DOI: 10.5604/01.3001.0014.8489

Volume 104 • Issue 2 • February 2021

Journal of Achievements in Materials and Manufacturing Engineering

International Scientific Journal published monthly by the World Academy of Materials and Manufacturing Engineering

Comparative study for deriving stage-discharge–sediment concentration relationships using soft computing techniques

P. Sihag [a,*], M.R. Sadikhani [b], V. Vambol [c], S. Vambol [d], A.K. Prabhakar [e], N. Sharma [a]

[a] Civil engineering Department, Shoolini University, Solan, Himachal Pradesh, India

[b] Soil science Department, College of Agriculture, Lorestan University, Iran

[c] Public Agency "National Scientific and Research Institute of Industrial Safety and Occupational Safety and Health", Kyiv, Ukraine

[d] Kharkiv Petro Vasylenko National Technical University of Agriculture, Kharkiv, Ukraine

[e] Civil Engineering Department, National Institute of Kurukshetra, Haryana, India

* Corresponding e-mail address: parveen12sihag@gmail.com

ORCID identifier: https://orcid.org/0000-0002-7761-0603 (P.S.);
https://orcid.org/0000-0002-8229-3956 (V.V.);
https://orcid.org/0000-0002-8376-9020 (S.V.)

ABSTRACT

Purpose: Knowledge of sediment load carried by any river is essential for designing and planning of hydro power and irrigation projects. So the aim of this study is to develop and evaluating the best soft-computing-based model with M5P and Random Forest regression-based techniques for computation of sediment using datasets of daily discharge, daily gauge and sediment load at the Champua gauging site of the Upper Baitarani river basin of India.

Design/methodology/approach: Last few decades, the soft computing techniques based models have been successfully used in water resources modelling and estimation. In this study, the potential of tree based models are examined by developing and comparing sediment load prediction models, based on M5P tree and Random forest regression (RF). Several M5P and RF based models have been applied to a gauging site of the Baitarani River at Odisha, India. To evaluate the performance of the selected M5P and RF-based models, three most popular statistical parameters are selected such as coefficient of correlation, root mean square error and mean absolute error.

Findings: A comparison of the results suggested that RF-based model could be applied successfully for the prediction of sediment load concentration with a relatively higher magnitude of prediction accuracy. In RF-based models Q_t, $Q_{(t-1)}$, $Q_{(t-2)}$, $S_{(t-1)}$, $S_{(t-2)}$, H_t and $H_{(t-1)}$ combination based M10 model work superior than other combination based models. Another major outcome of this investigation is Q_t, $Q_{(t-1)}$ and $S_{(t-1)}$ based model M4 works better than other input combination based models using M5P technique. The optimum input combination is Q_t, $Q_{(t-1)}$ and $S_{(t-1)}$ for the prediction of sediment load concentration of the Baitarani River at Odisha, India.

Research limitations/implications: The developed models were tested for Baitarani River at Odisha, India.

Keywords: Sediment load concentration, Baitarani river, M5P, Random Forest

Reference to this paper should be given in the following way:

P. Sihag, M.R. Sadikhani, V. Vambol, S. Vambol, A.K. Prabhakar, N. Sharma, Comparative study for deriving stage-discharge–sediment concentration relationships using soft computing techniques, Journal of Achievements in Materials and Manufacturing Engineering 104/2 (2021) 57-76. DOI: https://doi.org/10.5604/01.3001.0014.8489

ANALYSIS AND MODELLING

1. Introduction

Soil erosion is one of the most serious concerns not even a country like India but all over the world today, as it will deteriorate the agricultural and natural environment. Soil degradation with time would cause a miserable situation when agricultural efforts are focussed on increasing food production. The assessment of the volume of sediments being transported by a river is important for the estimation of sediment transport in rivers, design of dams, reservoirs and channels, environmental impact assessment, and determination of the efficacy of watershed management and other catchment treatment. According to the Global Assessment of Human-Induced Soil Degradation, there are 1.9 billion hectares affected by soil degradation worldwide, 850 million hectares of which are within the Asia-Pacific region, accounting for about 24% of the total regional land area [1]. In India, there are 175 Mha constitutes about; around 53% of the total geographical area is affected by land degradation and soil erosion [2].

In General, Sediment erosion rate depends on runoff generation due to the consequence of rainfall occurring. The frequency and magnitude of extreme rain events in rising trends in most part of central India in monsoon seasons [3]; this would cause more sediment erosion. Sediment rating curves are quite useful to get the impression of sediment flows with the discharge through the river gauging site from the catchment area. Sediment rating curves relating instantaneous sediment flux to discharge were established by Van Dijk et al. in work [4] for suspended, bedload and total sediment by fitting a power equation to all water discharge–sediment-discharge data pairs and no extra variation in sediment load was explained by runoff stage (i.e. rising or falling) and, therefore, a single curve was used.

Estimate the quantity of daily sediment load discharge which moves in a river is very significant as an indicator to measure the soil erosion loss, water quality, Management and Planning of reservoir, irrigation system, dam etc. [5]. Several researchers were working in the field of sediment load prediction [6-10].

Last few decades soft computing based models such as artificial neural network (ANN), Support vector machine (SVM), Wavelet-based least square support vector machine model (WLSSVM), Multiple linear regression (MLR) were used for the prediction of sediment load concentration in river [11-17]. In work [7] Melessea et al. compared the performance of ANN and MLR based models for the prediction of sediment load. The artificial neural network is used, with three different learning algorithms to predict the quantity of sediment load discharge in river is: Broyden-Fletcher-Goldfarb-Shanno Quasi-Newton (BFGS), Scaled Conjugate Gradient (SCD) and Levenberg-Marquardt (LM). Comparison of results shows that ANN based model perform better than MLR based model. In research [10] Roushangar & Shahnazi has used Gaussian Process regression and support vector machine based models for the prediction of sediment load. In his study, it was found that Gaussian process regression shows better results as compare to support vector machine and other applied regression based models. From this study they conclude that Gaussian process regression based model is performing better than SVM based model for the prediction of sediment load in Gravel-bed Rivers in the United States. In research [6] Nagy et al. has compared the ANN based model with conventional models and found that due to some uncertainty and the stochastic nature of the sediment movement, ANN model was found outperforming than conventional models. Nhu et al. [18] compared the soft computing techniques such as M5P, Random forest and random tree to predict the daily water level of Zrebar Lake and came to conclusion that M5P shows better results compared to other techniques. Sharafati et al. [19] studied machine learning models to predict the suspended sediment load such as gradient boost regression, AdaBoost regression and random forest regression. Their predictions were compared and found that random forest regression model predict better outcome as compared to other regression methods. This paper demonstrates the applicability of a data-driven soft-computing-based approach in developing the gauge–discharge–sediment relationship. Keeping in the view improved performance of M5P and RF based models So the aim of this study is evaluating the best soft-computing-based model with M5P and Random Forest regression-based techniques for computation of sediment using datasets of daily discharge, daily gauge and sediment load at the Champua gauging site of the Upper Baitarani river basin of India.

2. Materials and methods

2.1. Soft computing techniques

M5P tree regression: Quinlan [20] initially introduced the M5P model tree. This model is a binary decision tree that has a linear function at the last nodes (leaf). M5P model tree is valuable in linear and nonlinear problem-solving. The divide-and-overcome scheme generates Tree-based models (Fig. 1). A tree model has been developed based on two different stages. The first stage includes splitting criteria to form a decision tree. Due to the splitting process, the information in secondary nodes has conformity as compared to the significant node [20]. After examining all the possible splits, M5P selected a model that has minimized the error to a more substantial extent. The reason for this isolation is to allow the growth of the stable structure of the tree, which may produce overfitting. Substituting a subtree with a leaf helps to overcome this issue. However, the tree should be pruned back for illustration of this process.

2.2. Random Forest Regression

Random Forest (Fig. 2) is an adaptable assembly of decision trees that performs smoothly for linear and nonlinear estimation by adjusting bias and variance [21,22]. This assembly learning process is identified as ‘bagging’ as it develops tree lines which don't support ancient trees to be a base for consecutive trees. Every tree undergoes a distinct estimation process using a bootstrap sample of the dataset. Consequently, a single generic vote stands out for the final prediction [23]. RF model requires two specific standard users to define parameters: the number of variables (m) selected at every node to develop a tree and the number of trees to be produced (*k*).

2.3. The goodness of fit evaluation parameters

To analyse the performance of various modelling approaches, correlation coefficient (CC), root mean square error (RMSE) and mean absolute error (MAE) were calculated using training the testing dataset.

