Forecasting Sediment Accumulation in the Southwest Pass with Machine-Learning Models

Magdalena Asborno, Ph.D.[1]; Jacob Broders[2]; Kenneth N. Mitchell, Ph.D.[3]; Michael A. Hartman, Ph.D., M.ASCE[4]; and Lauren D. Dunkin[5]

Abstract: Connecting the Mississippi River and the Gulf of Mexico, the Southwest Pass (SWP) is one of the most highly utilized commercial waterways in the United States. Hard-to-predict accumulation of sediments in the SWP affects the access of deep-draft vessels to four of the nation's top 15 ports measured by tonnage. The U.S. Army Corps of Engineers (USACE) spends approximately 100 Million USD annually on dredging operations to maintain SWP at a 14.2-meter (50-ft.) depth. Presently, USACE project managers rely on rules-of-thumb with seasonal river stage trends and thresholds to get 10–14 days of lead time for shoaling conditions at the SWP. This work presents the development of a machine learning modeling framework to increase lead times and accuracy of shoaling forecasts in the SWP. Within a multivariate multistep timeseries forecasting framework, several regression models, input variables, and forecasting days are explored. All multivariate machine learning models outperformed an univariate ARIMA model used as baseline. A multilayered perceptron regressor implemented on a 60-day in-lag scenario was found to be the best model to forecast shoaling in the upcoming 45 days. The proposed model may be applied to forecast dredging needs at other critical waterways. **DOI: 10.1061/JWPED5.WWENG-2009**.

Author keywords: Waterways; Sediment; Dredging; Machine learning; Timeseries; Multivariate forecasting.

Introduction

The Southwest Pass (SWP) is the primary outlet for the Mississippi River at the Gulf of Mexico, one of the most highly utilized commercial deep-draft waterways in the United States (Fig. 1). Disruptions in navigation due to hard-to-predict accumulation of sediments in the SWP affect the access of deep-draft vessels to four of the nation's top 15 ports measured by tonnage. The ports of South Louisiana, New Orleans, Plaquemines, and Baton Rouge connect the U.S. Midwest with global markets, and handle around 500 M short tons of cargo annually (DOT, Bureau of Transportation Statistics 2022). Between 2015 and 2019, navigation through the SWP was disrupted 93 days per year, in average, increasing shipping costs and causing delays (Hartman et al. 2022).

The U.S. Army Corps of Engineers (USACE 2022) maintains an estimated 40,200 km (25,000 mi) of channels in both coastal and riverine areas. The SWP has been maintained by the USACE New Orleans District (MVN) at a depth of 15.2 m (50 ft.) since the last quarter of 2020. Before, the depth was maintained at Congress' authorized depth of 14.8 m (48.5 ft.). Hydrographic surveys to inform SWP depth and dimensions are conducted on nearly a daily basis by MVN, to monitor channel conditions. Such surveys are conducted within the USACE eHydro program, which provides a standardized public archive of available hydrographic surveys from across the U.S. accessible from a single website (Kress 2021). Building upon, the Coastal and Hydraulics Laboratory (CHL) of the USACE Engineer Research and Development (ERDC) developed the Corps Shoaling Analysis Tool (CSAT), which calculates daily channel shoaling volumes using historical channel surveys stored in eHydro (Dunkin et al. 2018). CSAT daily shoaling volumes are used in this work as a proxy for dredging needs.

An average of 25 M cubic yards of sediment was dredged per year in the SWP between 2015 and 2019. The USACE spent an average of $88 M annually on dredging operations to maintain a reliable shipping channel throughout the SWP (Hartman et al. 2022), and the unpredictability of rapid-onset shoaling has been known to drive annual costs to more than twice that amount. Presently, MVN project managers rely on rules-of-thumb with seasonal river stage trends and thresholds to get 10–14 days of lead time for shoaling conditions at the SWP. Short lead-times of only 10–14 days prevent a timely mobilization of dredging equipment.

This paper presents the development of a machine learning regression model to increase both the lead times for and accuracy of shoaling forecasts and associated dredging requirements in a 56-km (35-mi) stretch of the Mississippi River (Mile 13.4 Above Head of Passes to Mile 22 Below Head of Passes), including the SWP. In the absence of metrics to evaluate the state of the practice, a univariate Autoregressive Integrated Moving Average (ARIMA)

[1]Research Civil Engineer, Applied Research Associates, Inc. for Coastal and Hydraulics Laboratory, U.S. Army Engineer Research and Development Center, 3909 Halls Ferry Rd., Vicksburg, MS 39180 (corresponding author). ORCID: https://orcid.org/0000-0003-2289-9080. Email: masborno@ara.com

[2]Applied Research Associates, Inc., 119 Monument Pl., Vicksburg, MS 39180.

[3]Research Civil Engineer, Coastal and Hydraulics Laboratory, U.S. Army Engineer Research and Development Center, 3909 Halls Ferry Rd., Vicksburg, MS 39180.

[4]Research Civil Engineer, Coastal and Hydraulics Laboratory, U.S. Army Engineer Research and Development Center, 3909 Halls Ferry Rd., Vicksburg, MS 39180.

[5]Research Civil Engineer, Coastal and Hydraulics Laboratory, U.S. Army Engineer Research and Development Center, 3909 Halls Ferry Rd., Vicksburg, MS 39180.

Note. This manuscript was submitted on March 13, 2023; approved on November 13, 2023; published online on December 29, 2023. Discussion period open until May 29, 2024; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Waterway, Port, Coastal, and Ocean Engineering***, © ASCE, ISSN 0733-950X.

Fig. 1. Southwest Pass (thicker line) geographical context, Louisiana, USA. (Map © 2023 Google, and 2022 World Light Gray Base by Esri, HERE, Garmin, © OpenStreetMap contributors, and the GIS User Community.)

model is used as baseline. The regression models are embedded in a multivariate, multistep timeseries forecasting framework. Results obtained with a Random Forest (RF), and two Artificial Neural Networks (ANNs) (a Multilayered Perceptron (MLP) and a Long Short-Term Memory (LSTM)) are compared. In particular, this work forecasts 45-day channel shoaling volumes (i.e., CSAT output). An increase trend in daily shoaling values and its increment above a threshold indicate a need to mobilize dredges to the SWP area.

The background section of this paper reviews literature on the use of machine learning models to forecast dredging and sedimentation, as well as potential variables that affect shoaling. The methodology section presents an overview of the multivariate, multistep timeseries forecasting framework implemented. The results and discussion focus on the comparison of three different machine learning regression models used within the framework, namely a multilayered perceptron neural network, a long short-term memory neural network, and a random forest.

Background

Machine Learning Models Applied to Sediment and Dredging Modeling

Researchers have implemented machine learning approaches with relation to dredging operations and sediment forecasting. Afan et al. (2016) made a comprehensive review of artificial intelligence (AI)-based models to predict sediment transport. The authors focused on papers that forecasted suspended sediment concentration (SSC) and suspended sediment load (SSL). Throughout the literature, they found that AI models are superior to classical modeling. In particular, ANNs, neurofuzzy, wavelet-based models, and some evolutionary algorithms are the most effective methods that have been applied in the sediment forecasting field. Input variables for ANN models incorporated water level, discharge, SSL, SSC, rainfall, runoff, reservoir capacity, and sediment volume. The most highly utilized performance criteria to evaluate ANN models were the Root Mean Square Error (RMSE) and the coefficient of determination (R^2) (Afan et al. 2016).

Bhattacharya and Solomatine (2006) presented two machine learning (ML) models that predict sedimentation in the harbor basin of the Port of Rotterdam. Factors affecting the sedimentation process used as inputs were wave energy (WE), river discharge, and sedimentation potential (SP). SP acts as a memory of the sedimentation process, and is a function of the amount of sediments brought to suspension by the waves in the recent past, minus the actual sedimentation that has occurred at the port. WE is a function of wave height. Missing wave height data were estimated from wind data with an ANN model. Sedimentation is modeled with two ML methods: MLP ANN and M5 model tree. The latter is a collection of linear regression models, each being an expert for a particular region of the input space. The models are trained on data collected during the years 1992–1998 and tested with data from years 1999–2000. Model accuracy is evaluated with RMSE, Normalized Mean Squared Error (NMSE), and absolute error in prediction, which remained within 10%–30%. MLP ANN had slightly better results than the M5 tree model (Bhattacharya and Solomatine 2006).

More recently, Mateo-Pérez et al. (2021) determined water depth in two Spanish ports (Luarca and Candás) using ML algorithms and satellite data. The algorithms used were Support Vector

Machine (SVM), Random Forest (RF), and Multiadaptive Regression Splines (MARS). Results were validated with data collected in situ using a single-beam sonar. SVM and RF outperformed MARS, but different algorithms performed better depending on the location. RF performed better in the port of Candás, while SVM performed better in Luarca. Both ports are located in Asturias, Spain, at similar latitude, distanced approximately 80 km (50 mi) (Mateo-Pérez et al. 2021).

For dredging applications, Chou and Lin (2020) built a deterministic ML model to predict the duration and probability of risk of dredging projects. Monte Carlo simulation was used to establish the probabilistic distribution of the project duration based on historical patterns. Data from 48 projects were used, normalized to account for differences in scope that had an impact on duration. Ten-fold cross-validation was applied to account for the relatively low number of projects. Inputs were the percentage of sand, soil, and gravel; average soil and gravel price, total dredging volume, road transport cost, rainfall days, project cost, and project duration (days). Several model types were explored, namely Support Vector Regression (SVR), Classification and Regression Tree (CART), ANN, and linear regression. SVR showed the highest performance. Models were evaluated and compared by Mean Absolute Percentage Error (MAPE).

