing the regularized function $R(C)$, the coefficients w and b are estimated:

$$R(C) = C \frac{1}{N} \sum_{i=1}^N L_{\epsilon}(d_i, y_i) + \frac{1}{2} \|w\|^2 \tag{8}$$

$$L_{\epsilon}(d_i, y_i) = \begin{cases} |d_i - y_i| - \epsilon & \text{If } |d_i - y_i| \geq \epsilon \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

where N represents the total training dataset and d represents the target value of regression task. $L_{\epsilon}(d_i, y_i)$ stands for the ϵ -insensitive loss function, which return zero when the forecast value within the error tolerance ϵ . C represents the penalized coefficient in the optimization problem that

measures the trade off between the loss function and regularized term. And then Eq. (8) is transformed into the constraint function given by Eq. (10) by adding the following positive slack variables ξ_i and ξ_i^* as follows:

$$\text{Minimize : } R(C) = C \left(\sum_{i=1}^N (\xi_i + \xi_i^*) \right) + \frac{1}{2} \|w\|^2 \tag{10}$$

$$\text{Subject to: } \begin{cases} d_i - w\phi(x_i) - b_i \leq \epsilon + \xi_i, & i = 1, 2, \dots, N \\ w\phi(x_i) + b_i - y_i \leq \epsilon + \xi_i^*, & i = 1, 2, \dots, N \\ \xi_i, \xi_i^* \geq 0 & i = 1, 2, \dots, N \end{cases}$$

The first term ($\frac{1}{2} \|w\|^2$) is the weights vector norm. The parameters that governs the throughout regression process are the variables C , ϵ , and $\phi(x)$. When the x does not falls within the tube, there is bound to form ξ_i or ξ_i^* , and there a minimization is required. The SVM is then trained by minimizing the loss error $C \left(\sum_{i=1}^N (\xi_i + \xi_i^*) \right)$ and regularized penalty term $\frac{1}{2} \|w\|^2$. Since the kernel function establishes the feature space, selecting the appropriate kernel function is crucial in SVM regression. Four common types of SVM kernel functions were used as follows:

$$\text{Linear Kernel : } K(x_i, x_j) = x_i^T \times x_j \tag{11}$$

$$\text{Polynomial Kernel : } K(x_i, x_j) = (\gamma x_i^T \times x_j + r)^d, \quad \gamma > 0 \tag{12}$$

$$\text{Radial Basis Kernel : } K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right), \quad \gamma > 0 \tag{13}$$

$$\text{Sigmoid Kernel : } K(x_i, x_j) = \tanh(x_i^T \times x_j + r) \tag{14}$$

Herer, and d are kernel parameters.

Evaluation metrics

In this study, three different statistical metrics were used to validate the reliability of the proposed model: root mean square error (RMSE), coefficient of determination (R^2), and mean absolute error (MAE). RMSE and MAE are the two most commonly used metrics for determining the accuracy of continuous variables. They are excellent for estimating the average size of errors and penalizing values that deviate from observed values. Lower numbers are desired since they are negative-oriented scores. R^2 is a goodness-of-fit measure that provides a fitted regression line. This metric's value runs from 0 to 1, with 0 indicating no connection and 1 indicating predictions are identical to observed values. The following statistical indicators were used:

- The mean absolute error MAE:

$$\text{MAE} = \frac{\sum_{i=1}^n \text{ABS}(y_i - \lambda(x_i))}{n} \quad (15)$$

- The root mean square error RMSE:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\text{So} - \text{Sp})^2} \quad (16)$$

- The coefficient of determination R^2 :

$$R^2 = \left\{ (1/N) * \sum [(x_i - X) * (y_i - Y)] / (\sigma_x * \sigma_y^2) \right\} \quad (17)$$

