ICCEPM 2020
The 8th International Conference on Construction Engineering and Project Management
Dec. 7-8, 2020, Hong Kong SAR

Real-time prediction on the slurry concentration of cutter suction dredgers using an ensemble learning algorithm

Shuai Han[1], Mingchao Li[1]*, Heng Li[2], Huijing Tian[3], Liang Qin[3], Jinfeng Li[3]

[1] *State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin,* E-mail address: hs2015205039@tju.edu.cn
[2] *State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin,* E-mail address: lmc@tju.edu.cn
[3] *Department of Building and Real Estate, Hong Kong Polytechnic University, Hong Kong SAR,* E-mail address: heng.li@polyu.edu.hk
[4] *Tianjin Dredging Company Limited, China Communications Construction Company, Tianjin,* E-mail address: tiantian904@163.com
[5] *Tianjin Dredging Company Limited, China Communications Construction Company, Tianjin,* E-mail address: hydro_taitan@126.com
[6] *Tianjin Dredging Company Limited, China Communications Construction Company, Tianjin,* E-mail address: lijinfeng137@126.com
* Corresponding author: Mingchao Li, lmc@tju.edu.cn

Abstract: Cutter suction dredgers (CSDs) are widely used in various dredging constructions such as channel excavation, wharf construction, and reef construction. During a CSD construction, the main operation is to control the swing speed of cutter to keep the slurry concentration in a proper range. However, the slurry concentration cannot be monitored in real-time, i.e., there is a "time-lag effect" in the log of slurry concentration, making it difficult for operators to make the optimal decision on controlling. Concerning this issue, a solution scheme that using real-time monitored indicators to predict current slurry concentration is proposed in this research. The characteristics of the CSD monitoring data are first studied, and a set of preprocessing methods are presented. Then we put forward the concept of "index class" to select the important indices. Finally, an ensemble learning algorithm is set up to fit the relationship between the slurry concentration and the indices of the index classes. In the experiment, log data over seven days of a practical dredging construction is collected. For comparison, the Deep Neural Network (DNN), Long Short Time Memory (LSTM), Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Decision Tree (GBDT), and the Bayesian Ridge algorithm are tried. The results show that our method has the best performance with an R^2 of 0.886 and a mean square error (MSE) of 5.538. This research provides an effective way for real-time predicting the slurry concentration of CSDs and can help to improve the stationarity and production efficiency of dredging construction.

Keywords: Cutter suction dredger; Slurry concentration; Real-time prediction; Ensemble learning

1. INTRODUCTION

Dredging construction is the process of excavating and removing sediments and debris from below water level. The primary purposes of dredging include [1]: (1) to deepen the cross-section of channels to improve water transport capacity, flood discharge capacity, and irrigation capacity of the channels; (2) to deepen the cross-section of bays to meet the requirement of navigation, wharf construction, and ship docking; (3) to collect and bring up valuable substance from the bed of a river, sea, etc.; (4) to reclaim land from the sea.

Cutter suction dredgers (CSD) is one of the most widely used kinds of dredgers in dredging construction [2-4]. At present, most of the dredging constructions were completed with CSDs. The

reasons are: (1) compared with other kinds of dredgers (such as chain bucket dredger, drag suction dredger, grab dredger), the cost of using CSD is lower; (2) CSDs are suitable for various working conditions. Especially for the land reclamation constructions and dyke strengthening constructions, CSD is the only choice.

Because of the robust adaptability of CSD, researchers always pay much attention to CSD, no matter in academic and engineering. Tang and Wang [5] proposed an online fault diagnosis system for CSD. Ni et al. [6] studied the characteristics of CSD, simulated the construction process, and discussed the critical problems of CSD construction. Henriksen et al. [7] analyzed the underlying laws for soil disturbances exerted by cutter heads and proposed a near-field resuspension model. Zhang et al. [8] studied the flow law of the slurry in CSD slurry transportation systems based on numerical simulations. Li et al. [9] put forward a dynamic evaluation method on the efficiency of CSD constructions based on the real-time monitoring data.

There are various factors on CSD operation [10], such as the soil conditions and the performance of the dredger. Directly speaking, the CSD construction process is a controlling process to keep the slurry concentration within a proper range [11, 12]. However, because of the characteristics of the structure of the slurry transportation system, the slurry concentration cannot be measured in real-time, i.e., the monitoring data of the slurry concentration are time-lagged. Operators can only guess the current slurry concentration according to the time-lagged values and some other indices that can be measured in real-time, such as vacuum and swing speed. Obviously, this method is subjective and is not accurate. Although many studies have been done on the optimization of CSD operations, there is still not an effective and universal method to predict the slurry concentration with high accuracy. Miedema [13] designed an automatic control system to dynamically determine the boundary conditions of slurry flowrate based on mathematical derivation. Tang et al. [14, 15] presented an automation control system of CSD based on an expert system. Ye et al. [16] proposed a dredger cutter motor synchronous speed control system. Jiang et al. [17] studied the swing process of CSD and proposed to use an RBF-ARX model to optimize the swing process. Li et al. [18] presented a machine learning-based method to predict the construction productivity of CSD. However, even though the expert system proposed by Tang et al. [15] described how they predict the slurry concentration, however, Influencing factors (voltage and slurry concentration) considered in their research were not enough. Besides, most of the researches focuses on how other factors influence slurry concentration, but they still cannot solve the real-time measuring problem.

Real-time data prediction is research hotspot in intelligent construction, especially in the field of tunneling, excavation, earthwork, etc. Many key indicators can not be monitored in real time due to the variable geological conditions and complicated parameters of equipment. Data mining algorithms provide an effective solution for this problem [19-20]. For instance, Chen et al. [21] put forward a LSTM based method for predicting TBM parameters. Zhang et al. [22] presented a real-time analysis and regulation method for automatically steering Earth Pressure Balanced (EPB) based on Particle Swarm Optimization (PSO) and Random Forest (RF). Gao et al. [23] compared the performances of Long Short Time Memory (LSTM) neural network, recurrent neural network (RNN), gated recurrent unit (GRU), and some classical regression algorithms on predicting the parameters of TBM, and found that RNN-based predictors could usually make the best real-time predictions. Jing et al. [24] presented a TBM performance prediction model for limestone strata by establishing the relationship between penetration and normal force of single cutter. Gao et al. [25] used LSTM to predict tunnel boring machine (TBM) penetration rate. Leng et al. [26] developed a hybrid data mining method to predict TBM penetration rate based on convolutional neural network (CNN) and classification and regression tree (CART). From the perspective of construction, CSD dredging is similar to tunnel boring, and the fruitful progress of the real-time data prediction of TBM is a good reference for intelligent construction of CSDs.