Methodology

Variable to Predict

The objective of this work is to increase lead times and accuracy in forecasts of shoaling accumulation on the Southwest Pass to facilitate faster and more reliable mobilization of dredges to the area. In particular, the variable to predict is based on daily volumes of sediment accumulated in the 56-km (35-mi) stretch of the Mississippi River between Mile 13.4 Above Head of Passes and Mile 22 Below Head of Passes (Fig. 1). Past values of the sediment accumulated on the SWP are obtained from the Corps Shoaling Analysis Tool (CSAT), version 2.5 (Dunkin et al. 2018). Among others, CSAT estimates shoaling rates and daily volumes of sediment above a user-defined datum, for each point on a 10-by-10 ft. grid for each of the channels of the National Channel Framework (NCF). CSAT estimates are based on historical data from surveys conducted on U.S. waterways, made available through the USACE eHydro program (Kress 2021). Thirteen NCF channels constitute SWP (USACE 2022).

The historical timeseries of sediment accumulation on SWP was obtained from CSAT using the datum corresponding to the Congress-authorized depth, plus 1.8 m (6 ft.) of potential advanced dredging. Since the Congress-authorized depth changed from 14.8 m (48.5 ft.) to 15.2 m (50 ft.) in the fourth quarter of 2020, adjustments were made accordingly. Moreover, the daily timeseries of sediment accumulation on SWP is not stationary, thus CSAT estimates were transformed to 7-day rolling averages. The resulting timeseries of sediment accumulation in SWP, used both as variable to predict and input to the forecasting model, is depicted in Fig. 2.

Notably, CSAT produces shoaling forecasts at high spatial resolutions. However, such forecasts are based on a linear projection of sediment volumes (i.e., a single variable), using the latest-available shoaling rate for a given location (Dunkin et al. 2018). Building upon CSAT, this work explores the use of nonlinear forecasting methods (i.e., neural networks and random forest machine learning regression algorithms), within a multivariate framework to forecast sediment accumulation on SWP.

Input Data

The multivariate data collected to feed the framework to forecast sediment accumulation at SWP follows a physics-informed approach, in the sense that the data preselected to potentially feed the forecasting model are in line with variables that may affect shoaling. According to the literature, such variables may be sediment load, water velocity, sediment particle size, and water temperature (USGS 2018) (Nel et al. 2018). Measures of discharge (in cubic feet per second, cft/s), water level (in ft.), turbidity (formazin nephelometric units), and water temperature (in degrees Fahrenheit) are collected at USGS stations and used as potential input variables for the models presented in this paper (USGS 2022). Water level is currently used as the single rule-of-thumb indicator to predict dredging needs at the Southwest Pass within 10–14 days of lead time. In temporal terms, input data values represent daily readings of data collected during almost 10 years, between October 2012 and February 2022, prepared for quality control and consistency.

The magnitude of sediment production and its delivery to the coastal ocean is influenced by river source basin characteristics and processes that operate within basins (Wheatcroft and Sommerfield

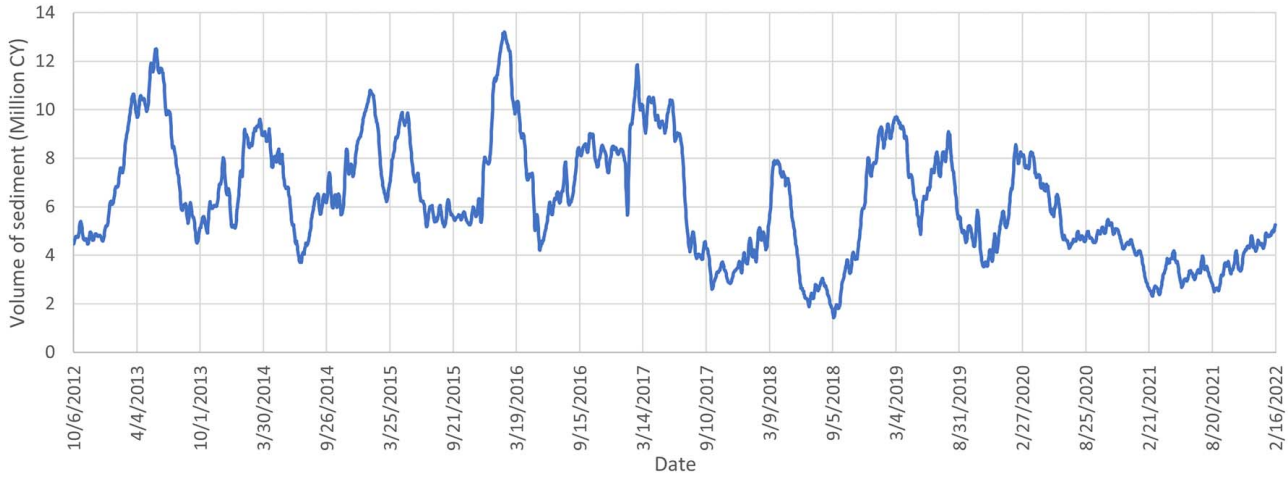

Fig. 2. Historical values used to forecast the variable to predict: volume of sediment accumulation in SWP over a datum corresponding to Congress-authorized depth plus 1.8 m (6 ft.) of potential advanced dredging (7-day rolling averages). Dates appear in the format MM/DD/YYYY.

2005). For that reason, data for this work are collected from a variety of stations located on the Mississippi River, expanding from the river mouth in the Southwest Pass, up river to its encounter with the Ohio River and the Upper Mississippi River (Fig. 3).

Seven variables feed the machine learning regression model. One of such variables is the past values of the variable to predict (i.e., daily sediment accumulation in SWP, in cubic yards). The remaining variables that feed the forecasting model are selected by applying a decision-tree based gradient boosting regressor (XGB). The top-6 features ranked by importance are fed into the sediment forecasting model.

For the selection of the above-mentioned top-6 ranked variables, 99 variables were considered. One of those variables was the week-of-year, which may take values between 1 and 52 (before data preparation), and indicate the sequenced-number of the week of the year when the data were collected. In addition, a pool of 98 variables from 57 stations located along the Mississippi and Ohio Rivers are automatically collected from USGS and RiverGages websites through APIs. The location of such stations is indicated in Fig. 3, as dots. Type of variables include: Stage (ft.), Discharge (cubic feet per second), Turbidity (FNU), Water Temperature (Fahrenheit), Precipitation (in.), Relative Humidity (%), and Air Temperature (Fahrenheit). A correlation analysis evidenced that correlation coefficients between the variable to predict and independent variables were in the range [−0.38 to 0.46]. Relatively low correlation coefficient values between influencing variables and sediment in machine learning modeling was also found by Bhattacharya and Solomatine (2006). Most importantly, the correlation analysis performed for this work evidenced that precipitation and relative humidity were not correlated with the variable to predict (i.e., coefficients between −0.03 and 0.07), probably because their connection with sediment accumulation in riverine settings is too indirect. Thus, 24 variables were removed from the pool of potential model inputs. Then, only variables with daily data available for more than 90% of the 10-year historical period of record were used to feed the feature selection model. This constitutes 39 variables from 35 stations, including Stage (32 variables), Discharge (four variables), Turbidity (two variables), and Water Temperature (one variable).

These 39 variables are subject to feature selection through a decision-tree based gradient boosting regressor (XGB). The top-6 variables ranked by importance are used to feed the sediment forecasting regression model. The calculation of XGB feature importance considers how each variable contributes to the changes in prediction performance (during training) when such variable is replaced with random noise (Zheng et al. 2017). Variables used the most to make key decisions during the training of boosted trees will have higher feature importance. This feature selection submodel provides flexibility to the automated rolling implementation of the model, so that as time goes by, and more data become available, the framework may automatically select different variables that better accommodate to data availability. For example, if one of the