Results and discussion

Assessing the best input and data partition

A preliminary study of input selection and data partition was performed prior to model development. Understanding the input selection is critical for model development. Also, due to the nature of a data-driven approach, determining input lag is merely another hyperparameter to be optimized (Wunsch et al. 2021). In the hydrological process, the significance of history is related to the time lag between input and output reactions, such as river discharge and sediment transport rate response, as reported in the previous study by Malik et al. (2019). Intriguingly, the analysis yielded the following input selection: A time lag of 1 was adequate as input to the model. Malik et al. (2019) performed a Gamma test to test out the impact of the input features on daily SSC. They demonstrated that models based on input lag 1 and 2 performed much better than other input lags

among the combinations tested. More crucially, input lag 1 corresponded to the variable sediment concentration, whereas input lag 2 corresponded to the variable streamflow. Following that, Idrees et al. (2021) confirmed that an input selection analysis based on input lag up to lag-1 was sufficient. They also indicated that the input combination would change depending on the techniques utilized. In agreement with them, this study also demonstrated that a time lag of 1 was adequate input to the model through another input feature selection known as autocorrelation function (ACF). It might appear to reveal that ACF may be an alternative in selecting the input lag variable to determine the magnitude of input lag despite the linearity property. Table 2 shows the considered parameter as input to the model. The selected input lag was sufficient for machine learning and deep learning models to learn from past data successfully. Under the same input lag, the data were divided into 80% training and 20% testing data.

Training for the machine learning and deep learning models

Different model architectures of LSTM, ANN, and SVM were developed to determine the best optimum set of hyperparameters for predicting sediment transport in the Johor river in Malaysia. A 200 number of hidden neurons selected was due to improved performance of ANN and LSTM at an iteration of 250. As for the SVM model, a linear support vector regression was used in this study, considering that the task required was a regression type of machine learning. The optimal hyperparameters for SVM were: capacity = 10.0, epsilon = 0.1, and RBF was the chosen kernel function with gamma equal to 1. These simulations were conducted on a workstation equipped with an Intel Core(TM) i7-7700HQ CPU and 16 GB of RAM. Intel®.

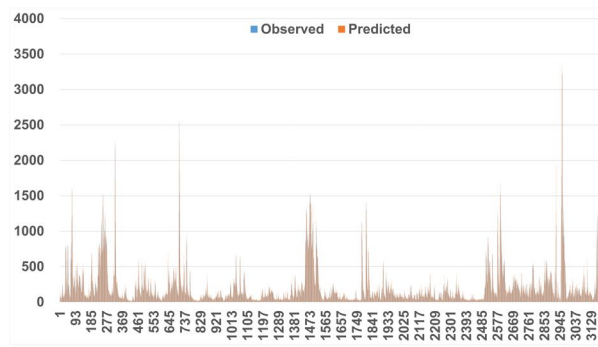
Evaluation of model performances

An infographic in the form of a line chart and scatter plot (Fig. 7) was generated to display the acquired findings of sediment transport prediction, ranging from standard machine learning such as ANN and SVM to deep learning LSTM. Figure 7 also shows how AI-based models could be employed effectively as an alternative model to predict sediment transport, with the estimated series displaying a significant correlation with the actual series. The low value of R^2 obtained ANN was also seen in a study by Yadav et al. (2022). It may imply that ANN requires some form of hybridization to achieve satisfactory results. Therefore, to improve the model performance, they adopted the GA to optimize the architecture of the ANN

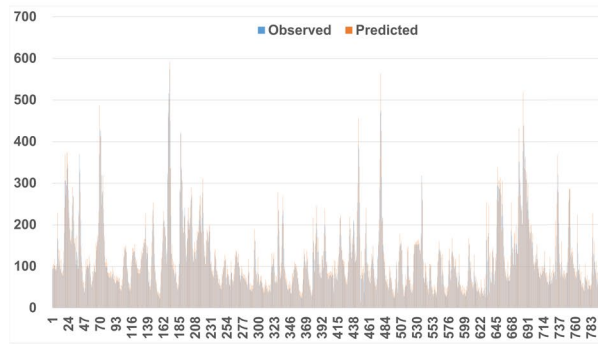
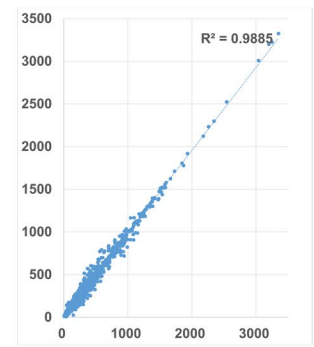
Table 2 Lag time input to the models

