International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 25s, 2025
https://theaspd.com/index.php

Hybrid Ensemble AND Deep Learning Models FOR Suspended Sediment Concentration Prediction IN THE Subarnarekha River Basin

Pratik Ranjan[1*], Ramakar Jha[2]
[1]Research Scholar, Department of Civil Engineering, National Institute of Technology Patna, Patna – 800005, India pratikranjan151@gmail.com
[2]HAG Professor, Department of Civil Engineering, National Institute of Technology Patna, Patna – 800005, India

Abstract:
Accurate prediction of suspended sediment concentration (SSC) is critical for water resources management, sediment control, and reservoir operation. This paper compares the performance of the Random Forest (RF), XGBoost, Long Short-Term Memory (LSTM), and Stacked Ensemble in predicting SSC on a daily basis with different lag-values (1-5 days). Model performance was assessed using the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and Nash–Sutcliffe efficiency (NSE). The results show that the inclusion of antecedent conditions enhances the predictive ability considerably, with a 2-day lag configuration producing the best predictive accuracy (R^2 = 0.96, NSE = 0.95; N). Tree-based methods (both RF and XGBoost) were more effective in capturing nonlinear responses and extremes compared to baseline models, whereas LSTM was fairly good at capturing sequential dependencies and essentially smooth peaks. The Stacked Ensemble model performed consistently better than all standalone models with respect to accuracy, stability, and variance reproduction between training and testing phases, as shown in scatter plots, Taylor diagrams and time-series simulation. The above results emphasize the need to apply hybrid methods to SSC prediction. The suggested ensemble framework would provide a robust and transferrable approach to the hydrological and water management community to anticipate and manage sediment and ensure that floods are controlled, and basin-scale systems sustainable.
Keywords: *Machine Learning, Random Forest, Extreme Gradient Boosting, Long Short-Term Memory, Stacked Ensemble, suspended sediment concentration prediction*

1. INTRODUCTION

The transportation of sediment in rivers is a serious issue for water resources management as it is associated with water quality degradation, reservoir siltation, dam safety, navigability, aquatic habitats, hydropower efficiency, and soil erosion (Kaveh et al., 2017; Francke et al., 2008). Suspended sediment concentration (SSC), especially, is a matter of great concern, as it is both a physical and chemical pollutant through increasing the turbidity, as well as transported adsorbed contaminants (Doğan et al., 2007). Accurate estimation of SSC is vital in hydraulic project planning, sustainability, and planning of the watershed (Kisi & Zounemat-Kermani, 2016). Direct measurements of SSC are limited by the intensive and expensive sampling exercises, leading to scarce coverage in most areas (Al-Mukhtar & Al-Yaseen, 2019). In contrast, streamflow and water level data are largely accessible, and they are key determinants of sediment movement (Vafakhah, 2013), forming favourable predictors of SSC estimation.

Conventional methods of SSC prediction, which include sediment rating curves, regression models and process-based hydrological models, have notable shortcomings. Physically based and conceptual models often demand extensive input data that exceed availability in many basins (Kalbus et al., 2012). Similarly, empirical approaches cannot represent dynamic, nonlinear, and stochastic behaviors of sediment transport, which results in a poor general prediction (Afan et al., 2016). Shiau and Chen (2015) emphasized that sediment rating curves are inadequate to capture the observed dispersion between sediment and discharge, which reduces their applicability. Recent advances on sediment modeling confirm these drawbacks and note that more flexible models are necessary that can address nonlinearity and uncertainty (Yue et al., 2024, Wang et al., 2025, Szalińska et al., 2024). These difficulties highlight the requirement of more comprehensive, flexible modelling frameworks that will be able to capture the dynamics of sediment processes across a spectrum of hydrological conditions.

Data-driven methods, and in particular Artificial Intelligence (AI), have become strong alternatives as they have the capability of capturing non-linear behavior and are able to process noisy hydrological data. Machine learning algorithms like Random Forest (RF), Extreme Gradient Boosting (XGBoost) and model

structures like Long Short-Term Memory (LSTM) have been used successfully to perform hydrological forecasting, including streamflow and water quality prediction (Breiman, 2001; Chen & Guestrin, 2016; Kratzert et al., 2019; Khosravi et al., 2024; Fan et al., 2025; Li et al., 2021). The LSTM networks are well adapted to time-series forecasting problems due to the gated design that suppresses the short-term dependencies (Feng et al., 2020), whereas RF and XGBoost are considered to be robust, scalable, and easy to understand.

Although this advanced, majority of previous studies were based on single-model framework, restricted datasets, or the use of a small number of input variables. Very few have systematically compared deep learning, ensemble machine learning, and hybrid approaches for sediment forecasting in data-scarce, monsoon-driven river systems. Other recent efforts incorporated tree-based models with optimization (Mirzakhani et al., 2022) or hybridized LSTM with boosting methods on extreme events (Slater et al., 2023; Fan et al., 2025), however, both used shorter datasets and did not provide strong quantification of uncertainty. Recent reviews stress that ensemble and hybrid frameworks can offer resilience by combining complementary model strengths (Dehghan-Souraki et al., 2024).