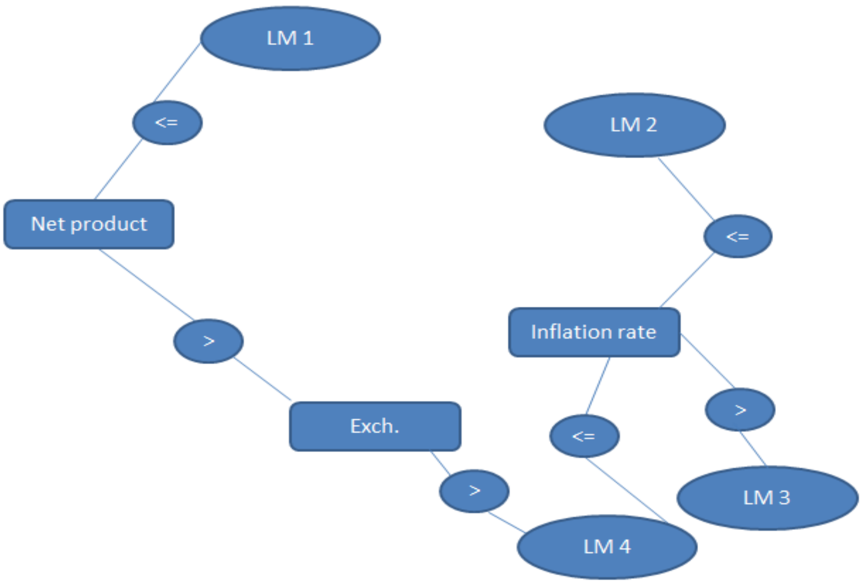


Fig. 1. General structure of M5P

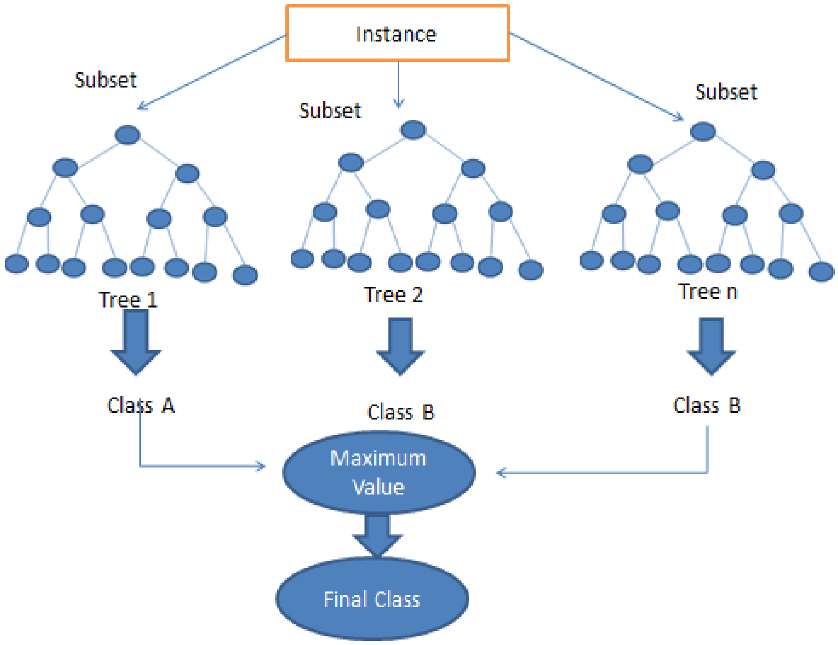


Fig. 2. General structure of Random forest

<u>Coefficient of Correlation</u>

The coefficient of correlation is used to measure the success of numeric prediction. The coefficient of correlation (CC) is computed as

$$CC = \frac{n\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{\sqrt{n(\sum_{i=1}^{n} {x_i}^2) - (\sum_{i=1}^{n} x_i^2)}\sqrt{n(\sum_{i=1}^{n} {y_i}^2) - (\sum_{i=1}^{n} y_i^2)}}, \quad (1)$$

where x_i = observed values, y_i = predicted values, n = number of observations.

<u>Root Mean Square Error (RMSE)</u>

Mean-square error is the most commonly used measure of success of numeric prediction, and root mean-squared error is the square root of mean-squared-error after we give it the same dimensions as the predicted values themselves. This method exaggerates the prediction error – the difference between prediction value and actual value. The root mean squared error (RMSE) is computed as:

$$RMSE = \sqrt{\frac{1}{n}(\sum_{i=1}^{n}(x_i - y_i)^2} \quad (2)$$

<u>Mean Absolute Error (MAE)</u>

The mean absolute error is used to measure of success of numeric estimation. The mean absolute error (MAE) is computed as:

$$MAE = \frac{1}{n}(\sum_{i=1}^{n}|x_i - y_i| \quad (3)$$

2.4. Study area

The study area is part of the Upper Baitarani river basin located in Odisha state of India. The study area is about 1815 km^2 and lies in between 8509'42.66" to 85044'10.42" E longitude and 2106'52.92" to 22011'51.65" N latitude in the Baitarani River basin. The location map of the area is given in Figure 3.The rainfall received in the basin is mainly from southwest monsoon and lasts from June to October. About 78% of annual precipitation occurs during these months. The annual average rainfall over the study is about 1438 mm [24]. Most of the living population is tribal communities. Cultivation is the basic source of livelihood for rural masses, and the majority of soil in the study area is sandy loam in texture.

2.5. Model development

Selection of input variables is the initial step in developing soft computing based models. Several researchers have developed model using discharge, gauge and sediment load concentration at time step. In this study, model was developed using time series of discharge, time series of sediment load concentration and time series of gauge.

<u>Discharge dependent models</u>

$M1: \; S_t = f(Q_t, Q_{t-1})$
$M2: \; S_t = f(Q_t, Q_{t-1}, Q_{t-2})$

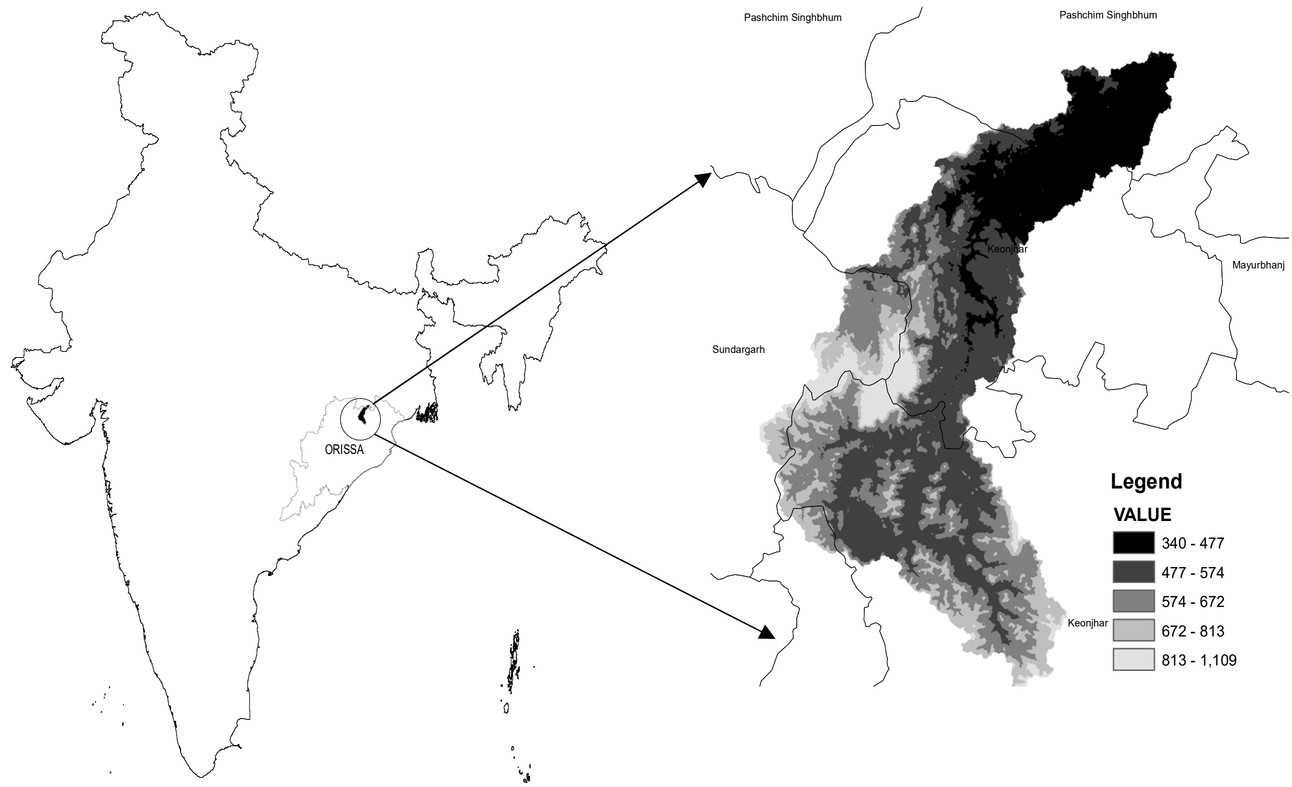


Fig. 3. Location map of study area (Upper Baitarani River Basin)

<u>Discharge and sediment load concentration dependent models</u>

$M3$: $S_t = f(Q_t, S_{t-1})$
$M4$: $S_t = f(Q_t, Q_{t-1}, S_{t-1})$
$M5$: $S_t = f(Q_t, Q_{t-1}, S_{t-1}, S_{t-2})$
$M6$: $S_t = f(Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}, S_{t-2})$
$M7$: $S_t = f(Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}, S_{t-2}, S_{t-3})$

<u>Discharge, sediment load concentration and gauge dependent models</u>

$M8$: $S_t = f(Q_t, S_{t-1}, H_t)$
$M9$: $S_t = f(Q_t, Q_{t-1}, S_{t-1}, S_{t-2}, H_t, H_{t-1})$
$M10$: $S_t = f(Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}, S_{t-2}, H_t, H_{t-1})$
$M11$: $S_t = f(Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}, S_{t-2}, S_{t-3}, H_t, H_{t-1})$
$M12$: $S_t = f(Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}, S_{t-2}, S_{t-3}, H_t, H_{t-1}, H_{t-2})$

2.6. Implementation of machine learning methods

Three popular standard statistical measures: CC, RMSE and MAE were selected as performance evaluation parameters to judge the performance of the machine learning-based models. A large number of trials were performed to find optimum value of user defined parameters. Higher values of CC and lesser values of RMSE and MAE suggest that better estimation accuracy of the models. Trees quantity in the forest (k) and the number of variables used (m) at every node to generate a tree are the two standard users define parameters essential for RF. In M5P, calibration of models were done using changing the value of number of instances allowed at each node (m) is the only user define parameter in M5P tree model.