Model	Input	Output
Model I	S_{t-1}	S_t

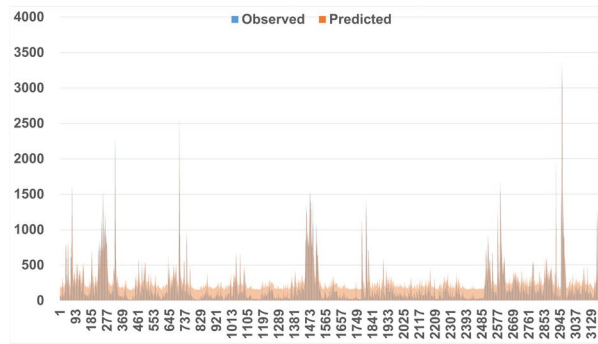
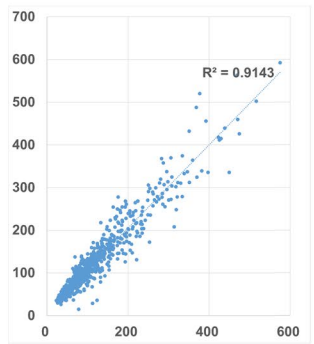
Fig. 7 Training and testing performance of the developed model



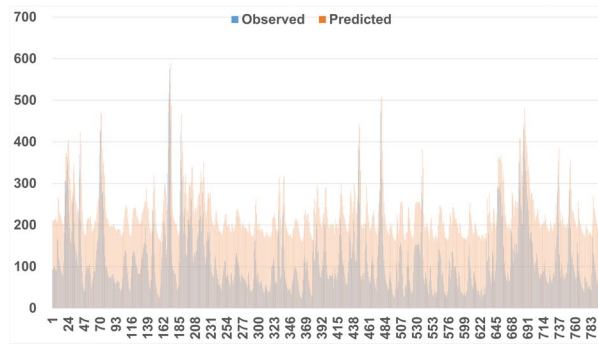
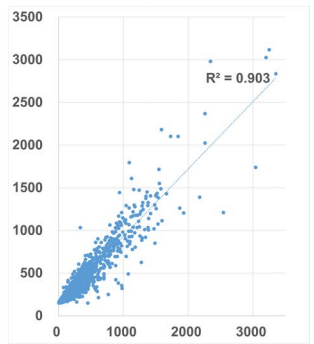
A) Training of LSTM