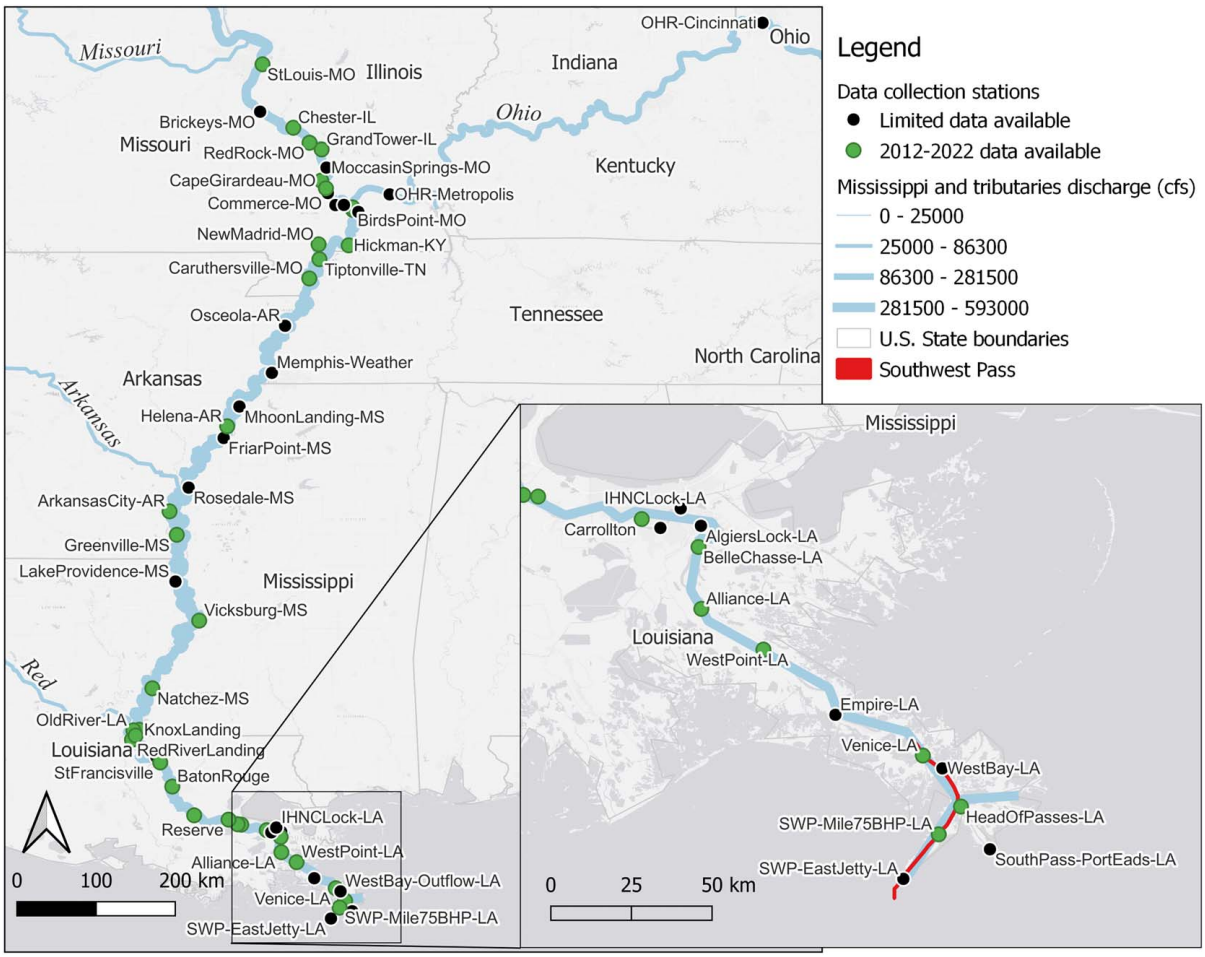

Fig. 3. Mississippi River and tributaries discharge, and location of data collection stations considered for model input. Stations with limited data available did not have sufficient data during the 10-year historical period to constitute an input to the model. (Map from 2022 World Light Gray Base by Esri, HERE, Garmin, © OpenStreetMap contributors, and the GIS User Community.)

stations that collects stage data becomes unavailable for a period of time, the framework will automatically select data from a different highly ranked station instead. As a result, each time series window for which the model is run will have different features selected as the most important. While the feature selection submodel results are not the focus of this paper, it was found that the variables selected the most by an XGB with 10 trees were the week of year, turbidity at Belle Chasse, and stage at Venice, Chester, Grand Tower, and Red Rock.

For each potential model input variable, data were prepared by replacing erroneous or missing values by linear interpolation, followed by a minimum-maximum scaling of values in the range [0–1]. The scaling was applied individually to each feature.

Forecasting Framework Overview

Data Parsing and Lags

A multivariate, multistep timeseries forecasting framework is applied to estimate future values of daily sediment accumulation on SWP. The target forecast (i.e., multistep) is 45 days. The last 45 days of the 10-year historical data are separated for model evaluation. Machine learning models have the ability to learn repetitive patterns from relatively big data. To allow for a single historical timeseries to be used for supervised machine learning, the 10 years of historical data are broken down into smaller timeseries, hereafter referred to as "instances." All instances are the same size. The size of the instances considers a number of days used by the model as input, or "in-lag," plus the 45-day prediction. For this work, models were applied and compared with six scenarios of in-lags, namely: 45, 60, 90, 120, 150, and 180 days. Thus, each instance has a size of 85, 105, 135, 165, 195, and 225 days, respectively. As many as possible instances are created, each starting one day apart from each other. In this way, the machine learning regressor benefits for an increased number of elements (i.e., instances) to learn from, without the need to generate synthetic data.

Model Evaluation Framework

For each in-lag scenario and regressor (RF, MLP, LSTM), the iterative framework can be briefly explained as follows (note that features had been previously obtained and prepared as described in the "Input Data" section. One of the data preparation steps consisted of scaling each input variable, including the variable to predict, to a [0–1] range using each variable's minimum and maximum values).

1. First, the first three years plus M months of data (from the total of 10 years) are selected. In the first iteration, $M = 2$. (This step is necessary to allow for an initial amount of sufficient data to be used for training purposes.)
2. Then, the last 45 days to be forecasted are separated from the selection made in Step 1, for out-of-sample model evaluation (i.e., not used until Step 6).
3. The remaining historical data (i.e., three years $+M$ months minus 45 days) are broken down into as many timeseries instances as the data allow for. The number of instances will depend on the size of the instances, corresponding to each in-lag scenario (45, 60, 90, 120, 150, and 180 days). The start of the timeseries instances is one day apart from each other.
4. The XGB estimates the relative importance of dozens of input variables. It selects the six most important variables, and together with the variable to predict, feeds the machine learning regressor.
5. From the timeseries instances created in Step 3, 80% are used to train the regressor, and the remining 20% are used to test the training process.
6. Once the regressor is trained, it is validated by forecasting the sediment volume in the last 45 days that were separated in Step 2.
7. The forecasted variable is scaled-up to reestablish the original values of the variable to predict. The original values allow practitioners to have a better idea of the volume of sediment expected, and are used for plotting.
8. Lastly, the framework calculates the normalized Root Mean Squared Error ($n\text{RMSE}_t$) for the last 45 days. The $n\text{RMSE}_t$ is calculated by comparing the observed shoaling volumes and the forecasted results obtained in Step 6. The $n\text{RMSE}_t$ calculated in each iteration is stored and used later for model evaluation and comparison.
9. The model resets itself (forgets all data previously seen).

The process above is iterative, with each iteration differing in the data selection in Step 1. For example, in Step 1 of the next iteration, the model would select three years plus $M = 4$ months of data. M increases in multiples of two until the complete 10-year historical timeseries is complete. (Multiples of +2 months provide for an adequate number of timeseries to evaluate the framework with, but significantly reduces the time to run with respect to a multiple of one). The last iteration considers all the 10 years of historical records in its Step 1. In this context, the framework evaluates the model over 36 iterations. After all iterations run, the average nRMSE is calculated (from the nRMSE stored in Step 8 of each iteration), and used to evaluate the model.

The iterative process described above was created with the purpose of selecting a model architecture (from MLP, RF, or LSTM type) and in-lag scenario (45, 60, 90, 120, 150, and 180 days) that better adapts to different shoaling conditions, observed during the 10 years of historical timeseries. However, for model implementation purposes (i.e., to forecast 45 days in the future), the iterative process is not necessary; only Steps 3 to 7 need to run once, considering the complete historical timeseries. Notably, in each iteration during model development phase (and every time the model is run during implementation phase), the regression model architecture is the same, but the selected input variables (output of Step 4) and the trained model parameters (output of Step 5) may differ, as identified automatically by the model, to better fit the data of the specific historical time-period being used.

Model Evaluation Metrics

Regression models are typically evaluated with the Mean Squared Error (MSE), RMSE, and the Mean Absolute Error (MAE). All these metrics compute the difference between observed and forecasted values. Unlike the MSE, the RMSE and MAE have the advantage that they are expressed in the same metric that the variable predicted (cubic yards of sediment, in our case). In addition, by using squares, the RMSE allows for extreme errors to have more influence on the result, which in our case is perceived as an advantage. Notably, since data have previously been cleared from outliers and anomalies, it is not unreasonable to allow for penalties to extreme or gross errors (Shcherbakov et al. 2013). The main disadvantage of RMSE is that it is scale-dependent, and as such, should not be used to compare models applied to data expressed in different scales (which is the case in this paper, when different timeseries instances are compared). To overcome scale-dependency issues, researchers have used normalized versions of RMSE, where the normalization factor takes different forms (such as average, maximum, range, standard deviation, etc. of observed values in the forecasting horizon) (Shcherbakov et al. 2013; Gupta and Kling 2011; Wei et al. 2018). For this work, the RMSE is normalized by the total sediment observed

during the 45 predicted days. Lower values of the evaluation metric indicate the model produces forecasts with smaller errors.

The normalized RMSE is calculated in each of the iterations of the model evaluation framework ($n\text{RMSE}_t$) [Eq. (1)]. For each in-lag scenario, the $n\text{RMSE}_t$ of all iterations are averaged [Eq. (2)], and further calculated over a subset N_{inc} timeseries with increasing shoaling trends [Eq. (3)]. In addition, it is desirable to select a model with low variability of forecasted results. To compare models in such terms, the standard deviation of nRMSE and confidence interval at 95% are calculated for each in-lag scenario [Eq. (4)] (Washington et al. 2020).

$$n\text{RMSE}_t = \frac{\sqrt{\frac{1}{S}\sum_{i=1}^{S}(\hat{y}_i - y_i)^2}}{\sum_{i=1}^{S} y_i} \tag{1}$$

$$n\text{RMSE} = \frac{\sum_{1}^{N} n\text{RMSE}_t}{N} \tag{2}$$

$$n\text{RMSE}_{\text{inc}} = \frac{\sum_{1}^{N_{\text{inc}}} n\text{RMSE}_t}{N_{\text{inc}}} \tag{3}$$

$$\text{confidence interval} = n\text{RMSE} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{N}} \tag{4}$$

where S = sample size (i.e., number of days forecasted = 45); $\hat{y}_i$ = predicted value for the ith element in the dataset; y_i = observed value for the ith element in the dataset; N = number of model evaluation instances; N_{inc} = number of model evaluation instances with increasing shoaling trend; σ = standard deviation; and $Z_{\alpha/2} = 1.96$ for 95% confidence level.

Implementation

The framework was written in python, using open-source software. Main libraries used were TensorFlow version 2.8.0 and kerastuner 1.1.2 for LSTM (Abadi et al. 2015); Scikit-Learn 1.0.2 for RF and MLP (Pedregosa, et al. 2011); statsmodels 0.12.2 for ARIMA (Seabold and Perktold 2010); PyPI for feature selection (PyPI 2021); and Matplotlib 3.5.0 for plotting (Hunter 2007). Other libraries used were pandas and numpy.

