Application of Newly Developed Ensemble Machine Learning Models for Daily Suspended Sediment Load Prediction and Related Uncertainty Analysis

Ahmad Sharafati[1,2,3]*, Seyed Babak Haji Seyed Asadollah[3], Davide Motta [4], Zaher Mundher Yaseen[5]

[1]. Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam.
[2]. Faculty of Civil Engineering, Duy Tan University, Da Nang 550000, Vietnam
[3] Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
[4.] Department of Mechanical and Construction Engineering, Northumbria University, Wynne Jones Building, Newcastle upon Tyne NE1 8ST, United Kingdom
[5] Faculty of Civil Engineering, Ton Duc Thang University, Ho Chi Min City, Vietnam

Corresponding author: Ahmad Sharafati

Corresponding email: ahmadsharafati@duytan.edu.vn

Abstract

Ensemble machine learning models have been widely used in hydro-systems modeling as robust prediction tools that combine multiple decision trees. In this study, three newly developed ensemble machine learning models, namely Gradient Boost Regression (GBR), Ada Boost Regression (ABR) and Random Forest Regression (RFR) are proposed for prediction of Suspended Sediment Load (SSL), and their prediction performance and related uncertainty are assessed. The Suspended Sediment Load (SSL) of the Mississippi River, which is one of the major world rivers and is significantly affected by sedimentation, is predicted based on daily values of river Discharge (Q) and Suspended Sediment Concentration (SSC). Based on performance metrics and visualization, the RFR model shows a slight lead in prediction performance. The uncertainty analysis also indicates that the input variable combination has more impact on the obtained predictions than the model structure selection.

Keywords: Suspended Sediment Load, Ensemble Machine Learning, Prediction, Uncertainty Analysis

1. Introduction

The quantitative evaluation of sediment load is important for ecosystem analysis and management as well as hydraulic structure design, operation and maintenance, particularly for dams and channels (Nourani & Andalib, 2015). Sediment load and associated deposition may contribute to reducing reservoir volume, obstructing dam outlets and reducing channel carrying capacity (Buyukyildiz & Kumcu, 2017). Sediment also affects water quality, by transporting contaminants and potentially leading to reduction of dissolved oxygen concentrations (Shiau & Chen, 2015).

Suspended Sediment Load (SSL) and Bed Load (BL) are the two components of the Total Sediment Load (TSL), and the former part has generally more complex characteristics compared to BL (Zounemat-Kermani *et al.*, 2016). The spatial and temporal characteristics of SSL generally depend non-linearly on hydrologic conditions (Frings & Kleinhans, 2008).

Typically, SSL is estimated using empirical equations based on data obtained in the laboratory, (i.e. Baosheng *et al.*, 2008, Dorrell *et al.*, 2018, Tang & Knight, 2006), which makes their derivation costly and time-consuming. Furthermore, the empirical equations are strictly applicable only for the conditions they were derived for and may not provide accurate estimates outside those conditions (Bhattacharya *et al.*, 2005). In addition, they generally do not provide an estimate of the prediction uncertainty associated with their parameters (Shamaei & Kaedi, 2016).

Mathematical methods are an alternative to the empirical equations for computing SSL. They are either numerical (i.e. Cao & Carling, 2003, Mohammadian *et al.*, 2004, Wu, 2004) or analytical (i.e. Gill, 1983a, b, Zhang & Kahawita, 1987). The former are generally more broadly applicable than the latter, which rely on several simplistic assumptions (Bor, 2008). Mathematical numerical methods, however, take a significant time to set up and are computationally burdensome.

In recent years, Artificial Intelligence (AI) techniques, such as Artificial Neural Networks (ANNs) and fuzzy methods have been increasingly used to predict hydrological variables such as temperature, evaporation, rainfall and runoff (Ghorbani *et al.*, 2017, Kuok *et al.*, 2018, T.-Y. Pan *et al.*, 2013, Ahmad Sharafati, Khosravi, *et al.*, 2019, L. Wang *et al.*, 2017, Z.M. Yaseen, Awadh, *et al.*, 2018, Zaher Mundher Yaseen *et al.*, 2017, Zaher Mundher Yaseen, Fu, *et al.*, 2018). AI models are generally very efficient in predicting hydrological phenomena because of their simple structure, fewer input parameters and lower computation time than

mathematical numerical models (Adnan, Liang, Heddam, *et al.*, 2019, Adnan, Liang, Trajkovic, *et al.*, 2019, Adnan, Liang, Yuan, *et al.*, 2019, Adnan, Malik, Kumar, *et al.*, 2019).

AI models have also been specifically applied to the evaluation of sediment concentration and load. Following is a brief review of the related research developments in the last decade. (Senthil Kumar *et al.*, 2011) modelled Suspended Sediment Concentration (SSC) using ANNs with Back Propagation (BP) and Levenberg-Marquardt (LM) algorithms, Adaptive Neuro-Fuzzy Inference System (ANFIS) and decision tree models (such as M5 and REPTree), with the M5 model showing to have the best SSC prediction performance among the various techniques considered. (Zounemat-Kermani *et al.*, 2016) used three ANN and four Support Vector Regression (SVR) models to estimate SSC and compared those with models based on conventional Multi Linear Regression (MLR) and Sediment Rating Curve (SRC). Based on the Root Mean Square Error (RMSE) index, ANN-BFGS (Broyden–Fletcher–Goldfarb–Shanno) and SRV-RBF (Radial Basis Function) showed the best performance among the different models considered. (Kumar *et al.*, 2016) compared six soft computing models - ANNs, Radial Basis Function Neural Networks (RBFNNs), Least Squares-Support Vector Regression (LS-SVR), MLR and the tree-based models Classification And Regression Tree (CART) and M5 - for predicting SSL using hydro-meteorological variables; LS-SVR and ANN models had the best prediction performance, with M5 being the best performing among the tree-based models. (Buyukyildiz & Kumcu, 2017) compared seven different models, based on Support Vector Machine (SVM), ANN and ANFIS, in predicting SSL, with results showing the SVR model having the best prediction performance (coefficient of determination $R^2 = 0.868$). (Himanshu *et al.*, 2017a) used a Wavelet-SVM combined model to estimate SSL from sediment concentration, flow rate and precipitation, and compared it with the SVM and Wavelet separate models, observing a better prediction performance for the combined model (coefficient of determination $R^2 = 0.94$ and Nash-Sutcliffe Efficiency coefficient $NSE = 0.94$). (Talebi *et*

al., 2017) used tree-based methods such as Regression Trees (RTs) and Model Trees (MTs) to predict SSL and compared their results with ANN and SRC models; the tree-based methods, especially MT (coefficient of determination $R^2 = 0.98$) showed a performance close to the ANN models and better than the other models considered. (Yilmaz *et al.*, 2018) used several machine learning models such as Artificial Bee Colony (ABC), Teaching-Learning-Based Optimization (TLBO) and Multivariate Adaptive Regression Splines (MARS) to calculate SSL, with the MARS model having the best prediction performance. (Choubin *et al.*, 2018) evaluated CART model application to predict SSL based on meteorological data and compared the model with ANFIS, Multi-Layer Perceptron (MLP), RBF-SVM and Proximal Support Vector Machine (PSVM); the CART model (NSE = 0.77) and the RBF-SVM model (NSE = 0.68) showed the best prediction performance. (Nourani *et al.*, 2019) introduced a Wavelet procedure based on data mining named Wavelet-M5 (WM5) to predict SSL; comparing this model with Wavelet-ANN (WANN) and M5 tree models they observed the WM5 model to provide a better prediction (NSE = 0.94 compared to NSE = 0.89 for WANN and NSE = 0.77 for M5). (Hassanpour *et al.*, 2019) used a hybrid Fuzzy C-Means (FCM-SVR) model to assess SSL, showing better prediction performance than SRC, ANN, ANFIS and SVR. (Adnan, Liang, El-Shafie, *et al.*, 2019) developed a Dynamic Evolving Neural-Fuzzy Inference System (DENFIS) to estimate SSL and compared the results with ANFIS-FCM and MARS based on RMSE, Mean Absolute Error (MAE) and NSE coefficient; the NSE performance was increased by 4% and 15% using DENFIS when compared with ANFIS-FCM and MARS, respectively. Tables 1-3 list the most recent AI-based studies on SSL prediction, whether they are based on ANN (Table 1), fuzzy logic (Table 2) or other AI techniques (Table 3), specifying in each case the best predictive model, the input variables, the study area and the SSL timescale. Tables 1-3 show the ANN-based models as the most popularly used for SSL prediction, discharge as the most used input variable for prediction and daily as the most considered timescale for sediment

load. River discharge is a key variable to predict the SLL where most of the previous investigations (Tables 1-3) used it as a predictive variable. Furthermore, various studies used rainfall depth and SSL itself as input variables for SSL prediction, and a few investigations used water stage (Choubin *et al.*, 2018, Jain, 2001, Lohani *et al.*, 2007, Sari *et al.*, 2017), temperature (Demirci & Baltaci, 2013) and turbidity (Sari *et al.*, 2017). In general, daily data on discharge and SSC are typically the most commonly available, as observed for instance in the U. S. Geological Survey (USGS) datasets: they are therefore used as input variables for SSL prediction in this study, as illustrated in detail later on. However, appropriated data with a suitable sample period is a major concern in sediment modeling. According to the available data in USGS daily datasets, the daily discharge (Q), suspended sediment concentration (SSC) are found suitable as predictive variables for predicting the SSL.

[Tables 1,2, and 3]

In general, machine learning algorithms can enhance the prediction performance offered by traditional empirical equations, because they can more easily consider and derive patterns in complex datasets. ANN, SVM and Extreme Learning Machine (ELM) algorithms have been increasingly applied to the evaluation of hydrologic and hydraulic phenomena (Alizadeh et al., 2018, Adnan et al., 2019). Despite being more and more commonly used, ANN and SVM models also present drawbacks: they usually require preprocessing of the input data and the kernel functions of the SVM models or the complex structures of the ANN models require time-consuming training. On the other hand, decision tree models do not require input data preprocessing and more easily map the data features to the data target values; however, less research has been done on single-decision tree models due to their weak prediction performance (Gong *et al.*, 2020, Song *et al.*, 2019).

Tree-based ensemble learning algorithms generally offer better and more robust prediction performance, obtained by combining multiple decision trees. The use of multiple decision trees simultaneously decreases variance and bias associated with forecasting, avoids overfitting and improves the prediction performance compared with single learning algorithms (Rokach, 2010, Zhao *et al.*, 2018, Zhou *et al.*, 2019). Ensemble learning models are generally understood to provide a better prediction performance than single algorithms (Opitz & Maclin, 1999, Polikar, 2006, Rokach, 2010). Results have shown that Boosting and Bagging, two among the "classic" tree-based ensemble models, outperform SVM- and ANN-based models (Alizadeh et al., 2017, Shamshirband et al., 2019).

Several boosting and bagging ensemble algorithms with regression capability are available, such as Gradient Boost Regression (GBR), Ada Boost Regression (ABR) and Random Forest Regression (RFR), MadaBoost, LogitBoost, BrownBoost and LP-Boost (Domingo & Watanabe, 2000, Saffari *et al.*, 2010, Yu *et al.*, 2013, G. Zhang & Fang, 2007). Among these models, GBR, ABR and RFR are vastly used in various scientific fields (Afifi & Abdelhamed, 2019, Beaulac & Rosenthal, 2019, Georganos *et al.*, 2019, Huang *et al.*, 2019, Y. Pan *et al.*, 2019, Samadi *et al.*, 2019, Tama & Rhee, 2019), because they have a high prediction accuracy, are computationally efficient, do not need data pre-processing, can handle missing data and can be optimized through the use of different loss functions.

Although there has been an increasingly broad implementation of machine learning models in the literature for the estimation of SSL in rivers, to the authors knowledge there hasn't been any study on the application of these novel ensemble learning methods to issues related to water resources management and engineering, and to the estimation of SSL in particular. These models have the potential to provide river managers and researchers with fast and accurately predicting tools for quantifying SSL. The Mississippi river is selected as the case study for this research, given its availability of SSL, sediment concentration and flow data and its importance

not just as the second longest river in the United States but also for its challenging sedimentation problems. Hence, the Mississippi River presents a very good case study to investigate the SSL prediction performance of the proposed machine learning models. Specifically, this study aims to predict daily SSL for various values of lead time (time span between input variables for prediction and corresponding SSL output) and to quantify the relative uncertainty associated with model structure and input variable selection.

2. Methodology

2.1. Study Area and Data Collection

Though rivers like Missouri, as the longest river in U.S., or Colorado also can be selected as the case study for suspended sediment prediction, the Mississippi River has been much investigated regarding various aspects of its sediment transport and sedimentation (Meade & Moody, 2010, Mossa, 1996, Nagy *et al.*, 2002). It is the second largest river in the United States and is also ranked as the 4^{th} longest and 5^{th} largest catchment in the world. It originates from Lake Itasca and flows for about 3,730 kilometers through ten U. S. states until it finally reaches the Gulf of Mexico. Its catchment covers nearly 3,220,000 km^2, which completely or partially include 32 U. S. states. Previous investigations on the Mississippi River focused on the prediction of various hydrological variables, i.e. river discharge and sedimentation load (Melesse *et al.*, 2011, Nourani & Andalib, 2015, Sivapragasam *et al.*, 2014) and water quality parameters (Rodriguez & Sérodes, 2004). Due to the heavy sediment transport, the Mississippi River is prone to significant changes in its physical and chemical characteristics, such as water depth and width, bed elevation, planform configuration, sediment concentration and water quality, with consequences on both terrestrial and aquatic life. In order to mitigate these

negative effects, solutions such as dams and dikes have been implemented, with general success but also serious sedimentation problems over time (Julien & Vensel, 2005).

The data required for our analysis are obtained from the USGS station 05587455 (Mississippi river below Grafton, IL), downstream of the confluence between Mississippi River and Illinois River, about 40 km northwest of St. Louis and 32 km upstream of the confluence of the Missouri River with the Mississippi River, at the eastern border of Illinois and Missouri states (Figure 1). The USGS daily data employed in this study refer to three variables, namely Q (provided in units of $\frac{ft^3}{s}$), SSC $\left(\frac{milligrams}{liter}\right)$ and SSL $\left(\frac{ton}{day}\right)$, for the selected period from 2007 to 2015 (with a total of 3000 daily SSL data), for which the presence of missing daily data is minimal compared to other periods. Obviously a longer historical period with adequately complete data for analysis would have been ideal; however, the period selected still allows for obtaining reliable results.

Table 4 summarizes the main statistical parameters for Q, SSC and SSL for the data sample considered, including maximum, minimum and mean value (Xmax, Xmin and Xmean), standard deviation (Sx) and coefficient of skewness (Csx). From Table 4, Q, SSC and SSL are in the range of 13,400 - 400,000 ft^3/s, 6 - 589 mg/L and 489 - 387,000 ton/day, respectively. Furthermore, the skewness value of Q, SSC and SSL is 0.717, 1.486 and 1.488, respectively.

[Figure 1]

[Table 4]

The dataset used in this study was randomly split into two groups, with proportion of 75 to 25, for training and testing phases respectively.

2.2. Ensemble Learning Methods

As seen, among the most commonly applied ensemble learning algorithms are GBR, ABR and RFR, which have been used in many fields of science and engineering such as traffic prediction (Lopez-Martin *et al.*, 2019), air quality index computation (Miskell *et al.*, 2019, Y. Wang *et al.*, 2019), failure detection in robotics (Costa *et al.*, 2019), estimation of agricultural soil pollution (Tan *et al.*, 2020), prediction of sea surface temperature (Xiao *et al.*, 2019) and computation of seismic indicators (Asim *et al.*, 2018).

Boosting algorithms, applied by GBR and ABR, are based on the principle that the combination of multiple experts' judgements/decisions is more reliable than the judgment/decision of a single expert. This, in the context of prediction performance, allows boosting algorithms to improve on the performance of weak regression algorithms. GBR has shown a better prediction performance than single regression tree models (J. H. Friedman, 2001) and ABR has shown low bias error and to avoid overfitting in training (Schapire *et al.*, 1998).

RFR applies bagging algorithms. Though they have been widely used, several studies have highlighted their limitations (Dudoit *et al.*, 2002, Larivière & Van den Poel, 2005). Among the bagging algorithms, the most well-known is the Classification and Regression Tree (CART), which is used by RFR. Specifically, RFR creates a set of CART models based on the training subset to then generate a forest of tree models (Bienvenido-Huertas *et al.*, 2019).

2.2.1. Gradient Boost Regression (GBR)

GBR (Figure 2) is a machine learning technique for regression and classification problems, in which the main prediction model is the combination of several weak prediction models. This technique is based on the progressive strengthening of a prediction function F_m through the addition of an estimator E. The training algorithm is such that E is fitted to the $(y - F_m)$

residual and, through each iteration, F_{m+1} is corrected to minimize the residual value (J. H. Friedman, 2001). In symbols:

$$F_{m+1}(x) = F_m(x) + E = y \;\rightarrow\; E = y - F_m(x) \tag{1}$$

For this purpose, a series of input values or $x = \{x_1,\cdots,x_k\}$ and a series of output variables or y are considered and a loss function or $\Psi(y, F(x))$ is defined. The prediction modeling initiates by computing the $F_0(x)$ as follows:

$$F_0(x) = \arg\min_{\delta} \sum_{i=1}^{k} \Psi(y_i, \delta) \tag{2}$$

where, the m^{th} pseudo-residual amount for i^{th} data sample , δ_{im} , is calculated by the following:

$$\delta_{im} = -\left[\frac{\partial \Psi\big(y_i, F(x_i)\big)}{\partial\big(F(x_i)\big)}\right]_{F(x_i)=F_{m-1}(x_i)}, \qquad for\ i = 1,\cdots,k \tag{3}$$

Then, a weak learner function such as a decision tree ($E_m(x_i)$) is fitted to δ_{im} and trained based on the $\{(x_i, \delta_{im})\}_{i=1}^{k}$ training set. By solving a one-dimensional optimization relation, the multiplier δ_m is calculated as follows:

$$\delta_m = \arg\min_{\delta} \sum_{i=1}^{n} \Psi\big(y_i, F_{m-1}(x_i) + \delta h_m(x_i)\big) \tag{4}$$

where, h_m is a new tree. The $F_m\ (x_i)$ function is then taken equal to $F_{m-1}\ (x) + \delta_m E_m(x\,)$ and the process is repeated until the second term of the sum, $\delta_m E_m(x\,)$, is minimized at iteration M where the final output $F_M\ (x_i)$ is obtained (J. H. Friedman, 2002).