In this paper, we proposed an ensemble learning-based method to establish the relationship between slurry concentration and some other indices that can be measured in real-time, and thus to predict the real-time slurry concentration of CSD dredging constructions. In the rest of the paper, the overall flow of the study is first presented. Then the construction technology of CSD is introduced. After that, the details of the methodology are described. Finally, a case study is conducted to test the method. Our methods, including the preprocessing, index selection, and the core algorithm, are tailored to the slurry concentration prediction problem. The results show that our algorithm is better than almost all the common regression algorithms, even including the Random Forest, the Deep Neural Network, and the Long Short Time Memory algorithm.

2. OVERALL FRAMEWORK

This flowchart of this research is as shown in Figure 1. First, the construction technology of CSD is briefly discussed in Section 3. Readers who familiar with it can skip this section. The preprocessing methods for the raw data are then proposed (see Section 4.1), including the processing method on the time-lag effect of slurry concentration, the rules for selecting normal construction data, and the filtering rules of the raw data. Among them, the solution to the time-lag effect is a standard method in dredging construction, but we will still talk about it briefly in the next section because it is crucial for the analysis. After that, the feature selection is proposed to find out the factors that both can be monitored in real-time and are related to slurry concentration (see Section 4.2). Then an ensemble learning algorithm is presented to establish the relationships between the selected indices and the slurry concentration (see Section 4.3). After that, a case study is carried out. In the case study, six advanced regression algorithms are used for comparison (see Section 5.2), including the Deep Neural Network (DNN), Long Short Time Memory (LSTM), Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Decision Tree (GBDT), and the Bayesian ridge algorithm. Finally, the research is ended with a meaningful discussion (see Section 5.3).

The red dashed rectangle in Figure 1 illustrates the main innovations of this research, including the data preprocessing method, the index selection method, and the proposed ensemble learning algorithm.

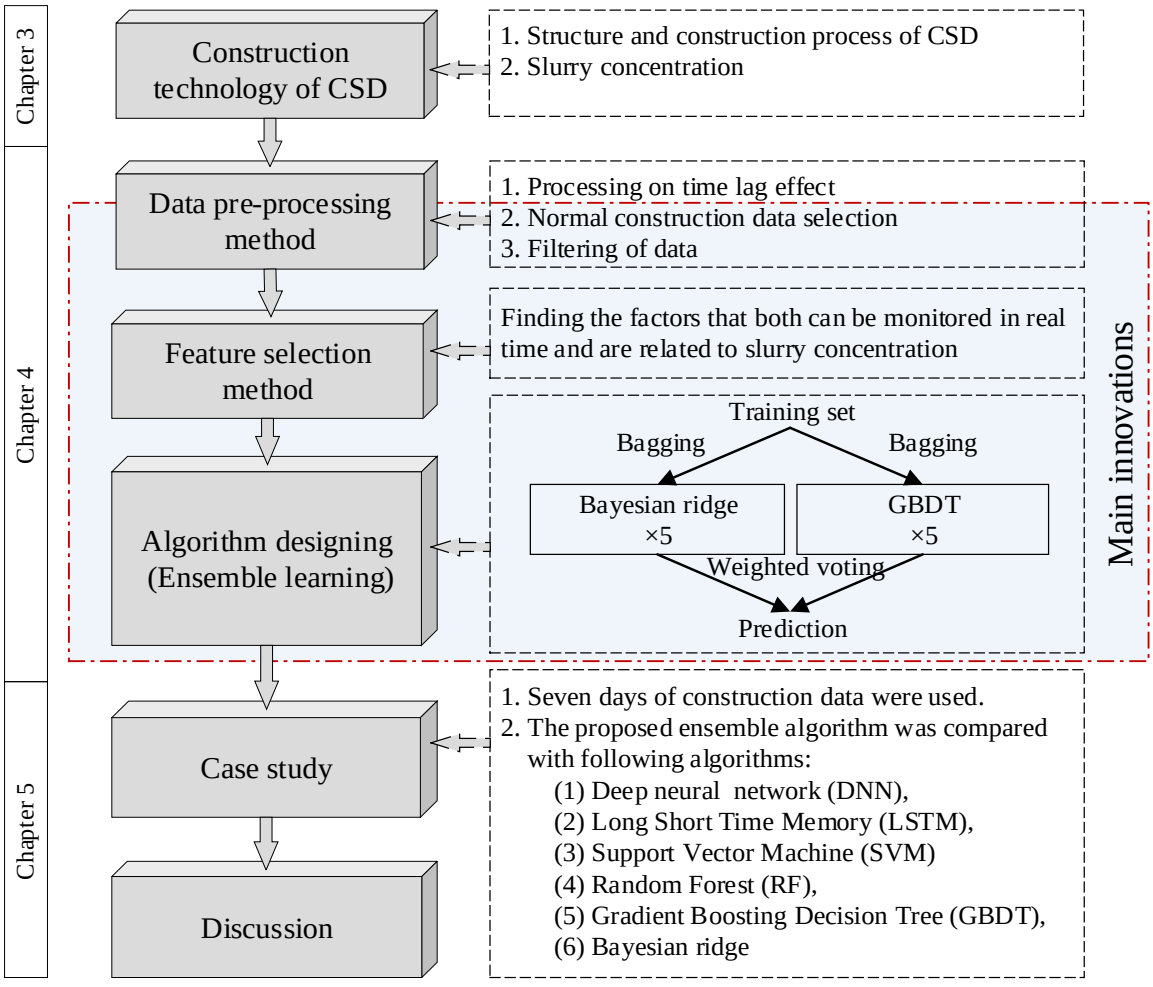


Figure 1. Flow chart of the research.

3. CONSTRUCTION TECHNOLOGY OF CSD

3.1. Structure and construction process of CSD

Figure 2 is the model of a typical CSD. The main components include a carriage, two spuds (primary spud and auxiliary spud), several pumps (underwater pump and carriage pumps), a ladder, two swing winches, two anchors, a cutter, and a series of pipelines. The primary spud (or working spud) is on the

carriage, and is used to fix the CSD; the auxiliary spud (or walking spud) is for assisting the primary spud in moving the CSD; the pumps are used to transport the slurry to a specific area through the discharge pipes, where the underwater pump is on the ladder, and the carriage pumps are in the carriage; the swing winches are for controlling the swing speed of the cutter; the cutter is on the end of the ladder, and is used to cut soil or rocks.