Regression Models

The objective of the broad exploratory work was to answer the question: Can machine learning models predict sediment accumulation within the next 45 days in SWP, as a proxy for dredging needs? The exploratory work considered several ML model families, namely K-nearest neighbors, support vector machines, decision trees, random forests, and artificial natural networks including multilayered perceptrons (MLP), long short-term memory neural networks, and convolutional neural networks, as well as ensemble models combining the methods mentioned above. The best results were obtained with a tree-based ensemble (RF) and ANN model families (MLP and LSTM), in line with previous research on ML models applied to sedimentation forecasting and dredging operations (please refer to the "Background" section). This paper focuses on the results obtained by the three most promising ML methods explored.

For each regression model type (RF, MLP, and LSTM), several architectures and parameter setups were tested. The selection of models presented in this paper resulted from applying the framework already explained to several different combinations of number of trees (for RF), layers (for MLP and LSTM), and parameter setups. The combination that produced the lowest average nRMSE for each type of regressor is discussed further in the remaining portion of this paper.

Autoregressive Integrated Moving Average Model (ARIMA)

To complement the comparison of models, in the absence of data to measure results from state-of-the-practice, a simple univariate ARIMA model was developed and used as baseline. ARIMA models have been used to forecast variables within natural processes, such as suspended sediment, water level, precipitation, and so forth (Pektas and Cigizoglu 2017; De Figueiredo and Cavalcante Blanco 2016; Zhang et al. 2020; Murat et al. 2018). ARIMA models forecast historical timeseries data by applying linear regression. The model is denoted with three parameters (p, d, q), each corresponding to a model component, namely:

1. p is the number of autoregressive terms (the model uses the dependent relation between past and current data values);
2. d is the degree of differencing, indicating the number of times the lagged indicators are subtracted to make the data stationary; and
3. q is the number of forecast errors, or the size of the moving average window.

The parameters were selected following the methodology of Washington et al. (2020). As a result, an ARIMA baseline model of order (1, 1, 5) was built and ran to forecast 45 days in each of the 36 iterations described in the "Model Evaluation Framework" section. The nRMSE of the ARIMA model was 0.34. However, regardless of the nRMSE value, and unlike results obtained with the machine learning regressors, all daily shoaling forecasts obtained with ARIMA tend to flatten out after the first handful of predicted days, indicating the univariate ARIMA model does not have a strong predictive power to forecast 45-day sediment accumulation in SWP (Fig. 4).

While the methods for choosing p, q, and d values give good starting points for the ARIMA baseline model, these values may be further fine-tuned to see if an improvement is made on the predictive strength of the ARIMA. The ARIMA gives a good baseline to compare the ML models and shows that the ML models do outperform the simpler, univariate linear regression model.

Random Forest Regressor

A Random Forest (RF) is a supervised ML algorithm consisting of an ensemble of decision-tree predictors, so that each tree depends on the values of a random vector sampled independently, and with the same distribution for all trees in the forest. To solve a regression task, the leaf nodes of each tree in the forest predict a numerical value that is the average of the training instances associated with that node. Then, these predictions are averaged to obtain the RF predictor result (Fig. 5). As the number of trees in the ensemble becomes large, the generalization error of the forest converges to a limit. The generalization error of an RF depends on the strength and correlation between its individual trees. Accurate RF regressors require low correlation between residuals and low-error trees (Breiman 2001). The predictive ability of the RF regressor is limited to the values of its training dataset (i.e., the model cannot predict values outside of the range of the direct observations within the training dataset) (Graw et al. 2021). RFs have been used for timeseries forecasting in varied subject topics, outperforming ARIMA (Tyralis 2017; Dudek 2015; Rady et al. 2021; Biswas et al. 2021).

The structure of the RF regressor model used in this work was obtained by testing several combinations of parameters and by the

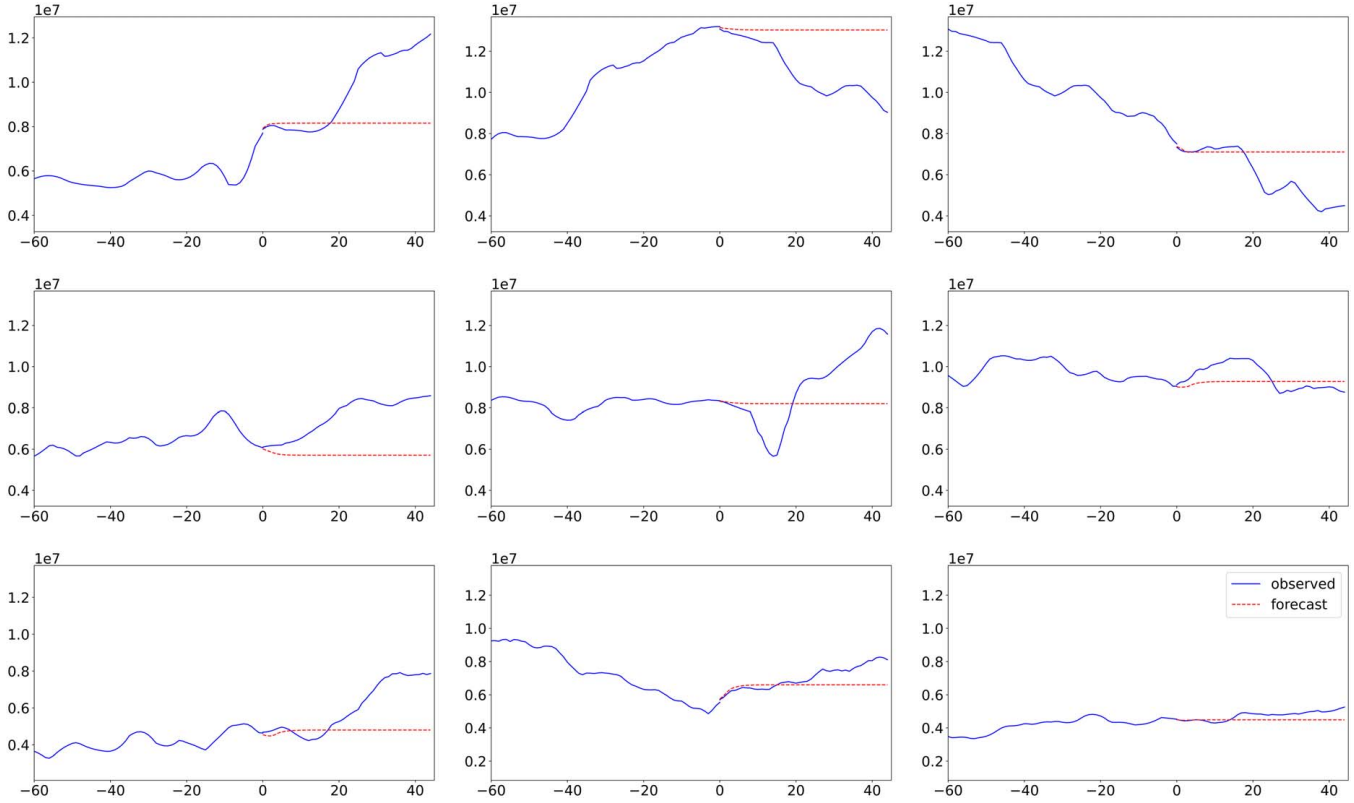

Fig. 4. ARIMA (1,1,5) model results for 45-day output for several evaluation instances. The continuous line represents observed values of the variable to predict. The dashed line is the ARIMA forecast of the variable to predict. Horizontal axis indicates number of days before (negative) and after (positive) forecast starts, set as 0. Vertical axis shows the volume of sediment accumulation in SWP (in million cubic yards) over a datum corresponding to Congress-authorized depth plus 1.8 m (6 ft) of potential advanced dredging.

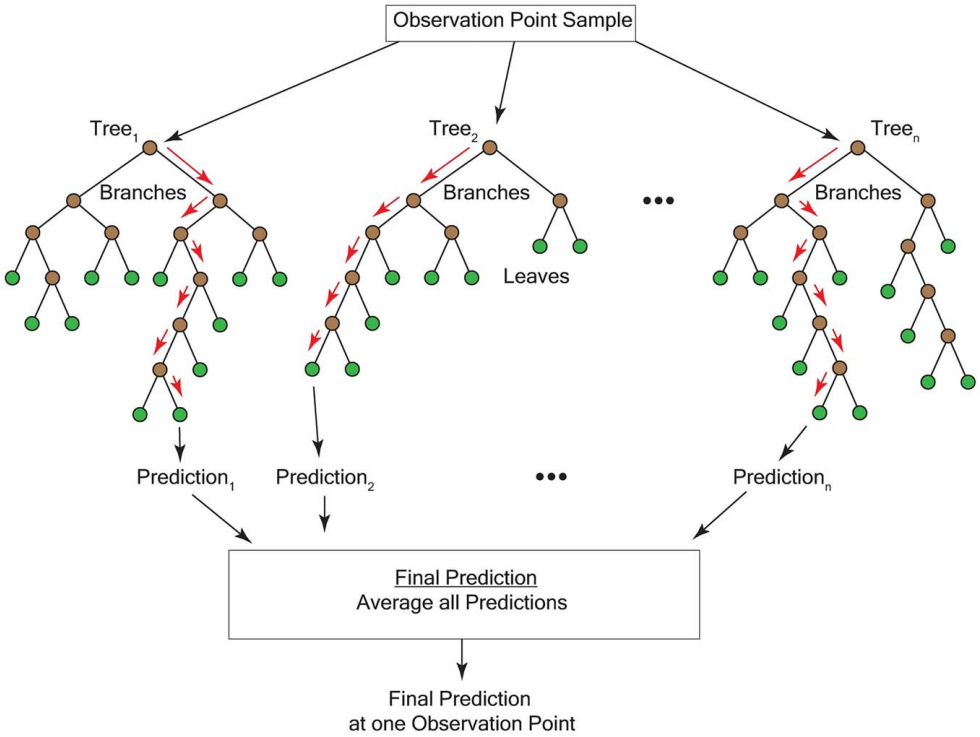

Fig. 5. Random Forest regressor prediction. Arrows indicate the prediction path in each decision tree of the RF. Dark nodes are new predictors and light nodes are the leaves. (Adapted from Graw et al. 2021.)