[Figure 2]

2.2.2. Ada Boost Regression (ABR)

Adaboost, the abbreviation of adaptive boosting (Figure 3), is one of the most applied models for machine learning and improves on the prediction performance of simpler learning models. This algorithm works by fitting a primary prediction function to the sum of the original data, calculating a prediction error, and applying a weighted vector to the data based on the prediction error. Given the error of the previous step, a series of additional models of primary function is applied on weighted data and the error of weighted data is calculated in this stage. It can be concluded that the error of each stage affects the next stage function in every replicate. When the weighted error reaches the minimum value, a weight is applied to each function and the result after summation is the final output (Freund & Schapire, 1997).

[Figure 3]

To illustrate the algorithm process, an additional model is considered whose components are functions of all input variables. Here, $f_m(x)$ is a weak primary learning function defined by a γ parameter and an additive factor β . The function $F_m(x)$ is expressed as an additional model of the ensemble of the weak functions:

$$F_m(x) = \sum_{m=1}^{M} f_m(x) = \sum_{m=1}^{M} \beta_m b(x, \gamma_m) \tag{5}$$

$$\{\beta_m, \gamma_m\} = arg \min_{\beta.\gamma} E \left[y - \sum_{k \neq m} \beta_k b(x, \gamma_k) - \beta b(x, \gamma) \right]^2 \tag{6}$$

where m = 1, 2, …, M is the number of replications of the algorithm until convergence is reached. Moreover, instead of the back-fitting method in the above function, it is possible to use a step-by-step forwarding greedy approach as a substitution method to solve the problem:

$$\{\beta_m, \gamma_m\} = arg \min_{\beta.\gamma} E\,[y - F_{m-1}(x) - \beta b(x;\gamma)]^2 \quad (7)$$

By computing the values γ_m and β_m, the value of $f_m(x)$ is determined and, consequently, the value of $F_m(x)$ is calculated at each iteration. Then, as in the GBR algorithm, a y_m residual value is obtained by subtracting the actual output from $F_m(x)$:

$$y_m = y - \sum_{k \neq m} f_m(x) \quad (8)$$

$$y_{m-1} = y_m - f_{m-1}(x) \quad (9)$$

In each iteration m, the value y_{m-1} is modified in such a way that the impact of y_m in each stage become lower and lower on the previous weak model $f_{m-1}(x)$. Generally, the Adaboost algorithm is a way to improve the weak learning algorithms $f_m(x)$ and prepares it to create a robust $F_m(x)$ model (J. Friedman *et al.*, 2000).

2.2.3. Random Forest Regression (RFR)

RFR (Figure 4) is a tree-based algorithm, which is widely used in a variety of areas of AI. This algorithm grows several predictor trees simultaneously and teach them separately. Ultimately, the result is obtained in classification phase by determining the final category using all the category modes and in regression phase by averaging the prediction of every individual tree (Barandiaran, 1998, Ho, 1995).

[Figure 4]

Applying methods such as bootstrap or bagging to tree teachers is the main component of the RFR algorithm. In the bootstrap method, the performance of the model is improved by reducing variance and without increasing bias. Numbering the algorithm replications as b = 1, …, B, if

training categories include $x = \{x_1, \cdots, x_k\}$ as input values and y as output, the bagging method repeatedly chooses X_b, Y_b random samples from X, Y training categories and fit them to a f_b tree. After having completed the learning, the prediction for a x' sample can be computed by averaging the predictions made by all trees on x' (Breiman, 2001):

$$f_b = \frac{1}{B}\sum_{b=1}^{B} f_b(x') \tag{10}$$

2.2.4. Ensemble Model Parameters

The development of the prediction models described in the previous sections is carried out in the Scientific Python Development Environment (Spyder), which is part of the Anaconda platform for the Python programming language. The SKLEARN Python library is utilized to develop and apply the ensemble algorithms (Pedregosa et al., 2011). Table 5 lists the default and adopted values of the various parameters of the GBR, ABR and RFR algorithms.

[Table 5]

The impact of the parameters in Table 5 on prediction performance and the identification of their optimal value is carried out through a sensitively analysis. Figure 5 shows an example of sensitivity analysis for the GBR algorithm, with the performance index R^2 evaluated and optimized for different values of the algorithm parameters n-estimator, max depth, loss function and learning rate.

[Figure 5]

Figure 5 shows that all the parameters can significantly affect the prediction performance, although R^2 values plateau for n-estimator greater than 50, max depth greater than 2 and learning rate lower than 0.4.

2.3. Evaluation of the Models Prediction Performance

A number of indicators, such as *KGE* (Kling-Gupta Efficiency), *PBIAS* (Percent Bias), *RSR* (*RMSE*-observations standard deviation ratio), *WI* (Willmott's Index of agreement), *MAPE* (Mean Absolute Percentage Error), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Correlation Coefficient (R), Coefficient of Determination (R^2) and Nash–Sutcliffe Efficiency (NSE) coefficient have been utilized in the literature to evaluate prediction performance (Abdulelah Al-Sudani *et al.*, 2019, A. Sharafati *et al.*, 2018, A. Sharafati & Zahabiyoun, 2013, Ahmad Sharafati, Tafarojnoruz, *et al.*, 2019, Z. Yaseen *et al.*, 2018, Z. M. Yaseen, Awadh, *et al.*, 2018). Having considered them and evaluated the similarity among some of them, the following indices are used in this study to evaluate the prediction performance of the proposed models:

$$MAE = \frac{1}{N}\sum_{i=1}^{n}|X_P - X_o| \tag{11}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(X_P - X_o)^2} \tag{12}$$

$$R^2 = \left(\frac{\sum_{i=1}^{N}(X_o - \overline{X_o})(X_P - \overline{X_P})}{\sqrt{\sum_{i=1}^{N}(X_o - \overline{X_o})^2 \sum_{i=1}^{N}(X_P - \overline{X_P})^2}}\right)^2 \tag{13}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(X_P - X_o)^2}{\sum_{i=1}^{n}(X_P - \overline{X_P})^2} \tag{14}$$

where $X_P, X_o, \overline{X_P}$ and $\overline{X_O}$ express predicted, observed, average predicted and average observed values, respectively.

The prediction performance is also visualized in this study using scatter plots, boxplots, normalized Taylor diagrams and heat maps. The scatter plot is the most common graph to make a direct comparison between predicted and observed outputs (Emamgholizadeh & Demneh, 2019, Hamaamin *et al.*, 2019, Hassanpour *et al.*, 2019, Nourani *et al.*, 2019, Sharghi *et al.*, 2019, Tabatabaei *et al.*, 2019). A normalized Taylor diagram plots together R, RMSE and normalized standard deviation; the closer the model performance point in the diagram to the "observed" point (RMSE = 0, R = 1, normalized standard deviation = 1), the better the performance of the model is (Taylor, 2001). A heat map shows at a glance the normalized performance indicators for each model. In addition to the plots employed in this study, the violin diagram was considered as visual tool of comparison; however, previous studies indicated a general consistency between the results obtained from violin diagrams and boxplots (Ahmad Sharafati, Khosravi, *et al.*, 2019) and therefore violin diagrams are not shown in this paper.

2.4. Uncertainty Analysis

The uncertainty in SSL prediction associated with the model structure and with the selection of the input variables is also evaluated. To evaluate the uncertainty associated with the model structure, for each observed SSL value the set of corresponding SSL values predicted by the three different models (GBR, ABR, RFR) considered in this study, for the same combination of input variables, is computed. In other words, a set of three predicted SSL values (predicted set) is assigned to each observed SSL. For each predicted set, the mean and standard deviation are computed to describe a normal distribution function. Using this distribution, 1000 values of SSL are generated through Monte Carlo simulation for each observed SSL. To quantify the corresponding uncertainty of SSL prediction, the 95% prediction confidence interval (interval between the 97.5% and the 2.5% quantiles), called the "95 percent prediction uncertainty" (95

PPU), is extracted using the generated SSLs for each observed SSL. Specifically, the uncertainty is measured using the $R-factor$ index as follows

$$R-factor=\frac{Sp}{So} \tag{15}$$

where So is the standard deviation of the observed data and Sp is computed as follows

$$Sp=\sum_{i=1}^{n}(U_{pi}-L_{pi})/n \tag{16}$$

where n is the number of observed data and U_{pi} and L_{pi} are the i^{th} values of upper quartile (97.5%) and lower quartile (2.5%) of the 95 PPU band, respectively. To assess the uncertainty associated with the input variables, the predicted SSL is computed for a single model but multiple input variable combinations, for each observed SSL. Then, the uncertainty associated with the input variables is quantified using the same $R-factor$ approach described above for the uncertainty associated with the model structure.

The Coefficient of Variation (CV) represents an alternative for quantifying prediction uncertainty. This index uses the mean and standard deviation of the predicted outputs while the $R-factor$ uses statistical indices obtained from both predicted and observed values, allowing for using each observed data as benchmark. For this reason several studies recommend the use of $R-factor$ instead of CV for uncertainty analysis (Abbaspour *et al.*, 2004, Kamali *et al.*, 2013, A Sharafati & Azamathulla, 2018, Ahmad Sharafati & Pezeshki, 2020).

3. Results and Discussion

The main goal of this study is to evaluate the performance of our newly developed ensemble machine learning models (ABR, GBR and RFR) in predicting SSL in rivers. The historical data on SSC and Q corresponding to several lead times starting from the an “origin” day (time “t”)

to five days earlier (time "t-5") are used as input for the models to predict SSL on the "origin" day (time "t") and up to three days ahead (time "t+1" and "t+3").

To quantify the level of correlation between SSL and input variables SSC and Q for different times, several methods were considered, such as Pearson correlation or more complex methods based on Auto-Correlation Function (ACF), Partial Auto-Correlation Function (PACF) and Cross-Correlation Function (CCF) (Buyukyildiz & Kumcu, 2017, Himanshu *et al.*, 2017a, Kisi & Yaseen, 2019, Nourani *et al.*, 2019). The Pearson correlation was finally selected, because of its simplicity and efficiency in evaluating the optimal set of input variables for AI modeling (Hai *et al.*, 2020, Malik *et al.*, 2020, Mohammed *et al.*, 2020, Salih *et al.*, 2019, Ahmad Sharafati, Tafarojnoruz, *et al.*, 2019). Table 6 reports the values of Pearson correlation coefficient between the target time series SSL (SSL(t), SSL(t+1) and SSL(t+3)) and the various times series of SSC and Q, computed based on the historically observed data at the USGS station considered in this study. From Table 6, for each of the three-time series SSL(t), SSL(t+1) and SSL(t+3), the SSC and Q time series for which the Pearson correlation coefficient is greater than 0.65 can be identified: these SSC and Q time series are considered in this study as possible inputs to predict the SSL time series. Table 6 shows a generally significant correlation between the input SSC and Q and the output SSL, with decreasing correlation for increasing time span between input and output. For instance, for SSL(t), the correlation coefficient varies between 0.915 (for SSC(t)) and 0.659 (for Q(t-4)). This is an expected pattern in hydrological process, where in general input variables with longer lead time have lower effect on output values. For this reason the prediction of SSL further in the future (e.g., "t+4" and "t+5") is not considered in this study. Table 6 already shows low correlation values between SSL and input variables for lead time equal to 4 days, and correlation values for longer lead time would be even lower. On the other hand, the dependence characteristics of the

targeted time series SSL with times series of SSC and Q have decreasing trend by increasing the lead time.

[Table 6]

Tables 7, 8 and 9 display the considered input combinations for predicting SSL(t), SSL(t+1) and SSL(t+3), respectively.

[Table 7, 8 and 9]

The prediction performance of the models developed in this study is evaluated based on the performance indicators described above (MAE, RMSE, R^2 and NSE) and graphically.

Tables 10, 11 and 12 display the prediction performance indices for the predictive models GBR, ABR and RFR, respectively. For the GBR model, the optimal SSL prediction is attained for input combination M7 (with performance R^2= 0.995, RMSE = 5512 ton/day), M5 (R^2 = 0.911, RMSE = 20734 ton/day) and M1 (R^2 = 0.71, RMSE = 36969 ton/day) for SSL(t), SSL(t+1) and SSL(t+3), respectively. The ABR model has its best performance for input combination M6 for both SSL(t) (R^2 = 0.995, RMSE = 5068 ton/day) and SSL(t+1) (R^2 = 0.887, RMSE = 23289 ton/day), and for input combination M1 (R^2 = 0.699, RMSE = 37672 ton/day) for SSL(t+3). The RFR model achieves its best performance for input combination M7 (R^2 = 0.996, RMSE = 4624 ton/day) for SSL(t), M5 (R^2 = 0.914, RMSE = 20378 ton/day) for SSL(t+1) and M1 (R^2 = 0.718, RMSE = 36089 ton/day) for SSL(t+3). The best prediction performance is obtained for short-term “one day ahead” SSL prediction; the prediction performance is noticeably reduced the longer the prediction span, although it remains

acceptable based on the acceptance ranges indicated by (Lawrence & Lin, 1989, Moriasi *et al.*, 2007, Ritter & Muñoz-Carpena, 2013, Willmott & Matsuura, 2005).

[Table 10, 11 and 12]

Table 13 summarizes the prediction performance indices for the best established input combinations for each SSL time series. As expected, the prediction performance is the best for SSL(t) and decreases, although remaining more than acceptable, for SSL(t+3). Prediction of SSL(t) requires four input time series, namely [SSC(t), SSC(t-1), SSC(t-2) and Q(t)], while SSL(t+1) and SSL(t+3) require three [SSC(t), SSC(t-1), Q(t)] and four [SSC(t), SSC(t-1), Q(t-1), Q(t)], respectively. It is noted that the SSC time series is a key input for SSL prediction for times "t" and "t+1" and the Q time series is for time "t+3". Overall, the developed GBR, ABR and RFR models produce comparable prediction performances, with RFR having a slight lead over the other two.

[Table 13]

Visual comparisons are also carried out to identify optimal prediction model(s). Figure 6 illustrates the heat map of the best selected predictive models based on normalized performance indicators for each time series. Again, for predicting SSL(t), RFR-M7 model has the best modeling performance. For SSL(t+3), RFR-M1 model is the most accurate predictive model and, for SSL(t+1), RFR-M6 and GBR-M5 have comparable performance and better prediction than ABR-M6.

[Figure 6]

Figure 7 compares the best predictive models on scatter plots. The figure shows an increasing spread of points for observed vs predicted SSL as the time considered for SSL changes from "t" to "t+3". Based on the coefficient of determination R^2, RFR-M7 ($R^2 = 0.996$), RFR-M6

($R^2 = 0.914$) and RFR-M1 ($R^2 = 0.718$) have the best prediction performance for SSL(t), SSL(t+1) and SSL(t+3), respectively.

The boxplots in Figure 8 reveal the marginally better performance of RFR-M7 ($Q_{50\%} = 26970$ ton/day vs observed $Q_{50\%} = 26900$ ton/day) and ABR-M6 ($Q_{50\%} = 29100$ ton/day vs observed $Q_{50\%} = 27500$ ton/day) for SSL(t) and SSL(t+1), respectively. For SSL(t+3), the GBR-M1 model ($Q_{50\%} = 27127.4$ ton/day vs observed data $Q_{50\%} = 27350$ ton/day) slightly outperforms the others.

[Figure 7 and 8]

The Taylor diagrams (Figure 9) confirm the slightly better performance of the RFR models (M7 for SSL(t), M6 for SSL(t+1) and M1 for SSL(t+3)) highlighted by heat maps and scatterplots. Additionally, in all cases the ABR algorithm shows the weakest performance in SSL prediction.

[Figure 9]

The prediction performance obtained from the ensemble machine learning models developed in this study reveals that they are remarkably capable of predicting SSL for different time spans between inputs and output. However, it is also important to compare the results obtained with reported modeling results from the literature for SSL prediction. Melesse *et al.* (2011) developed ANN, MLR and Auto Regressive Integrated Moving Average (ARIMA) models for predicting "one day ahead" SSL, attaining $R^2 = 0.96, 0.76$ and 0.98, respectively. Nourani & Andalib (2015) applied ANN and Least Squares Support Vector Regression (LSSVR) models for simulating "one day ahead" SSL, achieving $R^2 = 0.87$ and 0.92, respectively. (Olyaie *et al.*, 2015) predicted SSL in the Flathead River and in the Santa Clara River using ANN, ANFIS and WANN models, obtaining R^2 = 0.662, 0.683 and 0.894, respectively. (Rahgoshay *et al.*, 2018) modeled daily SSL using M5, SVM-GA (Genetic Algorithm) and MARS models, with

NSE results of 0.91, 0.90 and 0.96, respectively. (Kisi & Yaseen, 2019) predicted SSL with ANFIS-GP with a best obtained NSE value of 0.911. (Sharghi *et al.*, 2019) forecast daily SSL in the Upper Rio Grande and Lighvanchai Rivers using ANN, Emotional -ANN (EANN), Wavelet-Emotional-ANN (WEANN) and Wavelet-ANN (WANN) models, obtaining NSE of 0.948, 0.947, 0.989 and 0.987, respectively. It can be therefore seen that the predictive models presented in this study attain a better prediction performance compared to the models in the literature.

The SSL predictions provided by our newly developed ensemble machine learning models are associated, as is case with any model, with a certain degree of uncertainty that needs to be quantified. The uncertainty associated with the model structure is evaluated for the three ensemble machine learning models considered (GBR, ABR, RFR) for the best input combination M7, M6, and M1 for SSL(t), SSL(t+1) and SSL(t+3), respectively. As regards the uncertainty associated with the input variables, the generally best performing model RFR is evaluated for the input combinations $M1$ to $M10$, $M1$ to $M8$ and $M1$ to $M4$ for SSL(t), SSL(t+1) and SSL(t+3), respectively.

Figures 10, 11, and 12 show the 95PPU band created for SSL(t), SSL(t+1) and SSL (t+3), respectively, in each case considering both model structure and input variable uncertainties and comparing with the observed SSL values. For SSL(t) and SSL(t+1), the model structure $R-factor$ (0.15 and 0.28, respectively) is lower than the input variable $R-factor$ (0.66 and 0.67, respectively), especially for SSL(t), for which the difference between the two R-factor's is significant. Hence, the prediction results are more sensitive to the input variables than they are to the model structure. On the contrary, for SSL(t+3), the input variable $R-factor$ (0.11) is lower than the model structure $R-factor$ (0.48), which means that the model selection has larger impact on prediction than the selection of the input variables. In other words, the uncertainty analysis indicates that the input combination for short term prediction (less than

two days ahead) of SSL has a significant impact on the results obtained, while the model structure is more important than the input combination for the performance of longer term predictions of SSL. It must be noted that we exclusively focus here on the uncertainty associated with input variables and model structure; other sources of uncertainty, related to measurement errors, data handling and inadequate sampling may be important but are assumed to be negligible for the purposes of this study.