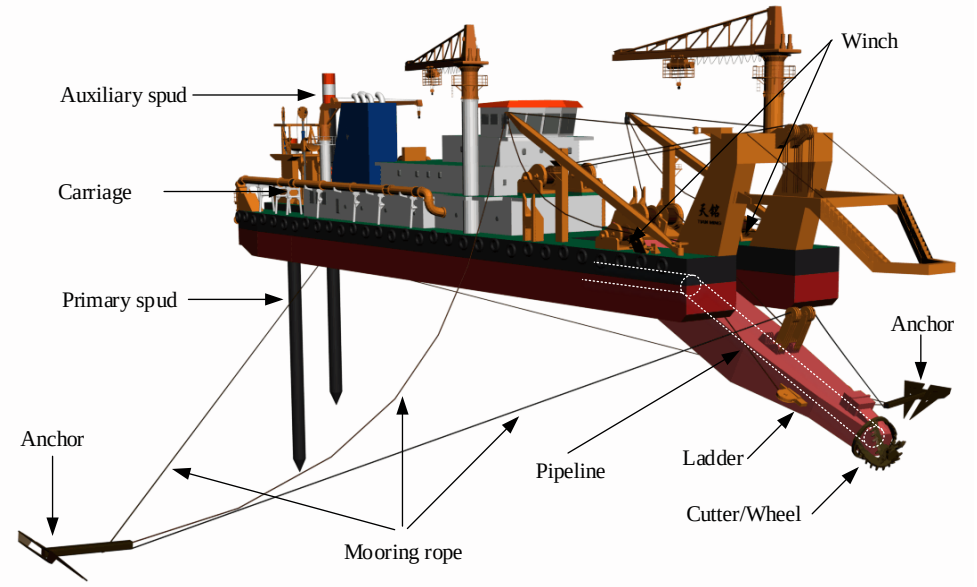


Figure 2. Structure of a CSD.

The construction process of CSD is as shown in Figure 3 and Figure 4. The first step is to fix the primary spud, and then retract or release the mooring ropes to make the cutter swing around the primary spud to dredge from one sideline of the channel to the opposite sideline. After that, the carriage will be moved forward for a step (a certain distance), and then by operating the mooring ropes and the ladder, the cutter will be swung to the reverse direction. The two swing processes can be regarded as a dredging cycle. After several cycles, the movement distance of the carriage will reach its limit, and then the whole CSD should be moved forward by shifting the spuds.

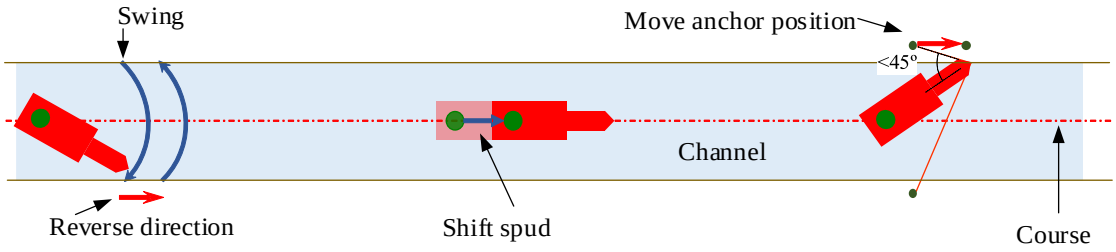


Figure 3. The construction process of CSD.

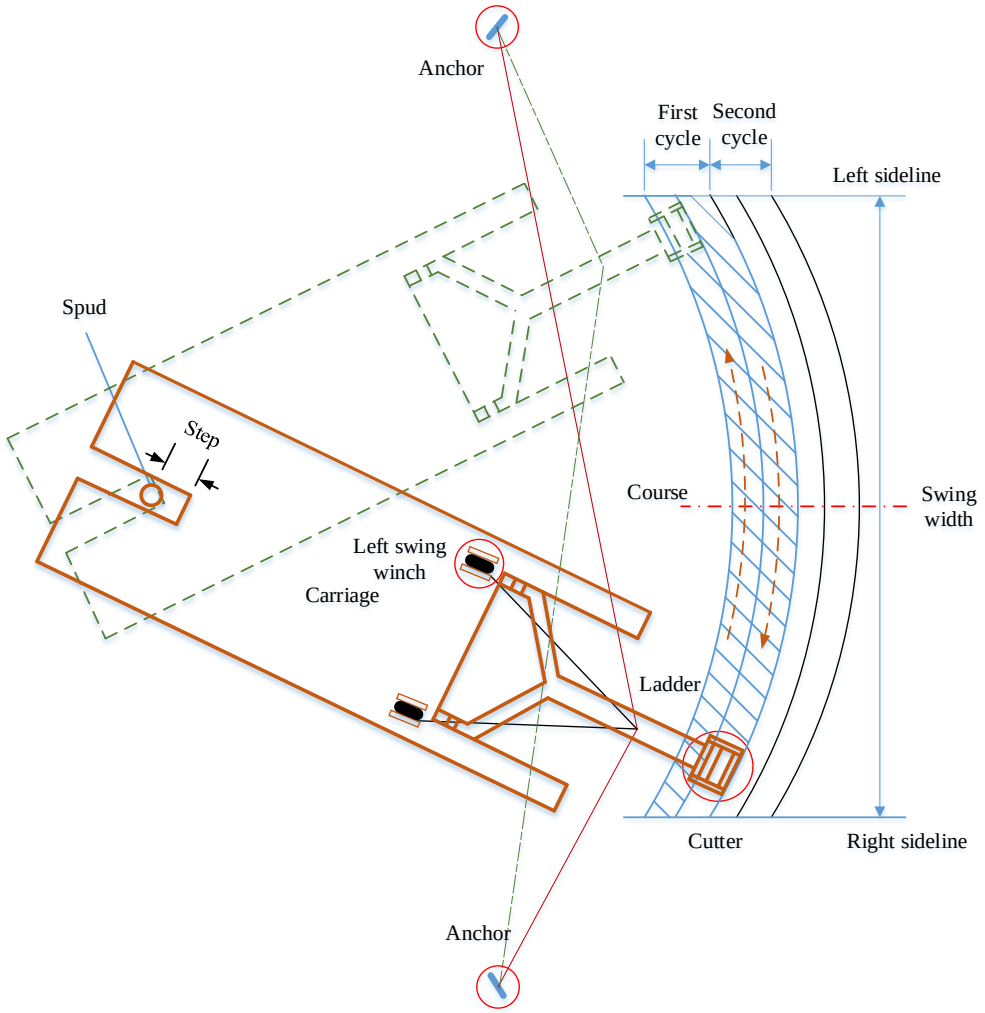


Figure 4. Swing process of CSD.