B) Testing of LSTM



C) Training of SVM



D) Testing of SVM

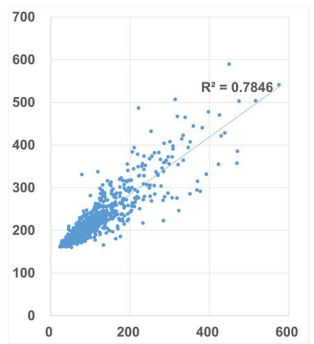
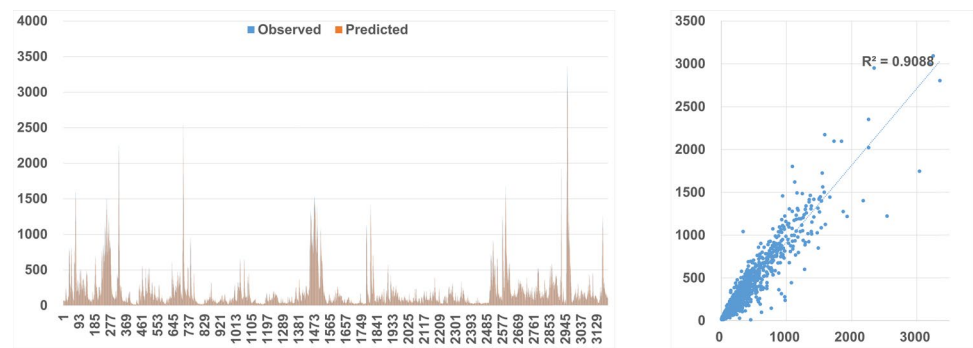
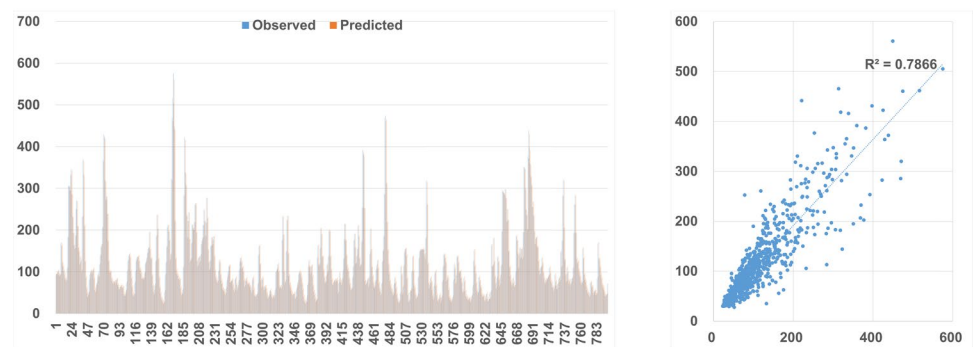


Fig. 7 (continued)



E) Training of ANN-MLP



F) Testing of ANN-MLP

model. While scatter plot has been used in most research studies, one minor downside of such graphical representation is that it is challenging to quantify how well or poorly the models perform.

On the contrary, the violin plots (Fig. 8) reveal a much clearer distinction between LSTM (deep learning), ANN, and SVM (machine learning). Except for the SVM model, all developed algorithms could consistently predict the sediment transport in terms of 4 measures statistics: minimum, first quartile, median, and third quartile. Among the developed models, LSTM was able to predict the maximum sediment transport more accurately than other algorithms. Besides, a violin plot also provides additional information, which is the distribution of the estimated data. Generally, the shape of the violin plot of LSTM and ANN-MLP is more similar to the measured value. Compared to them, the SVM model had higher dispersion of estimated data based on the violin plot than in the measured value. And worse yet, the model also failed to predict the maximum, lowest, and median sediment transport values.

Other comprehensive findings of the quantitative measures utilized to evaluate the performance of the constructed model (Table 3) are tabulated. Overall, LSTM scored the best in all three performance criteria, indicating the importance of such a model in estimating sediment transport. The

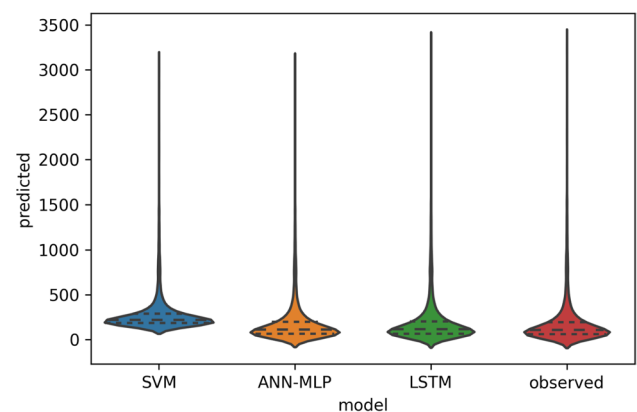


Fig. 8 Violin plot of predicted against. Measured of sediment transport

significance of these results in study employing such networks resides in their memory-storing capabilities as well as their universality in a sequence data application. ANN performed moderately in between LSTM and SVM. Finally, SVM had the lowest accuracy in terms of all the measure metrics. Such evidence was also shown in a study by Al-Mukhtar (2019), in which they found that ANN was a better choice than SVM in their comparison, therefore recognizing