[Figure 10, 11 and 12]

4. Conclusion

This study evaluates the prediction performance of models based on the application of the ensemble learning algorithms ABR, GBR and RFR (subcategories of machine learning algorithms) in predicting suspended sediment load in rivers. For this purpose, the Mississippi River in USA is selected as a case study and 3000 daily SSL data in the period 2007-2015 are considered. Daily discharge and suspended sediment concentration are considered as input variables for prediction of SSL. SSL is predicted for an "origin" day (t) and following days (t+1) and (t+3) and the input variables Q and SSC are considered for day (t) down to day (t-4). The prediction performance of the ABR, GBR and RFR algorithms is evaluated the performance indicators RMSE, MAE, R^2 and NSE coefficient for different combinations of input variables. Visual comparisons, in the form of heat maps, scatter plots, boxplots and Taylor diagrams, are also used to identify the best predictive models. The results show RFR as the generally leading algorithm. The optimal input variable combination is [*Q(t), SSC(t), SSC(t-1), SSC(t-2)*], [*Q(t), SSC(t), SSC(t-1)*] and [*Q(t), Q(t-1), SSC(t), SSC(t-1)*] for SSL(t), SSL(t+1) and SSL(t+3), respectively. Overall, our uncertainty analysis shows that the SSL prediction is significantly more affected by the selection of the input variables than it is by the model

structure (i.e. the type of algorithm). The prediction performance provided by the proposed algorithms for SSL up to three days after the current day makes them a potentially useful tool for planners and civil engineers involved in sedimentation mitigation.

One limitation of this study is the consideration of a single case study (Mississippi River below Grafton, IL). Future work will involve testing the proposed methodology on several other rivers. We also plan to evaluate the use of input variables for prediction generated from weather data using numerical rainfall-runoff modeling, as an alternative to input variables measured in situ, to facilitate the adoption of our proposed predictive models (especially where in situ data are limited) and possibly further improve the SSL prediction performance. In addition, satellite data can be incorporated as external informative attribute to the prediction matrix where more accurate prediction accuracy can be attained. Thus, this is also another vital exploration can be devoted as future research as an extension for the current study.

References

Abbaspour, K. C., Johnson, C. A. & Genuchten, M. T. van. (2004) Estimating Uncertain Flow and Transport Parameters Using a Sequential Uncertainty Fitting Procedure. *Vadose Zo. J.* **3**(4), 1340–1352. doi:10.2113/3.4.1340

Abdulelah Al-Sudani, Z., Salih, S. Q., sharafati, A. & Yaseen, Z. M. (2019) Development of multivariate adaptive regression spline integrated with differential evolution model for streamflow simulation. *J. Hydrol.* **573**(February), 1–12. Elsevier. doi:10.1016/j.jhydrol.2019.03.004

Adib, A. & Mahmoodi, A. (2017) Prediction of suspended sediment load using ANN GA conjunction model with Markov chain approach at flood conditions. *KSCE J. Civ. Eng.* **21**(1), 447–457. doi:10.1007/s12205-016-0444-2

Adnan, R. M., Liang, Z., El-Shafie, A., Zounemat-Kermani, M. & Kisi, O. (2019) Prediction of Suspended Sediment Load Using Data-Driven Models. *Water* **11**(10), 2060. Multidisciplinary Digital Publishing Institute.

Adnan, R. M., Liang, Z., Heddam, S., Zounemat-Kermani, M., Kisi, O. & Li, B. (2019) Least square support vector machine and multivariate adaptive regression splines for streamflow prediction in mountainous basin using hydro-meteorological data as inputs. *J. Hydrol.* 124371. Elsevier.

Adnan, R. M., Liang, Z., Trajkovic, S., Zounemat-Kermani, M., Li, B. & Kisi, O. (2019) Daily streamflow prediction using optimally pruned extreme learning machine. *J. Hydrol.* **577**, 123981. Elsevier.

Adnan, R. M., Liang, Z., Yuan, X., Kisi, O., Akhlaq, M. & Li, B. (2019) Comparison of LSSVR, M5RT, NF-GP, and NF-SC models for predictions of hourly wind speed and wind power based on cross-validation. *Energies* **12**(2), 329. Multidisciplinary Digital Publishing Institute.

Adnan, R. M., Malik, A., Kumar, A., Parmar, K. S. & Kisi, O. (2019) Pan evaporation modeling by three different neuro-fuzzy intelligent systems using climatic inputs. *Arab. J. Geosci.* **12**(20), 606. doi:10.1007/s12517-019-4781-6

Afan, H. A., El-Shafie, A., Yaseen, Z. M., Hameed, M. M., Wan Mohtar, W. H. M. & Hussain, A. (2014) ANN Based Sediment Prediction Model Utilizing Different Input Scenarios. *Water Resour. Manag.* **29**(4), 1231–1245. doi:10.1007/s11269-014-0870-1

Afifi, M. & Abdelhamed, A. (2019) AFIF4: deep gender classification based on AdaBoost-based fusion of isolated facial features and foggy faces. *J. Vis. Commun. Image Represent.* **62**, 77–86. Elsevier.

Agarwal, A., Singh, R. D., Mishra, S. K. & Bhunya, P. K. (2005) ANN-based sediment yield models for Vamsadhara river basin (India). *Water Sa* **31**(1), 85–100. Water Research Commission (WRC).

Alizadeh, M. J., Kavianpour, M. R., Danesh, M., Adolf, J., Shamshirband, S. & Chau, K.-W. (2018) Effect of river flow on the quality of estuarine and coastal waters using machine learning models. *Eng. Appl. Comput. Fluid Mech.* **12**(1), 810–823. Taylor & Francis.

Alizadeh, M. J., Nodoushan, E. J., Kalarestaghi, N. & Chau, K. W. (2017) Toward multi-day-ahead forecasting of suspended sediment concentration using ensemble models. *Environ. Sci. Pollut. Res.* **24**(36), 28017–28025. Springer.

Asim, K. M., Idris, A., Iqbal, T. & Martínez-Álvarez, F. (2018) Seismic indicators based earthquake predictor system using Genetic Programming and AdaBoost classification. *Soil Dyn. Earthq. Eng.* **111**, 1–7. Elsevier.

Baosheng, W. U., Maren, D. S. Van & Lingyun, L. I. (2008) Predictability of sediment transport in the Yellow River using selected transport formulas. *Int. J. Sediment Res.* **23**(4), 283–298. Elsevier.

Barandiaran, I. (1998) The random subspace method for constructing decision forests. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(8).

Beaulac, C. & Rosenthal, J. S. (2019) Predicting University Students' Academic Success and Major Using Random Forests. *Res. High. Educ.* **60**(7), 1048–1064. Springer.

Bhattacharya, B., Price, R. K. & Solomatine, D. P. (2005) Data-driven modelling in the context of sediment transport. *Phys. Chem. Earth, Parts A/B/C* **30**(4–5), 297–302. Elsevier.

Bienvenido-Huertas, D., Rubio-Bellido, C., Pérez-Ordóñez, J. L. & Oliveira, M. J. (2019) Automation and optimization of in-situ assessment of wall thermal transmittance using a Random Forest algorithm. *Build. Environ.* 106479. Elsevier.

Bor, A. (2008) Numerical modeling of unsteady and non-equilibrium sediment transport in rivers. İzmir Institute of Technology.

Breiman, L. (2001) Random forests. *Mach. Learn.* **45**(1), 5–32. doi:10.1023/A:1010933404324

Buyukyildiz, M. & Kumcu, S. Y. (2017) An Estimation of the Suspended Sediment Load Using Adaptive Network Based Fuzzy Inference System, Support Vector Machine and Artificial Neural Network Models. *Water Resour. Manag.* **31**(4), 1343–1359. Water Resources Management. doi:10.1007/s11269-017-1581-1

Cao, Z. & Carling, P. A. (2003) On evolution of bed material waves in alluvial rivers. *Earth*

Surf. Process. Landforms J. Br. Geomorphol. Res. Gr. **28**(4), 437–441. Wiley Online Library.

Chen, X. Y. & Chau, K. W. (2016) A Hybrid Double Feedforward Neural Network for Suspended Sediment Load Estimation. *Water Resour. Manag.* **30**(7), 2179–2194. doi:10.1007/s11269-016-1281-2

Choubin, B., Darabi, H., Rahmati, O., Sajedi-Hosseini, F. & Kløve, B. (2018) River suspended sediment modelling using the CART model: A comparative study of machine learning techniques. *Sci. Total Environ.* **615**, 272–281. Elsevier B.V. doi:10.1016/j.scitotenv.2017.09.293

Cigizoglu, H Kerem. (2004) Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. *Adv. Water Resour.* **27**(2), 185–195. Elsevier.

Cigizoglu, Hikmet Kerem & Alp, M. (2006) Generalized regression neural network in modelling river sediment yield. *Adv. Eng. Softw.* **37**(2), 63–68. Elsevier.

Cobaner, M., Unal, B. & Kisi, O. (2009) Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural network approaches using hydro-meteorological data. *J. Hydrol.* **367**(1–2), 52–61. Elsevier.

Costa, M. A., Wullt, B., Norrlöf, M. & Gunnarsson, S. (2019) Failure detection in robotic arms using statistical modeling, machine learning and hybrid gradient boosting. *Measurement* **146**, 425–436. Elsevier.

Demirci, M. & Baltaci, A. (2013) Prediction of suspended sediment in river using fuzzy logic and multilinear regression approaches. *Neural Comput. Appl.* **23**(1), 145–151. Springer.

Domingo, C. & Watanabe, O. (2000) MadaBoost: A modification of AdaBoost. *COLT*, 180–189. Citeseer.

Dorrell, R. M., Amy, L. A., Peakall, J. & McCaffrey, W. D. (2018) Particle size distribution controls the threshold between net sediment erosion and deposition in suspended load dominated flows. *Geophys. Res. Lett.* **45**(3), 1443–1452. Wiley Online Library.

Dudoit, S., Fridlyand, J. & Speed, T. P. (2002) Comparison of discrimination methods for the classification of tumors using gene expression data. *J. Am. Stat. Assoc.* **97**(457), 77–87. Taylor & Francis.

Emamgholizadeh, S. & Demneh, R. K. (2019) A comparison of artificial intelligence models for the estimation of daily suspended sediment load: a case study on the Telar and Kasilian rivers in Iran. *Water Sci. Technol. Water Supply* **19**(1), 165–178. IWA Publishing.

Firat, M. & Güngör, M. (2010) Monthly total sediment forecasting using adaptive neuro fuzzy inference system. *Stoch. Environ. Res. Risk Assess.* **24**(2), 259–270. Springer.

Freund, Y. & Schapire, R. E. (1997) A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **55**(1), 119–139. Elsevier.

Friedman, J. H. (2001) Greedy function approximation: a gradient boosting machine. *Ann. Stat.* 1189–1232. JSTOR.

Friedman, J. H. (2002) Stochastic gradient boosting. *Comput. Stat. Data Anal.* **38**(4), 367–378. Elsevier.

Friedman, J., Hastie, T. & Tibshirani, R. (2000) Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *Ann. Stat.* **28**(2), 337–407. Institute of Mathematical Statistics.

Frings, R. M. & Kleinhans, M. G. (2008) Complex variations in sediment transport at three large river bifurcations during discharge waves in the river Rhine. *Sedimentology* **55**(5), 1145–1171. Wiley Online Library.

Georganos, S., Grippa, T., Niang Gadiaga, A., Linard, C., Lennert, M., Vanhuysse, S., Mboga, N., et al. (2019) Geographical random forests: a spatial extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population

modelling. *Geocarto Int.* 1–16. Taylor & Francis.
Ghorbani, M. A., Deo, R. C., Yaseen, Z. M. & Kashani, M. H. (2017) Pan evaporation prediction using a hybrid multilayer perceptron-firefly algorithm (MLP-FFA) model : case study in North Iran. Theoretical and Applied Climatology. doi:10.1007/s00704-017-2244-0
Gill, M. A. (1983) Diffusion model for aggrading channels. *J. Hydraul. Res.* **21**(5), 355–367. Taylor & Francis Group.
Gill, M. A. (1983) Diffusion model for degrading channels. *J. Hydraul. Res.* **21**(5), 369–378. Taylor & Francis Group.
Gong, M., Bai, Y., Qin, J., Wang, J., Yang, P. & Wang, S. (2020) Gradient boosting machine for predicting return temperature of district heating system: A case study for residential buildings in Tianjin. *J. Build. Eng.* **27**, 100950. Elsevier.
Goyal, M. K. (2014) Modeling of sediment yield prediction using M5 model tree algorithm and wavelet regression. *Water Resour. Manag.* **28**(7), 1991–2003. Springer.
Guven, A. & Kişi, Ö. (2011) Estimation of suspended sediment yield in natural rivers using machine-coded linear genetic programming. *Water Resour. Manag.* **25**(2), 691–704. Springer.
Hai, T., Sharafati, A., Mohammed, A., Salih, S. Q., Deo, R. C., Al-Ansari, N. & Yaseen, Z. M. (2020) Global solar radiation estimation and climatic variability analysis using extreme learning machine based predictive model. *IEEE Access* **8**, 12026–12042. IEEE.
Hamaamin, Y. A., Nejadhashemi, A. P., Zhang, Z., Giri, S., Adhikari, U. & Herman, M. R. (2019) Evaluation of neuro-fuzzy and Bayesian techniques in estimating suspended sediment loads. *Sustain. Water Resour. Manag.* **5**(2), 639–654. Springer.
Hassanpour, F., Sharifazari, S., Ahmadaali, K., Mohammadi, S. & Sheikhalipour, Z. (2019) Development of the FCM-SVR Hybrid Model for Estimating the Suspended Sediment Load. *KSCE J. Civ. Eng.* **23**(6), 2514–2523. Springer.
Himanshu, S. K., Asce, S. M., Pandey, A. & Yadav, B. (2016) Ensemble Wavelet-Support Vector Machine Approach for Prediction of Suspended Sediment Load Using Hydrometeorological Data **22**(3), 1–14. doi:10.1061/(ASCE)HE.1943-5584.0001516.
Himanshu, S. K., Pandey, A. & Yadav, B. (2017) Ensemble wavelet-support vector machine approach for prediction of suspended sediment load using hydrometeorological data. *J. Hydrol. Eng.* **22**(7), 5017006. American Society of Civil Engineers.
Himanshu, S. K., Pandey, A. & Yadav, B. (2017) Assessing the applicability of TMPA-3B42V7 precipitation dataset in wavelet-support vector machine approach for suspended sediment load prediction. *J. Hydrol.* **550**, 103–117. Elsevier B.V. doi:10.1016/j.jhydrol.2017.04.051
Ho, T. K. (1995) Random decision forests. *Proc. 3rd Int. Conf. Doc. Anal. Recognit.*, Vol. 1, 278–282. IEEE.
Huang, Q., Chen, Y., Liu, L., Tao, D. & Li, X. (2019) On combining biclustering mining and AdaBoost for breast tumor classification. *IEEE Trans. Knowl. Data Eng.* IEEE.
Jain, S. K. (2001) Development of integrated sediment rating curves using ANNs. *J. Hydraul. Eng.* **127**(1), 30–37. American Society of Civil Engineers.
Julien, P. Y. & Vensel, C. W. (2005) Review of sedimentation issues on the Mississippi River. *Draft Rep. Present. to UNESCO ISI. Color. State Univ.*
Kamali, B., Mousavi, S. J. & Abbaspour, K. C. (2013) Automatic calibration of HEC-HMS using single-objective and multi-objective PSO algorithms. *Hydrol. Process.* **27**(26), 4028–4042. Wiley Online Library.
Kisi, O. (2005) Suspended sediment estimation using neuro-fuzzy and neural network approaches/Estimation des matières en suspension par des approches neurofloues et à base de réseau de neurones. *Hydrol. Sci. J.* **50**(4). Taylor & Francis.

Kisi, O. (2012) Modeling discharge-suspended sediment relationship using least square support vector machine. *J. Hydrol.* **456**, 110–120. Elsevier.

Kişi, Ö. (2010) River suspended sediment concentration modeling using a neural differential evolution approach. *J. Hydrol.* **389**(1–2), 227–235. Elsevier.

Kisi, O., Dailr, A. H., Cimen, M. & Shiri, J. (2012) Suspended sediment modeling using genetic programming and soft computing techniques. *J. Hydrol.* **450**, 48–58. Elsevier.

Kisi, O. & Shiri, J. (2012) River suspended sediment estimation by climatic variables implication: comparative study among soft computing techniques. *Comput. Geosci.* **43**, 73–82. Elsevier.

Kisi, O. & Yaseen, Z. M. (2019) The potential of hybrid evolutionary fuzzy intelligence model for suspended sediment concentration prediction. *Catena* **174**(May 2018), 11–23. Elsevier. doi:10.1016/j.catena.2018.10.047

Kisi, O., Yuksel, I. & Dogan, E. (2008) Modelling daily suspended sediment of rivers in Turkey using several data-driven techniques/Modélisation de la charge journalière en matières en suspension dans des rivières turques à l'aide de plusieurs techniques empiriques. *Hydrol. Sci. J.* **53**(6), 1270–1285. Taylor & Francis.

Kisi, O. & Zounemat-Kermani, M. (2016) Suspended Sediment Modeling Using Neuro-Fuzzy Embedded Fuzzy c-Means Clustering Technique. *Water Resour. Manag.* **30**(11), 3979–3994. Water Resources Management. doi:10.1007/s11269-016-1405-8

Kumar, D., Pandey, A., Sharma, N. & Flügel, W.-A. (2016) Daily suspended sediment simulation using machine learning approach. *Catena* **138**, 77–90. Elsevier.

Kuok, K. K., Kueh, S. M. & Chiu, P. C. (2018) Bat optimisation neural networks for rainfall forecasting: case study for Kuching city. *J. Water Clim. Chang.* jwc2018136. IWA Publishing.

Larivière, B. & Poel, D. Van den. (2005) Predicting customer retention and profitability by using random forests and regression forests techniques. *Expert Syst. Appl.* **29**(2), 472–484. Elsevier.

Lawrence, I. & Lin, K. (1989) A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 255–268. JSTOR.

Lisle, T. E., Cui, Y., Parker, G., Pizzuto, J. E. & Dodd, A. M. (2001) The dominance of dispersion in the evolution of bed material waves in gravel-bed rivers. *Earth Surf. Process. Landforms* **26**(13), 1409–1420. Wiley Online Library.