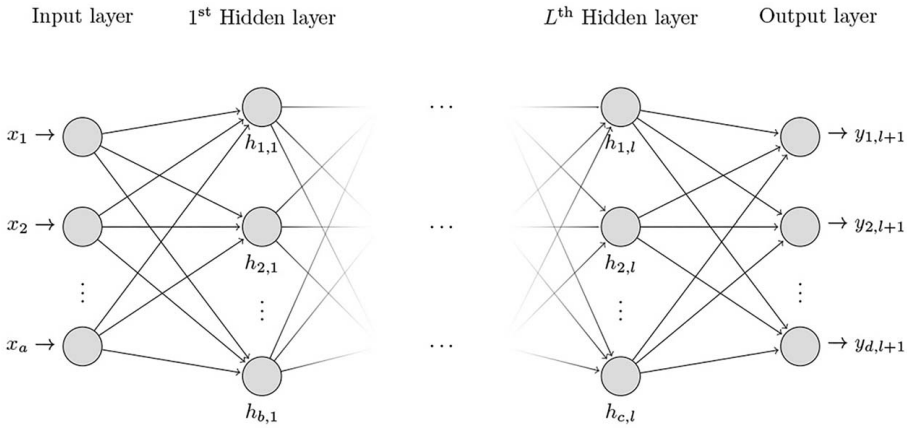

Fig. 6. Sample MLP architecture. The input, hidden, and output variables are represented by circular nodes, and the weight parameters are represented by links connecting the nodes. (Reprinted with permission from Castro et al. 2017 © PLOS One.)

adoption of measures to prevent overfitting. In particular, the RF used in this work had 75 trees in the ensemble, and the MSE was the loss function. To prevent overfitting, the minimum number of samples required to split an internal node was set to two, and the maximum depth is limited to three layers. Experiments conducted with trees allowed to grow further during training (maximum depth up to 10) evidenced that the latter produced lower training errors, but higher out-of-sample errors, possibly indicating overfitting in the presence of high variability of data. The combination of parameters adopted reduces out-of-sample errors.

MLP Neural Network

MLP models were found to be an accurate, fast, and cost-efficient method to forecast natural processes, such as water level fluctuation in navigable rivers and rainfall (Zhou et al. 2020; Esteves et al. 2019). An MLP is a feedforward ANN that is fully connected. The architecture of an MLP contains an input layer, one or more hidden layers, and an output layer. The input layer takes all data in a vector and connects these values to a hidden layer. The hidden layer adds these input values together using biases and weights which then feeds to the output layer (Fig. 6). The hidden layers weights are adjusted and trained using backpropagation, minimizing a training metric such as the MSE or MAE. The output layer creates another vector of a length corresponding to the shape of output data (Bishop 2006). For this work, the input layer had a size corresponding to the number of input days (45, 60, 90, 120, 150, 180) and input variables (7). The MLP regressor was constructed with two hidden layers of 10 and 5 nodes, respectively, using the hyperbolic tan activation function and stochastic gradient descent for optimization of weights. During implementation, it was found that adding a multioutput regression support improved results when compared with adopting a single MLP regressor per timeseries instance. The multioutput regressor fits one regressor per target (i.e., forecasted days) (Pedregosa, et al. 2011). The selection of this architecture was made by comparing forecasting framework results (nRMSE) obtained with different MLP architectures including one, two, and three hidden layers of different sizes.

LSTM Neural Network

In a feedforward ANN, like MLP, data flow in only one direction. In contrast, recurrent neural networks (RNNs) have a feedback loop that allows data to be reprocessed into the model (Brezak et al. 2012). The LSTM is a form of RNN, known by its ability to solve complex, long-time-lag tasks (Hochreiter and Schmidhuber 1997). LSTMs have outperformed ARIMA and Random Forests for timeseries forecasting applications (Siami-Namini et al. 2018; Mussumeci and Codeco Coelho 2020). LSTMs attempt to overcome the problems of exploding and vanishing gradients that arise from back propagation, or the systems layers becoming so disconnected that the network does not update (Gonzalez and Yu 2018). The LSTM solves this problem through the design of its cell state including an input gate, a forget gate, and an output gate. The input gate adds new information to the cell, and by using both tanh and sigmoid filters, we ensure redundant information is not added. The forget gate takes the hidden state and output state from the previous cell and creates bias from 0 to 1 on how much the information should be remembered. The final gate is the output, this gate ensures that all information leaving the cell is good information for both the output and the next cell's input (Fig. 7).

The LSTM model used to forecast shoaling was built with the Keras library, a library built on Tensorflow that allows for easily constructing ML models. In this model, a sequential LSTM was made with a dimensionality of 50–75 LSTM layers all with the hyperbolic tan activation function. The final layer is a dense layer with a linear activation. During the training phase of each epoch, the model has dropout layers after each LSTM layer to prevent

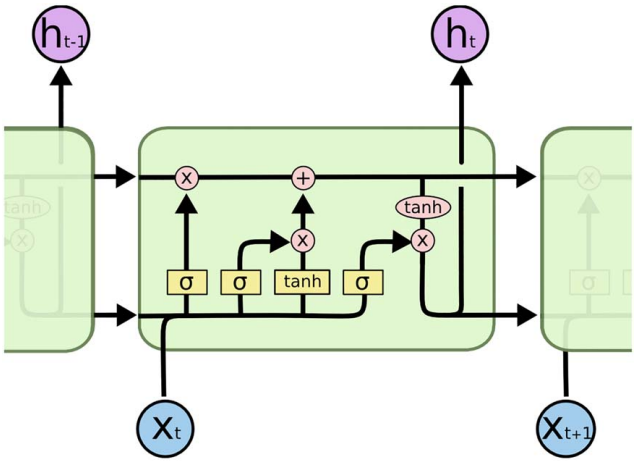

Fig. 7. Sample LSTM model schema. (Reprinted with permission from Springer Nature: Springer, Deep Learning for Natural Language Processing, "Unfolding recurrent neural networks," P. Goyal, S. Pandey, and K. Jain, © 2018.)

Table 1. Average *n*RMSE results for different in-lag scenarios and machine learning regressors. Forty-five-day forecasts

	RF			MLP			LSTM		
	All instances		Subset[a]	All instances		Subset[a]	All instances		Subset[a]
In-lag (days) scenario	*n*RMSE	CI	$n\text{RMSE}_{inc}$	*n*RMSE	CI	$n\text{RMSE}_{inc}$	*n*RMSE	CI	$n\text{RMSE}_{inc}$
45	**0.408**	0.113	0.308	**0.317**	0.064	0.328	0.430	0.082	0.389
60	0.422	0.119	**0.302**	0.320	0.069	**0.313**	0.453	0.072	0.387
90	0.433	0.127	0.310	0.337	0.068	0.354	**0.416**	0.089	**0.333**
120	0.445	0.138	0.300	0.429	0.089	0.462	0.438	0.085	0.416
150	0.469	0.141	0.337	0.366	0.078	0.336	0.460	0.076	0.464
180	0.477	0.139	0.328	0.384	0.106	0.345	0.462	0.078	0.411

Note: Bold = Lowest *n*RMSE, $n\text{RMSE}_{inc}$ for a given regressor considering all instances or subset.
Best model considering all regressors and scenarios is MLP with 60-day in-lag.
[a]Subset of model evaluation timeseries (instances) that evidence an increase in sediment accumulation.

overfitting on the training data. The activation function, number of LSTM layers, and dropout density were determined using the Keras tuner to optimize the model. The input used the first 80% of the data set as training, with a buffer of 195 days (the sum of input and forecasted days for a single timeseries instance), and the remaining data in the test set. For example, the model for the scenario of 90-day input day had 52,820 trainable parameters.

Results and Discussion

Comparison of Models and In-Lag Scenarios

Table 1 and Fig. 8 summarize the results, in terms of average *n*RMSE, $n\text{RMSE}_{inc}$, and confidence intervals, obtained for each regression model type and in-lag scenario.

From Table 1 we observe that, in terms of the *n*RMSE calculated on all instances, the MLP outperforms the RF and LSTM for all in-lag scenarios [Table 1 and Fig. 8(a)] in terms of both *n*RMSE and CI. A priori, scenarios of longer in-lags may be perceived as more beneficial to generate forecasts. However, the longer the in-lag days, the smaller the number of timeseries instances available to train the model. For the RF and LSTM, a recognition of patterns that occur within timeseries instances of 85 and 135 days (respectively, considering 45 and 90 in-lags, respectively, plus 45 forecasted days) is found more beneficial (Table 1). For the MLP, the 45-day in-lag scenario leads to lower *n*RMSE. For the MLP and RF framework, additional training sets (i.e., lower in-lag day scenarios) are found to lower the *n*RMSE by approximately 0.07; while for LSTM, the *n*RMSE improvement is by 0.05. We can also conclude that all machine learning regressors outperform the univariate ARIMA model, due to the latter not being able to capture any changes in shoaling passed the first five days of forecasting (Fig. 4).