Data set: Data set was collected from Champua gauging site in the Upper Baitarani basin, India from 8 September, 2001 to 31 December, 2011. Total 3797 observations were collected from the metrological station, out of which 1166 observations (monsoon period data) considered for the analysis, in which, 795 observations, randomly taken from the total data were used for preparing the models and rest 371 were used to validate/test the developed model performance. Discharge (m^3/s), gauge stage (m) and sediment load concentration (mg/l) are available in the total dataset. Table 1 records the statistical features of training and testing data set that were taken in this investigation.

Table 1.
Features of the data set used for model development and validation

Range	Training dataset			Testing dataset		
	Q(t)	H(t)	S(t)	Q(t)	H(t)	S(t)
Minimum	2.59	3.79	0.00	2.75	3.80	1.20
Maximum	362.99	7.59	183.70	310.78	6.83	183.50
Mean	63.83	4.81	37.65	64.61	4.82	36.90
Standard Deviation	61.84	0.63	33.42	61.51	0.61	30.61
Kurtosis	4.40	0.85	2.59	3.53	0.18	2.35
Skewness	1.90	0. 70	1.60	1.82	0.53	1.45

Table 2.
Performance evaluation parameters of various M5P based models

Models M5P	Training data set			Testing data set		
	CC	RMSE	MAE	CC	RMSE	MAE
M1	0.66	25.42	19.18	0.53	25.98	20.19
M2	0.72	23.77	17.78	0.54	25.69	19.66
M3	0.80	20.18	12.40	0.75	20.37	13.47
M4	0.85	17.78	10.97	**0.79**	**18.72**	**12.38**
M5	0.85	17.72	10.87	0.79	18.95	12.43
M6	0.86	16.95	10.40	0.78	19.20	12.41
M7	0.87	16.78	10.33	0.78	19.21	12.30
M8	0.82	19.32	11.83	0.76	20.10	13.15
M9	0.86	17.01	10.37	0.79	18.98	12.56
M10	0.87	16.44	10.03	0.78	19.18	12.42
M11	0.88	16.20	9.81	0.78	19.33	12.43
M12	0.88	16.01	9.76	0.79	18.99	12.17

3. Results

The optimal number is found 4 of the user-defined parameter (m), and for all models, this number has been kept constant for achieving an unbiased assessment among M5P based various models. Table 2 shows that the M4-M5P model performs better than the other M5P based models for the prediction of sediment load concentration of the Baitarani River. In Figures 4-7, the agreement plot among actual and predicted values of sediment load concentration is shown for the testing stage. Figures 4-7 suggests that M4 based M5P models work better than other input combination based M5P models with minimum deviation. Table 2 also concludes that M4 based M5P models work better than other input combination based M5P models with higher values of CC (0.7920), R^2 (0.6270) and lower value of RMSE (18.7207) and MAE(12.3833) in the testing stage.

3.1. Results of RF-based models

RF-based model development is also a trial and error process. In Random Forest Regression-based models only two user-defined parameters are needed to optimize. These are: how many trees are there to be grown (*k*) and how many variables utilized (*m*) for generating a tree at every node. The optimum value of user-defined parameters of RF is found *m* as 1 and *k as*10 and these values remained constant in all models for the fair comparison among various input combination based models. It is clear from Table 3 that M10RF-based model performs better for the Baitarani River than other RF-based models. Table 3, shows the performance values viz. CC, RMSE and MAE of M10-RF based model are (0.9811, 7.7251 and 4.9472) and (0.8075, 18.1455 and 12.7289) for training and testing stages, respectively for Baitarani River. Figures 8-11 indicates the agreement plot among actual and predicted values of sediment load concentration in Baitarani River by using various input combination based RF models for Baitarani River with a testing dataset. Figures 8-11 and Table 3 indicate that M10 work is better than other RF-based models for the prediction of sediment load concentration in the Baitarani River.

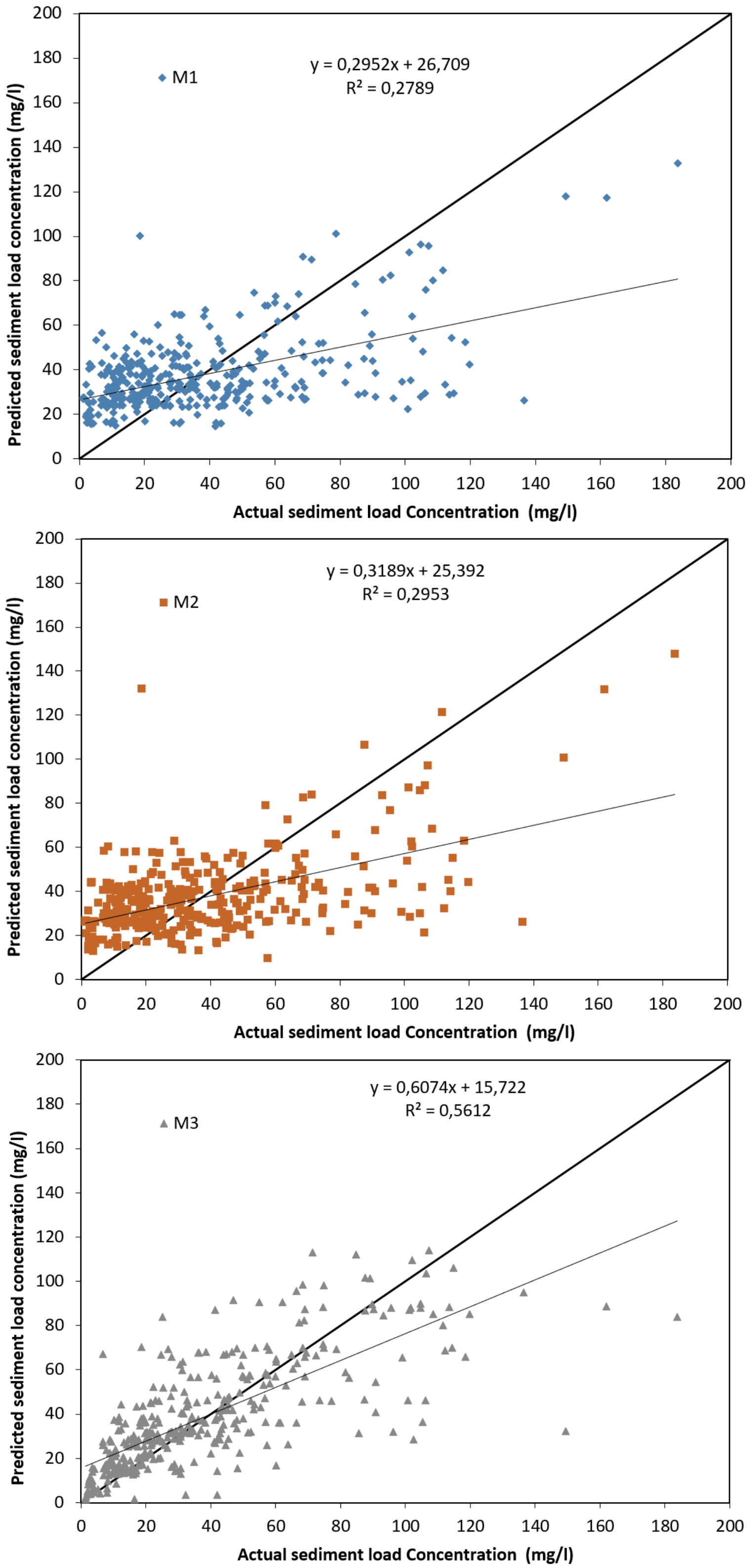


Fig. 4. Performance of M5P based models for the prediction of sediment load concentration using testing data set (M1-M3)