Lohani, A. K., Goel, N. K. & Bhatia, K. K. S. (2007) Deriving stage–discharge–sediment concentration relationships using fuzzy logic. *Hydrol. Sci. J.* **52**(4), 793–807. Taylor & Francis.

Lopez-Martin, M., Carro, B. & Sanchez-Esguevillas, A. (2019) Neural network architecture based on gradient boosting for IoT traffic prediction. *Futur. Gener. Comput. Syst.* Elsevier.

Lyn, D. A. & Altinakar, M. (2002) St. Venant–Exner equations for near-critical and transcritical flows. *J. Hydraul. Eng.* **128**(6), 579–587. American Society of Civil Engineers.

Malik, A., Kumar, A., Kim, S., Kashani, M. H., Karimi, V., Sharafati, A., Ghorbani, M. A., et al. (2020) Modeling monthly pan evaporation process over the Indian central Himalayas: application of multiple learning artificial intelligence model. *Eng. Appl. Comput. Fluid Mech.* **14**(1), 323–338. Taylor & Francis.

Malik, A., Kumar, A. & Piri, J. (2017) Daily suspended sediment concentration simulation using hydrological data of Pranhita River Basin, India. *Comput. Electron. Agric.* **138**, 20–28. Elsevier B.V. doi:10.1016/j.compag.2017.04.005

Meade, R. H. & Moody, J. A. (2010) Causes for the decline of suspended-sediment discharge in the Mississippi River system, 1940–2007. *Hydrol. Process. An Int. J.* **24**(1), 35–49.

Wiley Online Library.
Melesse, A. M., Ahmad, S., McClain, M. E., Wang, X. & Lim, Y. H. (2011) Suspended sediment load prediction of river systems: An artificial neural network approach. *Agric. Water Manag.* **98**(5), 855–866. Elsevier.
Miskell, G., Pattinson, W., Weissert, L. & Williams, D. (2019) Forecasting short-term peak concentrations from a network of air quality instruments measuring PM2. 5 using boosted gradient machine models. *J. Environ. Manage.* **242**, 56–64. Elsevier.
Moeeni, H. & Bonakdari, H. (2018) Impact of Normalization and Input on ARMAX-ANN Model Performance in Suspended Sediment Load Prediction. *Water Resour. Manag.* **32**(3), 845–863. Water Resources Management. doi:10.1007/s11269-017-1842-z
Mohammadian, A., Tajrishi, M. & Azad, F. L. (2004) Two dimensional numerical simulation of flow and geo-morphological processes near headlands by using unstructured grid. *Int. J. Sediment Res* **19**(4), 258–277.
Mohammed, M., Sharafati, A., Al-Ansari, N. & Yaseen, Z. M. (2020) Shallow Foundation Settlement Quantification: Application of Hybridized Adaptive Neuro-Fuzzy Inference System Model. *Adv. Civ. Eng.* **2020**. Hindawi.
Moriasi, D. N., Arnold, J. G., Liew, M. W. Van, Bingner, R. L., Harmel, R. D. & Veith, T. L. (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* **50**(3), 885–900. American society of agricultural and biological engineers.
Mossa, J. (1996) Sediment dynamics in the lowermost Mississippi River. *Eng. Geol.* **45**(1–4), 457–479. Elsevier.
Mustafa, M. R., Rezaur, R. B., Saiedi, S. & Isa, M. H. (2012) River suspended sediment prediction using various multilayer perceptron neural network training algorithms—a case study in Malaysia. *Water Resour. Manag.* **26**(7), 1879–1897. Springer.
Nagy, H. M., Watanabe, K. & Hirano, M. (2002) Prediction of sediment load concentration in rivers using artificial neural network model. *J. Hydraul. Eng.* **128**(6), 588–595. American Society of Civil Engineers.
Nivesh, S. & Kumar, P. (2018) River suspended sediment load prediction using neuro-fuzzy and statistical models : Vamsadhara river basin , India (April).
Nourani, V., Alizadeh, F. & Roushangar, K. (2016) Evaluation of a Two-Stage SVM and Spatial Statistics Methods for Modeling Monthly River Suspended Sediment Load. *Water Resour. Manag.* **30**(1), 393–407. doi:10.1007/s11269-015-1168-7
Nourani, V. & Andalib, G. (2015) Daily and monthly suspended sediment load predictions using wavelet based artificial intelligence approaches. *J. Mt. Sci.* **12**(1), 85–100. doi:10.1007/s11629-014-3121-2
Nourani, V., Molajou, A., Tajbakhsh, A. D. & Najafi, H. (2019) A wavelet based data mining technique for suspended sediment load modeling. *Water Resour. Manag.* **33**(5), 1769–1784. Springer.
Olyaie, E., Banejad, H., Chau, K.-W. & Melesse, A. M. (2015) A comparison of various artificial intelligence approaches performance for estimating suspended sediment load of river systems: a case study in United States. *Environ. Monit. Assess.* **187**(4), 189. doi:10.1007/s10661-015-4381-1
Opitz, D. & Maclin, R. (1999) Popular ensemble methods: An empirical study. *J. Artif. Intell. Res.* **11**, 169–198.
Özger, M. & Kabataş, M. B. (2015) Sediment load prediction by combined fuzzy logic-wavelet method. *J. Hydroinformatics* **17**(6), 930–942. doi:10.2166/hydro.2015.148
Pan, T.-Y., Yang, Y.-T., Kuo, H.-C., Tan, Y.-C., Lai, J.-S., Chang, T.-J., Lee, C.-S., et al. (2013) Improvement of watershed flood forecasting by typhoon rainfall climate model with an ANN-based southwest monsoon rainfall enhancement. *J. Hydrol.* **506**, 90–100.

Elsevier.
Pan, Y., Chen, S., Qiao, F., Ukkusuri, S. V & Tang, K. (2019) Estimation of real-driving emissions for buses fueled with liquefied natural gas based on gradient boosted regression trees. *Sci. Total Environ.* **660**, 741–750. Elsevier.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., et al. (2011) Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**(Oct), 2825–2830.
Pektas, A. O. & Cigizoglu, H. K. (2017) Investigating the extrapolation performance of neural network models in suspended sediment data. *Hydrol. Sci. J.* **62**(10), 1694–1703. Taylor & Francis. doi:10.1080/02626667.2017.1349316
Polikar, R. (2006) Ensemble based systems in decision making. *IEEE Circuits Syst. Mag.* **6**(3), 21–45. IEEE.
Rahgoshay, M., Feiznia, S., Arian, M. & Hashemi, S. A. A. (2018) Modeling daily suspended sediment load using improved support vector machine model and genetic algorithm. *Environ. Sci. Pollut. Res.* **25**(35), 35693–35706. Springer.
Rai, R. K. & Mathur, B. S. (2008) Event-based sediment yield modeling using artificial neural network. *Water Resour. Manag.* **22**(4), 423–441. Springer.
Rajaee, T., Mirbagheri, S. A., Zounemat-Kermani, M. & Nourani, V. (2009) Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models. *Sci. Total Environ.* **407**(17), 4916–4927. Elsevier.
Ramezani, F., Nikoo, M. & Nikoo, M. (2014) Artificial neural network weights optimization based on social-based algorithm to realize sediment over the river. *Soft Comput.* **19**(2), 375–387. doi:10.1007/s00500-014-1258-0
Rashidi, S., Vafakhah, M., Lafdani, E. K. & Javadi, M. R. (2016) Evaluating the support vector machine for suspended sediment load forecasting based on gamma test. *Arab. J. Geosci.* **9**(11). Arabian Journal of Geosciences. doi:10.1007/s12517-016-2601-9
Ritter, A. & Muñoz-Carpena, R. (2013) Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *J. Hydrol.* **480**, 33–45. Elsevier.
Rodriguez, M. J. & Sérodes, J.-B. (2004) Application of back-propagation neural network modeling for free residual chlorine, total trihalomethanes and trihalomethanes speciation. *J. Environ. Eng. Sci.* **3**(S1), S25–S34. NRC Research Press.
Rokach, L. (2010) Ensemble-based classifiers. *Artif. Intell. Rev.* **33**(1–2), 1–39. Springer.
Saffari, A., Godec, M., Pock, T., Leistner, C. & Bischof, H. (2010) Online multi-class LPBoost. *2010 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 3570–3577. Ieee.
Salih, S. Q., sharafati, A., Khosravi, K., Faris, H., Kisi, O., Tao, H., Ali, M., et al. (2019) River suspended sediment load prediction based on river discharge information: application of newly developed data mining models. *Hydrol. Sci. J.* (just-accepted). Taylor & Francis.
Samadi, S. H., Ghobadian, B. & Nosrati, M. (2019) Prediction of higher heating value of biomass materials based on proximate analysis using gradient boosted regression trees method. *Energy Sources, Part A Recover. Util. Environ. Eff.* 1–10. Taylor & Francis.
Samantaray, S. & Ghose, D. K. (2018) Evaluation of suspended sediment concentration using descent neural networks. *Procedia Comput. Sci.* **132**(Iccids), 1824–1831. Elsevier B.V. doi:10.1016/j.procs.2018.05.138
Sari, V., Reis Castro, N. M. dos & Pedrollo, O. C. (2017) Estimate of Suspended Sediment Concentration from Monitored Data of Turbidity and Water Level Using Artificial Neural Networks. *Water Resour. Manag.* **31**(15), 4909–4923. Water Resources Management. doi:10.1007/s11269-017-1785-4

Schapire, R. E., Freund, Y., Bartlett, P. & Lee, W. S. (1998) Boosting the margin: A new explanation for the effectiveness of voting methods. *Ann. Stat.* **26**(5), 1651–1686. Institute of Mathematical Statistics.

Senthil Kumar, A. R., Ojha, C. S. P., Goyal, M. K., Singh, R. D. & Swamee, P. K. (2011) Modeling of suspended sediment concentration at Kasol in India using ANN, fuzzy logic, and decision tree algorithms. *J. Hydrol. Eng.* **17**(3), 394–404. American Society of Civil Engineers.

Shamaei, E. & Kaedi, M. (2016) Suspended sediment concentration estimation by stacking the genetic programming and neuro-fuzzy predictions. *Appl. Soft Comput. J.* **45**, 187–196. Elsevier B.V. doi:10.1016/j.asoc.2016.03.009

Shamshirband, S., Jafari Nodoushan, E., Adolf, J. E., Abdul Manaf, A., Mosavi, A. & Chau, K. (2019) Ensemble models with uncertainty analysis for multi-day ahead forecasting of chlorophyll a concentration in coastal waters. *Eng. Appl. Comput. Fluid Mech.* **13**(1), 91–101. Taylor & Francis.

Sharafati, A., Yasa, R. & Azamathulla, H. M. (2018) Assessment of stochastic approaches in prediction of wave-induced pipeline scour depth. *J. Pipeline Syst. Eng. Pract.* **9**(4). doi:10.1061/(ASCE)PS.1949-1204.0000347

Sharafati, A. & Zahabiyoun, B. (2013) Stochastic generation of storm pattern. *Life Sci. J.* **10**(1), 1575–1583.

Sharafati, A & Azamathulla, H. M. (2018) Assessment of Dam Overtopping Reliability using SUFI Based Overtopping Threshold Curve. *Water Resour. Manag.* **32**(7), 2369–2383. Springer.

Sharafati, Ahmad, Khosravi, K., Khosravinia, P., Ahmed, K., Salman, S. A., Mundher, Z. & Shamsuddin, Y. (2019) The potential of novel data mining models for global solar radiation prediction. *Int. J. Environ. Sci. Technol.* (0123456789). Springer Berlin Heidelberg. doi:10.1007/s13762-019-02344-0

Sharafati, Ahmad & Pezeshki, E. (2020) A strategy to assess the uncertainty of a climate change impact on extreme hydrological events in the semi-arid Dehbar catchment in Iran. *Theor. Appl. Climatol.* **139**(1–2), 389–402. Springer.

Sharafati, Ahmad, Tafarojnoruz, A., Shourian, M. & Yaseen, Z. M. (2019) Simulation of the depth scouring downstream sluice gate: The validation of newly developed data-intelligent models. *J. Hydro-environment Res.* Elsevier.

Sharghi, E., Nourani, V., Najafi, H. & Gokcekus, H. (2019) Conjunction of a newly proposed emotional ANN (EANN) and wavelet transform for suspended sediment load modeling. *Water Supply* **19**(6), 1726–1734. IWA Publishing.

Shiau, J. T. & Chen, T. J. (2015) Quantile Regression-Based Probabilistic Estimation Scheme for Daily and Annual Suspended Sediment Loads. *Water Resour. Manag.* **29**(8), 2805–2818. doi:10.1007/s11269-015-0971-5

Singh, A., Imtiyaz, M., Isaac, R. K. & Denis, D. M. (2012) Comparison of artificial neural network models for sediment yield prediction at single gauging station of watershed in eastern India. *J. Hydrol. Eng.* **18**(1), 115–120. American Society of Civil Engineers.

Sivapragasam, C., Vanitha, S., Muttil, N., Suganya, K., Suji, S., Selvi, M. T., Selvi, R., et al. (2014) Monthly flow forecast for Mississippi River basin using artificial neural networks. *Neural Comput. Appl.* **24**(7–8), 1785–1793. Springer.

Song, Y., Zhou, H., Wang, P. & Yang, M. (2019) Prediction of clathrate hydrate phase equilibria using gradient boosted regression trees and deep neural networks. *J. Chem. Thermodyn.* **135**, 86–96. Elsevier.

Tabatabaei, M., Jam, A. S. & Hosseini, S. A. (2019) Suspended sediment load prediction using non-dominated sorting genetic algorithm II. *Int. Soil Water Conserv. Res.* Elsevier.

Talebi, A., Mahjoobi, J., Dastorani, M. T. & Moosavi, V. (2017) Estimation of suspended

sediment load using regression trees and model trees approaches (Case study: Hyderabad drainage basin in Iran). *ISH J. Hydraul. Eng.* **23**(2), 212–219. Taylor & Francis. doi:10.1080/09715010.2016.1264894

Tama, B. A. & Rhee, K.-H. (2019) An in-depth experimental study of anomaly detection using gradient boosted machine. *Neural Comput. Appl.* **31**(4), 955–965. Springer.

Tan, K., Wang, H., Chen, L., Du, Q., Du, P. & Pan, C. (2020) Estimation of the spatial distribution of heavy metal in agricultural soils using airborne hyperspectral imaging and random forest. *J. Hazard. Mater.* **382**, 120987. Elsevier.

Tang, X. & Knight, D. W. (2006) Sediment transport in river models with overbank flows. *J. Hydraul. Eng.* **132**(1), 77–86. American Society of Civil Engineers.

Taylor, K. E. (2001) Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **106**(D7), 7183–7192. doi:10.1029/2000JD900719

Tfwala, S. S. & Wang, Y. M. (2016) Estimating sediment discharge using sediment rating curves and artificial neural networks in the Shiwen River, Taiwan. *Water (Switzerland)* **8**(2). doi:10.3390/w8020053

Wang, L., Kisi, O., Zounemat-Kermani, M., Zhu, Z., Gong, W., Niu, Z., Liu, H., et al. (2017) Prediction of solar radiation in China using different adaptive neuro-fuzzy methods and M5 model tree. *Int. J. Climatol.* **37**(3), 1141–1155. doi:10.1002/joc.4762

Wang, Y., Du, Y., Wang, J. & Li, T. (2019) Calibration of a low-cost PM2. 5 monitor using a random forest model. *Environ. Int.* **133**, 105161. Elsevier.

Wilcock, P. R., Kenworthy, S. T. & Crowe, J. C. (2001) Experimental study of the transport of mixed sand and gravel. *Water Resour. Res.* **37**(12), 3349–3358. Wiley Online Library.

Willmott, C. J. & Matsuura, K. (2005) Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **30**(1), 79–82.

Wu, W. (2004) Depth-averaged two-dimensional numerical modeling of unsteady flow and nonuniform sediment transport in open channels. *J. Hydraul. Eng.* **130**(10), 1013–1024. American Society of Civil Engineers.

Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J. & Chen, Z. (2019) Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. *Remote Sens. Environ.* **233**, 111358. Elsevier.

Yaseen, Z., Ehteram, M., Sharafati, A., Shahid, S., Al-Ansari, N. & El-Shafie, A. (2018) The Integration of Nature-Inspired Algorithms with Least Square Support Vector Regression Models: Application to Modeling River Dissolved Oxygen Concentration. *Water* **10**(9), 1124. Multidisciplinary Digital Publishing Institute.

Yaseen, Z. M., Awadh, S. M., Sharafati, A. & Shahid, S. (2018) Complementary data-intelligence model for river flow simulation. *J. Hydrol.* **567**, 180–190. doi:10.1016/j.jhydrol.2018.10.020

Yaseen, Z. M., Ebtehaj, I., Bonakdari, H., Deo, R. C., Mehr, A. D., Hanna, W., Wan, M., et al. (2017) Novel approach for streamflow forecasting using a hybrid ANFIS-FFA model. *J. Hydrol.*

Yaseen, Z. M., Fu, M., Wang, C., Mohtar, W. H. M. W., Deo, R. C. & El-shafie, A. (2018) Application of the Hybrid Artificial Neural Network Coupled with Rolling Mechanism and Grey Model Algorithms for Streamflow Forecasting Over Multiple Time Horizons. *Water Resour. Manag.* 1–17. Springer.

Yilmaz, B., Aras, E., Nacar, S. & Kankal, M. (2018) Estimating suspended sediment load with multivariate adaptive regression spline, teaching-learning based optimization, and artificial bee colony models. *Sci. Total Environ.* **639**, 826–840. Elsevier B.V. doi:10.1016/j.scitotenv.2018.05.153

Yu, D.-J., Hu, J., Tang, Z.-M., Shen, H.-B., Yang, J. & Yang, J.-Y. (2013) Improving protein-ATP binding residues prediction by boosting SVMs with random under-sampling. *Neurocomputing* **104**, 180–190. Elsevier.

Zhang, G. & Fang, B. (2007) LogitBoost classifier for discriminating thermophilic and mesophilic proteins. *J. Biotechnol.* **127**(3), 417–424. Elsevier.

Zhang, H. & Kahawita, R. (1987) Nonlinear model for aggradation in alluvial channels. *J. Hydraul. Eng.* **113**(3), 353–368. American Society of Civil Engineers.

Zhao, J., Jiao, L., Xia, S., Fernandes, V. B., Yevseyeva, I., Zhou, Y. & Emmerich, M. T. M. (2018) Multiobjective sparse ensemble learning by means of evolutionary algorithms. *Decis. Support Syst.* **111**, 86–100. Elsevier.