Table 3 Summary of the developed model performances

Models	type	Evaluation metrics		
		MAE	RMSE	R2
LSTM	Training	20.105	31.436	0.989
	Testing	18.094	11.395	0.914
ANN-MLP	Training	39.260	85.945	0.909
	Testing	24.794	16.998	0.914
SVM	Training	118.404	135.080	0.904
	Testing	115.093	53.340	0.914

that SVM performs well in sediment transport prediction. However, research conducted by Shafaghat and Dezvareh (2020) discovered that ANN outperformed SVM, albeit by a modest margin of roughly 4%. As stated in the introduction, because they work as universal function approximations based on their compliance with the nature of the problem, it is generally preferable to make a comparison in each conducted experiment. Another key statistic used to quantify the model's capacity to respond to new data is the generalizability of the developed model. It is frequently described as the variation in a model's performance between training and test sets of data derived from the same probability. Since AI-based model used in this work is a supervised learning approach, overfitting might be a possible concern. Not only does LSTM have better training dataset accuracy, but it also performs better in the testing dataset assessment metric. It demonstrates how well LSTM is in predicting the unforeseen dataset. It ought to also be noted that the testing error is lower than the training error, which might be attributable to sampling bias in the testing dataset.

The prior section's adopted measure discovered the best model for predicting river sediment. However, without benchmarking criteria, determining which model performs better is difficult. While RMSE and MAE are crucial measurements for comparing models, they provide little to no information on which model is best suited for a specific job. Therefore, an RMSE-Standard Deviation Ratio (RSR) was employed and defined in Eq. 1 to deliver a far more meaningful comparison result. Table 4 shows the RSR index ranges in terms of performance rate and class. The findings were tabulated; Fig. 9 reveals that the application of the LSTM model is related to the lowest RSR value, followed by ANN and SVM. It is discovered that the SVM model for this dataset is incapable of extracting the underlying relationship between input and output parameters.

$$RSR = \frac{\sqrt{\sum_{i=1}^N (S_i^{obs} - S_i^{sim})^2}}{\sqrt{\sum_{i=1}^N (S_i^{obs} - S_i^{mean})^2}} \tag{18}$$

where S_i^{mean} is the mean derived from the simulation dataset (Ehteram et al. 2020).

Conclusions

Predictions of sediment transport are vital for forecasting flood events, monitoring coastal erosion, water resource planning, and irrigation management. Three models of AI, such as LSTM, ANN, and SVM, were developed and then validated to assess their performance in sediment transport prediction. The findings in this study showed that the input lag selection based on ACF showed that a lag value of 1 was adequate for the model to train. It implies that ACF can work as an alternative input lag selection aside from Gamma tests. Despite the fact that all of the developed models had a high R2 value, ranging from 0.79 to 0.91, giving the impression that they performed similarly, they were rather diverse. The outperform of ANN and SVM might be attributed to information loss while dealing with the long-term dependency on sediment transport time series. While this study may confirm or contradict some previous findings, the general trend revealed that they operate effectively in sediment transport, provided that both are being utilized and compared simultaneously. Another importance finding in this study is that, in practice, the developed LSTM deep learning model may be used as an alternative model for predicting the time-varying and spatially distributed sediment transport as it outperforms better than either SVM or ANN. Deep learning models with