In a bid to bridge these gaps, the study compares and evaluates four state-of-the-art AI frameworks (LSTM, RF, XGBoost, and a Stacked Ensemble) in predicting suspended sediment concentration (SSC) in the Subarnarekha River at Jamshedpur, India. The analysis leverages a rare 51-year dataset (1972–2023) of daily discharge and water level observations, combined with lagged predictor selection (via auto- and partial correlation), k-fold cross-validation, and uncertainty quantification. By systematically benchmarking these approaches, the study makes three distinctive contributions: (i) it provides one of the most comprehensive long-term evaluations of AI-based SSC forecasting in an Indian river system, (ii) it shows the potential of stacked ensembles to exceed individual models by capturing complementary strengths, and (iii) it provides pragmatic information to the management of the sediment issue in monsoon-influenced rivers, wherein precise prediction is a key element in reducing the risk of floods, operating reservoirs and in the control of water quality.

2. Study Area

The Subarnarekha is one of the longest east-flowing interstate river. It originates at Nagri village in Ranchi district, Jharkhand, at an elevation of 997 m. The river is approximately 395 kilometers long. The river's primary tributaries are Kanchi, Kharkai, and Karkari. The basin location is in the north east region of India, between latitudes 21°33' 0" N to 23°32' 0" N and longitudes 85°09' 0" E to 87°27' 0" E. The basin is enclosed by the Chhotnagpur Plateau in the north-west, the Brahmani basin in the south-west, the Burhabalang basin in the south, and the Bay of Bengal in the South-East. The basin has a total catchment area of 18,951 square kilometers.

The Jamshedpur gaging station, founded in 1972, is located on the Subarnarekha River in Jamshedpur, India at latitude 22°49'00" N and longitude 86°12'39" E. The drainage area of the Subarnarekha River up to Jamshedpur station is 12649 km^2 (Figure 1). Based on the DEM, the elevation ranges between -64 m and 997 m.

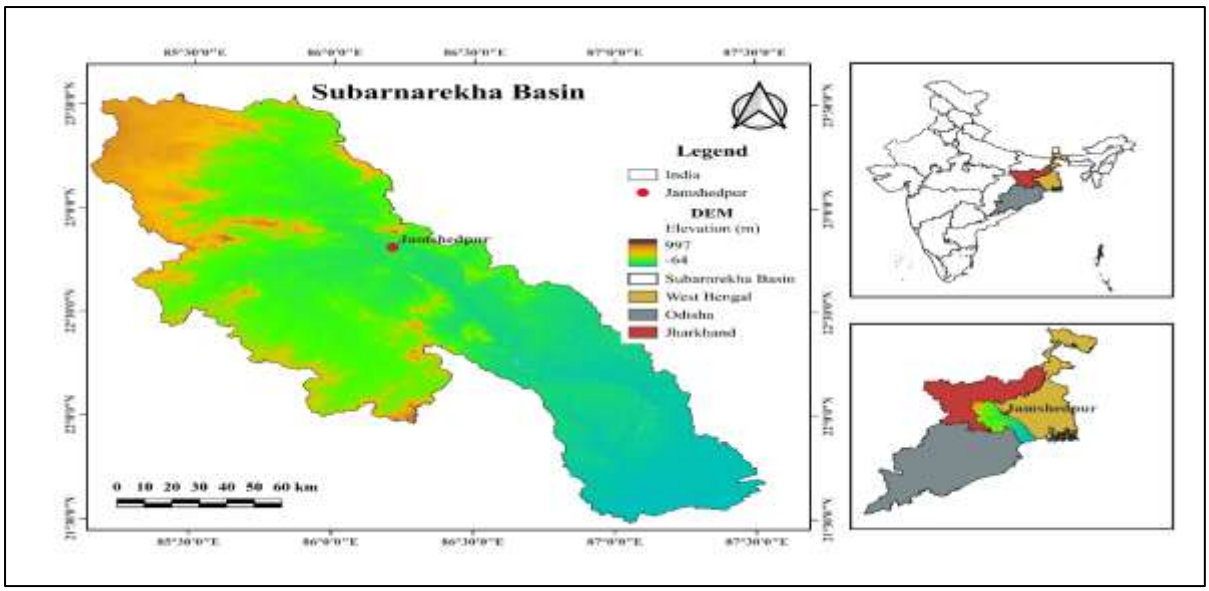

Figure 1- Location of Jamshedpur Station in Subarnarekha Basin

3. Dataset and Statistical Analysis

For the purpose of this study, the data collected for the Jamshedpur Station were daily discharge (m^3/s), water level (m) and their corresponding suspended sediment concentrations (g/l) from CWC, Bhubaneshwar, Odisha, India. Daily discharge (m^3/s), water level (m) and suspended sediment concentrations (g/l) data were used to predict suspended sediment concentrations (g/l) during the period 1972–2023. The dataset spans a duration of 51 years, from 27 November 1972 to 31 March 2023. The raw data indicate large outliers in discharge (Q), water level (WL), and suspended sediment concentration (SSC), which were reduced using IQR and Isolation Forest filtering. The cleaned dataset more accurately depicts the underlying variability and is more suited for model training and interpretation. It was divided into the training (80% of dataset) and test (20% of dataset) purposes to build the predictive SSC models. Table 1 summarizes the descriptive statistics of discharge (Q), water level (WL), and suspended sediment concentration (SSC). Discharge shows high variability (CV = 141.6%) and strong positive skewness (C_{sx} = 2.10), indicating that occasional high-flow events strongly influence the distribution. WL is comparatively stable (CV = 0.35%), with only minor deviations around the mean stage. SSC exhibits low mean values (0.055 g/L) but substantial variability (CV = 83.6%) and moderate positive skewness (C_{sx} = 1.42), reflecting its sensitivity to episodic transport processes. Overall, the statistics indicate that while stage remains stable, discharge variability governs SSC dynamics, underlining the importance of incorporating nonlinear and event-responsive predictors in the modeling framework.