Zhou, F., Zhang, Q., Sornette, D. & Jiang, L. (2019) Cascading logistic regression onto gradient boosted decision trees for forecasting and trading stock indices. *Appl. Soft Comput.* **84**, 105747. Elsevier.

Zounemat-Kermani, M., Kişi, Ö., Adamowski, J. & Ramezani-Charmahineh, A. (2016) Evaluation of data driven models for river suspended sediment concentration modeling. *J. Hydrol.* **535**, 457–472. Elsevier B.V. doi:10.1016/j.jhydrol.2016.02.012

Table 1: Summary of Artificial Neural Network (ANN) models for Suspended Sediment Load (SSL) prediction.

Scholars	Best predictive model	Input variable(s)	Study area	Time scale
(Jain, 2001)	FFNN	Discharge, Water stage, SSL	USA	Daily
(H Kerem Cigizoglu, 2004)	MLP	Discharge, SSL	USA	Daily
(Agarwal *et al.*, 2005)	FFNN	Discharge, Rainfall, SSL	India	Daily, Monthly
(Hikmet Kerem Cigizoglu & Alp, 2006)	FFNN	Discharge, SSL	USA	Daily
(Rai & Mathur, 2008)	FFNN	Discharge, Rainfall, SSL	USA	Daily
(Kisi *et al.*, 2008)	MLP	Discharge, SSL	Turkey	Daily
(Melesse *et al.*, 2011)	FFNN	Discharge, Rainfall, SSL	USA	Daily, Weekly
(Mustafa *et al.*, 2012)	MLP	Discharge, SSL	Malaysia	Daily
(Singh *et al.*, 2012)	FFNN	Discharge, Rainfall	India	Monthly
(Afan *et al.*, 2014)	FFNN	Discharge, SSL	Malaysia	Daily
(Ramezani *et al.*, 2014)	FF-SBA	Discharge, Debit, River length	Iran	Monthly
(Zounemat-Kermani *et al.*, 2016)	FFNN	Discharge, SSL	USA	Daily
(Tfwala & Wang, 2016)	MLP	Discharge, SSL	Taiwan	Hourly
(Chen & Chau, 2016)	HDFNN	Discharge, SSL	USA	Daily
(Sari *et al.*, 2017)	MLP	Turbidity, Water stage	Brazil	Monthly
(Adib & Mahmoodi, 2017)	MLP-GA	Discharge	Iran	Daily
(Pektas & Cigizoglu, 2017)	FFBP	SSL	USA	Daily
(Samantaray & Ghose, 2018)	NNFIT	Discharge, SSL	India	Monthly

FFBP: feed forward back propagation, FFNN: feed forward neural network, FF-SBA: feed forward social based algorithm, HDFNN: hybrid double feedforward neural network, MLP: multi-layer back prorogation, NNFIT: neural network fitting.

Table 2: Summary of fuzzy logic-based models for Suspended Sediment Load (SSL) prediction

Scholars	Best predictive model	Input variable(s)	Study area	Time scale
(Kisi, 2005)	Neuro-fuzzy	Discharge, SSL	USA	Daily
(Lohani *et al.*, 2007)	Fuzzy logic	Discharge, Water stage, , SSL	India	Daily
(Firat & Güngör, 2010)	ANFIS	Discharge, SSL	Turkey	Monthly
(Rajaee *et al.*, 2009)	Neuro-fuzzy	Discharge, SSL	USA	Daily
(Cobaner *et al.*, 2009)	Neuro-fuzzy	Discharge, Rainfall, SSL	USA	Daily
(Demirci & Baltaci, 2013)	Fuzzy logic	Discharge, Temperature, SSL	USA	Daily
(Özger & Kabataş, 2015)	Fuzzy logic	SSL	Turkey	Monthly
(Kisi & Zounemat-Kermani, 2016)	ANFIS-FCM	Discharge	USA	Daily
(Malik *et al.*, 2017)	CANFIS	Discharge, SSL	India	Daily
(Nivesh & Kumar, 2018)	ANFIS	Discharge, Rainfall, SSL	India	Daily
(Kisi & Yaseen, 2019)	EF	Discharge	USA	Daily

ANFIS: adaptive neuro fuzzy inference system, ANFIS-FCM: adaptive neuro-fuzzy embedded fuzzy c-means clustering, CANFIS: co-active neuro-fuzzy inference system , EF: evolutionary fuzzy

Table 3: Summary of other artificial intelligence models for Suspended Sediment Load (SSL) prediction

Scholars	Best predictive model	Input variable(s)	Study area	Time scale
(Guven & Kişi, 2011)	LGP	Discharge, SSL	USA	Daily
(Kişi, 2010)	NDE	Discharge, SSL	USA	Daily
(Kisi, 2012)	LSSVM	Discharge, SSL	USA	Daily
(Kisi *et al.*, 2012)	GP	Discharge, SSL	USA	Daily
(Kisi & Shiri, 2012)	GEP	Discharge, Rainfall, SSL	USA	Daily
(Goyal, 2014)	MT	Discharge, Rainfall, SSL	India	Monthly
(Nourani & Andalib, 2015)	WLSSVM	Discharge, SSL	USA	Daily, Monthly
(Shiau & Chen, 2015)	QR	Discharge	Taiwan	Daily
(Nourani *et al.*, 2016)	SVM	Discharge, SSL	Iran	Monthly
(Rashidi *et al.*, 2016)	GT-SVM	Discharge, SSL	Iran	Monthly
(Shamaei & Kaedi, 2016)	LGP	Discharge, SSL	USA	Daily
(Himanshu *et al.*, 2016)	WLSSVM	Discharge, Rainfall, SSL	India	Daily
(Buyukyildiz & Kumcu, 2017)	SVR	Discharge, SSL	Turkey	Daily
(Himanshu *et al.*, 2017b)	WLSSVM	Discharge, Rainfall, SSL	India	Daily, Monthly
(Talebi *et al.*, 2017)	MT	Discharge	Iran	Daily
(Moeeni & Bonakdari, 2018)	ARMAX-ANN	Discharge, SSL	USA	Daily
(Choubin *et al.*, 2018)	CART	Discharge, Rainfall, Water stage, SSL	Iran	Daily
(Yilmaz *et al.*, 2018)	MARS	Discharge, SSL	Turkey	Daily
(Emamgholizadeh & Demneh, 2019)	GEP	Discharge	Iran	daily
(Hassanpour *et al.*, 2019)	FCM-SVR	Discharge, SSL	Iran	daily
(Nourani *et al.*, 2019)	Wavelet-M5	Discharge, SSL	Iran, USA	Daily, Monthly
(Tabatabaei *et al.*, 2019)	NSGA-II	Discharge, SSL	Iran	Monthly

ARMAX: autoregressive-moving average with exogenous terms, , CART: classification and regression tree, GEP: gene expression programming, GP: genetic programming, GT-SVM: gamma test support vector machine, LGP: linear genetic programming, LSSVM: least square support vector machine, MARS: multivariate adaptive regression splines, MT: model trees, NDE: neural differential evolution, QR: quantile regression, SVM: support vector machine, WLSSVM: wavelet least square support vector machine, NSGA-II: non-dominated sorting Genetic Algorithm II

Table 4: Statistical parameters for Daily Discharge (Q), Suspended Sediment Concentration (SSC) and Suspended Sediment Load (SSL) at USGS station 05587455 (Mississippi river below Grafton, Illinois) for the period 2007-2015

Statistic	Q (ft^3/s)	SSC (mg/L)	SSL (ton/day)
Xmax	400000	589	387000
Xmin	13400	6	489
Xmean	145346.7	118.6	61538.6
Sx	92695.5	102.8	70827.8
Csx	0.717	1.486	1.488

Table 5: Ensemble model parameters for Gradient Boost Regression (GBR), Ada Boost Regression (ABR) and Random Forest Regression (RFR) models

	Gradient Boost Regression		**Ada Boost Regression**		**Random Forest regression**	
Parameter	**Default**	**Used**	**Default**	**Used**	**Default**	**Used**
n Estimator	100	200	50	30	100	200
Max Depth	3	4	10	100	None	None
Min Sample Split	2	2	---	---	2	5
Learning Rate	0.1	0.1	1	0.1	---	---
Loss Function	Ls Lad Huber Quantile	Ls	Linear Square exponential	Linear	---	---

Table 6 – Pearson correlation between selected input and predicted variables

	SSC(t)	**Q(t)**	**SSC(t-1)**	**Q(t-1)**	**SSC(t-2)**	**Q(t-2)**	**SSC(t-3)**	**Q(t-3)**	**SSC(t-4)**	**Q(t-4)**
SSL(t)	0.915	0.781	0.880	0.754	0.810	0.721	0.744	0.689	0.684	0.659
SSL(t+1)	0.881	0.754	0.810	0.721	0.744	0.689	0.684	0.659		
SSL(t+3)	0.744	0.689	0.684	0.659						

Table 7 – Best input combinations considered for prediction of the SSL(t) time series

Input Combinations	Predictive variables
M1	SSC(t), SSC(t-1), SSC(t-2), Q(t), Q(t-1), SSC(t-3), Q(t-2), Q(t-3), SSC(t-4), Q(t-4)
M2	SSC(t), SSC(t-1), SSC(t-2), Q(t), Q(t-1), SSC(t-3), Q(t-2), Q(t-3), SSC(t-4)
M3	SSC(t), SSC(t-1), SSC(t-2), Q(t), Q(t-1), SSC(t-3), Q(t-2), Q(t-3)
M4	SSC(t), SSC(t-1), SSC(t-2), Q(t), Q(t-1), SSC(t-3), Q(t-2)
M5	SSC(t), SSC(t-1), SSC(t-2), Q(t), Q(t-1), SSC(t-3)
M6	SSC(t), SSC(t-1), SSC(t-2), Q(t), Q(t-1)
M7	SSC(t), SSC(t-1), SSC(t-2), Q(t)
M8	SSC(t), SSC(t-1), SSC(t-2)
M9	SSC(t), SSC(t-1)
M10	SSC(t)

Table 8 – Best input combinations considered for prediction of the SSL(t+1) time series

Input Combinations	Predictive variables
M1	SSC(t), SSC(t-1), Q(t), SSC(t-2), Q(t-1), Q(t-2), SSC(t-3), Q(t-3)
M2	SSC(t), SSC(t-1), Q(t), SSC(t-2), Q(t-1), Q(t-2), SSC(t-3)
M3	SSC(t), SSC(t-1), Q(t), SSC(t-2), Q(t-1), Q(t-2)
M4	SSC(t), SSC(t-1), Q(t), SSC(t-2), Q(t-1)
M5	SSC(t), SSC(t-1), Q(t), SSC(t-2)
M6	SSC(t), SSC(t-1), Q(t)
M7	SSC(t), SSC(t-1)
M8	SSC(t)

Table 9 – Best input combinations considered for prediction of the SSL(t+3) time series

Input Combinations	Predictive variables
M1	SSC(t), Q(t), SSC(t-1), Q(t-1)
M2	SSC(t), Q(t), SSC(t-1)
M3	SSC(t), Q(t)
M4	SSC(t)

Table 10– Prediction performance indices for SSL(t), SSL(t+1) and SSL(t+3) obtained using Gradient Boost Regression (GBR)

Target Variable	Input Combination	MAE (ton/day)	RMSE (ton/day)	R^2	NSE
SSL(t)	M1	2305.232	5748.586	0.994	0.994
	M2	2515.935	6783.551	0.992	0.991
	M3	2304.048	6216.710	0.993	0.993
	M4	2411.340	6659.683	0.992	0.991
	M5	2456.772	7249.205	0.990	0.990
	M6	2463.932	5999.943	0.994	0.993
	M7	2195.211	5512.286	0.995	0.994
	M8	16773.242	28013.148	0.852	0.848
	M9	17517.357	29449.214	0.835	0.832
	M10	18480.468	32026.136	0.803	0.802
SSL(t+1)	M1	10253.613	21292.751	0.906	0.905
	M2	10292.561	21783.516	0.901	0.901
	M3	10129.111	20750.383	0.911	0.910
	M4	10020.559	21136.775	0.907	0.906
	M5	9815.220	20734.854	0.911	0.910
	M6	9654.667	21096.573	0.908	0.907
	M7	19403.365	32208.402	0.787	0.783
	M8	20417.279	32631.537	0.778	0.777
SSL(t+3)	M1	18711.426	36969.555	0.710	0.702
	M2	19047.896	37560.660	0.701	0.692
	M3	19623.782	37704.309	0.697	0.690
	M4	26913.895	46211.982	0.548	0.534

Table 11 – Prediction performance indices for SSL(t), SSL(t+1) and SSL(t+3) obtained using Ada Boost Regression (ABR)

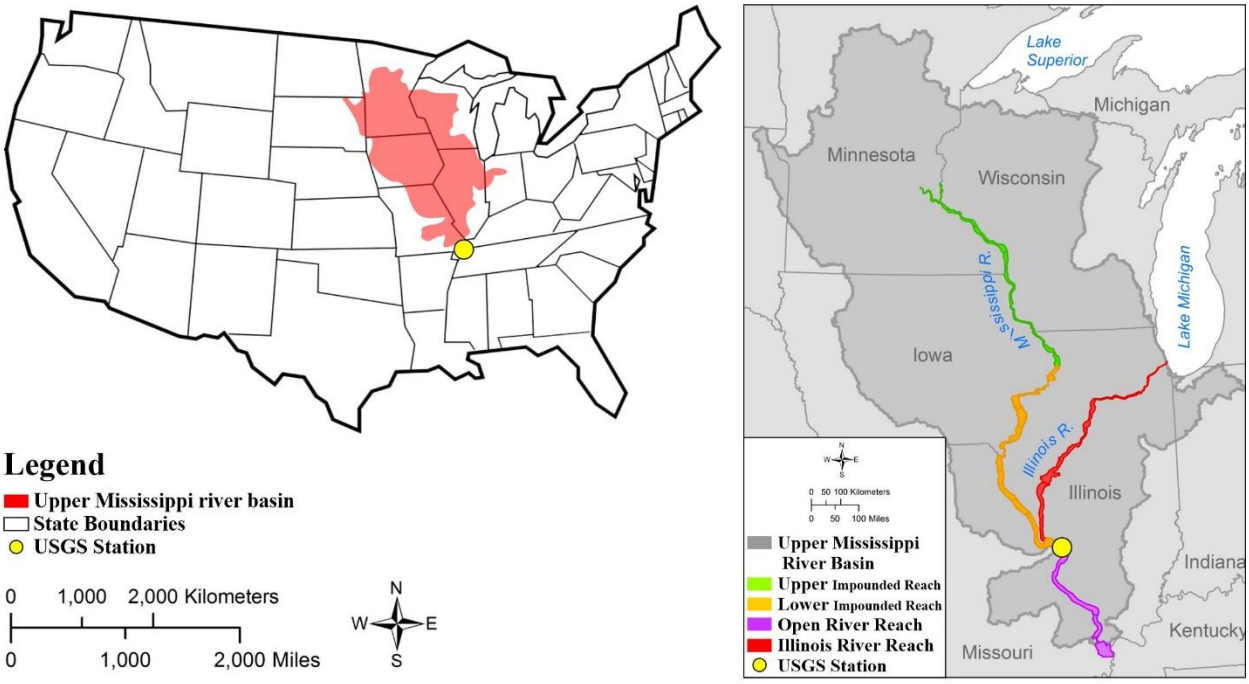
Target Variable	Input Combination	MAE (ton/day)	RMSE (ton/day)	R^2	NSE
SSL(t)	M1	2074.539	5366.950	0.995	0.994
	M2	2263.754	6055.567	0.993	0.993
	M3	2072.264	5256.417	0.995	0.995
	M4	2101.006	5394.456	0.995	0.994
	M5	2104.699	5466.907	0.995	0.994
	M6	2000.133	5068.103	0.995	0.995
	M7	1929.674	5195.675	0.995	0.995
	M8	18267.302	31510.344	0.811	0.808
	M9	19431.838	33675.553	0.786	0.781
	M10	20424.620	34252.543	0.777	0.773
SSL(t+1)	M1	11857.817	27279.099	0.855	0.844
	M2	11443.277	26003.969	0.865	0.858
	M3	11160.600	25098.966	0.873	0.868
	M4	10917.098	24609.220	0.876	0.873
	M5	11356.666	25792.813	0.865	0.861
	M6	10786.019	23289.028	0.887	0.886
	M7	23583.464	39062.095	0.693	0.680
	M8	24856.751	38938.279	0.698	0.682
SSL(t+3)	M1	19136.383	37672.409	0.699	0.691
	M2	19763.037	38037.871	0.691	0.685
	M3	23105.003	44204.423	0.600	0.574
	M4	32846.963	53401.274	0.448	0.378

Table 12 – Prediction performance indices for SSL(t), SSL(t+1) and SLL(t+3) obtained using Random Forest Regression (RFR)

Target Variable	Input Combination	MAE (ton/day)	RMSE (ton/day)	R^2	NSE
SSL(t)	M1	1701.143	4883.321	0.996	0.995
	M2	1705.770	4899.723	0.996	0.995
	M3	1705.422	4847.778	0.996	0.995
	M4	1687.944	4846.538	0.996	0.995
	M5	1679.124	4834.228	0.996	0.995
	M6	1662.183	4789.289	0.996	0.996
	M7	1575.734	4624.332	0.996	0.996
	M8	17481.476	28386.176	0.845	0.844
	M9	17952.052	29003.889	0.838	0.837
	M10	18767.744	31116.474	0.813	0.813
SSL(t+1)	M1	10851.661	22030.972	0.902	0.898
	M2	10839.857	21912.638	0.902	0.899
	M3	10765.171	21580.923	0.905	0.902
	M4	10509.656	21218.148	0.907	0.906
	M5	10733.093	21199.616	0.908	0.906
	M6	10431.096	20378.909	0.914	0.913
	M7	20822.172	32635.966	0.778	0.777
	M8	22734.544	34818.599	0.750	0.746
SSL(t+3)	M1	20387.187	36089.119	0.718	0.716
	M2	21017.470	37041.385	0.704	0.701
	M3	22451.284	38874.833	0.676	0.671
	M4	30851.522	49265.926	0.495	0.471

Table 13 – Prediction performance indices for the best predictive models

Target Variable	Predictive Model	MAE (ton/day)	RMSE (ton/day)	R^2	NSE
SSL(t)	ABR-M6	2000.133	5068.103	0.995	0.995
	GBR-M7	2195.211	5512.286	0.995	0.994
	RFR-M7	1575.734	4624.332	0.996	0.996
SSL(t+1)	ABR-M6	10786.019	23289.028	0.887	0.886
	GBR-M5	9815.220	20734.854	0.911	0.910
	RFR-M6	10431.096	20378.909	0.914	0.913
SSL(t+3)	ABR-M1	19136.383	37672.409	0.699	0.691
	GBR-M1	18711.426	36969.555	0.711	0.702
	RFR-M1	20387.187	36089.119	0.718	0.716