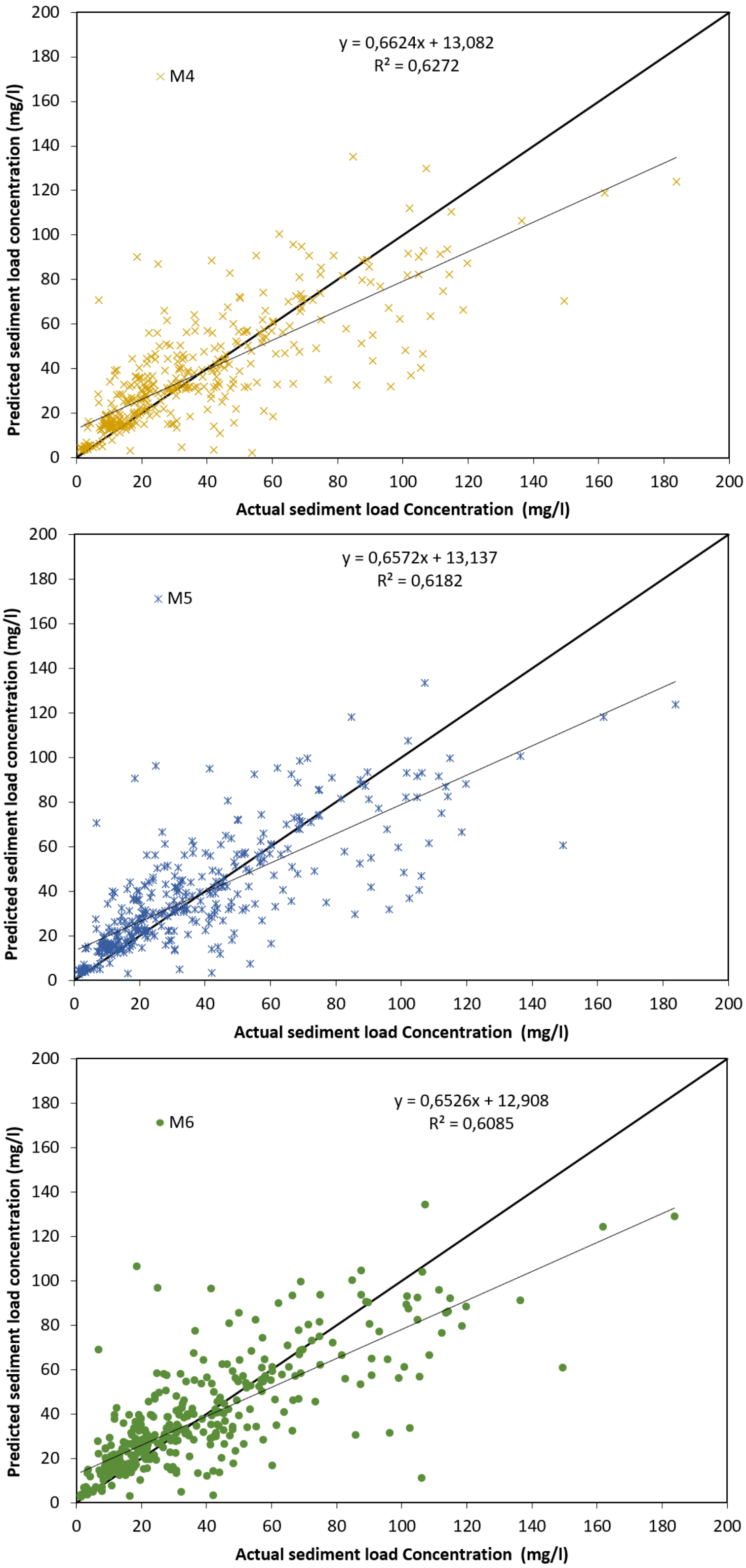
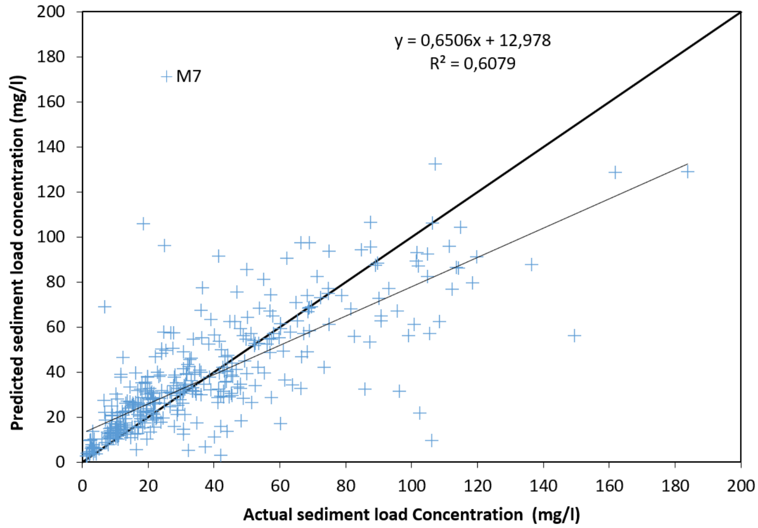
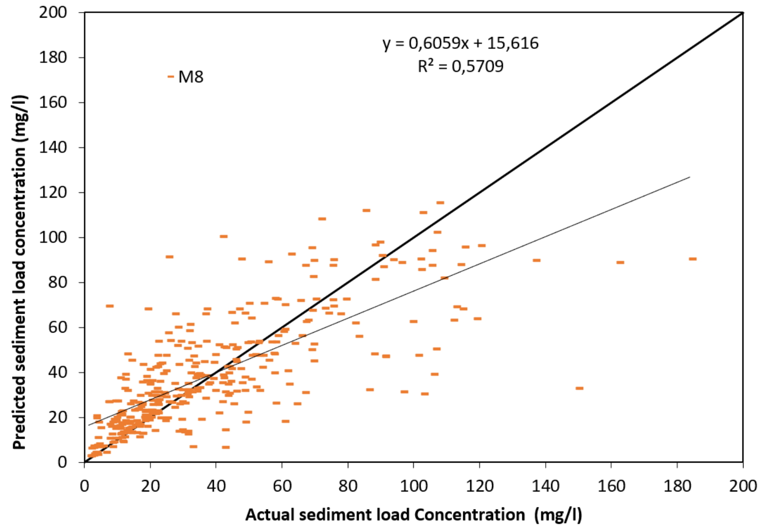


Fig. 5. Performance of M5P based models for the prediction of sediment load concentration using testing data set (M4-M6)





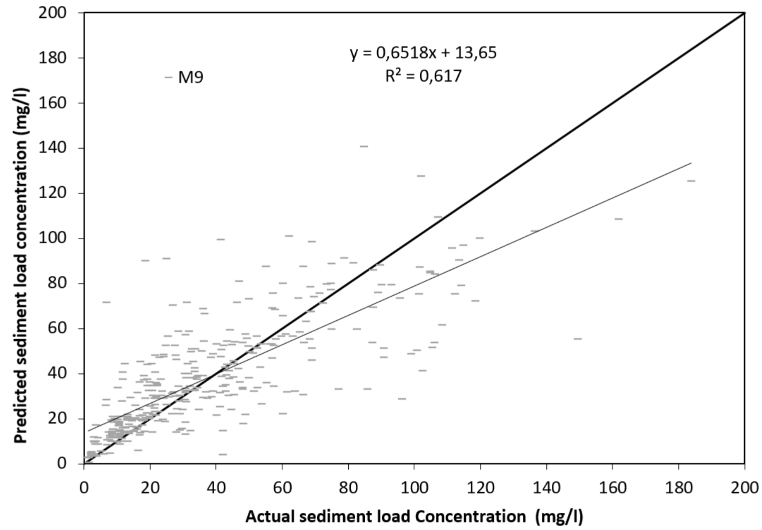


Fig. 6. Performance of M5P based models for the prediction of sediment load concentration using testing data set (M7-M9)