The *n*RMSE model evaluation metric analyzed in the precedent paragraph is calculated by comparing daily observed and forecasted volumes in each of the timeseries instances used for model evaluation. However, the *n*RMSE value by itself does not provide insight into the characteristics of the instances with poorer or better metric results. For the purpose of this work, which is to forecast dredging needs (using sediment accumulation as proxy), it is particularly interesting to observe the behavior of model predictions in conditions of increasing shoaling. A model that forecasts decreases in shoaling with little to no error but fails to identify increasing trends will not be as useful, despite the metric results. From the 36 model evaluation timeseries instances (with start date 2 months apart from each other), 13 have increasing shoaling trends, 11 have decreasing trends, and the remaining 12 have a relatively steady shoaling trend. The $n\text{RMSE}_{inc}$ was calculated on the subset of the 13 instances with increasing shoaling trends [Eq. (3)]. The lowest $n\text{RMSE}_{inc}$ for all in-lag scenarios was obtained with RF. However, by comparing RF values for *n*RMSE and $n\text{RMSE}_{inc}$ (Table 1), and observing the results of all testing instances, RF tends to forecast increasing shoaling trends

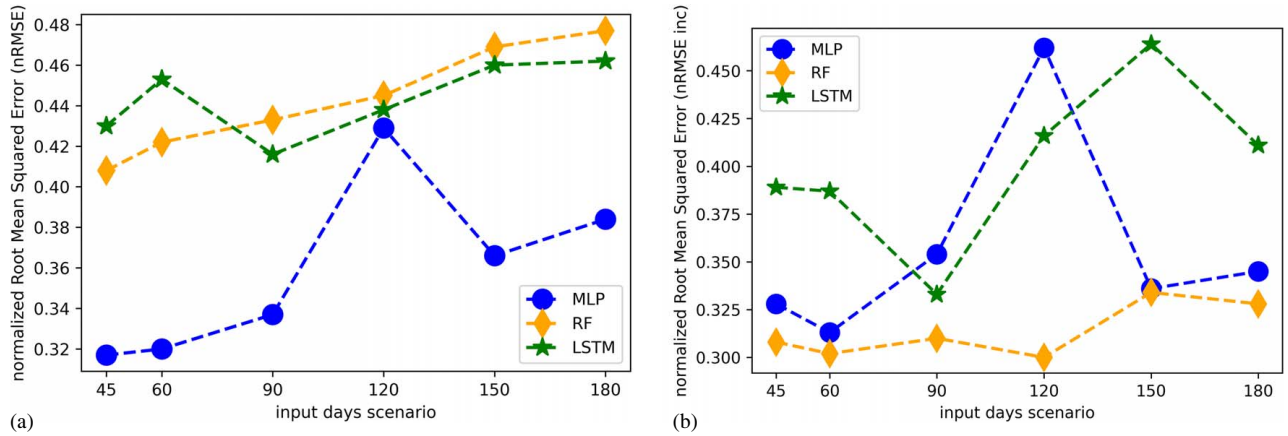

Fig. 8. Comparison of results obtained with RF, MLP, and LSTM model types and in-lag scenarios, based on *n*RMSE [Fig. 8(a)] and $n\text{RMSE}_{inc}$ [Fig. 8(b)]: (a) all testing timeseries. Each point represents the average *n*RMSE of all 36 evaluation instances tested for each combination of model type and in-lag scenario; and (b) testing timeseries with increasing shoaling trend. Each point represents the average *n*RMSEinc, of the 13 evaluation instances with increasing observed shoaling trend for each combination of model type and in-lag scenario.

even in cases when they are not observed. An algorithm that falsely predicts a need to mobilize dredges would lose credibility and may generate unnecessary expenditures. For the MLP, the lowest $n\text{RMSE}_{\text{inc}}$ value was obtained for the 60-day input scenario [Table 1, Fig. 8(b)], which is recommended by this paper.

As discussed, it is desirable to avoid models that incorrectly forecast a nonexistent increase in shoaling conditions. In this context, Fig. 9 depicts the daily forecasted volumes of sediment for different scenarios of in-lags for a sample timeseries instance used for model evaluation, with the MLP [Fig. 9(a)], RF [Fig. 9(b)], and LSTM [Fig. 9(c)]. The continuous line represents observed values of sediment accumulation, while the dashed line shows the forecasted values. While all the forecasts in Fig. 9 seem to correctly predict that there is an increase shoaling trend, some models do better than others in identifying when the peak occurs, and the order of magnitude of such an increase. The sample timeseries depicted in Fig. 9 corresponds to a forecasting date starting in November 2021, that is, instance "21_11" in Fig. 10. Fig. 10 exhibits the detailed *n*RMSE results (vertical axis values) for each of the model evaluation instances (horizontal axis values), regressors (one per subfigure), and in-lag scenarios (one per line). The background represents whether the observed shoaling trend for each instance is increasing (in grey, indicating a potential need to mobilize dredges), or decreasing/steady (in white, indicating no need to mobilize dredges within the next 45 days). From Fig. 10, it is noticeable that the RF is less sensitive to the number of in-lag days than the MLP and LSTM, especially for instances before December 2020, when the Southwest Pass was deepened from 14.8 m (48.5 ft.) to 15.2 m (50 ft.).

Building upon the *n*RMSE, another important metric to consider is the confidence interval of the *n*RMSE between all 36 model evaluation instances in a given model. A narrower confidence interval leads to higher confidence in the ability to use the models as a method for predicting future sediment volumes. The *n*RMSE of

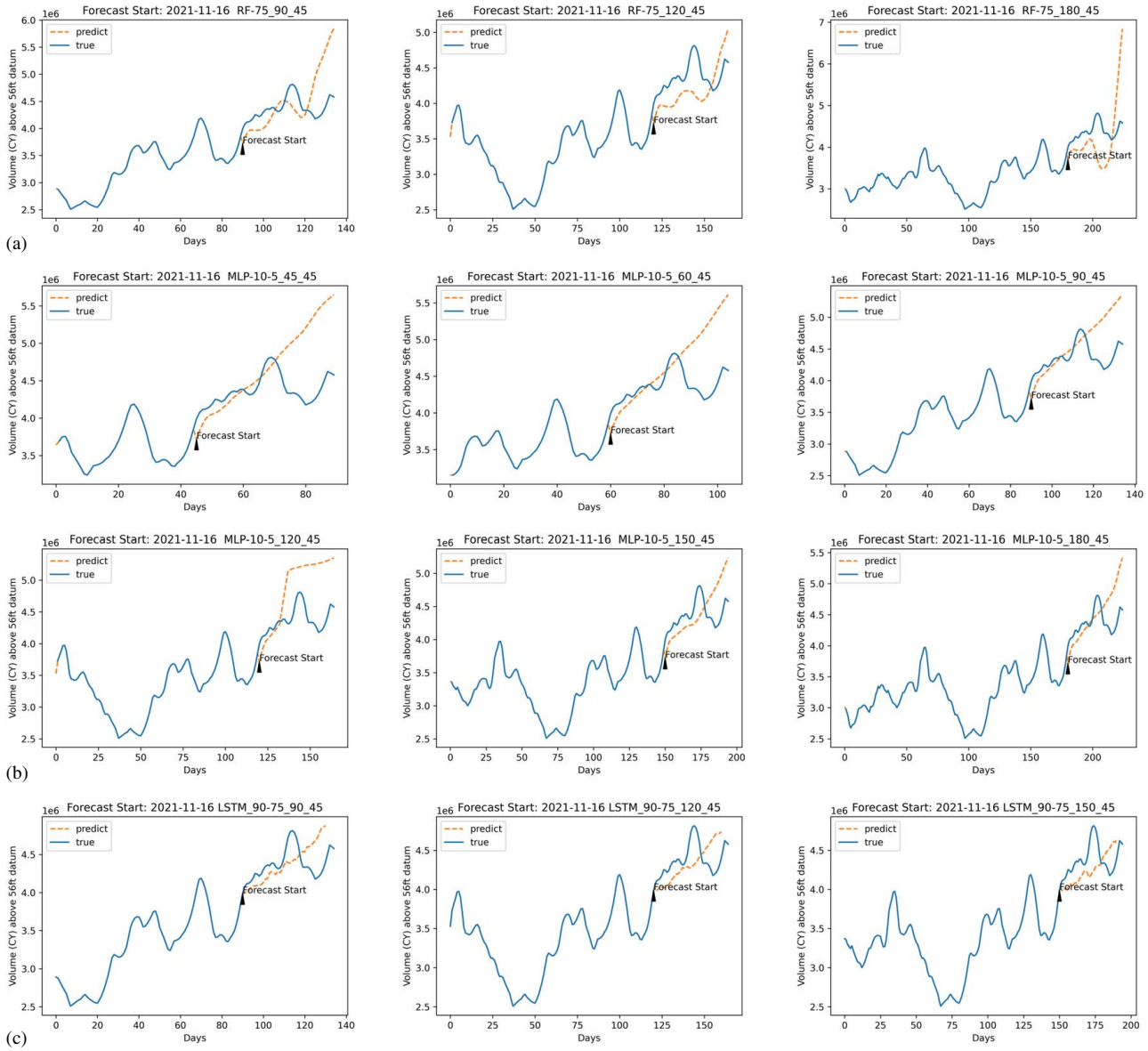

Fig. 9. Daily volume of sediment accumulation in SWP, observed (solid lines), and forecasted (dashed lines) for different in-lag scenarios with RF, MLP, and LSTM in a sample timeseries instance: (a) random forest. In-lag scenarios of 90, 120, and 150 days, from left to right; (b) multilayered perceptron neural network. In-lag scenarios of 45, 60, 90, 120, 150, and 180 days, from top-left to bottom-right; and (c) long short-term memory neural network. In-lag scenarios of 90, 120, and 150 days, from left to right.