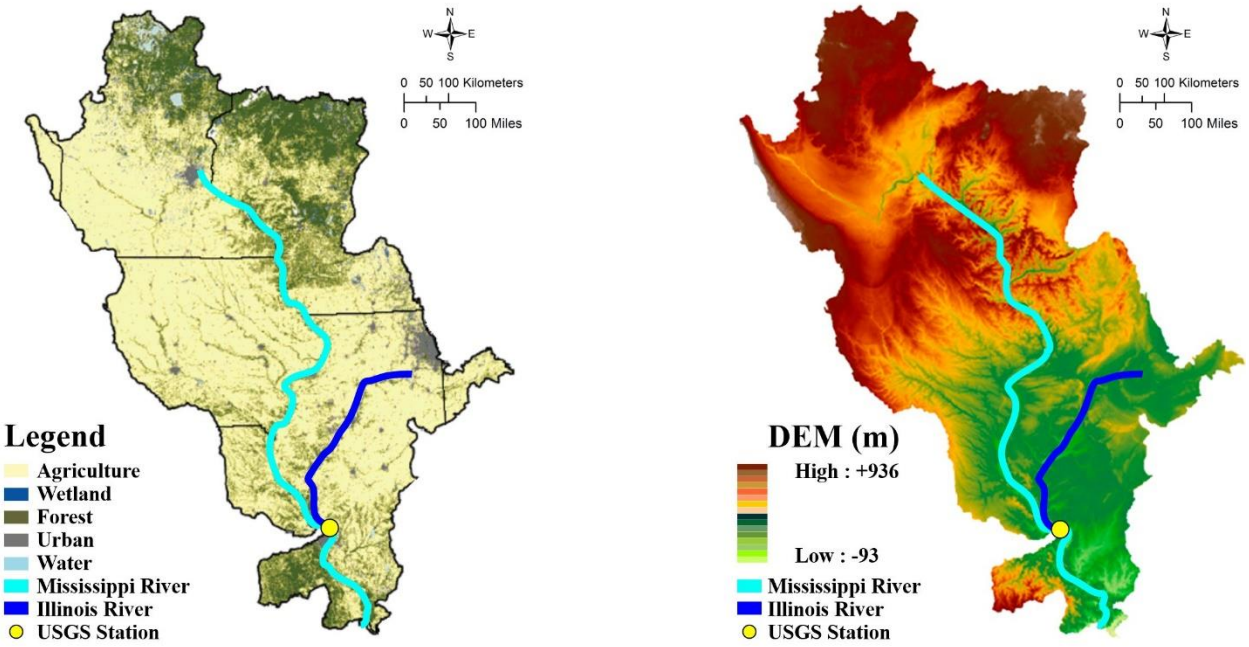


Figure 1: Location of the USGS station 05587455 (Mississippi river below Grafton, IL) and geographical information on the case study area.

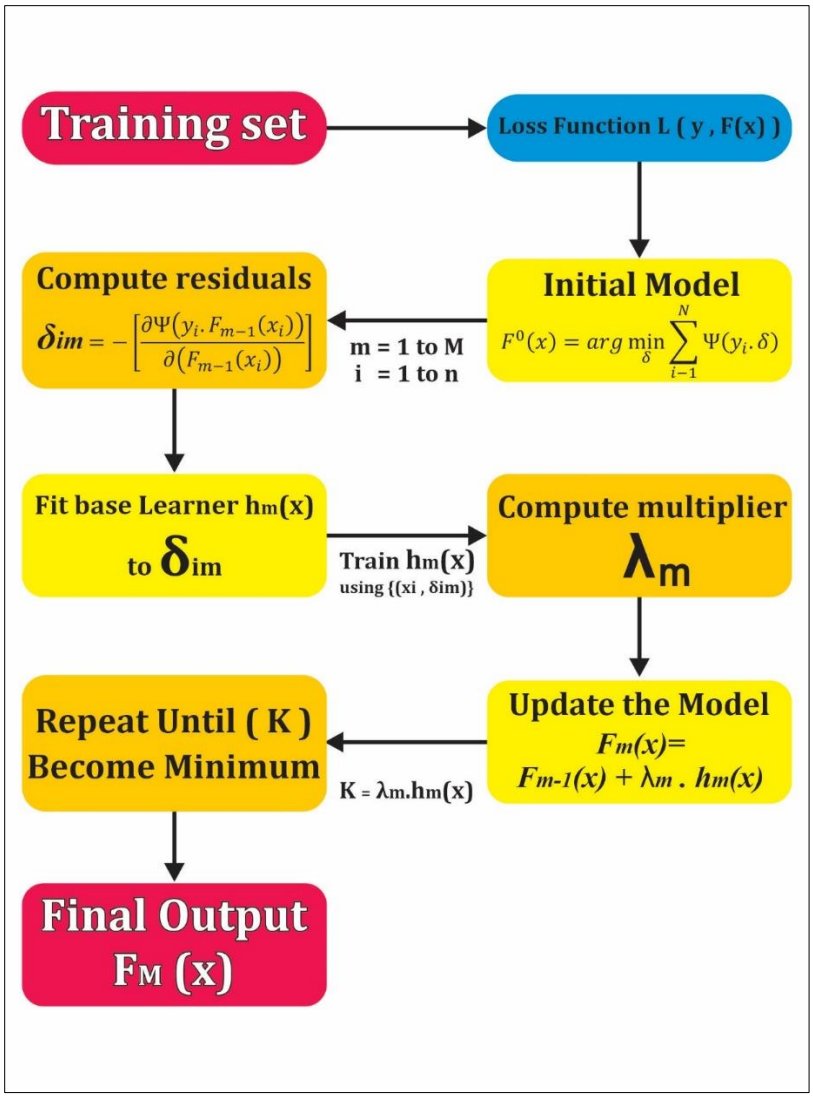


Figure 2: Conceptual representation of the Gradient Boost Regression (GBR) algorithm.

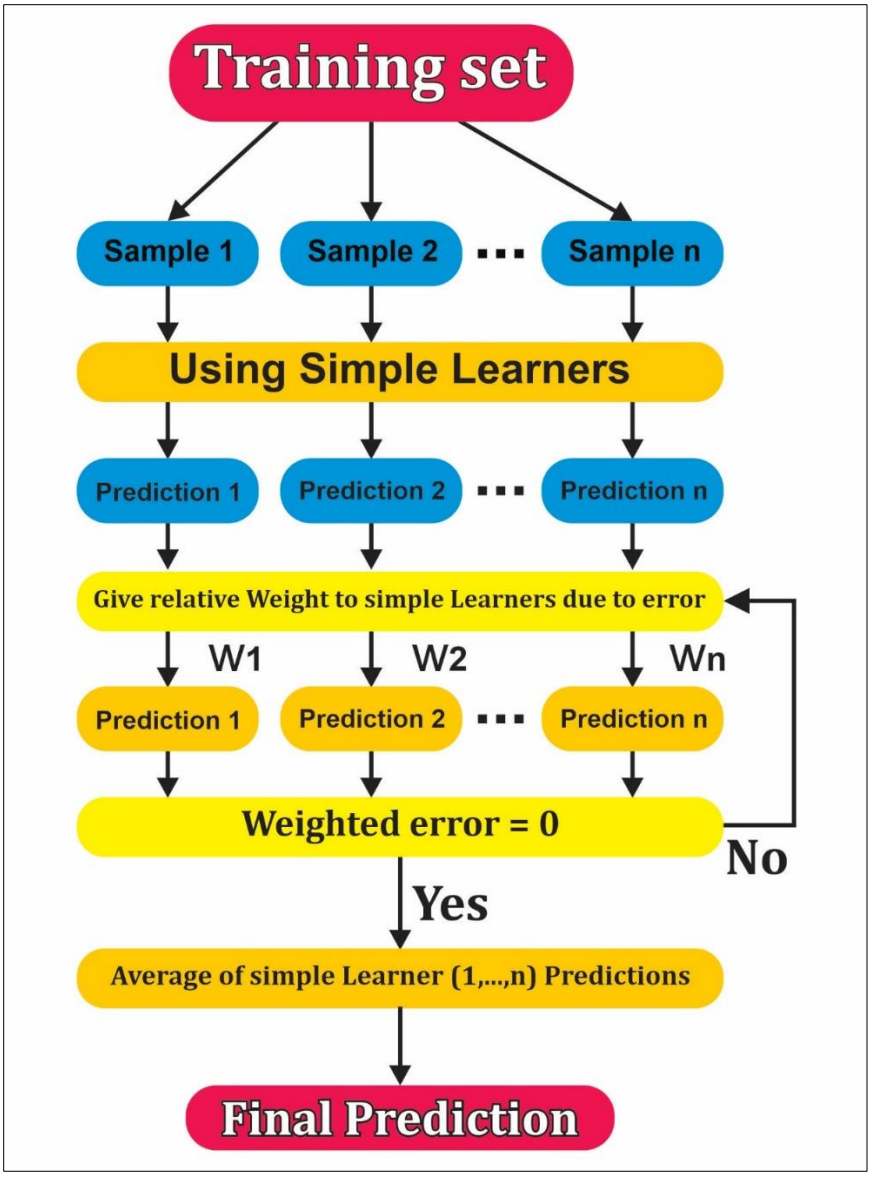


Figure 3: Conceptual representation of the Ada Boost Regression (ABR) algorithm.

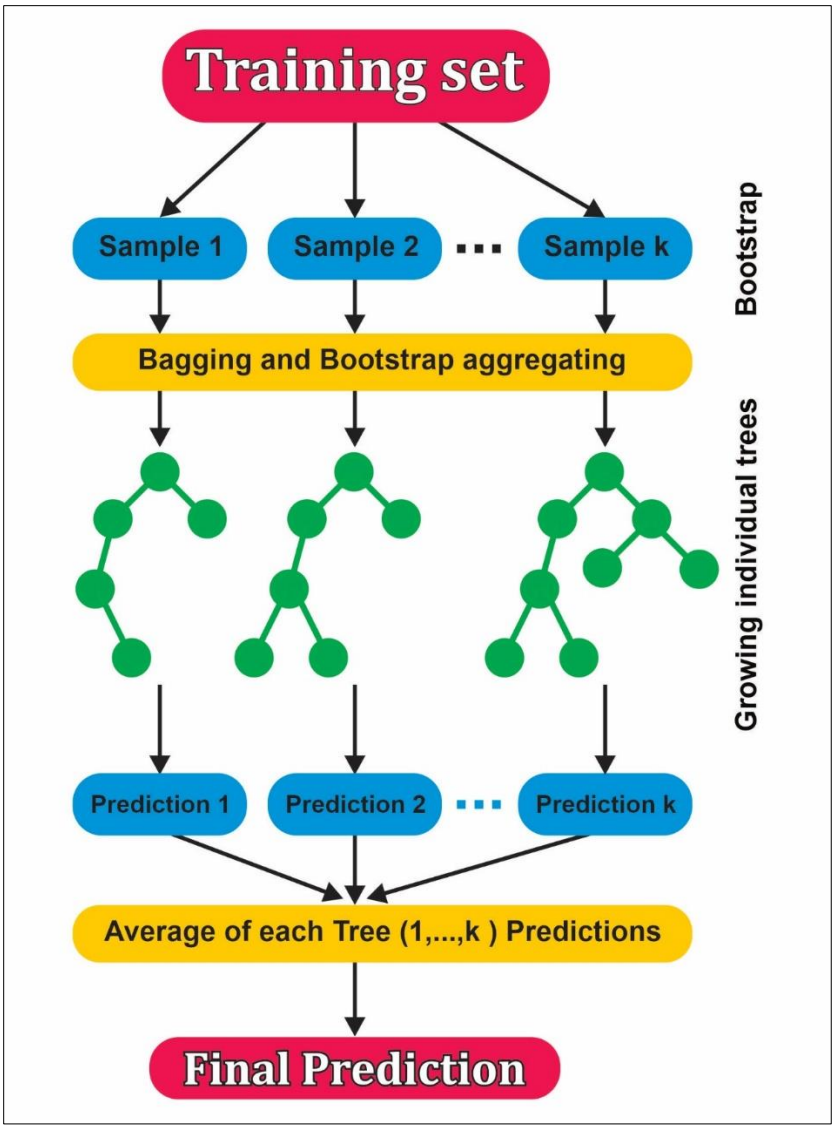


Figure 4: Conceptual representation of the Random Forest Regression (RFR) algorithm.

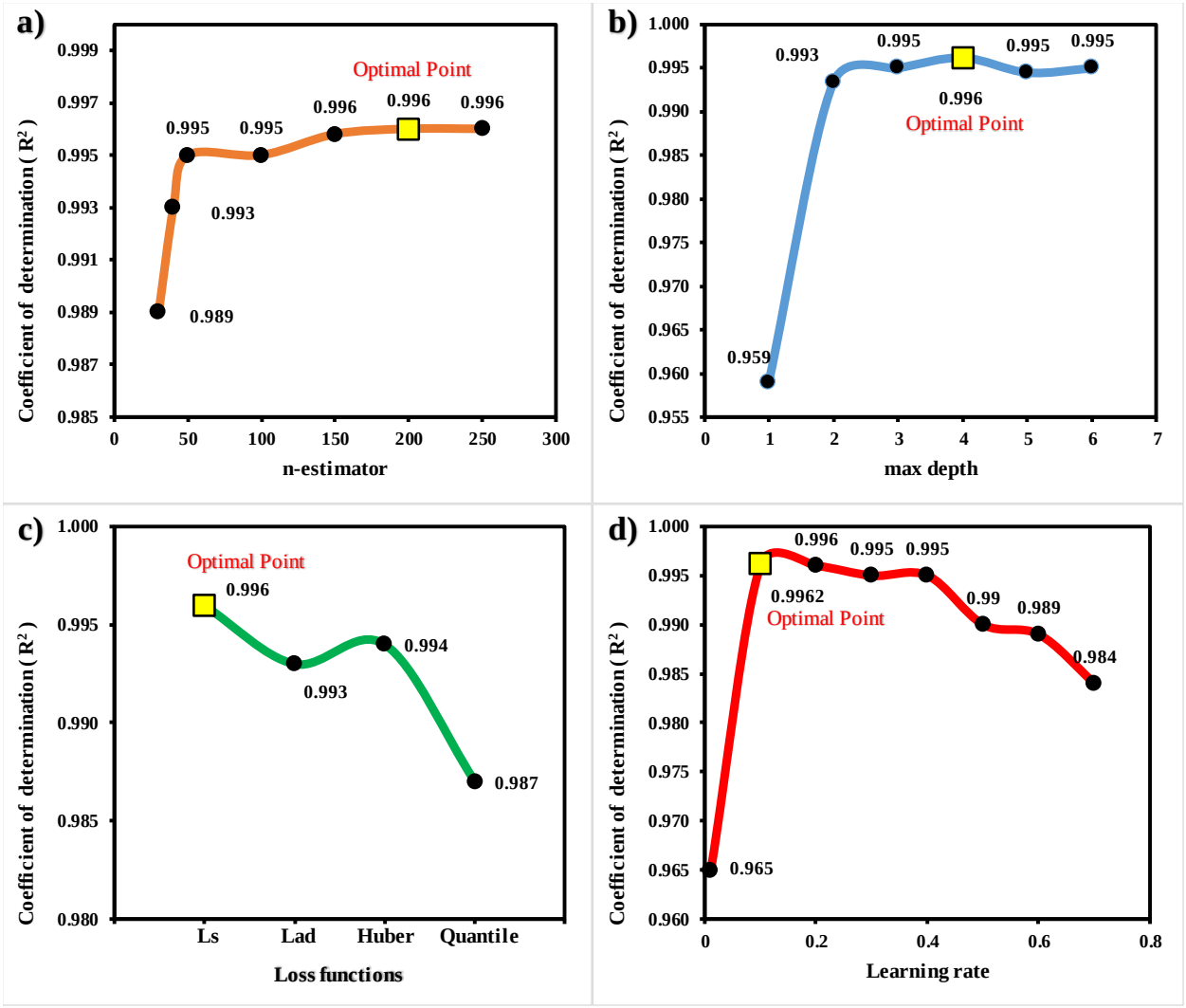
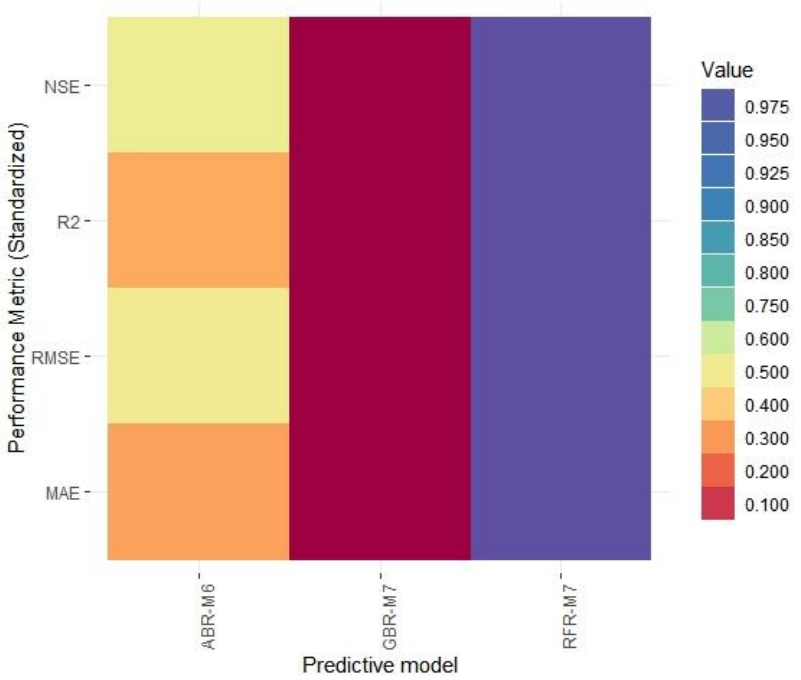
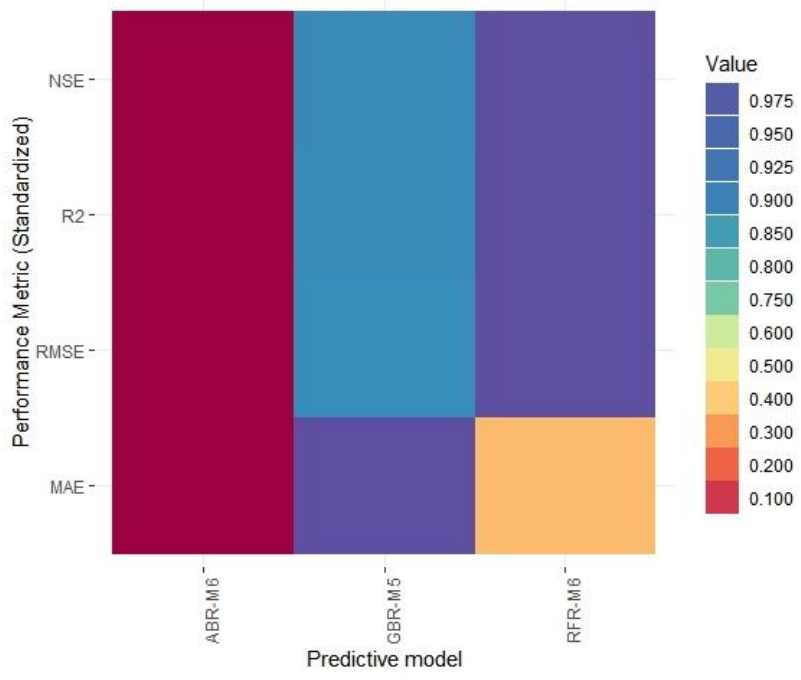