3.2. Slurry concentration

Slurry concentration in one of the most significant indicators in the dredging process [27]. As can be seen from Figure 3, the CSD is always in the swing process except shifting spuds. Therefore, swinging the cutter is the central part of CSD operations. So what is the major indicator for operators decide to accelerate or decelerate the swing speed? The answer is slurry concentration. An experienced operator can always keep the slurry concentration in a proper range by controlling swing speed. If the current slurry concentration is high, it will affect the normal operation of the devices (such as block the pipelines), and the swing speed should be slow down. On the other hand, if the current slurry concentration is low, the dredging productivity will be low, and usually, it's needed to accelerate the swing speed; but in some particular instances, low slurry concentration is caused by the soil conditions, for example, when dredging on a piece of rock, the slurry concentration will also be low, and some special measures will be needed. All in all, slurry concentration is the most important indicator of dredging.

However, during a swing process, it is impossible to obtain the real-time value of slurry concentrations because of the "time-lag effect". As is shown in Figure 5, in the pipeline, the slurry first flows through the vacuum meter, and then the flow rate meter and the slurry concentration meter. Generally, the slurry concentration meter is dozens of meters far from the vacuum meter. Therefore, the values of the slurry concentration meter are not synchronized with the dredging process. The flowrate values are also not real-time, but they can be converted into real-time values with the continuity equation of flow. However, the time lag problem of slurry concentration cannot be easily solved. In the real construction process, operators always try to guess the current concentration according to the time-lagged values and some other indicators that can be obtained in real-time.

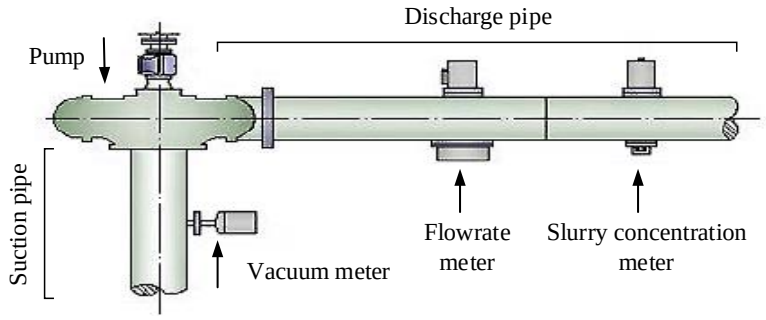


Figure 5. Schematic of time lag effect.

4. METHODOLOGY

4.1. Data preprocessing

4.1.1. Processing on time-lagged slurry concentration

The time lag effect, as mentioned in Section 3.2, is related to the distance between the vacuum meter and the slurry concentration meter. It is also related to the diameters of the suction pipe and discharge pipe, as well as the velocity of slurry. Generally, there are 10-30 seconds of lag. Only if time lag effect is eliminated we can establish the relationship be slurry concentration and other factors.

The time lag effect is eliminated with the continuity of flow [28]. The continuity of flow is defined as "the mean velocities at all cross-sections having equal areas are then equal, and if the areas are not equal, the velocities are inversely proportional to the areas of the respective cross-sections", as shown in Figure 6. The continuity equation of flow is:

$$A_1 v_1 = A_2 v_2 \tag{1}$$

where A_1 is the area of the cross-section of the first pipe, v_1 is the flow rate in the first pipe; A_2 is the area of the cross-section of the second pipe, v_1 is the flow rate in the second pipe;

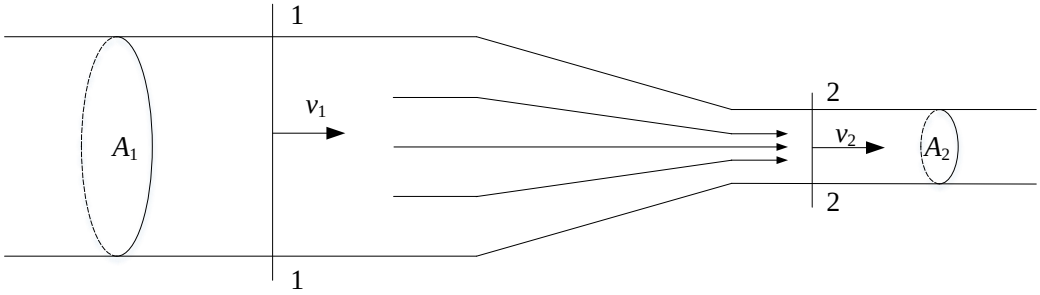


Figure 6. Schematic of the continuity of flow.

The integral of the slurry velocity over time is the distance that the slurry flows. Supposing that the diameter of the suction pipe is d_s and the diameter of the discharge pipe is d_d, the slurry velocity in the suction pipe is v_s, and the slurry velocity in the discharge pipe is v_d, then formula (2) can be derived according to the continuity equation of fluid.

$$v_s = \frac{d_d^2}{d_s^2} v_d \tag{2}$$

The d_s and d_d are constant, the v_s and v_d change over time, and the v_d can be monitored in real-time, as is shown in Figure 4. Supposing that the distance between the vacuum meter and the pump is l_s, and the distance between the slurry concentration meter and the pump is l_d, then the following equation set can be derived:

$$\begin{cases} \int_{t_0}^{t_1} v_s \, \mathrm{d}t = l_s \\ \int_{t_1}^{t_2} v_d \, \mathrm{d}t = l_d \end{cases} \tag{3}$$

where the l_s and l_d are constant. It can be seen that the slurry concentration at t_0 can be measured at t_2.

Base on the above principle, the time lag effect of the slurry concentration can be eliminated.

4.1.2. Normal construction data selection

All the non-construction data should be removed. After eliminating the time lag effect, the period that the slurry concentrations are 0 can be determined as non-construction data. However, not all the periods that the slurry concentrations are not 0 belong to construction periods, and we definite that the dredging process begins at the first wave of slurry flow through the vacuum meter. Figure 7 shows how to determine the construction period in a construction cycle. The t_b and t_e represent the beginning and the ending of the slurry concentration series respectively, the $t_{b\text{-}lag}$ and $t_{e\text{-}lag}$ represent the beginning and the ending of the time-lagged slurry concentration respectively, and the construction period is from $t_{b\text{-}lag}$ to t_e.