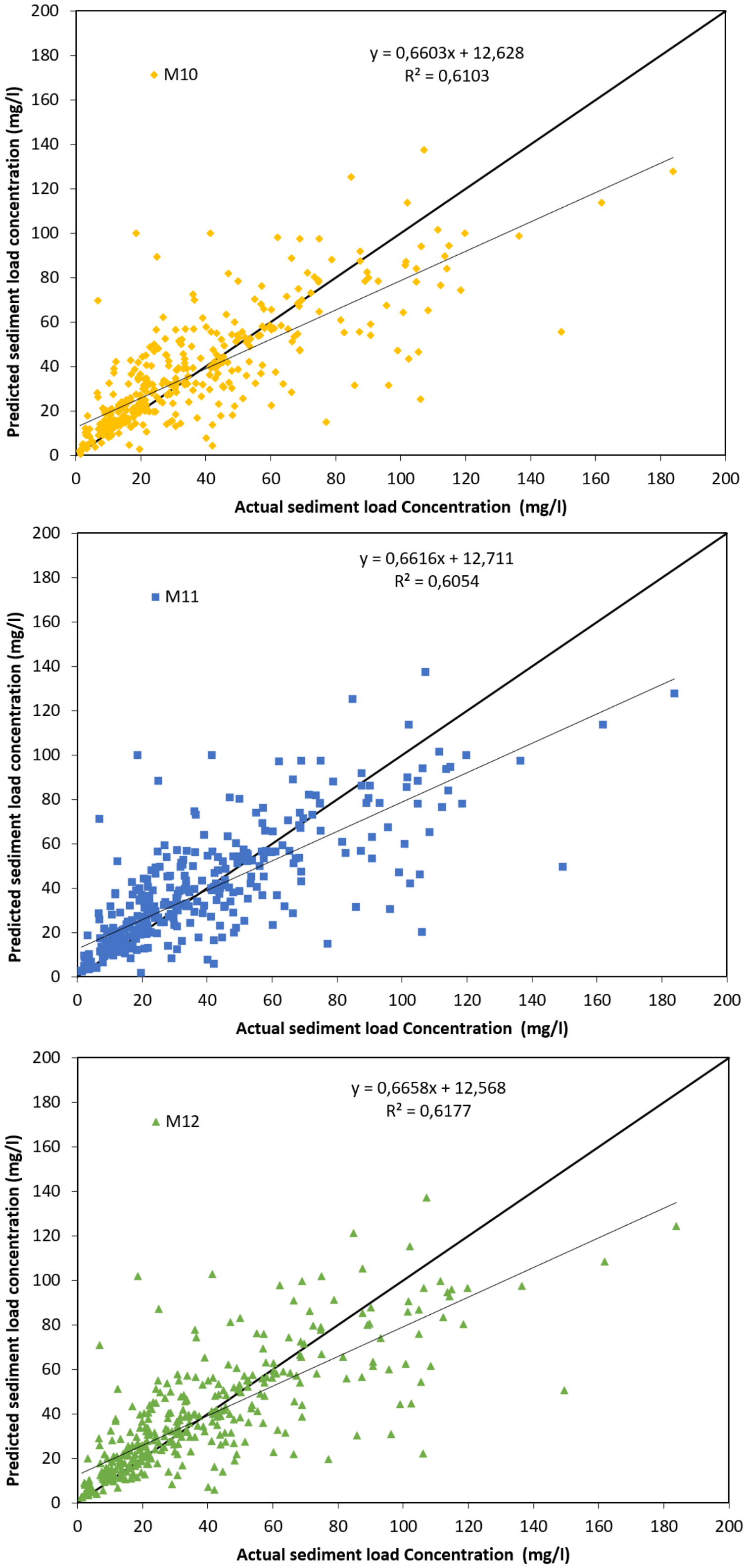


Fig. 7. Performance of M5P based models for the prediction of sediment load concentration using testing data set (M10-M12)

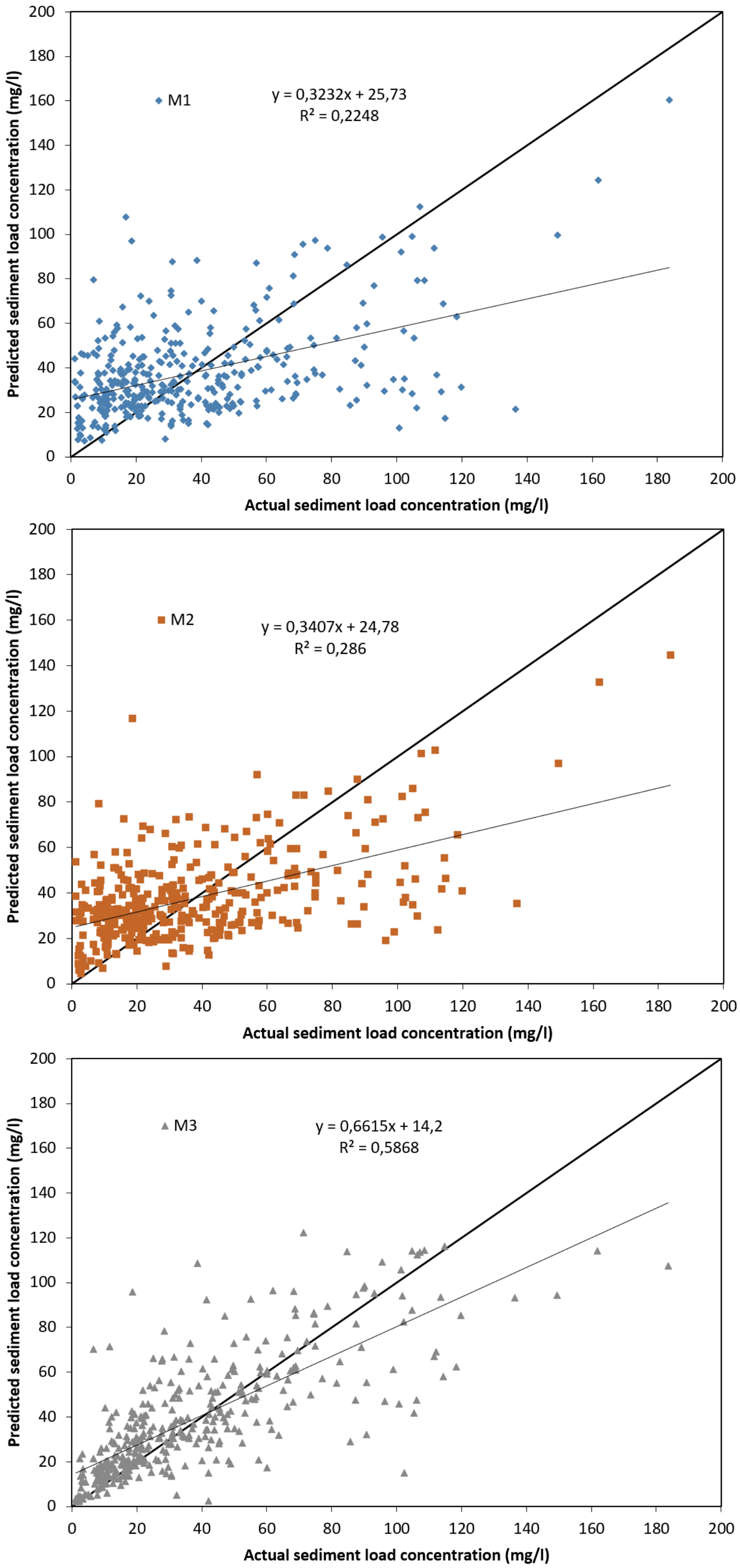


Fig. 8. Performance of RF based models for the prediction of sediment load concentration using testing data set (M1-M3)

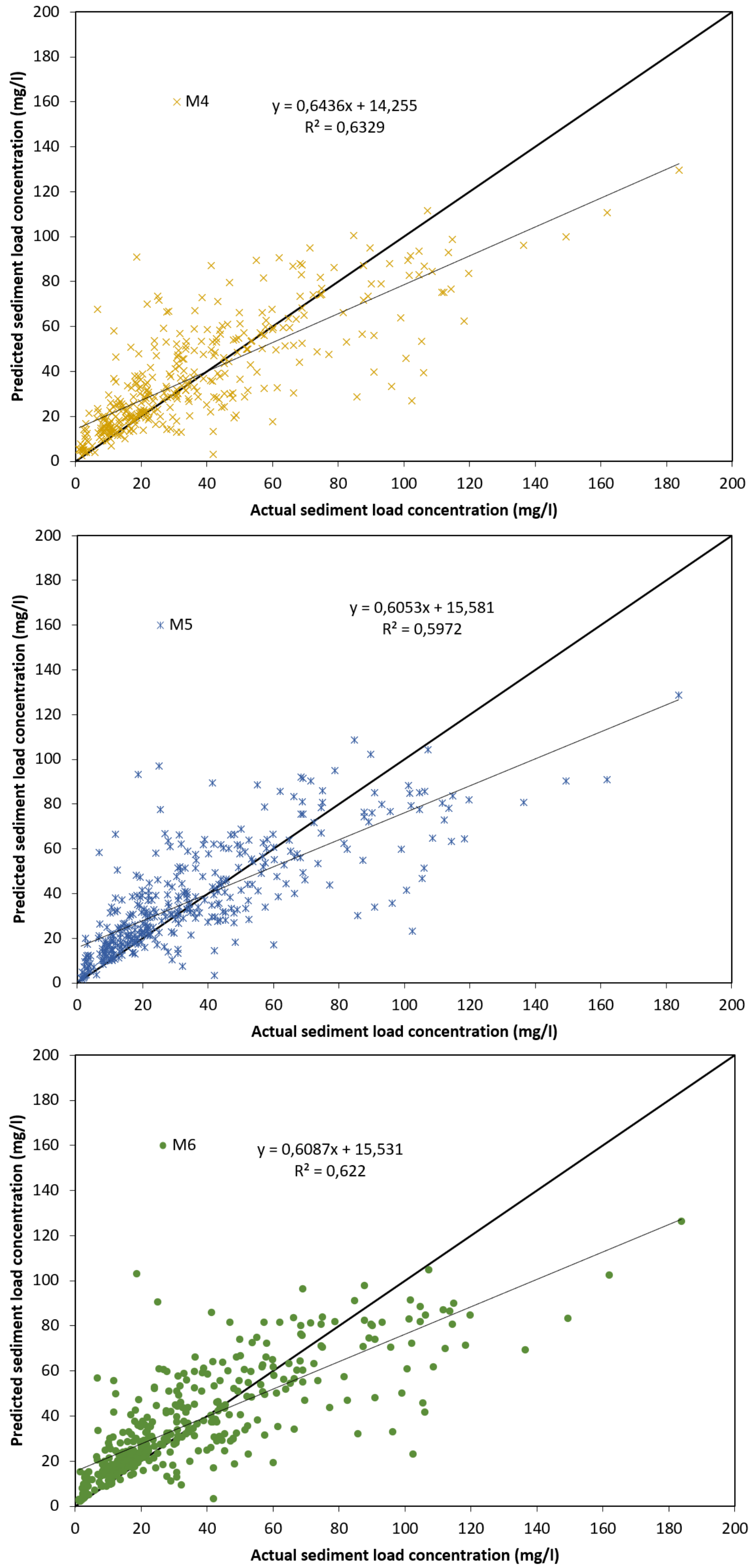


Fig. 9. Performance of RF based models for the prediction of sediment load concentration using testing data set (M4-M6)