J. Waterway, Port, Coastal, Ocean Eng., 2024, 150(2): 04023022

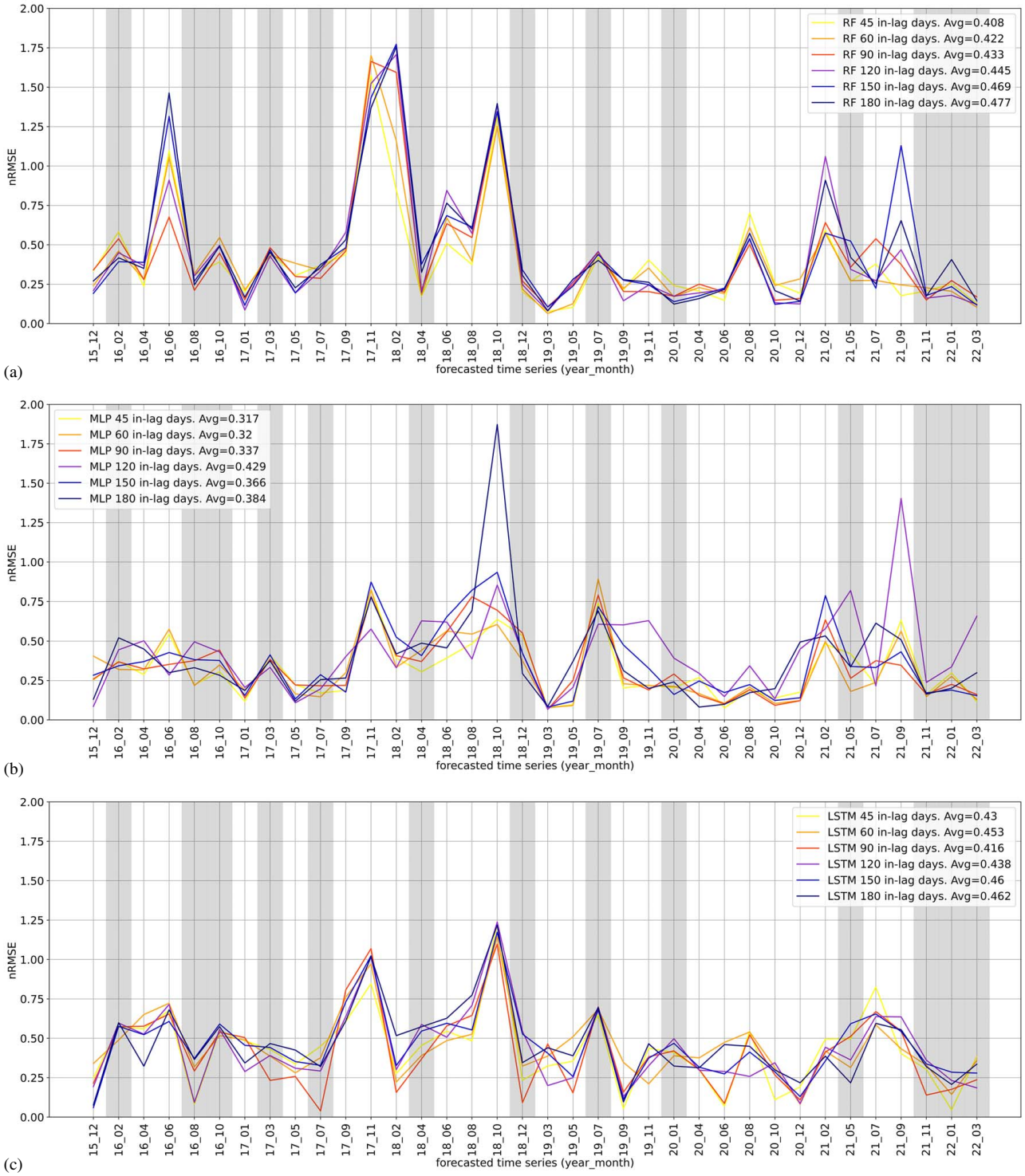

Fig. 10. Detailed *n*RMSE results (vertical axis) for all machine learning regressors (RF, MLP, LSTM in each subfigure), in-lag scenarios (lines), and model evaluation instances (horizontal axis). Background represents observed shoaling trend in each model evaluation instance; grey for increasing (i.e., need to mobilize dredges); white for decreasing or steady (i.e., no need to mobilize dredges): (a) random forest *n*RMSE results for all in-lag scenarios and model evaluation instances; (b) multilayered perceptron neural network *n*RMSE results for all in-lag scenarios and model evaluation instances; and (c) long short-term memory neural network *n*RMSE results for all in-lag scenarios and model evaluation instances.

each model is listed in Table 1. For example, for the MLP 60-45 model, the *n*RMSE is 0.320 with a standard deviation of 0.20, and a confidence interval of 0.069. Lower confidence intervals indicate a preferred model. Considering all model types and in-lag scenarios tested, the lowest confidence intervals were obtained with the MLP regressors, for scenarios of 90-day input days or less (Table 1).

Based on observations of graphs like those in Figs. 9 and 10, and considering the metrics presented in Table 1 and Fig. 8, it was concluded that *the most appropriate machine learning*

regressor model to forecast 45-day dredging needs in SWP is the MLP regressor that identifies shoaling patterns with a 60-day input window.

Future Work

Several lines of research may derive from the initial effort to forecast dredging needs at SWP presented in this paper. In the short term, the authors will focus on examining algorithms to detect and forecast timeseries peaks and valleys values. In addition, the efforts developed thus far consider the daily sediment accumulation along the 56-km (35-mi) SWP as a single variable. The framework may be adapted to break down SWP into three stretches or more. In the mid-term, the authors plan to explore the prediction of timeseries based on images, each image representing the two-dimensional sediment accumulation along the SWP in a color-coded fashion (i.e., darker indicating more presence of sediment). This future analysis would allow for a higher spatial resolution of results, directing dredging efforts to specific locations within SWP within a shorter time-period.

In the long-term, the work presented in this paper may constitute the basis to create a shoaling forecasting capability automatically implemented in a rolling basis (daily or weekly). Thus, one of the premises of this work was to use data that could be automatically obtained and collected with a daily/weekly frequency. As USACE develops further tools to share data through APIs, more data could be considered as potential inputs to the model. For example, the number of dredges in the SWP or the volume of sediment actually dredged at the SWP could constitute the dependent variable in the future. Building upon input variables, the authors will explore the incorporation of existing USACE model outputs, such as the Streamflow forecasting tool (Yeates et al. 2020), and implement lagged steps for input variables, as done by Bhattacharya and Solomatine (2006). Variables further up the Mississippi river will take longer to have influence on the shoaling occurring around the SWP, which could lead to certain variables not being used currently by some models having a stronger relation with the variable to predict. In this paper, it is assumed that the machine learning models are automatically recognizing and learning from such lagged-patterns. Currently, we are looking into what lags should be used for different stations based on discharge rates and other factors.

Lastly, alternative methods to evaluate results may better inform dredging needs. For example, the results obtained with this initial framework may be subject to a classification algorithm to evaluate the ability of the model to predict the shoaling trend, comparing the forecasted and observed shoaling slopes, and classifying them into categorical classes based on whether the prediction was correct.

Conclusion

This paper presents initial research steps in the development of a capability to increase response time and accuracy in forecasting dredging needs at the Southwest Pass (SWP). The SWP is a 56-km (35-mi) stretch of the Mississippi River Delta, in its mouth at the Gulf of Mexico, key for U.S. commerce. The daily volume of sediment accumulation in the SWP is used as proxy for dredging needs. Machine learning regression models are tested and compared within a multivariate, multistep timeseries forecasting framework. Forty-five days are forecasted. The framework collects several variables (such as stage, discharge, turbidity) from over 50 USGS sites. A Decision Tree-based Extreme Gradient Boosting algorithm (XGB) selects the six most important features from the variables collected and feeds multiple machine leaning regression models. The paper focuses on comparing results obtained by applying a Random Forest ensemble (RF), a Multilayered Perceptron (MLP), and a Long Short-Term Memory (LSTM) neural network. Sediment accumulation estimates obtained with a simple univariate ARIMA model are used as baseline. Different scenarios of data parsing are tested (45, 60, 90, 120, 150, and 180 days). In terms of the normalized RMSE (*n*RMSE), all three multivariate machine learning regressors perform better than the univariate baseline ARIMA model. In addition, the MLP outperforms the RF and LSTM for all in-lag scenarios. By comparing the *n*RMSE on the model evaluation instances that evidence an increase in sediment accumulation ($n\text{RMSE}_{\text{inc}}$), and considering that a better model prevents the false identification of dredging needs, an in-lag scenario of 60 days is found to lead to best results. The proposed framework is transferable to other river channels and inland waterways.

Data Availability Statement

Some or all data, models, or code used during the study were provided by a third party. Direct requests for these materials may be made to the provider as indicated in the Acknowledgements. Input data for the model were obtained from CSAT Tool (https://cirp.usace.army.mil/products/csat.php), RiverGages.com, and USGS National Water Information System (https://nwis.waterdata.usgs.gov/nwis).

Acknowledgments

This study was funded by the USACE Mississippi Valley Division and the USACE – Engineer Research and Development Center (ERDC) Dredging Innovations Group (DIG). The authors thank the USACE New Orleans District for their support and expert knowledge of the SWP, the team at the ERDC Coastal and Hydraulics Laboratory and Information Technology Laboratory, and data providers. Input data for the model were obtained from CSAT Tool (https://cirp.usace.army.mil/products/csat.php), RiverGages.com, and USGS National Water Information System (https://nwis.waterdata.usgs.gov/nwis).

References

Abadi, M., et al. 2015. "TensorFlow: Large-scale machine learning on heterogeneous distributed systems." *Preprint, submitted March 14, 2016*. https://doi.org/10.48550/arXiv.1603.04467.