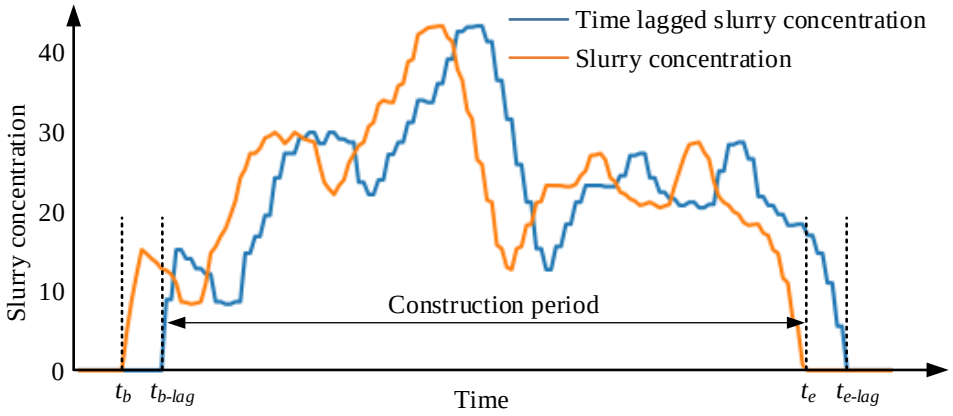


Figure 7. Construction data selection.

Besides, the abnormal data should also be removed: 1) some slurry concentration values may be extremely larger than others, and they should be removed because they may be caused by the sediment of the mud; 2) the electrical indices (such as the cutter motor power and winch power) contains a small amount of abnormal data, which should also be removed. In this research, the Box-plot was used to distinguish between normal and abnormal data.

4.1.3. Data filtering

In this research, three kinds of CSD were analyzed, and we found a common problem that even in the same CSD, the sampling frequencies of different indices are different. For example, the sampling frequency of the winch power meter is once every two seconds, while the sampling frequency of the vacuum meter is once every four seconds. However, the recording frequency of the whole monitoring system should be consistent with the maximum sampling frequency. Therefore, during the construction period, the records of the indicators with small sampling frequency will be copied several times to adapt the recording frequency, and it will lead to errors. One way to reduce these errors is to process the data with smooth filtering. Besides, some instruments are of low accuracy, filtering the data of them also helps for analysis. In this research, the Savitzky-Golay (S-G) filtering [29] was used.

The S-G filtering is commonly used in spectra pretreatment. Its main idea is to smooth a curve within a window by the polynomial fitting. Supposing that $X=(x_0, x_1, \ldots, x_n)$ is the series that need to be filtered, then the ith point x_i will be smoothed as follows:

Supposing the length of the window is $l=2m+1$, and the power of the polynomial is k-1, then our goal is the fit the point series, $(x_{i-m}, x_{i-m+1}, \ldots, x_i, \ldots, x_{i+m-1}, x_{i+m})$, with the formula:

$$y = a_0 + a_1 x + a_2 x^2 + \ldots + a_{k-1} x^{k-1} \quad (4)$$

The coefficients $(a_0, a_1, \ldots, a_{k-1})$ are determined by the least square method. More details about this method can be found in the paper [28].

4.2. Index selection

The goal of this research is to establish the relationship between the slurry concentration and other indices; however, it is difficult to decide which indices should be used because there are hundreds of indices in CSD control system [30], as shown in Table 1. In this study, an index selection method is presented based on the working principles of CSD.

Table 1. Structure of the logging file of the CSD monitoring system

Time (2012-7-7)	1	2	3	4	…	713	715
	Density (t/m^3)	Velocity (m/s)	Quantity of flow (m^3/h)	Cutter power (kW)	…	Cutting angle (°)	Cutter head depth (m)
6:43:35	1.03	5.14	9304.35	936.55	…	2.04	17.8
6:43:37	1.03	5.15	9325.29	902.33	…	2.01	17.85
6:43:39	1.03	5.15	9325.29	902.33	…	2.01	17.85
6:43:41	1.04	5.19	9398.6	910.99	…	2.03	17.78
6:43:43	1.04	5.19	9398.6	910.99	…	2.03	17.78
6:43:45	1.04	5.19	9398.6	910.99	…	2.03	17.78
6:43:47	1.05	5.21	9435.25	893.03	…	2.03	17.78
6:43:49	1.05	5.21	9435.25	833.66	…	2.1	17.82

Figure 8 is a typical slurry transportation system of CSD. The dredging process can be simplified as (1) the winches control the swing speed of the cutter head, (2) the cutter head cut the soil into the slurry, and (3) the pumps then suck the slurry away. Therefore, swinging, cutting, and pumping are the three main processes of CSD construction and are the main factors of slurry concentration, and they are also the basis of the principle of the proposed index selection method.

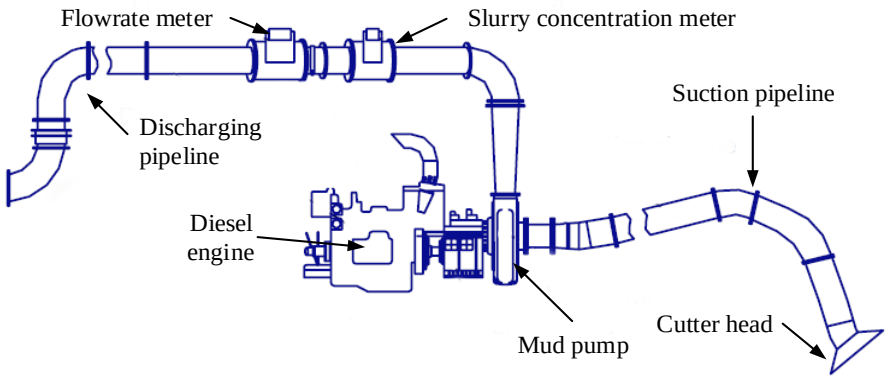


Figure 8. Slurry transportation system[14].

Considering that the indices of different CSDs are not quite the same, we proposed to use "index class" to select indices. Four index classes are proposed: (1) swing-related indices, (2) cutter-related indices, (3) pump-related indices, and (4) time-lagged slurry concentration. Each class contains several indices, as shown in Table 2.