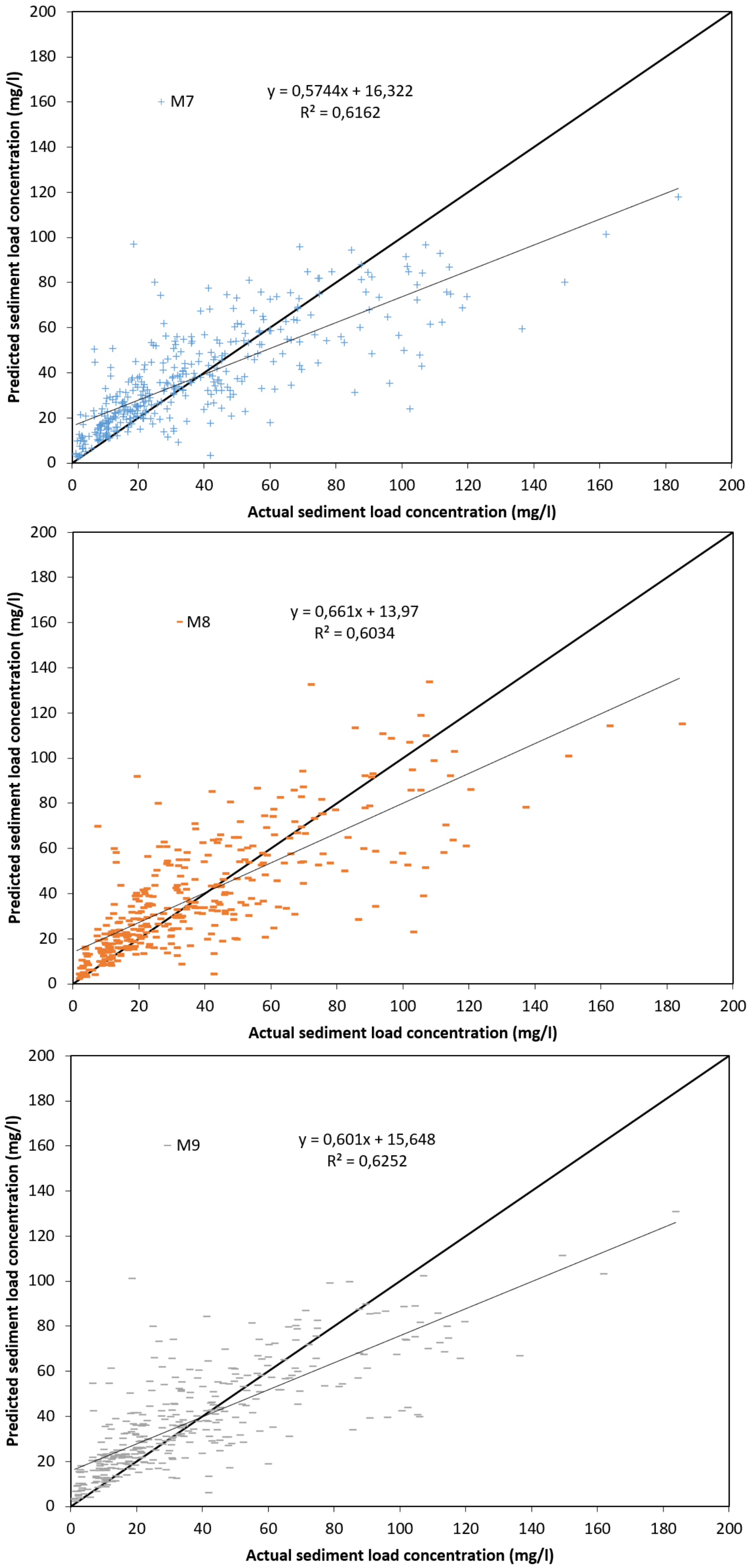


Fig. 10. Performance of RF based models for the prediction of sediment load concentration using testing data set (M7-M9)

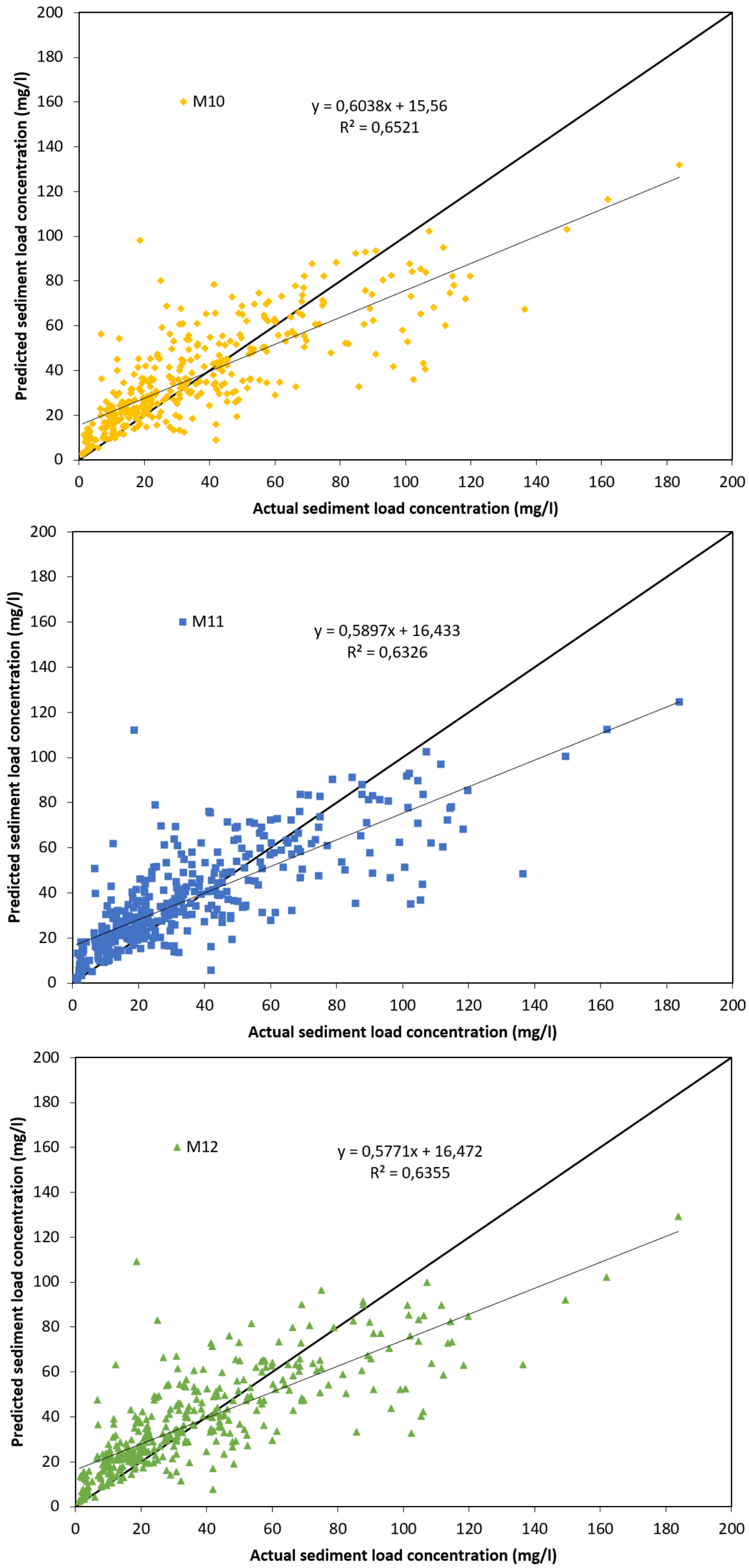


Fig. 11. Performance of RF based models for the prediction of sediment load concentration using testing data set (M10-M12)

Table 3.
Performance evaluation parameters of various RF based models

Models (RF)	Training data set			Testing data set		
	CC	RMSE	MAE	CC	RMSE	MAE
M1	0.96	11.12	8.12	0.47	27.66	20.72
M2	0.97	10.60	7.63	0.53	26.02	19.62
M3	0.97	9.03	5.50	0.77	19.95	13.31
M4	0.98	7.87	4.94	0.80	18.56	12.37
M5	0.98	8.21	5.17	0.77	19.43	13.21
M6	0.98	7.96	5.06	0.79	18.83	12.86
M7	0.98	8.30	5.27	0.79	19.02	13.00
M8	0.98	8.24	5.09	0.78	19.43	13.28
M9	0.98	8.08	5.10	0.79	18.76	12.86
M10	0.98	7.73	4.95	**0.81**	**18.15**	**12.73**
M11	0.98	8.07	5.21	0.80	18.65	12.70
M12	0.98	7.82	5.04	0.80	18.61	12.93