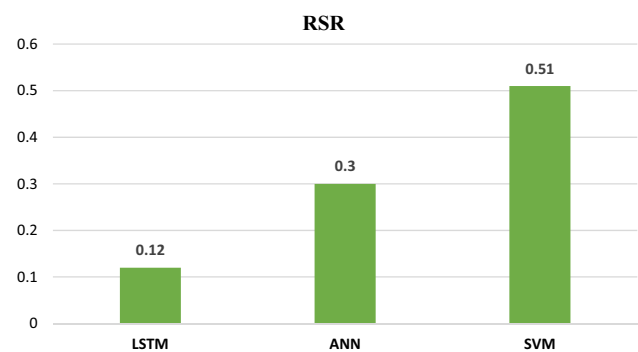


Fig. 9 RSR value for the computed models

Table 4 RSR range and the corresponding performance rate

Performance rating	Unsatisfactory	Satisfactory	Good	Very good
RSR value	$RSR > 0.7$	$0.6 \leq RSR \leq 0.7$	$0.5 \leq RSR \leq 0.6$	$0.00 \leq RSR \leq 0.5$

additional feature extraction capabilities, such as memory storage capability, might be thought of as a unique contribution to sediment transport prediction. Nonetheless, incorporating the geographical state of the area into deep learning algorithms such as LSTM in anticipating sediment migration might be a significant finding in future studies.

There is, however, still opportunity for improvement. This study primarily focuses on the standalone model, even though hybridization may be crucial. In fact, the inclusion of hybridization would frequently lead to model enhancement by introducing components, such as metaheuristic algorithms or data preprocessing, to complement its current model. These new components can improve the training procedure of the model through better searchability or reduce the complexity in the dataset for the model to learn the pattern easier. However, because the potential of deep learning LSTM for its memory-storing capacity to anticipate sediment movement is of particular relevance in the current investigation, these hybridization can be studied further. Also, it is worth mentioning that climate changes play a vital role in characterizing the variation in the river flow, which increase erosion, resulting in an increase in the sediment load movement. The absence of climatic components in the model will increase the uncertainty of the outcome. Last but not least, working on more than one station may give more substantial evidence to illustrate the capabilities of the developed models.

Acknowledgements This research was supported by the Ministry of Education (MOE) through Fundamental Research Grant Scheme (FRGS/1/2020/TK0/UNITEN/02/16).

Author contributions SDL was responsible for investigation, conceptualization, methodology, formal analysis, software, writing the original draft, and writing—reviewing and editing. KLC, ANA, and YF were involved in investigation, conceptualization, visualization, and writing—reviewing and editing. MS and AE-S contributed to visualization, validation, and writing—reviewing and editing.

Funding No funding was received for conducting this study.

Data availability The data that support the findings of this study can be obtained from the corresponding author upon requested.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical approval This article does not contain any studies involving human participants or animals performed by any of the authors.

Informed consent All of the authors have consented to submit the manuscript to Applied Water Sciences Journal.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes

were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Addo-Bediako A, Nukeri S, Kekana M (2021) Heavy metal and metalloid contamination in the sediments of the Spekboom River. *South Afr Appl Water Sci* 11(7):1–9
- Afan HA, El-Shafie A, Yaseen ZM, Hameed MM, Wan Mohtar WHM, Hussain A (2015) ANN based sediment prediction model utilizing different input scenarios. *Water Resour Manage* 29(4):1231–1245
- Al-Mukhtar M (2019) Random forest, support vector machine, and neural networks to modelling suspended sediment in Tigris River-Baghdad. *Environ Monit Assess* 191(11):1–12
- Asadi H, Dastorani MT, Sidle RC, Shahedi K (2021) Improving flow discharge-suspended sediment relations: intelligent algorithms versus data separation. *Water* 13(24):3650
- Bandini F, Sunding TP, Linde J, Smith O, Jensen IK, Köppl CJ, Butts M, Bauer-Gottwein P (2020) Unmanned aerial system (UAS) observations of water surface elevation in a small stream: comparison of radar altimetry, LIDAR and photogrammetry techniques. *Remote Sens Environ* 237:111487
- Chong KL, Lai SH, Yao Y, Ahmed AN, Jaafar WZW, El-Shafie A (2020) Performance enhancement model for rainfall forecasting utilizing integrated wavelet-convolutional neural network. *Water Resour Manage* 34(8):2371–2387
- Cortes C, Vapnik V (1995) Support-Vector Networks. *Mach Learn* 20(3):273–297
- Ebtehaj I, Bonakdari H, Sharifi A (2014) Design criteria for sediment transport in sewers based on self-cleansing concept. *J Zhejiang Univ Sci A* 15(11):914–924
- Ebtehaj I, Bonakdari H, Shamshirband S, Mohammadi K (2016) A combined support vector machine-wavelet transform model for prediction of sediment transport in sewer. *Flow Meas Instrum* 47:19–27
- Ehteram M, Ahmed AN, Ling L, Fai CM, Latif SD, Afan HA, Banadkooki FB, El-Shafie A (2020) Pipeline scour rates prediction-based model utilizing a multilayer perceptron-colliding body algorithm. *Water* 12(3):902
- Fassoni-Andrade AC, de Paiva RCD (2019) Mapping spatial-temporal sediment dynamics of river-floodplains in the Amazon. *Remote Sens Environ* 221:94–107
- Fiyadh SS, AlOmar MK, Binti Jaafar WZ, AlSaadi MA, Fayaed SS, Binti Koting S, Lai SH, Chow MF, Ahmed AN, El-Shafie A (2019) Artificial neural network approach for modelling of mercury ions removal from water using functionalized CNTs with deep eutectic solvent. *Int J Mol Sci* 20(17):4206
- Goldstein EB, Coco G, Plant NG (2019) A review of machine learning applications to coastal sediment transport and morphodynamics. *Earth Sci Rev* 194:97–108
- Hapsari D, Onishi T, Imaizumi F, Noda K, Senge M (2019) The use of sediment rating curve under its limitations to estimate the suspended load. *Rev Agric Sci* 7:88–101
- Harada E, Gotoh H, Ikari H, Khayyer A (2019) Numerical simulation for sediment transport using MPS-DEM coupling model. *Adv Water Resour* 129:354–364
- Idrees MB, Jehanzaib M, Kim D, Kim T-W (2021) Comprehensive evaluation of machine learning models for suspended sediment