Table 2. Details of index class

Index class	Indices
Swing-related	Swing speed; Swing direction; Ladder angle
Cutter-related	Motor power; Cutting angle;
Pump-related	Vacuum; Drive power of shaft; Power; Rotate speed; Motor/Diesel power; Motor/Diesel speed
Time-lagged concentration	Time-lagged concentration

The reason why the time-lagged slurry concentration is selected is: although there is a delay in the concentration measuring, its value will not significantly larger or smaller than the real value. In the real construction process, the time-lagged value is still an essential reference to operators. It should also be noted that the numbers of the pumps of different CSDs are different, some CSDs only have one underwater pump, while some CSDs have one underwater pump and 1~2 carriage pumps, and all the pump-related indices of all the pumps are required for the slurry concentration prediction.

4.3. Ensemble learning

4.3.1. Structure of the algorithm

Ensemble learning [31-33] is to combine multiple meta-learners (algorithms) into a stronger learner by a particular strategy, as shown in Figure 9. The meta-learners can be any algorithms, such as Decision

Tree, Support Vector Machine, Artificial Neural Network, or even other ensemble learning algorithms. Lots of studies have shown that by combing the meta-learners together, an ensemble learning will have a performance than its meta-learners. These three key points to establishing an ensemble learning algorithm: (1) choosing the meta-learners, (2) determining the sampling strategy, and (3) determining the combining strategy.

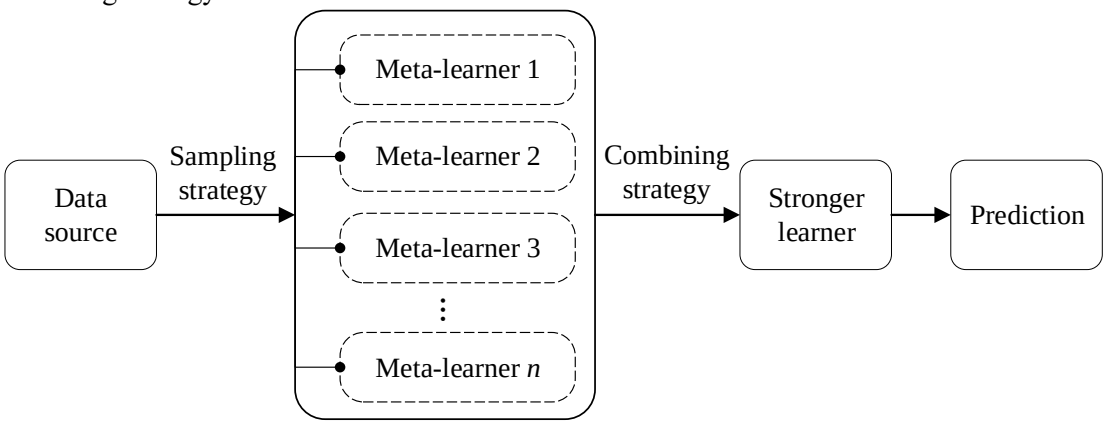


Figure 9. Structure of ensemble learning.

In this research, we used five Bayesian ridge learners and five Gradient Boosting Decision Tree (GBDT) learners as the meta-learners, and took Bagging as the sampling strategy, and used the weighted voting method as the combining strategy.

Bagging is the abbreviation of Bootstrap aggregating. According to the principle of Bagging, we train each meta-learner by random sampling from the training set with replacement. The numbers of samples in each random sampling are the same. After that, there will be ten well-trained learners. For a new sample, ten predictions will be made by the ten learners. The final prediction can be made by the formula:

$$p = \sum_{i=1}^{n} w_i p_i \tag{5}$$

where, p_i is the prediction of the ith meta-learner, w_i is the weight of the ith meta-learner, and p is the final prediction.

4.3.2. Meta-algorithm 1: Bayesian ridge

Bayesian ridge (BR) algorithm is a kind of Bayesian linear regression [34] and is based on the Bayesian inference. In the principle of Bayesian linear regression, the parameters of the linear model are regarded as random variables, and the posterior of the parameters can be calculated with prior knowledge. Supposing that $\boldsymbol{X}=\{x_1, x_2, \ldots, x_N\} \in \mathbf{R}^N$ and $\boldsymbol{y}=\{y_1, y_2, \ldots, y_N\}$ are the training set, then the Bayesian linear regression model is:

$$f(\boldsymbol{X}) = \boldsymbol{X}^T \boldsymbol{w}, \quad \boldsymbol{y} = f(\boldsymbol{X}) + \varepsilon \tag{6}$$

where, $\boldsymbol{w}$ is the weights (or parameters), ε is the residual. It is assumed that the residual obeys the Normal distribution, and the variance of the residual obeys the Inverse-Gamma distribution:

$$\begin{cases} p(\varepsilon) = \mathrm{N}(\varepsilon \mid \mu_n, \sigma_n^2) \\ \sigma_n^2 = \text{Inv-Gamma}(\sigma_n^2 \mid a, b) \end{cases} \tag{7}$$

where the mean of ε, μ_n, and the parameters, (a, b), should be determined by the prior knowledge. Generally, the μ_n is set as 0.

Because the $\boldsymbol{w}$ is independent of $\boldsymbol{X}$ and σ_n^2, the posterior of $\boldsymbol{w}$ can be derived formula (8) by the Bayes' theorem.

$$p(\boldsymbol{w} \mid \boldsymbol{X}, \boldsymbol{y}, \sigma_n^2) = \frac{p(\boldsymbol{y} \mid \boldsymbol{X}, \boldsymbol{w}, \sigma_n^2)\, p(\boldsymbol{w})}{p(\boldsymbol{y} \mid \boldsymbol{X}, \sigma_n^2)} \tag{8}$$

where, $p(\boldsymbol{y} \mid \boldsymbol{X}, \boldsymbol{w}, \sigma_n^2)$ is the likelihood, and is determined by the linear regression model; $p(\boldsymbol{y} \mid \boldsymbol{X}, \sigma_n^2)$ is the marginal likelihood of $\boldsymbol{y}$, and only related to the training set $\boldsymbol{X}$. Our goal is to maximize the likelihood. Generally, there are three solutions, including the Maximum A Posterior estimation (MAP), conjugate prior method, and numerical method. In this research, the MAP is used, as

follows:

The prior knowledge is supposed:

$$p(\boldsymbol{w}) = \mathrm{N}\left(\boldsymbol{w} \mid 0, \sigma_w^2\right) \tag{9}$$