Table 1 - Descriptive statistics of discharge (Q), water level (WL), and suspended sediment concentration (SSC) at the study site.

Cleaned Dataset	**Mean**	**SD**	**CV (%)**	**Skewness (C_{sx})**	**Max**	**Min**
Q(m^3/s)	57.658	81.625	141.568	2.103	426.897	0.04
WL(m)	115.231	0.4	0.347	1.065	116.8	113.93
SSC(g/l)	0.055	0.046	83.573	1.422	0.26	0

4. METHODS

The methodological approach to the prediction of suspended sediment concentration (SSC) involved a logical process of curation, feature engineering, and model building (Figure 2). Raw data were processed through imputation of missing data and removal of outliers, prediction variables were selected and lagged variables were created with normalization employed to ensure compatibility across feature sets. The dataset was partitioned into training (80%) and testing (20%) subsets. Four predictive models were developed: Random Forest (RF), Extreme Gradient Boosting (XGBoost), Long Short-Term Memory (LSTM), and a Stacked Ensemble integrating the base learners through a linear regression meta-learner using k-fold cross-validation. Model training involved systematic hyperparameter optimization, and their predictive accuracy was evaluated using multiple statistical indicators (R^2, RMSE, MAE, NSE). The best-performing model was identified as the final predictor of SSC.

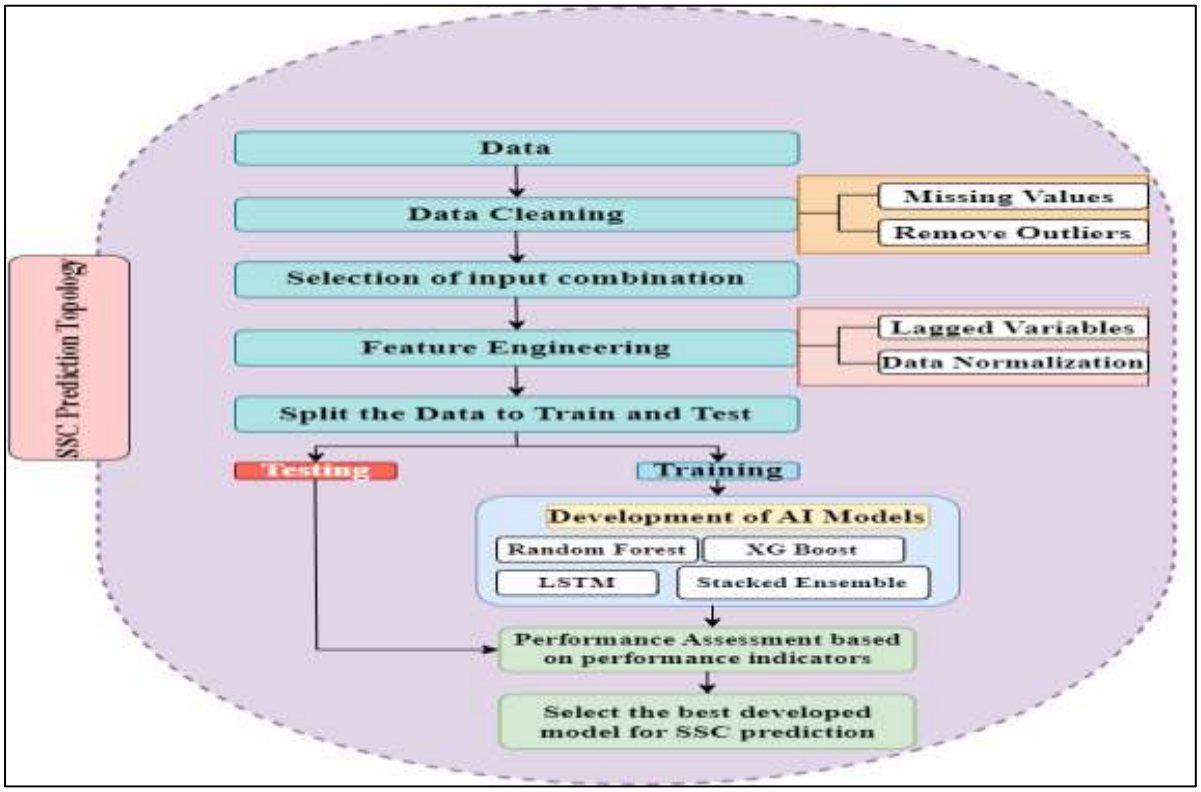

Figure 2- Illustrates the overall methodological framework

4.1. Deep learning models used to predict SSC

4.1.1. Long Short Term Memory (LSTM)

Hochreiter and Schmidhuber (1997) proposed the LSTM model, an extension of RNN that uses three components for prediction. The fundamental formulas for the LSTM structure are as follows:

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

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

$$o_t = \sigma(U_o x_t + W_o h_{t-1} + b_o) \quad (3)$$

$$h_t = \sigma_t * \tanh(c_t) \quad (4)$$

Where x_t, h_{t-1}, and o_t indicate input, hidden, and output layers respectively. U and W indicate the weights in the gates (input (i_t), forget (g_t), and output (o_t)). b and ht are the bias term and the hidden state, respectively. The sigmoid was chosen as the activation function.