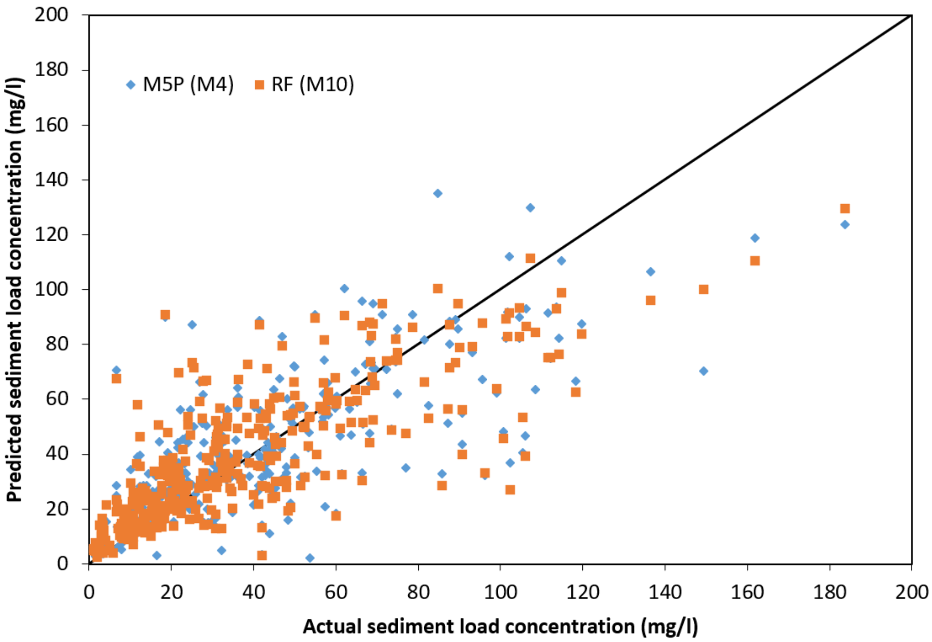


Fig. 12. Agreement plot among actual and predicted values of sediment load concentration using RF and M5P based models for testing data set

3.2. Intercomparison of soft computing based models

Table 4 proposes that Random forest based model work better than M5P based models. Figure 12 indicates that predicted values using M10 RF-based model are lies closer to the line of perfect agreement than the values predicted by M4 M5P based model. Performance plot among actual and predicted values using M5P and RF-based models are shown in Figure 13. In Figure 13 M10RF-based models follows the same bath as followed by the actual values in both training and testing stages. Overall performance of M10 based RF model is reliable and suitable for the prediction of sediment load concentration for the Baitarani River so $Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}, S_{t-2}, H_t$ *and* H_{t-1} input combination based RF model could be used for the prediction of sediment load concentration of rivers.

4. Discussion

The basic aim of the current study is to evaluate the performance of tree based modelling methods and give evidence that a specific model is suitable for the prediction of sediment load concentration of Upper Baitarani River. There are several soft computing based modelling methods. Inter-comparison among various models is necessary for the identification of superior models. Last few decades few artificial intelligence techniques were successfully used in the prediction of sediment load concentration. A direct and organized comparison of various studies is not fairly possible, because there is a large variation in input features, parameters, spatial variation and structure of the models. Table 5 contains the modelling studies addressing the performance of various soft computing models in predicting the sediment load in Rivers. The summary of these studies in Table 5indicates a broad dissimilarity for their main features but absolutely, there is no pointer towards any specific single most successful and better modelling most. Till date, it is not clear which technique is superior for the prediction of sediment load concentration. In present study Random forest works better than M5P based models and it could be successfully used for the prediction of sediment load of other Rivers.

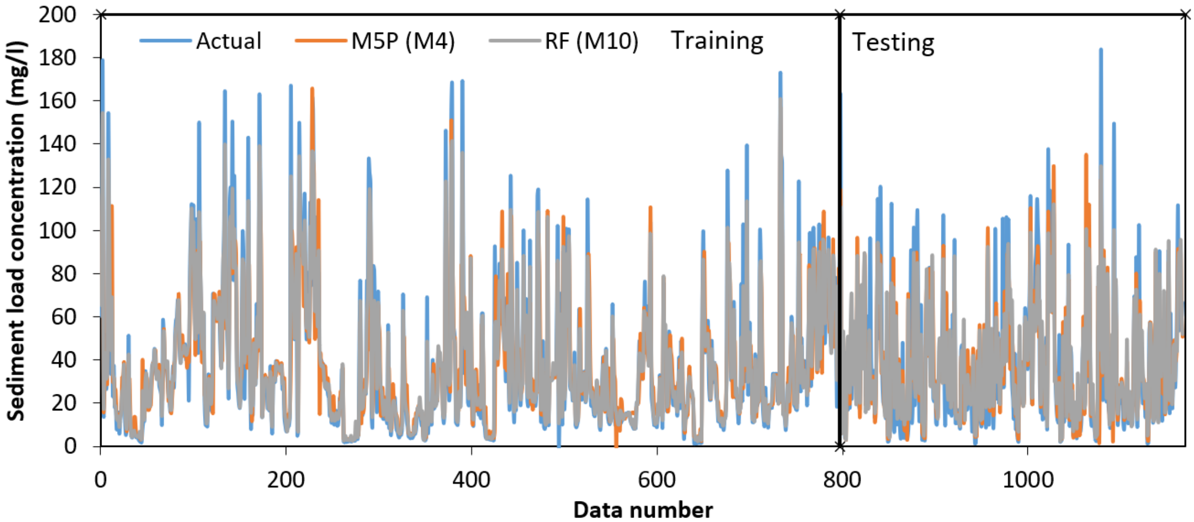


Fig. 13. Performance of best performing RF and M5P based model using training and testing stages

Table 4.
Performance evaluation parameters of best performing M5P and RF based model

	Training data set			Testing data set		
Models	CC	RMSE	MAE	CC	RMSE	MAE
M5P-M4	0.96	11.12	8.12	0.47	27.66	20.72
RF-M10	0.97	10.60	7.63	0.53	26.02	19.62

Table 5.
Comparison of soft computing based models for prediction of sediment load concentration

Sr.no.	Author's	Technique	Study area	Preferrable
1.	Boukhrissa et al. [8]	Artificial neural network (ANN) and sediment rating curve models	El Kebir catchment, Algeria	The neural network can be a potential estimation method that can be used for a better understanding of sediment flux that was considered high in the study catchment.

Sr.no.	Author's	Technique	Study area	Preferrable
2.	Bouzeria et al. [9]	Multilayer Perceptron (MLP) and Artificial neural network (ANN)	Mellah catchment north east of Algeria	The study suggested that the appropriate application of ANN other than MLPs to the sediment records may lead to solving several problems of water resources engineering and could provide a superior alternative for developing input-output simulations and estimation models in situations that do not require modelling of the internal structure of the catchment.
3.	Msadala & Basson [25]	Regional probabilistic method& empirical methods	South Africa	Where observed data is available, it is always recommended to use observed data, since the regional probabilistic and empirical methods have limited predictive capability due to the range of calibration data.
4.	Melessea et al. [7]	Artificial neural network (ANN) and multiple linear regression (MLR)	Three major rivers of USA (Mississippi, Missouri and Rio Grande	Comparison of the ANN and MLR approaches revealed that ANN performed better than MLR for all the rivers for both daily and weekly simulations. Comparison of the different input arrangements has shown that precipitation was not a significant parameter
5.	Mustafa et al. [26]	Radial basis function (RBF) neural network.	Pari River, in Perak, Malaysia	Results obtained from the RBF model are satisfactory and was found that RBF is able to predict the nonlinear behaviour of suspended sediment discharge of Pari River.
6.	Nagy et al. [6]	Artificial neural network(ANN)	Rio Grande, Mississippi, Sacramento	Approach gives better results compared to several commonly used formulas of sediment discharge.
7.	Roushangar & Shahnazi [10]	Gaussian process regression(GPR)& Support vector machine (SVM) models	Gravel-bed rivers in the United States	GPR models present better performance compared to the SVM model.
8.	Nourani & Andalib [27]	Wavelet-based least square support vector machine model (WLSSVM) & Artificial neural network (ANN).	Mississippi river	WLSSVM and wavelet-based ANN showed same consequences

Sr.no.	Author's	Technique	Study area	Preferrable
9.	Toriman et al. [28]	BFGS, Multiple Linear Regression and Artificial Neural Network	Jenderam catchment	BFGS model structure is the better and more accurate to prediction suspended sediment discharge
10.	Baniya et al. [29]	Artificial neural networks, multiple linear regression, nonlinear multiple regression, general power model, and Log transform models	Kali Gandaki River Basin, Himalaya, Nepal	Result was satisfactory compared to the multiple linear regression, nonlinear multiple regression, general power model, and Log transform models, including the sediment rating curve.
11.	Present study	M5P and Random Forest (RF)	Baitarani River at Odisha, India	RF based model is better than M5P based models

5. Conclusions

The performances of two popular tree based modelling techniques summarized by current investigation effort to give proof for most reliable and accurate modelling techniques for predicting sediment load concentration values. The tree based modelling techniques selected were M5P and Random Forest. The investigation selected a combination of stage, discharge and sediment load concentration for the model development of sediment load prediction of Baitarani River, India.