Because the marginal likelihood is independent of $\boldsymbol{w}$, maximizing the posterior is equivalent to maximizing the product of the likelihood and the prior:

$$\arg\max_{\boldsymbol{w}} p\left(\boldsymbol{w} \mid \boldsymbol{X}, \boldsymbol{y}, \sigma_n^2\right) \Leftrightarrow \arg\max_{\boldsymbol{w}} p\left(\boldsymbol{y} \mid \boldsymbol{X}, \boldsymbol{w}, \sigma_n^2\right) p(\boldsymbol{w}) \tag{10}$$

Then the following formula can be derived:

$$\begin{aligned}
&\arg\max_{\boldsymbol{w}} p\left(\boldsymbol{y} \mid \boldsymbol{X}, \boldsymbol{w}, \sigma_n^2\right) p(\boldsymbol{w}) \Leftrightarrow \arg\max_{\boldsymbol{w}} \log\left[p\left(\boldsymbol{y} \mid \boldsymbol{X}, \boldsymbol{w}, \sigma_n^2\right)\right] + \log\left[p(\boldsymbol{w})\right] \\
&\log\left[p\left(\boldsymbol{y} \mid \boldsymbol{X}, \boldsymbol{w}, \sigma_n^2\right)\right] + \log\left[p(\boldsymbol{w})\right] \\
&= \log\left[\frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{\left|\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{w}\right|^2}{2\sigma_n^2}\right)\right] + \log\left[\frac{1}{\sqrt{2\pi}\sigma_w} \exp\left(-\frac{\boldsymbol{w}^T\boldsymbol{w}}{2\sigma_w^2}\right)\right] \\
&= -\frac{1}{2\sigma_n^2}\left(\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{w}\right)^T\left(\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{w}\right) - \frac{1}{2\sigma_w^2}\boldsymbol{w}^T\boldsymbol{w} \\
&= -\frac{1}{2\sigma_n^2}\left\|\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{w}\right\|^2 - \frac{1}{2\sigma_w^2}\|\boldsymbol{w}\|^2
\end{aligned} \tag{11}$$

Because the coefficients of formula (11) are all negative, the maximization problem can be transformed into a minimization problem, and then the $\boldsymbol{w}$ can be calculated, as follows:

$$\begin{aligned}
&\arg\min_{\boldsymbol{w}} \left\|\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{w}\right\|^2 + \lambda\|\boldsymbol{w}\|^2, \quad \lambda = \frac{\sigma_n^2}{\sigma_w^2} \\
&\Rightarrow \boldsymbol{w} = \left(\boldsymbol{X}^T\boldsymbol{X} + \lambda\boldsymbol{I}\right)^{-1}\boldsymbol{X}^T\boldsymbol{y}
\end{aligned} \tag{12}$$

where, λ is the ratio of the variance of the residual and the variance of the weights, and can be calculated by the hyper-parameters; $\boldsymbol{I}$ is a unit matrix.

4.3.3. Meta-algorithm 2: Gradient Boosting Decision Tree

Gradient Boosting Decision Tree (GBDT) itself is a kind of ensemble learning algorithm [35]. Its main idea is to respectively train its meta-learners (regression tree) in a stage-wise fashion. Supposing that the training set is $\boldsymbol{T}=\{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, and the loss function is $\mathrm{L}(y, f(x))$, then a GBDT model can be established with the following steps:

(1) Initialization. Calculate a value, c, that can minimize the loss function $\mathrm{L}(\boldsymbol{y}, c)$. The first regression tree can be determined:

$$f_0(x) = \arg\min_c \sum_{i=1}^{N} L(y_i, c) \tag{13}$$

(2) Supposing that M is the number of the regression trees, then the mth (m=1, 2, …, M) regression tree can be trained by step (3)~(6).

(3) For the ith training sample (i=1, 2, …, N), calculate the negative gradient of the loss function of the current (the (m-1)th) regression tree:

$$r_{mi} = -\left[\frac{\partial L\left(y_i, f\left(x_i\right)\right)}{\partial f\left(x_i\right)}\right]_{f(x)=f_{m-1}(x)} \tag{14}$$

(4) Make a new dataset, $\{(x_1, r_{m1}), (x_2, r_{m2}), \ldots, (x_N, r_{mN})\}$, then use this dataset to determine the R_{mj} (j=1, 2, …, J). R_{mj} is the range of the leaf nodes of the mth regression tree, and J is the number of the leaf nodes.

(5) Calculate the c_{mj} that can minimize the loss function. This step is similar to step (1).

$$c_{mj} = \arg\min_c \sum_{x_i \in R_{mj}} L\left(y_i, f_{m-1}\left(x_i\right) + c\right) \tag{15}$$

(6) Update the $f(x)$:

$$f_m(x) = f_{m-1}(x) + \sum_{j=1}^{J} c_{mj} I\left(x \in R_{mj}\right) \tag{16}$$

(7) The final GBDT model can be described:

$$F(x)=\sum_{m=1}^{M}\sum_{j=1}^{J}c_{mj}I\left(x\in R_{mj}\right) \tag{17}$$

4.4. Algorithms for comparison

In related researches such as the prediction works of TBM parameters [21-26], algorithms such as RF, LSTM, SVM are usually used. In this research, six representative algorithms, including DNN, LSTM, SVM, RF, and GBDT, are tested for comparison.

(1) Deep Neural Network (DNN). DNN is attracting more and more attention in recent years and has been proven to have a tremendous non-linear mapping ability [36, 37]. Compared with traditional Artificial Neural Network (ANN), DNN has more hidden layers and more complex structures. In this study, we designed a DNN with eight hidden layers, as shown in Figure 10.