4.1.2. Random Forest (RF)

Random Forest (RF) (Breiman, 2001) is an ensemble of decision trees trained using bagging, with each tree growing on a bootstrapped sample of data. At each split, a random subset of characteristics is evaluated, introducing decorrelation between trees and improving resilience against overfitting. The ensemble prediction for input **x** is:

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \quad (5)$$

where B denotes the number of trees and $T_b(x)$ the prediction of the b-th tree. RF's inherent feature importance ranking (Gini impurity reduction) and scalability for high-dimensional data enable it to be an adaptable tool for classification and regression problems (Liaw & Wiener, 2002).

4.1.3. Extreme Gradient Boosting (XGBoost)

XGBoost (Chen & Guestrin, 2016) is a gradient-boosting framework that optimizes a regularized objective function using additive tree models, L2-norm penalty, and gradient-based methods. The objective at iteration t is:

$$\mathcal{L}^t = \sum_{i=1}^{n} l\left(y_i \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \quad (6)$$

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda\|w\|^2 \quad (7)$$

where f_t is the t-th tree, T the number of leaves, w leaf weights, and γ, λ regularization hyperparameters. XGBoost's handling of sparse data, parallelized tree construction, and early halting features have established it as a cutting-edge approach for structured data (Ke et al., 2017).

4.1.4. Stacked Ensemble

Stacked generalization (Wolpert, 1992) uses a meta-learner to best aggregate predictions from various base models (e.g., RF, XGBoost). To avoid data leakage, base learner outputs were generated using k-fold cross-validation, where out-of-fold predictions served as meta-features. The meta-learner f_{meta} was then trained on these features to yield the final stacked prediction: $\{h_k(x)\}_{k=1}^{K}$:

$$\hat{y}_{stack} = f_{meta}(h_1(x), \ldots\ldots, h_k(x)) \quad (8)$$

This framework takes advantage of the complementary nature of heterogeneous models and can outperform individual learners, especially on complex hydrological (van der Laan et al., 2007).

4.2. Hyperparameter, tunning, and optimization of four DL models used for the prediction of SSL

Hyperparameter tuning was performed to ensure robust SSC forecasts. For Random Forest (RF), parameters including n_estimators (100, 200), max_depth (5, 10), and min_samples_split (2, 5) were optimized via grid search with 3-fold cross-validation. XGBoost (XGB) was tuned over n_estimators (100, 200), max_depth (5, 10), and learning_rate (0.01, 0.1), targeting the highest R^2 and lowest error metrics. The LSTM model had 50 and 30 units at the first and the second layer, respectively, dropout (0.2), and was trained during 50 epochs with a 16-batches batch size via the Adam optimizer. Standardized and also reshaped input features were adjusted to fit into the requirements of sequential learning.

The Stacked Ensemble combined RF, XGB, and LSTM as base learners. Base model predictions were generated using k-fold cross-validation and transformed into meta-features. A linear regression meta-learner was then trained on these features to integrate outputs, ensuring a balanced trade-off between accuracy and generalizability

Table 2 - Tuned Hyperparameters and Methodological Steps for RF, XGB, LSTM, and Stacked Ensemble

Model	**Tuned Hyperparameters**
RF	*n_estimators*: 100, 200; *max_depth*: 5, 10; *min_samples_split*: 2, 5; grid search with 3-fold CV
XGB	*n_estimators*: 100, 200; *max_depth*: 5, 10; *learning_rate*: 0.01, 0.1; optimized for max R^2 & min error
LSTM	Two LSTM layers (50, 30 units), dropout = 0.2; Adam optimizer; 50 epochs; batch size = 16; standardized sequential inputs
Stacked Ensemble	Base learners: RF, XGB, LSTM; training with k-fold CV; meta-features → **Linear Regression meta-learner** for optimal model combination

5. RESULT AND DISCUSSION

The predictive performance of RF, XGBoost, LSTM, and their Stacked Ensemble was assessed for suspended sediment concentration (SSC) using lag times from 1 to 5 days. In this framework, lag days denote how many antecedent observations are incorporated as predictors; for example, a 2-day lag considers conditions from both the previous day and the day before to estimate today's SSC. This formulation captures short-term hydrological memory, which is critical for sediment mobilization.

Model performance across lag days

Table - 3 gives the statistical assessment measures (R^2, RMSE, MAE and NSE) of all models with different lag structures. As it can be seen, the 2-day lag yielded the best predictive precision, both R^2 and NSE were greater than 0.96 and 0.95 in the majority of cases. Among the models tested, Stacked Ensemble produced the best results during the training and testing process followed by RF and XGBoost, with LSTM performing a little bit less, especially reproducing extreme values of SSC. These numerical results add to the graphical patterns depicted in the following illustrations.

Table 3 - Statistical performance (R^2, RMSE, MAE, NSE) of Random Forest, XGBoost, LSTM, and Stacked Ensemble models for predicting suspended sediment concentration (SSC) across different lag times (1–5 days) during training and testing.