Afan, H. A., A. El-Shafie, W. H. M. W. Mohtar, and Z. M. Yaseen. 2016. "Past, present and prospect of an Artificial Intelligence (AI) based model for sediment transport prediction." *J. Hydrol.* 541 (B): 902–913. https://doi.org/10.1016/j.jhydrol.2016.07.048.

Bhattacharya, B., and D. P. Solomatine. 2006. "Machine learning in sedimentation modelling." *Neural Networks* 19 (2): 208–214. https://doi.org/10.1016/j.neunet.2006.01.007.

Bishop, C. M. 2006. *Pattern recognition and machine learning*. New York: Springer.

Biswas, A., S. Ahmed, T. Bankefa, P. Ranganathan, and H. Salehfar. 2021. "Performance analysis of short and mid-Term Wind Power prediction using ARIMA and Hybrid Models." In *Proc.,* 2021 *IEEE Power and Energy Conf.*, 1–7. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE).

Breiman, L. 2001. "Random forests." *Mach. Learn.* 45: 5–32. https://doi.org/10.1023/A:1010933404324.

Brezak, D., T. Bacek, D. Majetic, J. Kasac, and B. Novakovic. 2012. "A comparison of feed-forward and recurrent neural networks in time series forecasting." In Proc., *IEEE Conf. on Computational Intelligence for Financial Engineering & Economics*, 1–6. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE).

Castro, W., J. Oblitas, R. Santa-Cruz, and H. Avila-George. 2017. "Multilayer perceptron architecture optimization using parallel computing techniques." *PLoS One* 12 (12): e0189369. https://doi.org/10.1371/journal.pone.0189369.

Chou, J.-S., and J.-W. Lin. 2020. "Risk-Informed Prediction of Dredging Project Duration Using Stochastic Machine Learning." *Water* 12 (6): 1643. doi:https://doi.org/10.3390/w12061643.

De Figueiredo, N. M., and C. J. Cavalcante Blanco. 2016. "Water level forecasting and navigability conditions of the tapajós river – Amazon - Brazil." *La Houille Blanche* 102: 53–64. https://doi.org/10.1051/lhb/2016031.

DOT, Bureau of Transportation Statistics. 2022. *Port data catalog. Port performance freight statistics program: Annual report*. Washington, DC: DOT.

Dudek, G. 2015. "Short-term load forecasting using random forests." *Adv. Intell. Syst. Comput.* 323: 821–828. https://doi.org/10.1007/978-3-319-11310-4_71.

Dunkin, L. M., L. A. Coe, and J. J. Ratcliff. 2018. *Corps shoaling analysis tool: Predicting channel shoaling*. Vicksburg, MS: U.S. Army Engineer Research and Development Center.

Esteves, J. G., G. de Souza Rolim, and A. Ferraudo. 2019. "Rainfall prediction methodology with binary multilayer perceptron neural networks." *Clim. Dyn.* 52: 2319–2331. https://doi.org/10.1007/s00382-018-4252-x.

Gonzalez, J., and W. Yu. 2018. "Non-linear system modeling using LSTM neural networks." *IFAC* 51: 485–489.

Goyal, P., S. Pandey, and K. Jain. 2018. "Unfolding recurrent neural networks." *Springer*. Accessed January 19, 2023. https://link.springer.com/chapter/10.1007/978-1-4842-3685-7_3.

Graw, J. H., W. T. Wood, and B. J. Phrampus. 2021. "Predicting global marine sediment density using the random forest regressor machine learning algorithm." *J. Geophys. Res.: Solid Earth* 126. https://doi.org/10.1029/2020JB020135.

Gupta, H. V., and H. Kling. 2011. "On typical range, sensitivity, and normalization of Mean Squared Error and Nash-Sutcliffe Efficiency type metrics." *Water Resour. Res.* 47. https://doi.org/10.1029/2011WR010962.

Hartman, M. A., K. N. Mitchell, L. M. Dunkin, J. Lewis, B. Emery, N. F. Lenssen, and R. Copeland. 2022. "Southwest pass sedimentation and dredging data analysis." *J. Waterw. Port Coastal Ocean Eng.* 148 (2): 05021017. https://doi.org/10.1061/(ASCE)WW.1943-5460.0000684.

Hochreiter, S., and J. Schmidhuber. 1997. "Long short-term memory." *Neural Comput.* 9: 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Hunter, J. D. 2007. "Matplotlib: A 2D graphics environment." *Comput. Sci. Eng.* 9 (3): 90–95. https://doi.org/10.1109/MCSE.2007.55.

Kress, M. 2021. "Hydrosurvey data archive for federal navigation projects." *Data.gov* Accessed February 21, 2023. https://data.gov/maritime/hydrosurvey-data-archive-for-federal-navigation-projects/index.html.

Mateo-Pérez, V., M. Corral-Bobadilla, F. Ortega-Fernández, and V. Rodríguez-Montequín. 2021. "Determination of water depth in ports using satellite data based on machine learning algorithms." *Energies* 14 (9): 2486. https://doi.org/10.3390/en14092486.

Murat, M., I. Malinowska, M. Gos, and J. Krzyszczak. 2018. "Forecasting daily meteorological time series using ARIMA and regression models." *Int. Agrophys.* 32 (2): 253–264. https://doi.org/10.1515/intag-2017-0007.

Mussumeci, E., and F. Codeco Coelho. 2020. "Large-scale multivariate forecasting models for Dengue - LSTM versus random forest regression." *Spatial Spatio-temporal Epidemiol.* 35: 100372. https://doi.org/10.1016/j.sste.2020.100372.

Nel, H. A., T. Dalu, and R. J. Wasserman. 2018. "Sinks and sources: Assessing microplastic abundance in river sediment and deposit feeders in an Austral temperate urban river system." *Sci. Total Environ.* 612: 950–956. https://doi.org/10.1016/j.scitotenv.2017.08.298.

Pedregosa, F., et al. 2011. "Scikit-learn: Machine learning in Python." *J. Mach. Learn. Res.* 12 (85): 2825–2830.

Pektas, A. O., and H. K. Cigizoglu. 2017. "Long-range forecasting of suspended sediment." *Hydrol. Sci. J.* 62 (14): 2415–2425. https://doi.org/10.1080/02626667.2017.1383607.

PyPI. 2021. "xgboost 1.5.0." *pypi.org*. Accessed December 19, 2022. https://pypi.org/project/xgboost/1.5.0/.

Rady, E. A., H. Fawzy, and A. Fattah. 2021. "Time series forecasting using tree based methods." *J. Stat. Appl. Probab.* 10: 229–244. http://dx.doi.org/10.18576/jsap/100121.

Seabold, S., and J. Perktold. 2010. "Statsmodels: Econometric and statistical modeling with Python." In *Proc., of the 9th Python in Science Conf*. Austin, TX: Enthought.

Shcherbakov, M., A. Brebels, N. Shcherbakova, A. Tyukov, T. Janovsky, and V. Kamaev. 2013. "A survey of forecast error measures." *World Appl. Sci. J.* 24: 171–176.

Siami-Namini, S., N. Tavakoli, and A. Siami Namin. 2018. "A comparison of ARIMA and LSTM in forecasting time series." In *Proc.,* 17th *IEEE Int. Conf. on Machine Learning and Applications*, 1394–1401. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE).

Tyralis, H. 2017. "Variable selection in time series forecasting using random forests." *Algorithms* 10 (114): 114. https://doi.org/10.3390/a10040114.

USACE. 2022. "National channel framework." *USACE Geospatial*. Accessed October 19, 2022. https://geospatial-usace.opendata.arcgis.com/maps/9227967a2748410983352b501c0c7b39/about.

USGS. 2018. "Sediment and sustpended sediment." *Water Science School*. Accessed September 19, 2022 https://www.usgs.gov/special-topics/water-science-school/science/sediment-and-suspended-sediment.

USGS. 2022. "National water information system: Web interface." *Water Data for the Nation*. Accessed December 21, 2022. https://nwis.waterdata.usgs.gov/nwis.

Washington, S., M. Karlaftis, F. Mannering, and P. Anastasopoulos. 2020. *Statistical and econometric methods for transportation data analysis*. 3rd ed. New York: Taylor & Francis Group.

Wei, R., J. Wang, M. Su, E. Jia, S. Chen, T. Chen, and Y. Ni. 2018. "Missing value imputation approach for mass spectrometry-based metabolomics data." *Sci. Rep.* 8: 663. https://doi.org/10.1038/s41598-017-19120-0.

Wheatcroft, R. A., and C. K. Sommerfield. 2005. "River sediment flux and shelf sediment accumulation rates on the Pacific Northwest margin." *Cont. Shelf Res.* 25 (3): 311–332. https://doi.org/10.1016/j.csr.2004.10.001.

Yeates, E., A. Tavakoly, K. Mitchell, G. Dreaper, and S. Afshari. 2020. "Utilizing stream flows to forecast dredging requirements." *USACE-ERDC*. Accessed January 20, 2023. https://apps.dtic.mil/sti/pdfs/AD1105836.pdf.

Zhang, Y., H. Yang, H. Cui, and Q. Chen. 2020. "Comparison of the ability of ARIMA, WNN and SVM models for drought forecasting in the Sanjiang Plain, China." *Nat. Resour. Res.* 29: 1447–1464. https://doi.org/10.1007/s11053-019-09512-6.

Zheng, H., J. Yuan, and L. Chen. 2017. "Short-Term load forecasting using EMD-LSTM neural networks with a xgboost algorithm for feature importance evaluation." *Energies* 10 (8): 1168. https://doi.org/10.3390/en10081168.

Zhou, T., Z. Jiang, X. Liu, and K. Tan. 2020. "Research on the long-term and short-term forecasts of navigable river's water-level fluctuation based on the adaptive multilayer perceptron." *J. Hydrol.* 591: 125285. https://doi.org/10.1016/j.jhydrol.2020.125285.