- load inflow prediction in a reservoir. *Stoch Env Res Risk Assess* 35(9):1805–1823
- Katimon A, Shahid S, Mohsenipour M (2018) Modeling water quality and hydrological variables using ARIMA: a case study of Johor river. *Malays Sustain Water Resour Manag* 4(4):991–998
- Kaveh K, Kaveh H, Bui MD, Rutschmann P (2021) Long short-term memory for predicting daily suspended sediment concentration. *Eng Comput* 37(3):2013–2027
- Kumar M, Kumar P, Kumar A, Elbeltagi A, Kuriqi A (2022) Modeling stage–discharge–sediment using support vector machine and artificial neural network coupled with wavelet transform. *Appl Water Sci* 12(5):1–21
- Kuriqi A, Koçileri G, Ardiçlioğlu M (2020) Potential of Meyer-Peter and Müller approach for estimation of bed-load sediment transport under different hydraulic regimes. *Model Earth Syst Environ* 6(1):129–137
- Li Z, Tian X, Shu L, Xu X, Hu B (2017) Emotion recognition from EEG using RASM and LSTM. Springer, Berlin, pp 310–318
- Liu Y, Zarfl C, Basu NB, Cirpka OA (2019) Turnover and legacy of sediment-associated PAH in a baseflow-dominated river. *Sci Total Environ* 671:754–764
- Lu C-M, Chiang L-C (2019) Assessment of sediment transport functions with the modified SWAT-Twn model for a Taiwanese small mountainous watershed. *Water* 11(9):1749
- Ma H, Nittrouer JA, Wu B, Lamb MP, Zhang Y, Mohrig D, Fu X, Naito K, Wang Y, Moodie AJ (2020) Universal relation with regime transition for sediment transport in fine-grained rivers. *Proc Natl Acad Sci* 117(1):171–176
- Malik A, Kumar A, Kisi O, Shiri J (2019) Evaluating the performance of four different heuristic approaches with Gamma test for daily suspended sediment concentration modeling. *Environ Sci Pollut Res* 26(22):22670–22687
- Mao L, Comiti F, Carrillo R, Penna D (2019) Geomorphology of proglacial systems. Springer, Berlin, pp 199–217
- Matos T, Faria CL, Martins MS, Henriques R, Gomes P, Goncalves LM (2019) Development of a cost-effective optical sensor for continuous monitoring of turbidity and suspended particulate matter in marine environment. *Sensors* 19(20):4439
- Měžíně J, Ferrarin C, Vaičūtė D, Idzelytė R, Zemlys P, Umgieser G (2019) Sediment transport mechanisms in a lagoon with high river discharge and sediment loading. *Water* 11(10):1970
- Rahman KU, Pham QB, Jadoon KZ, Shahid M, Kushwaha DP, Duan Z, Mohammadi B, Khedher KM, Anh DT (2022) Comparison of machine learning and process-based SWAT model in simulating streamflow in the Upper Indus Basin. *Appl Water Sci* 12(8):1–19
- Roushangar K, Shahnazi S (2020) Prediction of sediment transport rates in gravel-bed rivers using Gaussian process regression. *J Hydroinf* 22(2):249–262
- Sathya K, Nagarajan K, Carlin Geor Malar G, Rajalakshmi S, Raja Lakshmi P (2022) A comprehensive review on comparison among effluent treatment methods and modern methods of treatment of industrial wastewater effluent from different sources. *Appl Water Sci* 12(4):1–27
- Schmidhuber J, Hochreiter S (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780
- Shafaghat M, Dezvareh R (2020) Predicting the sediment rate of Nakhilo Port using artificial intelligence. *Int J Coast Offshore Eng* 4(2):41–49
- Vo Q-H, Nguyen H-T, Le B, Nguyen M-L (2017) Multi-channel LSTM-CNN model for Vietnamese sentiment analysis. *IEEE, New York*, pp 24–29
- Wesselman D, De Winter R, Oost A, Hoekstra P, Van der Vegt M (2019) The effect of washer geometry on sediment transport during inundation events. *Geomorphology* 327:28–47
- Wilkes MA, Gittins JR, Mathers KL, Mason R, Casas-Mulet R, Vanzo D, Mckenzie M, Murray-Bligh J, England J, Gurnell A (2019) Physical and biological controls on fine sediment transport and storage in rivers. *Wiley Interdiscip Rev Water* 6(2):e1331
- Wunsch A, Liesch T, Broda S (2021) Groundwater level forecasting with artificial neural networks: a comparison of long short-term memory (LSTM), convolutional neural networks (CNNs), and non-linear autoregressive networks with exogenous input (NARX). *Hydrol Earth Syst Sci* 25(3):1671–1687
- Xing W, Qian Y, Guan X, Yang T, Wu H (2020) A novel cellular automata model integrated with deep learning for dynamic spatio-temporal land use change simulation. *Comput Geosci* 137:104430
- Yadav A, Joshi D, Kumar V, Mohapatra H, Iwendi C, Gadekallu TR (2022) Capability and robustness of novel hybridized artificial intelligence technique for sediment yield modeling in Godavari river. *India Water* 14(12):1917
- Żarczyński M, Szmańda J, Tylmann W (2019) Grain-size distribution and structural characteristics of varved sediments from Lake Żabińskie (Northeastern Poland). *Quaternary* 2(1):8
- Zhang X, Yang Y (2020) Suspended sediment concentration forecast based on CEEMDAN-GRU model. *Water Supply* 20(5):1787–1798
- Zounemat-Kermani M, Fadaee M, Adarsh SH, R, (2020) Predicting Sediment transport in sewers using integrative harmony search-ANN model and factor analysis. IOP Publishing, Bristol, p 012004
- Samal K, Babu K, Das S (2021) Spatio-temporal prediction of air quality using distance based interpolation and deep learning techniques. *EAI Endorsed Trans Smart Cities* 5(14)
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I (2017) Attention is all you need. *Adv neural inf Process syst*

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.