Model Name	**Lag (Days)**	**Training**	**Testing**

		R^2	RMSE	MAE	NSE	R^2	RMSE	MAE	NSE
Random Forest	1	0.958	0.009	0.005	0.958	0.897	0.015	0.007	0.897
	2	**0.973**	**0.007**	**0.003**	**0.973**	**0.937**	**0.012**	**0.005**	**0.937**
	3	0.898	0.015	0.007	0.898	0.902	0.015	0.007	0.902
	4	0.895	0.015	0.007	0.895	0.898	0.015	0.007	0.898
	5	0.897	0.015	0.007	0.897	0.885	0.016	0.007	0.885
XGBoost	1	0.951	0.010	0.005	0.951	0.899	0.015	0.007	0.899
	2	**0.963**	**0.009**	**0.004**	**0.963**	**0.940**	**0.011**	**0.005**	**0.940**
	3	0.939	0.011	0.005	0.939	0.905	0.014	0.007	0.905
	4	0.937	0.011	0.006	0.937	0.889	0.015	0.007	0.889
	5	0.936	0.012	0.006	0.936	0.882	0.016	0.007	0.882
LSTM	1	0.899	0.015	0.009	0.899	0.888	0.015	0.009	0.888
	2	**0.921**	**0.013**	**0.007**	**0.921**	**0.933**	**0.012**	**0.007**	**0.933**
	3	0.876	0.016	0.010	0.876	0.889	0.015	0.010	0.889
	4	0.886	0.015	0.008	0.886	0.886	0.016	0.008	0.886
	5	0.889	0.015	0.008	0.889	0.878	0.016	0.009	0.878
Stacked Ensemble	1	0.949	0.010	0.005	0.949	0.901	0.014	0.007	0.901
	2	**0.964**	**0.009**	**0.004**	**0.964**	**0.941**	**0.011**	**0.005**	**0.941**
	3	0.918	0.013	0.006	0.918	0.906	0.014	0.007	0.906
	4	0.915	0.013	0.006	0.915	0.896	0.015	0.007	0.896
	5	0.916	0.013	0.006	0.916	0.886	0.016	0.007	0.886

Across all models, the 2-day lag configuration yielded the best results, with the highest R^2 (>0.96) and NSE (>0.95) and the lowest error indices (RMSE ≈ 0.009–0.011). This demonstrates that sediment response in the basin is primarily governed by conditions in the preceding 48 hours, while longer lags (>2 days) added less relevant information and slightly reduced predictive accuracy. The line plots of statistical metrics (Figure 3) clearly illustrate this trend, with the Ensemble consistently outperforming individual models across all lag structures.

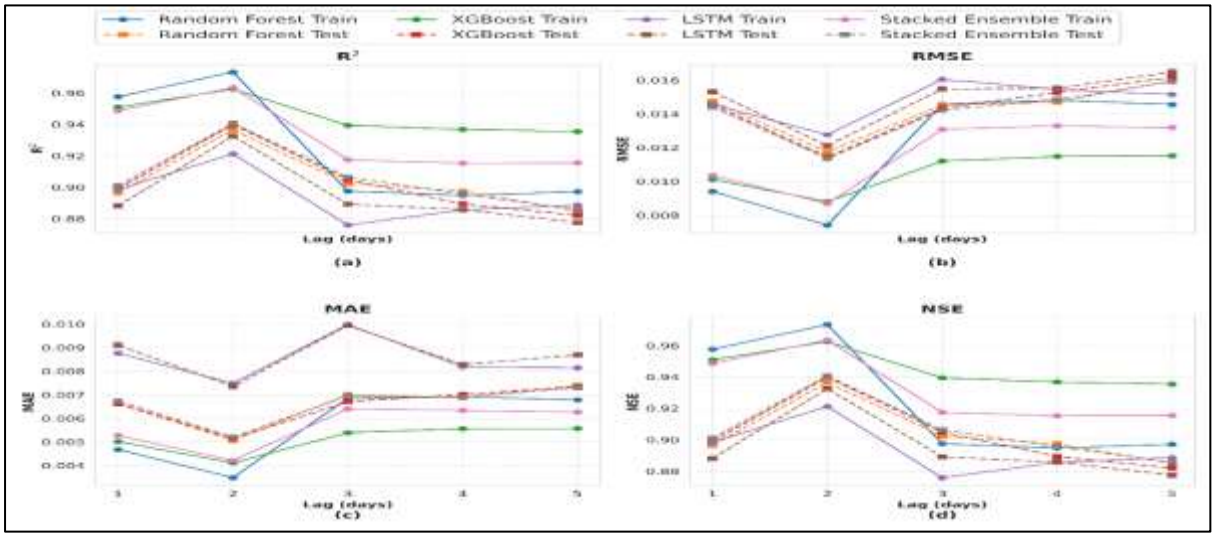

Figure 3- Training and testing performance of Random Forest (RF), XGBoost, LSTM, and Stacked Ensemble models across different lag days. Subplots show evaluation metrics: (a) R^2, (b) RMSE, (c) MAE, and (d) NSE. Solid lines represent training performance, while dashed lines indicate testing performance.

Observed vs. predicted relationships

Scatter plots (Figure - 4) of observed versus predicted SSC confirm these findings. All models showed strong positive correlation, but the regression trend line was most closely aligned with the observed data

in the case of the Ensemble. Unlike a 1:1 reference line, the plotted line represents the fitted regression trend line, which for the Ensemble had a slope near unity and $R^2 > 0.94$ during testing. RF and XGBoost also demonstrated good agreement, while LSTM showed a tendency to smooth peaks and underestimate higher SSC values.