Figure 5: Example of sensitivity analysis and identification of the optimal parameter value for the GBR algorithm. a) n-estimator, b) max depth, c) loss function and d) learning rate.

a)


Performance Metric (Standardized)
NSE
R2
RMSE
MAE
ABR-M6
GBR-M7
RFR-M7
Predictive model
Value
0.975
0.950
0.925
0.900
0.850
0.800
0.750
0.600
0.500
0.400
0.300
0.200
0.100

b)


Performance Metric (Standardized)
NSE
R2
RMSE
MAE
ABR-M6
GBR-M5
RFR-M6
Predictive model
Value
0.975
0.950
0.925
0.900
0.850
0.800
0.750
0.600
0.500
0.400
0.300
0.200
0.100

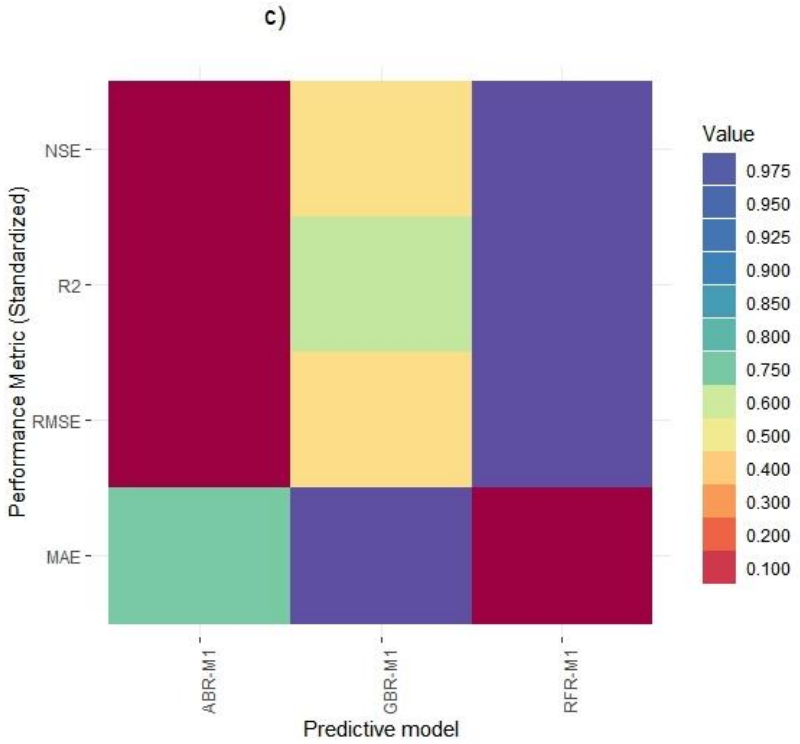
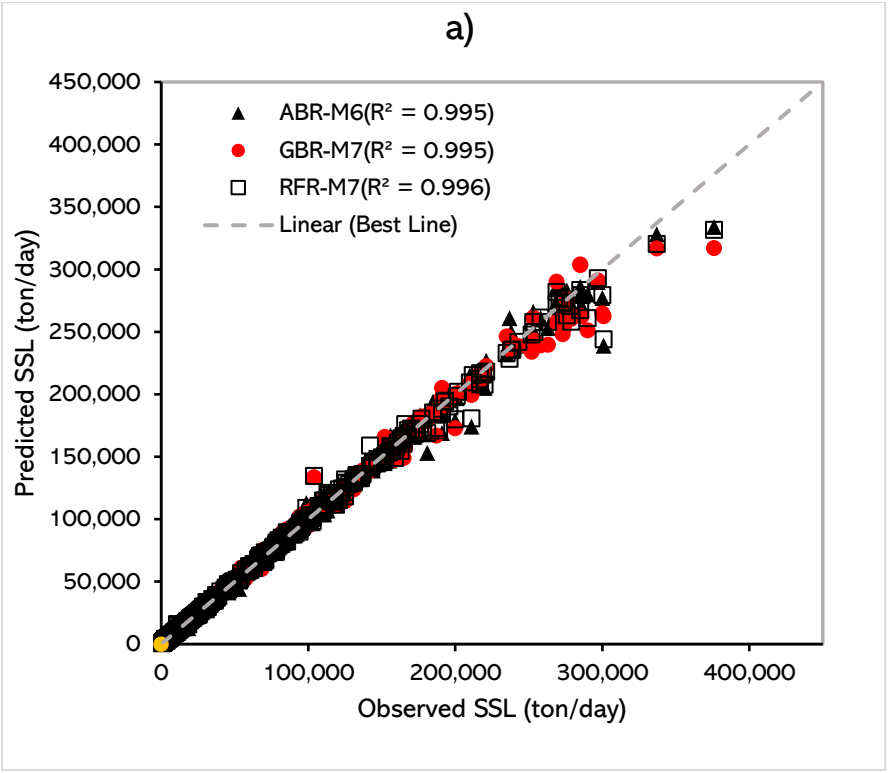
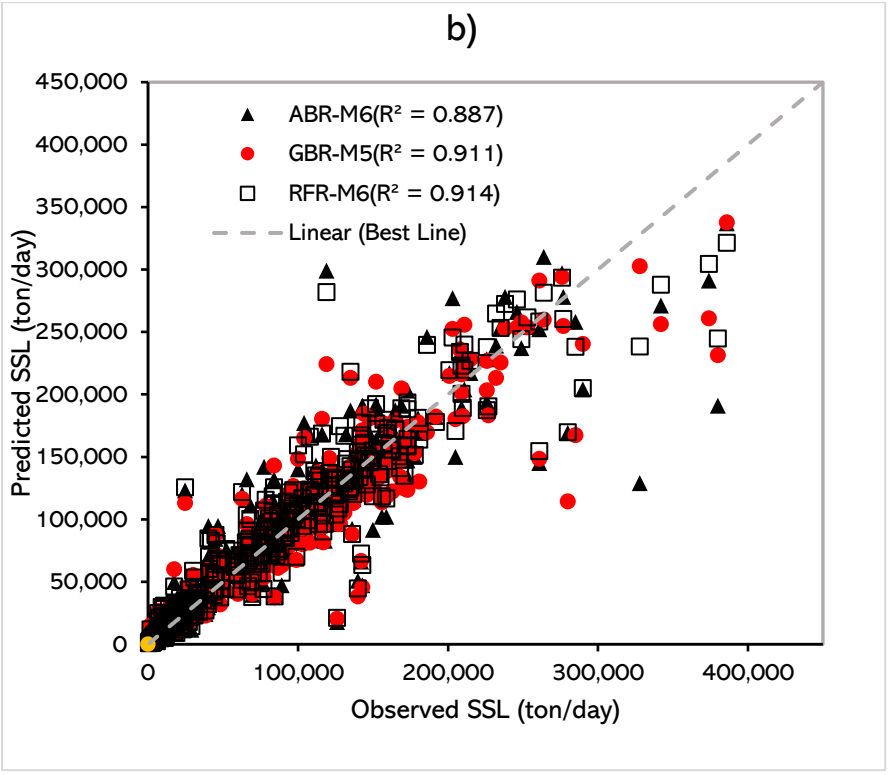


Figure 6: Heat maps of the best predictive models based on different standardized performance metrics for testing phase for SSL with different lead time. a) SSL(t), b) SSL(t+1) and c) SSL(t+3).


a)
ABR-M6(R² = 0.995)
GBR-M7(R² = 0.995)
RFR-M7(R² = 0.996)
Linear (Best Line)
Predicted SSL (ton/day)
Observed SSL (ton/day)
0
50,000
100,000
150,000
200,000
250,000
300,000
350,000
400,000
450,000


b)
ABR-M6(R² = 0.887)
GBR-M5(R² = 0.911)
RFR-M6(R² = 0.914)
Linear (Best Line)
Predicted SSL (ton/day)
Observed SSL (ton/day)
0
50,000
100,000
150,000
200,000
250,000
300,000
350,000
400,000
450,000

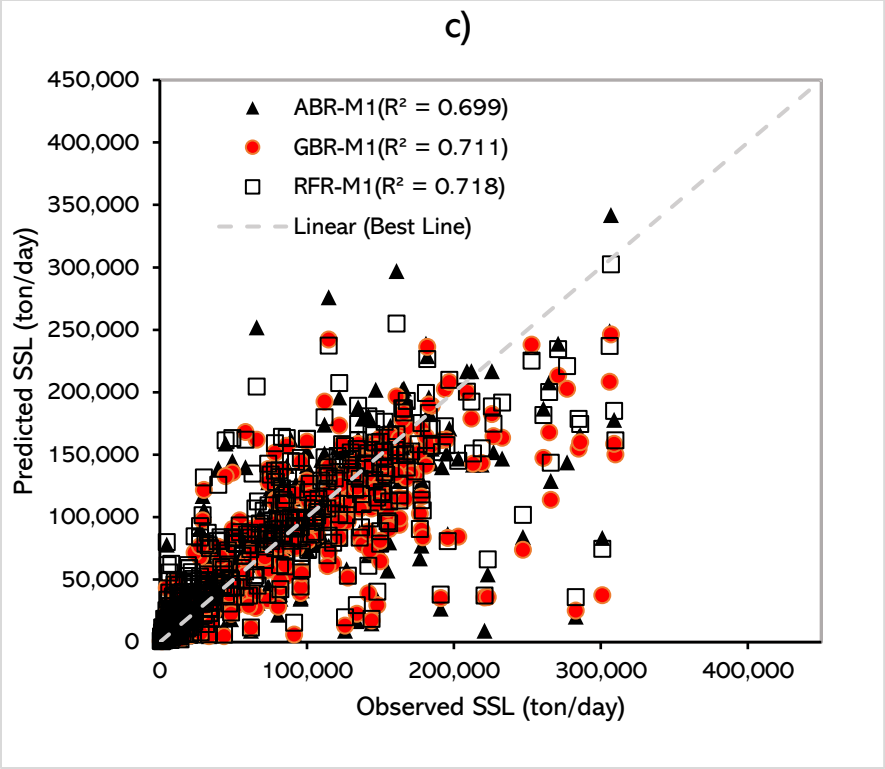
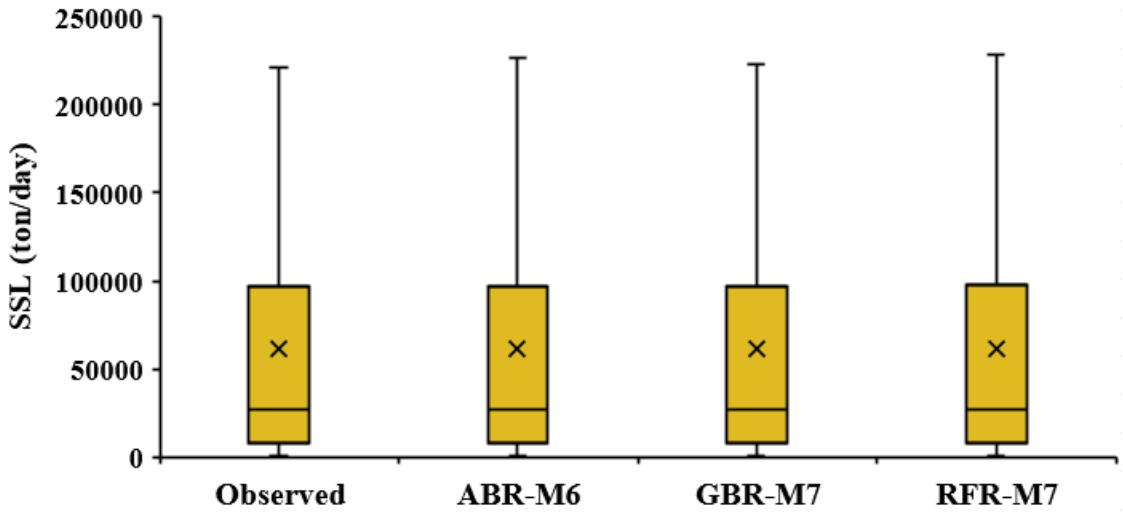


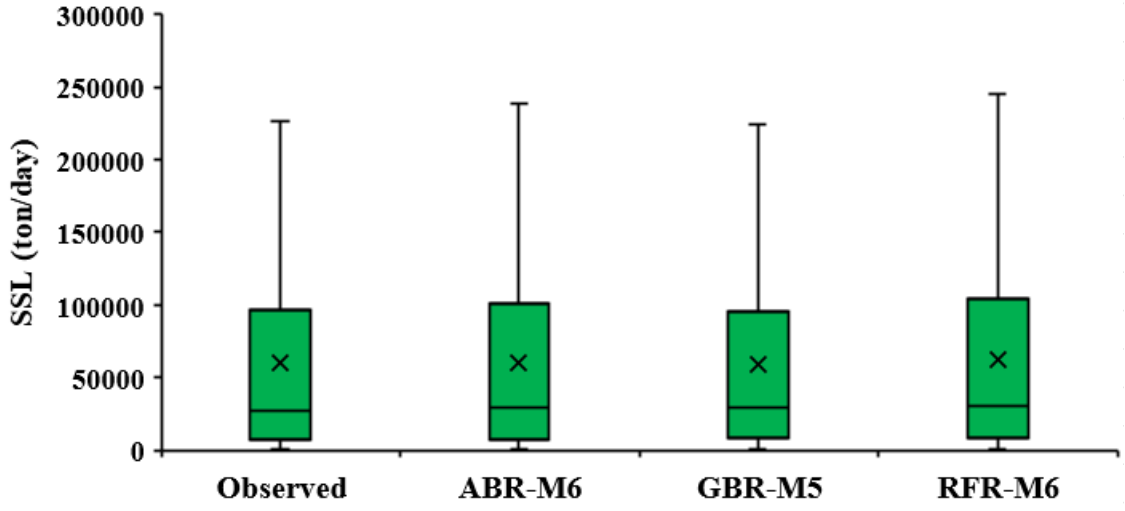
Figure 7: Scatter plots (predicted vs observed SSL) for the best predictive models for testing phase for SSL with different lead time. a) SSL(t), b) SSL(t+1), and (c) SSL(t+3).

a)

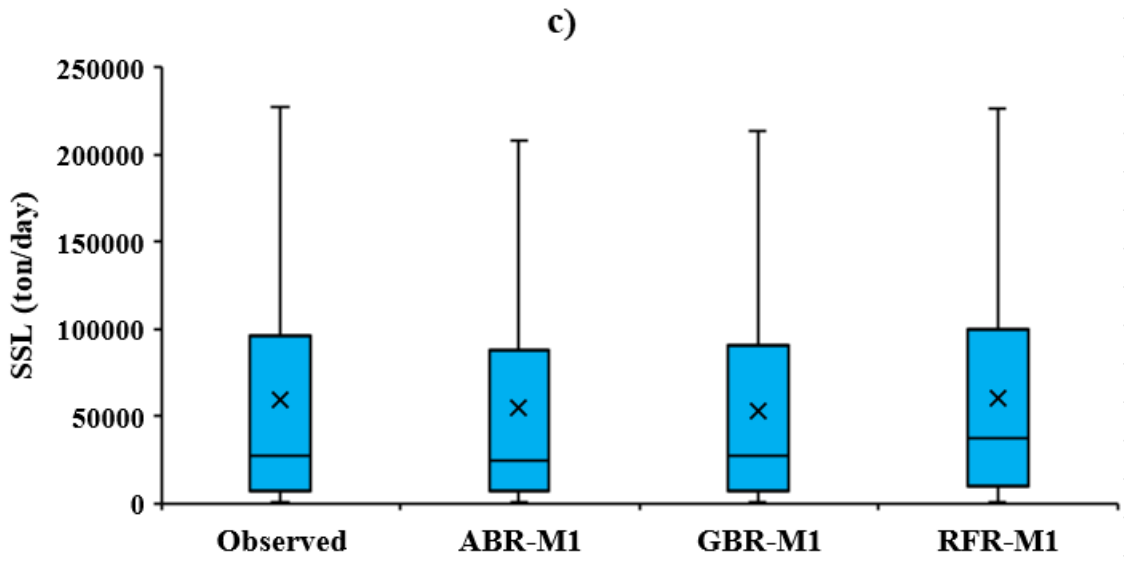


Q25%	7717.50	7482.50	7685.07	7564.28
Q50%	26900.00	27050.00	27154.00	26970.85
Q75%	96625.00	96900.00	97134.50	97328.38
IQR	88907.50	89417.50	89449.43	89764.10

b)