This investigation indicates that the RF-based model works better than M5P based models for the prediction of sediment load concentration. In RF-based models $Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}, S_{t-2}$, H_t and H_{t-1}combination based M10 model work superior than other combination based models. Another major outcome of this investigation is Q_t, Q_{t-1} and S_{t-1} based model M4 works better than other input combination based models using M5P technique. The optimum input combination is Q_t, Q_{t-1} and $S_{t-1}$for the prediction of sediment load concentration of the Baitarani River at Odisha, India.

Acknowledgements

The authors are thankful for ASTER Global Digital Elevation Model (GDEM) from the National Aeronautics and Space Administration (NASA), and India Water Resources Information System is popularly known as India-WRIS Website data used for analysis of the study area.

References

[1] United Nations, Dryland degradation keeping hundreds of millions in poverty, Press Release: Secretariat of the United Nations Convention to Combat Desertiûcation Geneva, Switzerland, 1997.

[2] P.K. Shit, A.S. Nandi, G.S. Bhunia, Soil erosion risk mapping using RUSLE model on Jhargram sub-division at West Bengal in India, Modeling Earth Systems and Environment 1/3 (2015) 28. DOI: https://doi.org/10.1007/s40808-015-0032-3

[3] B.N. Goswami, G. Wu, T. Yasunari, The annual cycle, intraseasonal oscillations, and roadblock to seasonal predictability of the Asian summer monsoon, Journal of Climate 19/20 (2006) 5078-5099. DOI: https://doi.org/10.1175/JCLI3901.1

[4] A.I.J.M. Van Dijk, L.A. Bruijnzeel, R.A. Vertessy, J. Ruijter, Runoff and sediment generation on bench terraced hillsides: measurements and upscaling of a field based model, Hydrological Processes: An International Journal 19/8 (2005) 1667-1685. DOI: https://doi.org/10.1002/hyp.5629

[5] O. Kisi, A. Guven, A machine code-based genetic programming for suspended sediment concentration estimation, Advances in Engineering Software 41/7-8 (2010) 939-945. DOI: https://doi.org/10.1016/j.advengsoft.2010.06.001

[6] H.M. Nagy, K.A.N.D. Watanabe, M. Hirano, Prediction of sediment load concentration in rivers using artificial neural network model, Journal of Hydraulic Engineering 128/6 (2002) 588-595. DOI:

https://doi.org/10.1061/(ASCE)0733-9429(2002)128:6(588)

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

[8] Z.A. Boukhrissa, K. Khanchoul, Y. Le Bissonnais, M. Tourki, Prediction of sediment load by sediment rating curve and neural network (ANN) in El Kebir catchment, Algeria, Journal of Earth System Science 122/5 (2013) 1303-1312. DOI: https://doi.org/10.1007/s12040-013-0347-2

[9] H. Bouzeria, A.N. Ghenim, K. Khanchoul, Using artificial neural network (ANN) for prediction of sediment loads, application to the Mellah catchment, northeast Algeria, Journal of Water and Land Development 33/1 (2017) 47-55. DOI: https://doi.org/10.1515/jwld-2017-0018

[10] K. Roushangar, S. Shahnazi, Prediction of sediment transport rates in gravel-bed rivers using Gaussian process regression, Journal of Hydroinformatics 22/2 (2020) 249-262. DOI: https://doi.org/10.2166/hydro.2019.077

[11] P. Sihag, V. Kumar, F.R. Afghan, S.M. Pandhiani, A. Keshavarzi, Predictive modeling of PM 2.5 using soft computing techniques: case study – Faridabad, Haryana, India, Air Quality, Atmosphere & Health 12/12 (2019) 1511-1520. DOI: https://doi.org/10.1007/s11869-019-00755-z

[12] B. Singh, P. Sihag, K. Singh, S. Kumar, Estimation of trapping efficiency of a vortex tube silt ejector. International Journal of River Basin Management (published online in 2018). DOI: https://doi.org/10.1080/15715124.2018.1476367

[13] B. Singh, P. Sihag, K. Singh, Modelling of impact of water quality on infiltration rate of soil by random forest regression, Modeling Earth Systems and Environment 3/3 (2017) 999-1004. DOI: https://doi.org/10.1007/s40808-017-0347-3

[14] M.M. Saritas, A. Yasar, Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification, International Journal of Intelligent Systems and Applications in Engineering 7/2 (2019) 88-91. DOI: https://doi.org/10.18201//ijisae.2019252786

[15] M. Kumar, P. Sihag, N.K. Tiwari, S. Ranjan, Experimental study and modelling discharge coefficient of trapezoidal and rectangular piano key weirs, Applied Water Science 10/1 (2020) 1-9. DOI: https://doi.org/10.1007/s13201-019-1104-8

[16] F.B. Banadkooki, M. Ehteram, A.N. Ahmed, F.Y. Teo, M. Ebrhimi, C.M. Fai, Y.F. Huang, A. El-Shafie, Suspended sediment load prediction using artificial neural network and ant lion optimization algorithm, Environmental Science and Pollution Research 27 (2020) 38094-38116. DOI: https://doi.org/10.1007/s11356-020-09876-w

[17] S.G. Meshram, V.P. Singh, O. Kisi, V. Karimi, C. Meshram, Application of Artificial Neural Networks, Support Vector Machine and Multiple Model-ANN to Sediment Yield Prediction, Water Resources Management 34 (2020) 4561-4575. DOI: https://doi.org/10.1007/s11269-020-02672-8

[18] V.H. Nhu, H. Shahabi, E. Nohani, A. Shirzadi, N. Al-Ansari, S. Bahrami, S. Miraki, M. Geertsema, H. Nguyen, Daily Water Level Prediction of Zrebar Lake (Iran): A Comparison between M5P, Random Forest, Random Tree and Reduced Error Pruning Trees Algorithms, ISPRS International Journal of Geo-Information 9/8 (2020) 479. DOI: https://doi.org/10.3390/ijgi9080479

[19] A. Sharafati, S.B.H.S. Asadollah, D. Motta, Z.M. Yaseen, Application of newly developed ensemble machine learning models for daily suspended sediment load prediction and related uncertainty analysis, Hydrological Sciences Journal 65/12 (2020) 2022-2042. DOI: https://doi.org/10.1080/02626667.2020.1786571

[20] J.R. Quinlan, Learning with continuous classes. Proceedings of Australian Joint Conference on Artificial Intelligence, Hobart 16-18 November 1992, vol. 92, 343-348.

[21] L. Breiman, Bagging predictors, Machine Learning 24/2 (1996) 123-140. DOI: https://doi.org/10.1007/BF00058655

[22] L. Breiman, Random forests, Machine Learning 45/1 (2001) 5-32. DOI: https://doi.org/10.1023/A:1010933404324

[23] A. Liaw, M. Wiener, Classification and regression by randomForest. R News 2/3 (2002) 18-22.

[24] A.K. Prabhakar, K.K. Singh, A.K. Lohani, S.K. Chandniha, Study of Champua watershed for management of resources by using morphometric analysis and satellite imagery, Applied Water Science

9/5 (2019) 127. DOI: https://doi.org/10.1007/s13201-019-1003-z

[25] V.C. Msadala, G.R. Basson, Revised regional sediment yield prediction methodology for ungauged catchments in South Africa, Journal of the South African Institution of Civil Engineering 59/2 (2017) 28-36. DOI: http://dx.doi.org/10.17159/2309-8775/2017/v59n2a4

[26] M.R.U. Mustafa, M.H. Isa, R.R. Bhuiyan, Prediction of river suspended sediment load using radial basis function neural network-a case study in Malaysia. Proceedings of the 2011 National Postgraduate Conference, Perak, Malaysia, 2011, 1-4. DOI: https://doi.org/10.1109/NatPC.2011.6136377

[27] V. Nourani, G. Andalib, Daily and monthly suspended sediment load predictions using wavelet based artificial intelligence approaches, Journal of Mountain Science 12/1 (2015) 85-100.
DOI: https://doi.org/10.1007/s11629-014-3121-2

[28] E. Toriman, O. Jaafar, R. Maru, A. Arfan, A.S. Ahmar, Daily Suspended Sediment Discharge Prediction Using Multiple Linear Regression and Artificial Neural Network, Journal of Physics: Conference Series 954/1 (2018) 012030. DOI: https://doi.org/10.1088/1742-6596/954/1/012030

[29] M.B. Baniya, T. Asaeda, K.C. Shivaram, S.M. Jayashanka, Hydraulic Parameters for Sediment Transport and Prediction of Suspended Sediment for Kali Gandaki River Basin, Himalaya, Nepal, Water 11/6 (2019) 1229.
DOI: https://doi.org/10.3390/w11061229