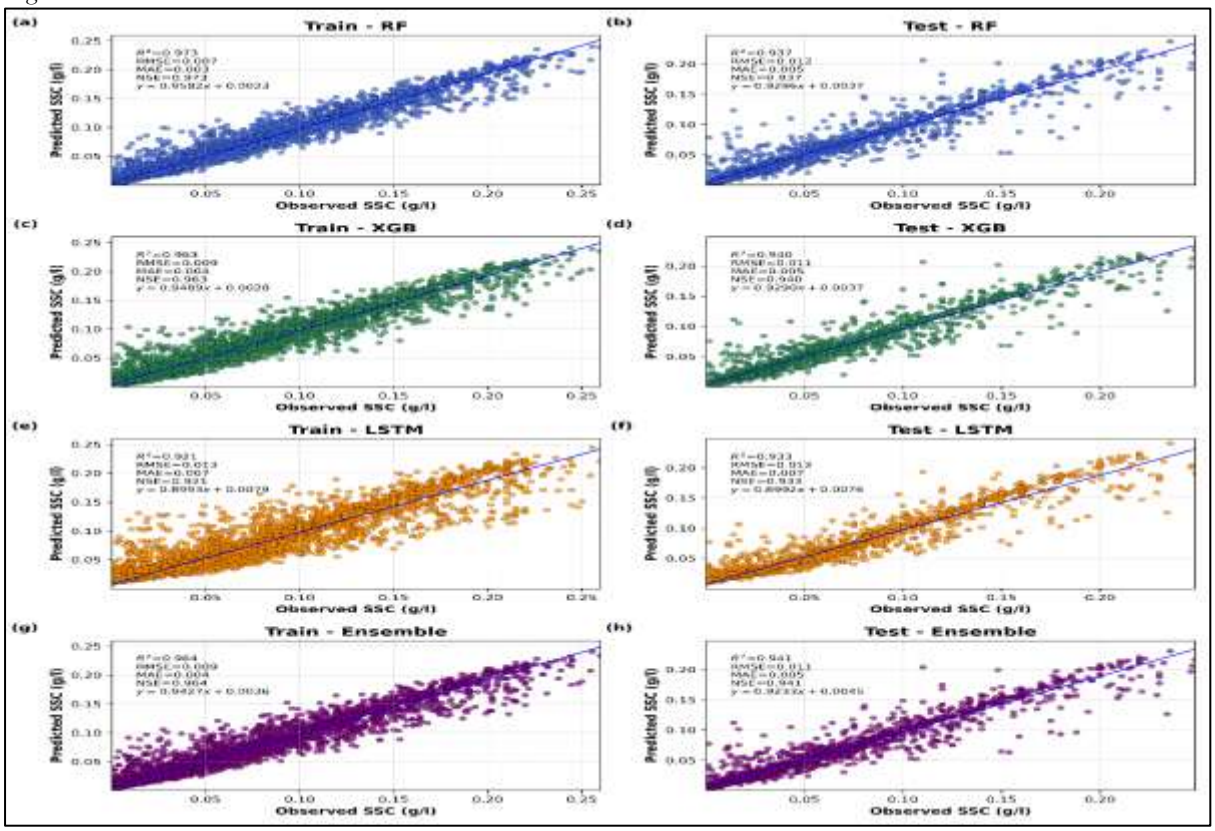

Figure 4- Scatter plots of observed versus predicted suspended sediment concentration (SSC) during training and testing phases for (a, b) Random Forest, (c, d) XGBoost, (e, f) LSTM, and (g, h) Stacked Ensemble models.

Taylor diagram analysis

Taylor diagrams provided further insights into model behaviour during both training (Figure 5) and testing (Figure 6). In the training phase, the Ensemble was located nearest to the reference point, with correlation coefficients exceeding 0.96 and nearly identical variance reproduction, confirming its structural fidelity. RF and XGBoost also clustered close to the reference, while LSTM, though correlated ($r \approx 0.92$), underestimated variance. In the testing phase, the Ensemble maintained the closest proximity to the observed reference point, whereas LSTM's variance deviation became more pronounced. The consistency of RF and XGBoost across both phases highlights their stability, but the Ensemble clearly provided the most balanced representation of correlation, variability, and error minimization

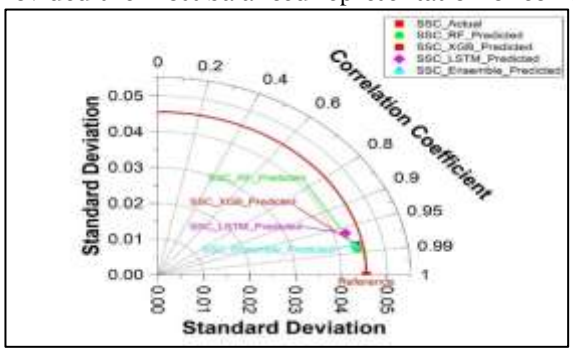

Figure 5 - Taylor diagram for training dataset showing comparative performance of RF, XGB, LSTM, and Stacked Ensemble models in predicting SSC

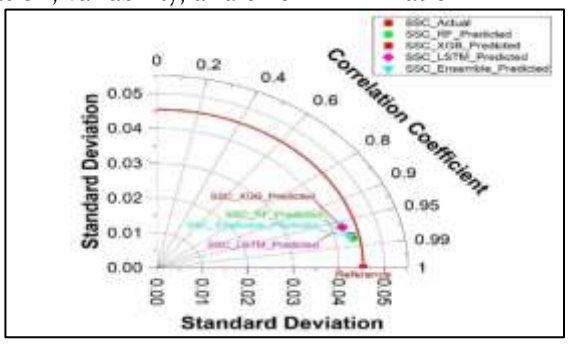

Figure 6 - Taylor diagram for testing dataset showing comparative performance of RF, XGB, LSTM, and Stacked Ensemble models in predicting SSC.