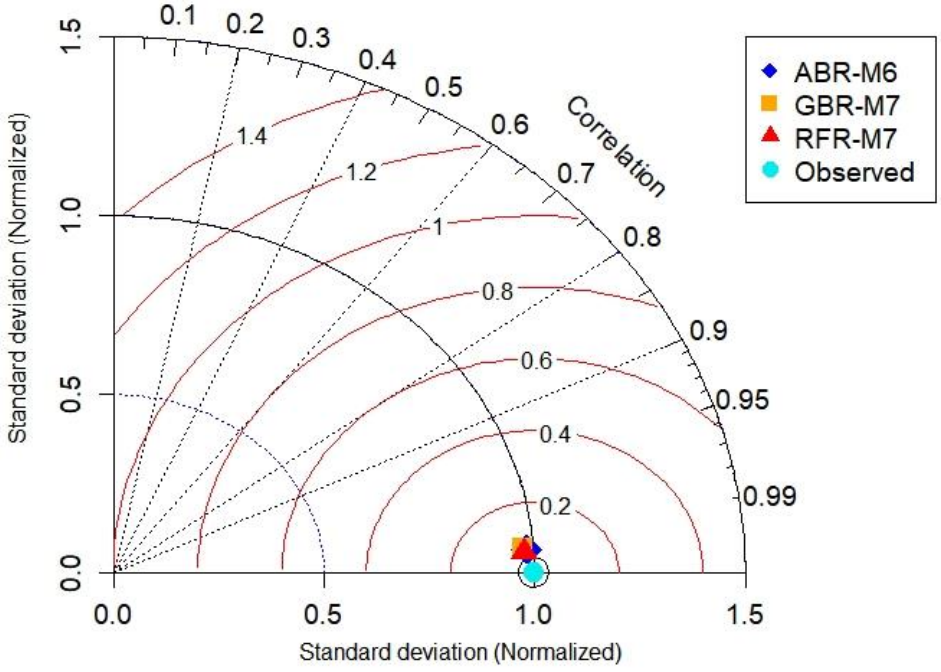
Q25%	7827.50	7595.00	8175.32	8278.14
Q50%	27500.00	29100.00	29663.70	30764.30
Q75%	96800.00	101000.00	94962.78	103707.50
IQR	88972.50	93405.00	86787.45	95429.36



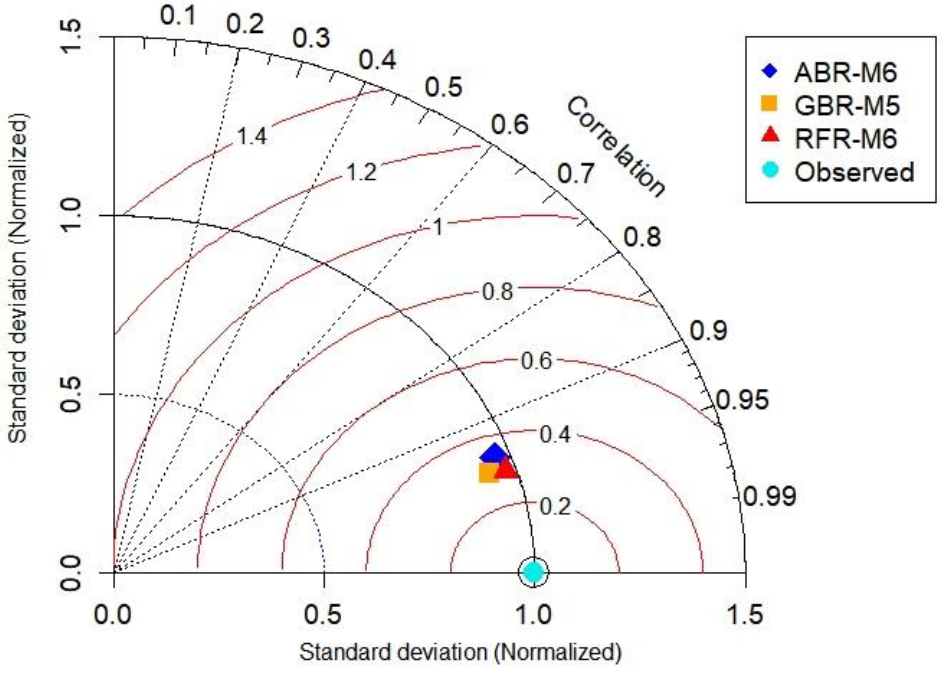
Q25%	**7137.50**	**7467.50**	**7443.70**	**9507.98**
Q50%	**27350.00**	**24600.00**	**27127.40**	**37120.60**
Q75%	**96150.00**	**88225.00**	**90798.35**	**99472.98**
IQR	**89012.50**	**80757.50**	**83354.65**	**89964.99**

Figure 8: Boxplot of observed and predicted SSL for the best predictive models for testing phase for SSL with different lead time. a) SSL(t), b) SSL(t+1) and c) SSL(t+3).

a)


Correlation
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0.95
0.99
1.4
1.2
1
0.8
0.6
0.4
0.2
0.0
0.5
1.0
1.5
Standard deviation (Normalized)
Standard deviation (Normalized)
ABR-M6
GBR-M7
RFR-M7
Observed

b)


Correlation
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0.95
0.99
1.4
1.2
1
0.8
0.6
0.4
0.2
0.0
0.5
1.0
1.5
Standard deviation (Normalized)
Standard deviation (Normalized)
ABR-M6
GBR-M5
RFR-M6
Observed

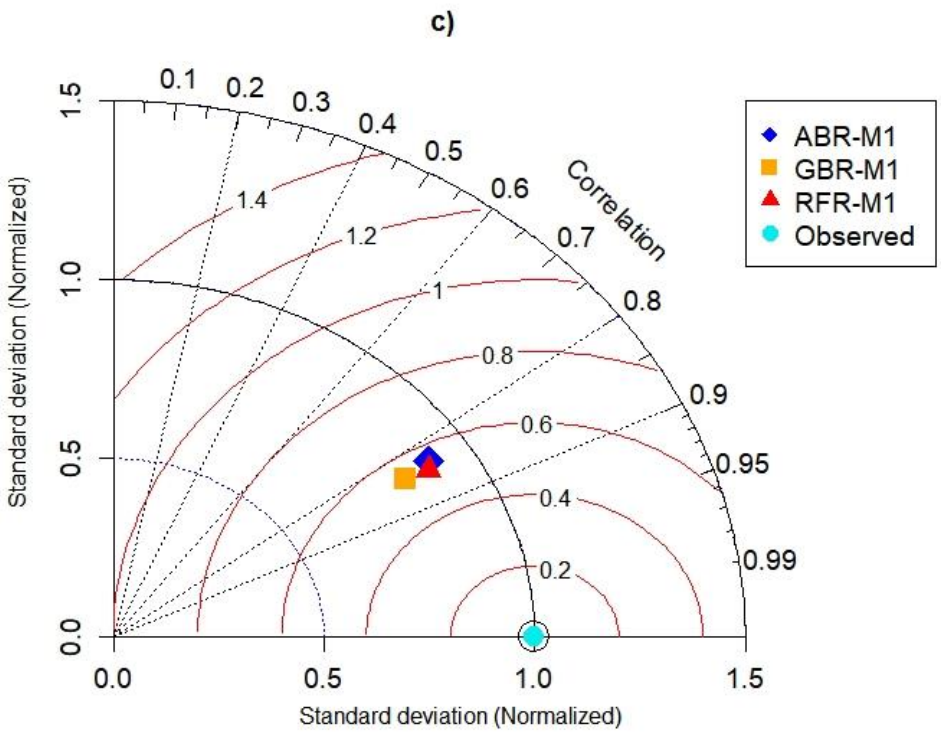


Figure 9: Normalized Taylor diagrams for the best predictive models for testing phase for SSL with different lead time. a) SSL(t), b) SSL(t+1) and c) SSL(t+3).

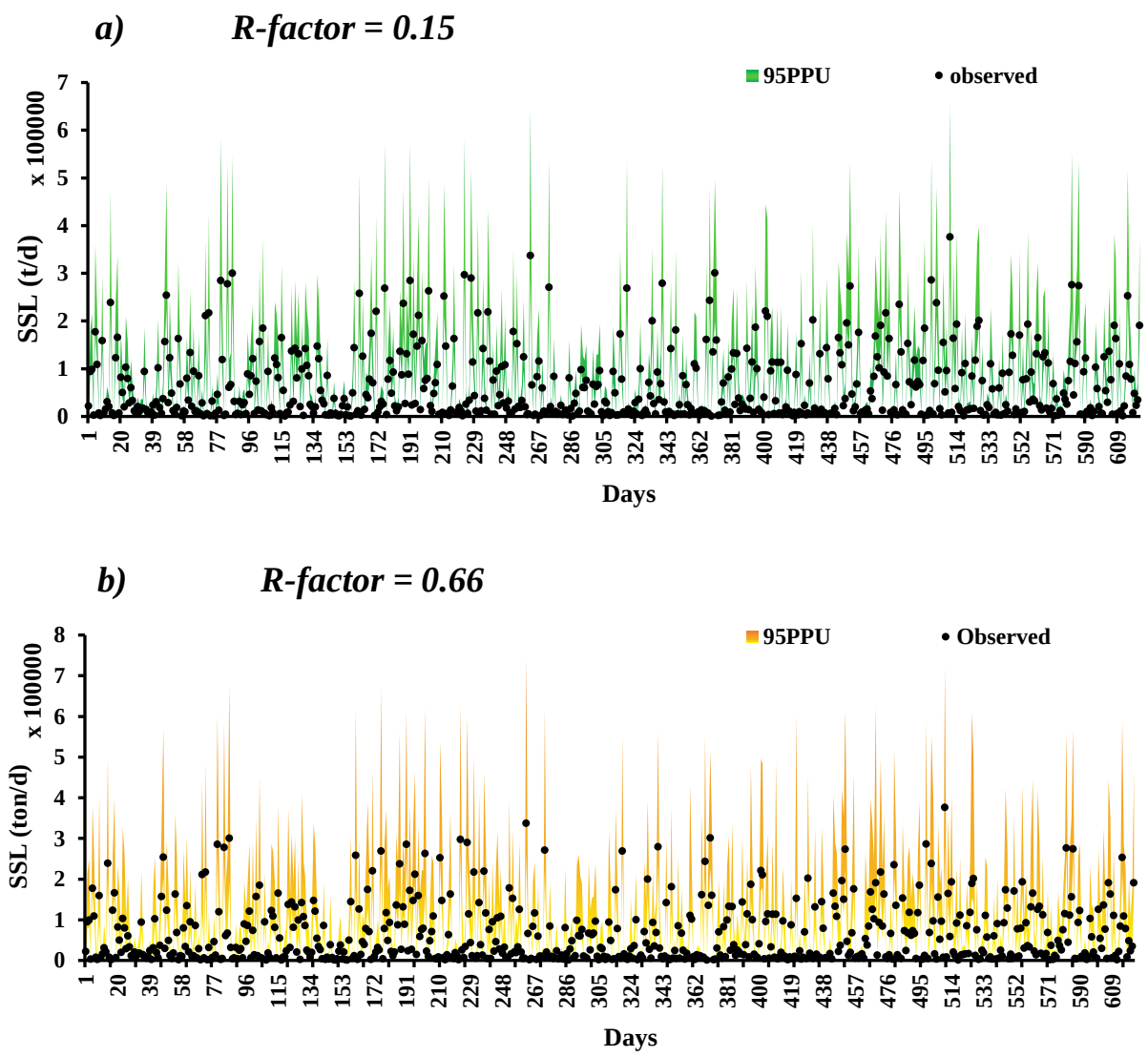


Figure 10: Generated 95PPU band for SSL(t) considering, a) model structure uncertainty and b) input variable uncertainty.

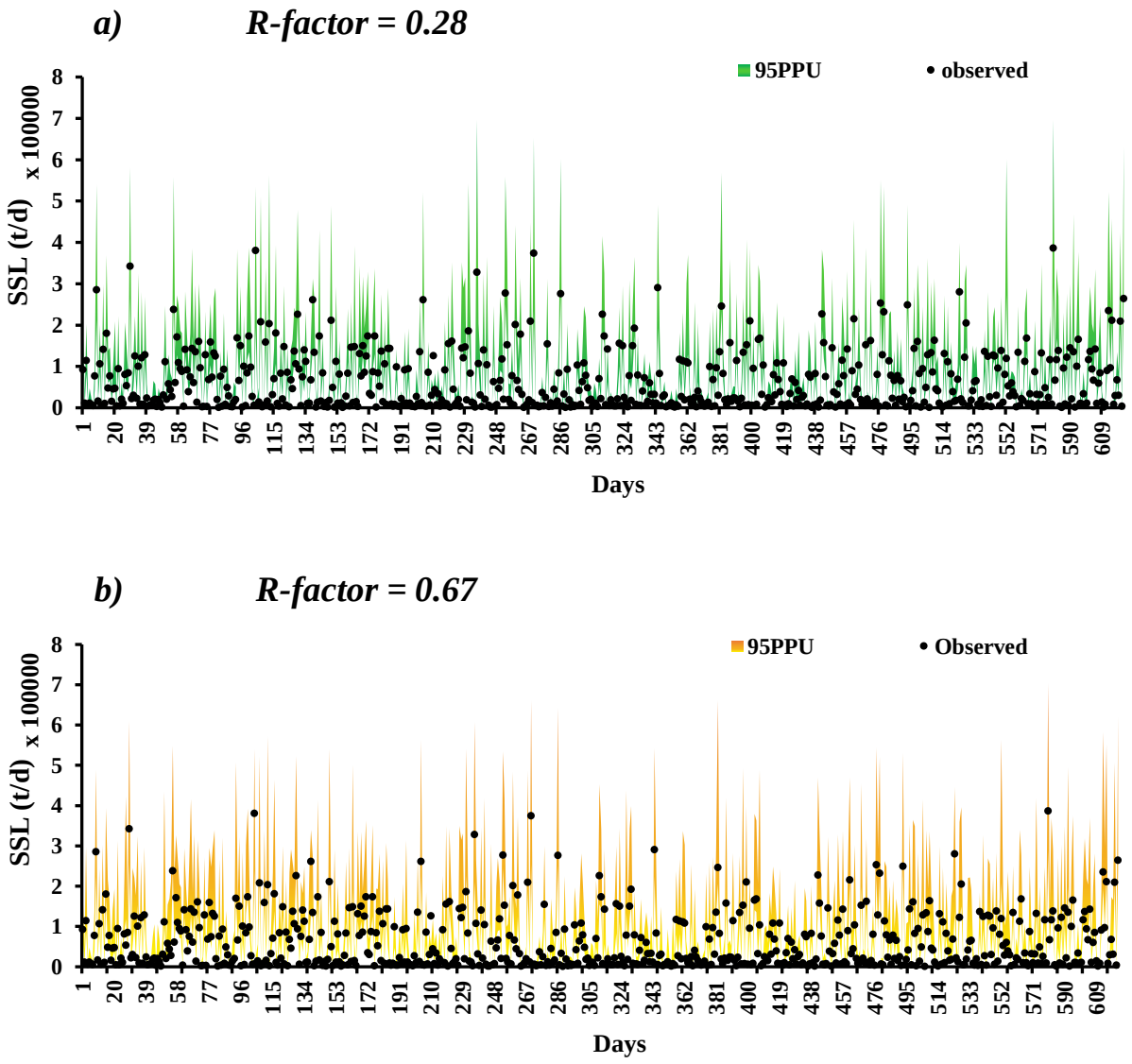


Figure 11: Generated 95PPU band for SSL(t+1) considering, a) model structure uncertainty and b) input variable uncertainty.

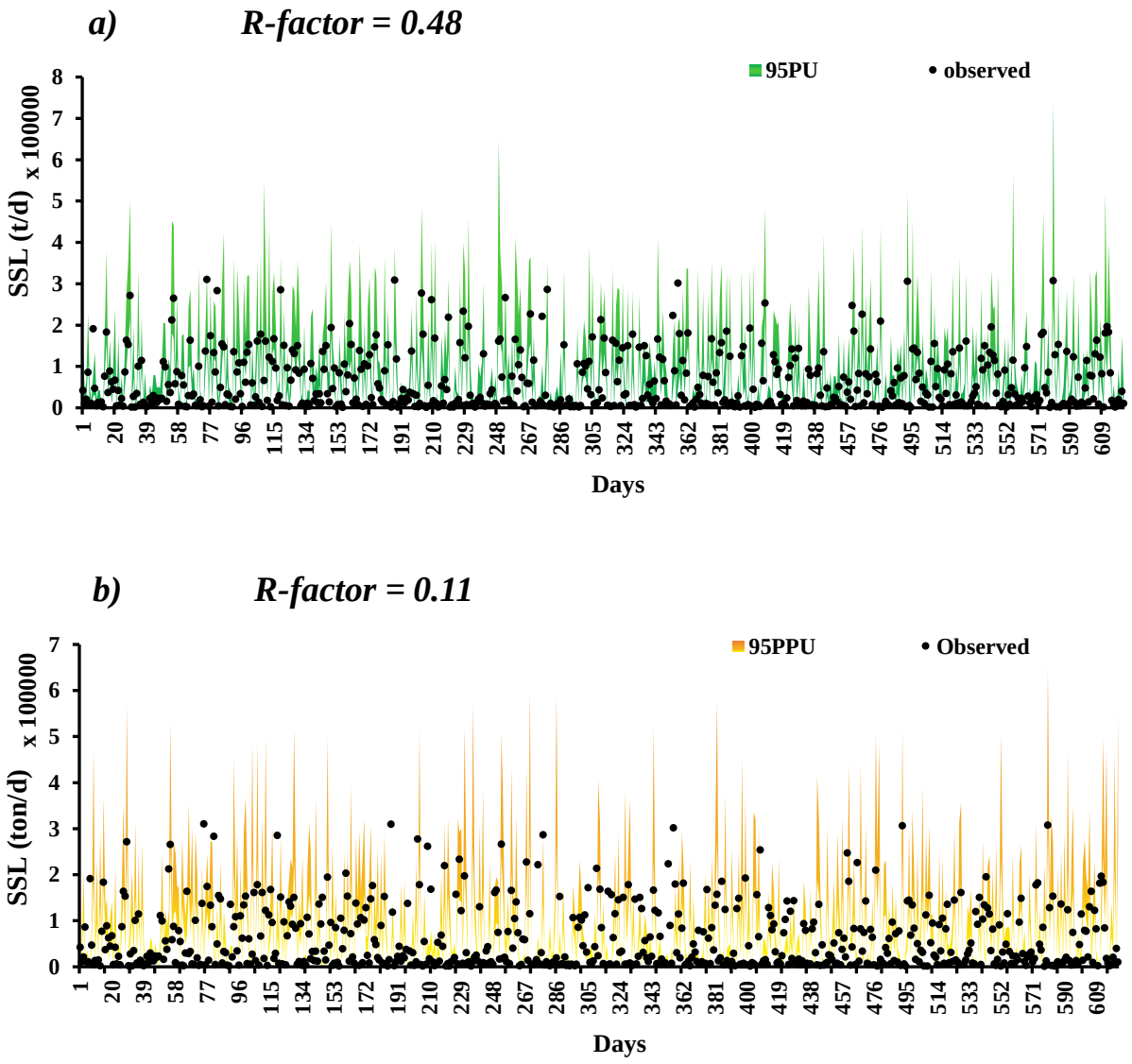


Figure 12: Generated 95PPU band for SSL(t+3) considering, a) model structure uncertainty and b) input variable uncertainty.