Temporal dynamics of SSC prediction

Time series plots (Figure - 7) of observed and predicted SSC further underline the differences among models. The Ensemble tracked both seasonal cycles and episodic high-flow sediment pulses with high fidelity, capturing both the magnitude and timing of peaks. RF and XGBoost also performed strongly, although with slight over- or under-estimation during extremes. LSTM predictions appeared smoother, effectively reproducing baseline sediment levels but underrepresenting sharp peaks, which aligns with its scatter and Taylor diagram patterns.

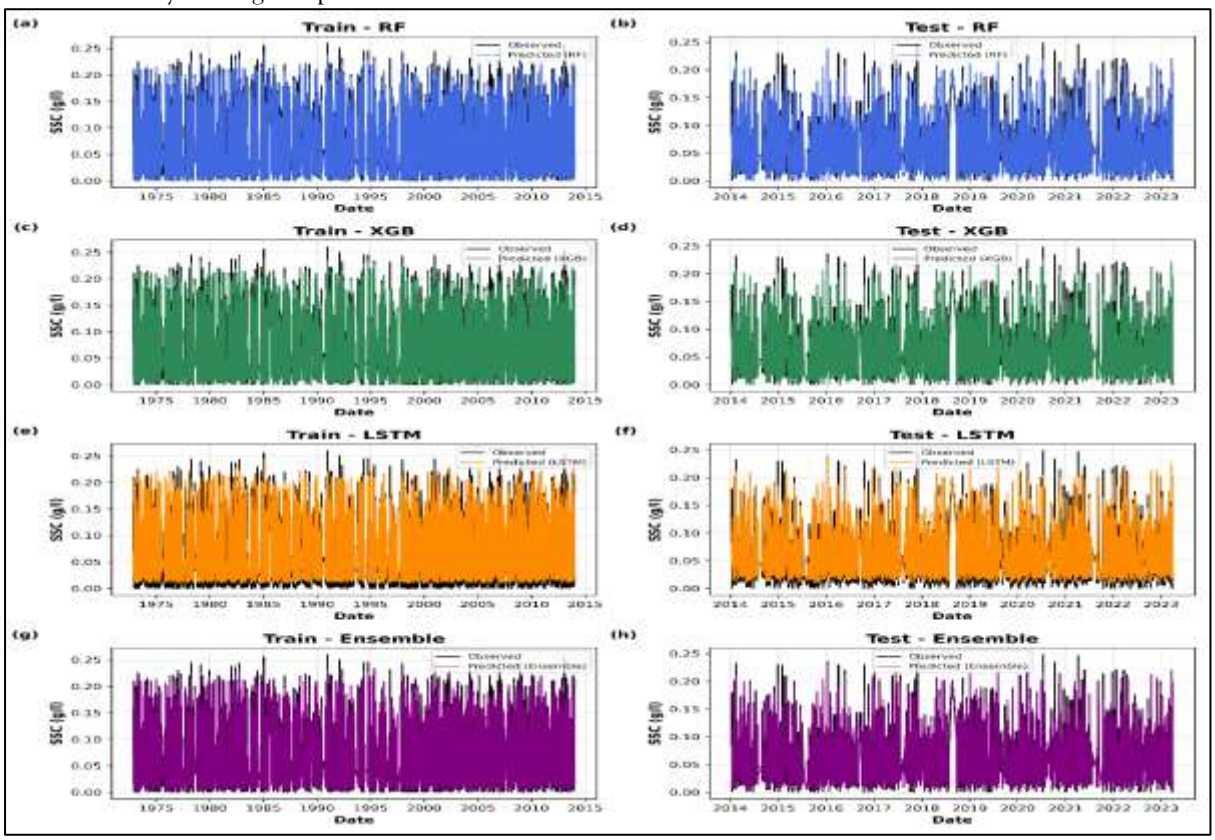

Figure 7 - Time series comparison of observed and predicted suspended sediment concentration (SSC) during training and testing phases for (a, b) Random Forest, (c, d) XGBoost, (e, f) LSTM, and (g, h) Stacked Ensemble models.

To summarize, it is possible to note three key things that the table and figures reveal: (i) a lagged antecedent condition is beneficial in increasing the predicting ability, with 2-day lag being the most optimal; (ii) tree-based models (RF, XGBoost) capture nonlinear responses and extremes, while LSTM emphasizes sequential dependencies but tends to smooth peaks; and, (iii) the Stacked Ensemble model effectively supplements these strengths to produce a balanced, robust, and explicable framework of SSC prediction.

Implications for Practice

With better performance outcomes, the Stacked Ensemble demonstrates the prospect to become a stable decision-support system in sediment forecasting. It allows water managers to predict more accurately the level of both baseline and peak sediment, and hence predict future surges of sediment, optimize reservoir management, and design erosion control. Its resistance to data noise and nonlinearities is also a strong reason to use it in basins where flood-driven sediment pulses and long-term accumulation endanger infrastructure and water security.

6. CONCLUSION

This study evaluated the predictive skill of RF, XGBoost, LSTM, and their Stacked Ensemble for SSC prediction under varying lag structures. The results highlight three major findings:

1. Lag day sensitivity - A 2-day lag provided the most informative predictor set, emphasize the short-term memory effect in sediment mobilization.
2. Model complementarity - RF and XGBoost captured nonlinear responses and extremes, LSTM addressed temporal dependencies, but the Ensemble integrated these capabilities to overcome individual limitations.

3. Ensemble superiority - Among all figures, metrics, and datasets, the Stacked Ensemble consistently outperformed standalone models, producing higher correlations, lower errors, and improved structural fidelity, as validated by scatter plots, Taylor diagrams and time series comparisons.

Overall, Stacked Ensemble framework provides a robust, accurate, and generalizable approach for SSC prediction. By harnessing lagged information on hydrological behaviour and by integrating the complementary performance characteristics of varied learning methods, it presents a realistic avenue toward enhanced sediment forecasting, basin-scale management of water resources, and design of sediment management strategies.

Acknowledgements: The authors wish to thanks to anonymous reviewers and to all the members of CWC, Bhubaneshwar, Odisha, India for providing data and other relevant information that adds in facilitating this study.

Conflict of Interest: The authors have no conflicts of interest to declare that are relevant to the content of this article.

Funding: No external funding has been received for this research work.

REFERENCES

4. Afan, H. A., El-shafie, A., Mohtar, W. H. M. W., & Yaseen, Z. M. (2016). Past, present and prospect of an Artificial Intelligence (AI) based model for sediment transport prediction. Journal of Hydrology, 541, 902-913.
5. Al-Mukhtar, M., & Al-Yaseen, F. (2019). Modeling water quality parameters using data-driven models, a case study Abu-Ziriq marsh in south of Iraq. Hydrology, 6(1), 24.
6. Breiman, L. (2001). Random forests. Machine learning, 45, 5-32.
7. Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).
8. Dehghan-Souraki, D., López-Gómez, D., Bladé-Castellet, E., Larese, A., & Sanz-Ramos, M. (2024). Optimizing sediment transport models by using the Monte Carlo simulation and deep neural network (DNN): A case study of the Riba-Roja reservoir. *Environmental Modelling & Software*, *175*, 105979.
9. Doğan, E., Yüksel, İ., & Kişi, Ö. (2007). Estimation of total sediment load concentration obtained by experimental study using artificial neural networks. Environmental fluid mechanics, 7, 271-288.
10. Fan, J., Li, R., Zhao, M., & Pan, X. (2025). A BiLSTM-Based Hybrid Ensemble Approach for Forecasting Suspended Sediment Concentrations: Application to the Upper Yellow River. *Land*, *14*(6), 1199.
11. Feng, D., Fang, K., & Shen, C. (2020). Enhancing streamflow forecast and extracting insights using long-short term memory networks with data integration at continental scales. Water Resources Research, 56(9), e2019WR026793.
12. Francke, T., López-Tarazón, J. A., & Schröder, B. (2008). Estimation of suspended sediment concentration and yield using linear models, random forests and quantile regression forests. Hydrological Processes: An International Journal, 22(25), 4892-4904.
13. Kaveh, K., Bui, M. D., & Rutschmann, P. (2017). A comparative study of three different learning algorithms applied to ANFIS for predicting daily suspended sediment concentration. International Journal of Sediment Research, 32(3), 340-350.
14. Kalbus, E., Kalbacher, T., Kolditz, O., Krüger, E., Seegert, J., Röstel, G., ... & Krebs, P. (2012). Integrated water resources management under different hydrological, climatic and socio-economic conditions. Environmental Earth Sciences, 65, 1363-1366.
15. Khosravi, K., Farooque, A. A., Bateni, S. M., Jun, C., Mohammadi, D., Kalantari, Z., & Cooper, J. R. (2024). Fluvial bedload transport modelling: advanced ensemble tree-based models or optimized deep learning algorithms?. *Engineering Applications of Computational Fluid Mechanics*, *18*(1), 2346221.
16. Kisi, O., & Zounemat-Kermani, M. (2016). Suspended sediment modeling using neuro-fuzzy embedded fuzzy c-means clustering technique. Water resources management, 30, 3979-3994.
17. Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., & Nearing, G. (2019). Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. Hydrology and Earth System Sciences, 23(12), 5089-5110.
18. Li, H. Y., Tan, Z., Ma, H., Zhu, Z., Abeshu, G., Zhu, S., ... & Leung, L. Y. R. (2021). A new large-scale suspended sediment model and its application over the United States. *Hydrology and Earth System Sciences Discussions*, *2021*, 1-44.
19. Mirzakhani, G., Ghanbari-Adivi, E., Fattahi, R., Ehteram, M., Mosavi, A., Ahmed, A. N., & El-Shafieg, A. (2022). Application of group method of data handling and new optimization algorithms for predicting sediment transport rate under vegetation cover. *arXiv preprint arXiv:2209.09623*.
20. Shiau, J. T., & Chen, T. J. (2015). Quantile regression-based probabilistic estimation scheme for daily and annual suspended sediment loads. Water Resources Management, 29, 2805-2818.
21. Slater, L. J., Arnal, L., Boucher, M. A., Chang, A. Y. Y., Moulds, S., Murphy, C., ... & Zappa, M. (2023). Hybrid forecasting: blending climate predictions with AI models. *Hydrology and earth system sciences*, *27*(9), 1865-1889.
22. Szalińska, E., Orlińska-Woźniak, P., Wilk, P., Jakusik, E., Skalák, P., Wypych, A., & Arnold, J. (2024). Sediment load assessments under climate change scenarios and a lack of integration between climatologists and environmental modelers. Scientific Reports, 14(1), 21727.
23. Vafakhah, M. (2013). Comparison of cokriging and adaptive neuro-fuzzy inference system models for suspended sediment load forecasting. Arabian Journal of Geosciences, 6, 3003-3018.
24. Wang, N., Wu, G., Wang, K., You, Z., & Zhuang, X. (2025). Explainable data-driven modeling of suspended sediment concentration at a deltaic marsh boundary under river regulation and storm events. *Coastal Engineering*, *198*, 104722.