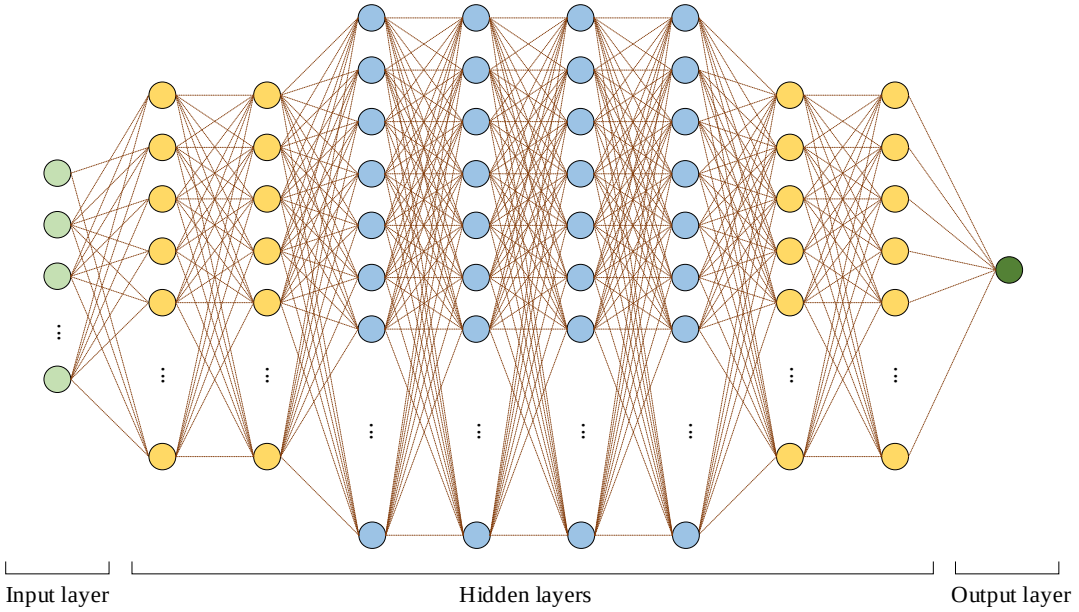


Figure 10. Structure of the DNN.

In the DNN model, the input layer were the indices shown in Table 2, and the output was the slurry concentration. The numbers of the neural cells in each hidden layers were 100, 100, 200, 200, 200, 200, 100, 100. We used the Relu function as the activation function. The learning rate is 0.003.

(2) Long Short Time Memory (LSTM). LSTM is a kind of recurrent neural network (RNN) and is good at solving time sequence problems [38, 39]. A typical LSTM model is shown in Figure 11.

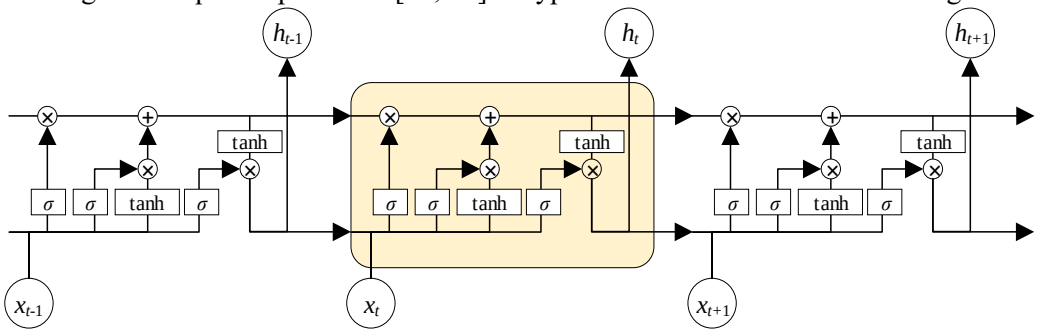


Figure 11. Structure of LSTM.

In this research, we designed an LSTM model with two hidden layers, and each hidden layer contained 50 cells. The time step was five. The learning rate was 0.0001. The batch size was 64. The number of epoch was 100.

(3-6) Support Vector Machine (SVM), Random Forest (RF), GBDT, and Bayesian ridge. SVM and RF are two of the most widely used algorithms, so they are not repeated here. The details of them can be found in the two publications [40, 41]. The GBDT and the Bayesian ridge algorithm have been described in Section 4.3

4.5. Evaluation methods

To quantitatively evaluate the performances of the algorithms, two evaluation indices were used in this study.

(1) Goodness of fit (R^2). The R^2 is calculated:

$$R^2 = \frac{(\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{y}^{'}))^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\hat{y}_i - \bar{y}^{'})^2} \tag{18}$$

(2) Mean square error (MSE). The MSE is calculated:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{19}$$

where, n is the number of samples, y_i is the target of the ith sample and $\hat{y}_i$ is the prediction of the ith sample.

5. CASE STUDY

5.1. Data collection

Totally seven days' CSD monitoring data were collected from a dredging construction of Tianjin, China. The sampling frequency of the monitoring system was once every two seconds. In total, there were 112,637 samples. Figure 12 shows the slurry concentrations of the CSD at different dredging positions in the dredging area.

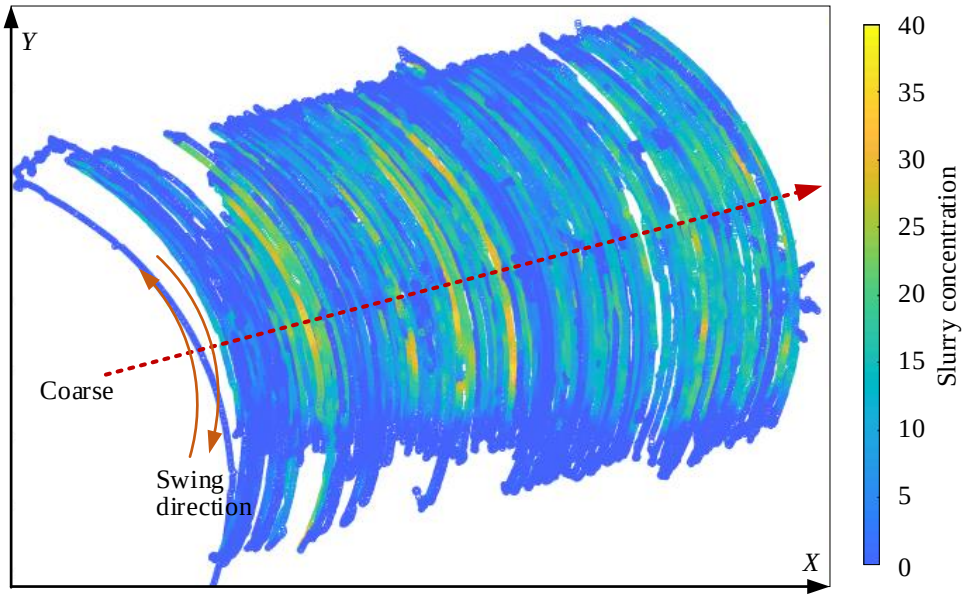


Figure 12. Distribution of the slurry concentration in the dredging area.

The control system of the CSD can be simplified as Figure 13, and 715 indices were monitored. Compared with the typical slurry transportation system, this CSD had three pumps, including an underwater pump and two carriage pumps. According to the proposed index selection method (Table 2), 23 indices were selected, including two swing-related indices, two cutter-related indices, 18 pump-related indices, and the time-lagged slurry concentration. It should be noted that the "ladder angle" was not selected because the data quality of this index was poor.

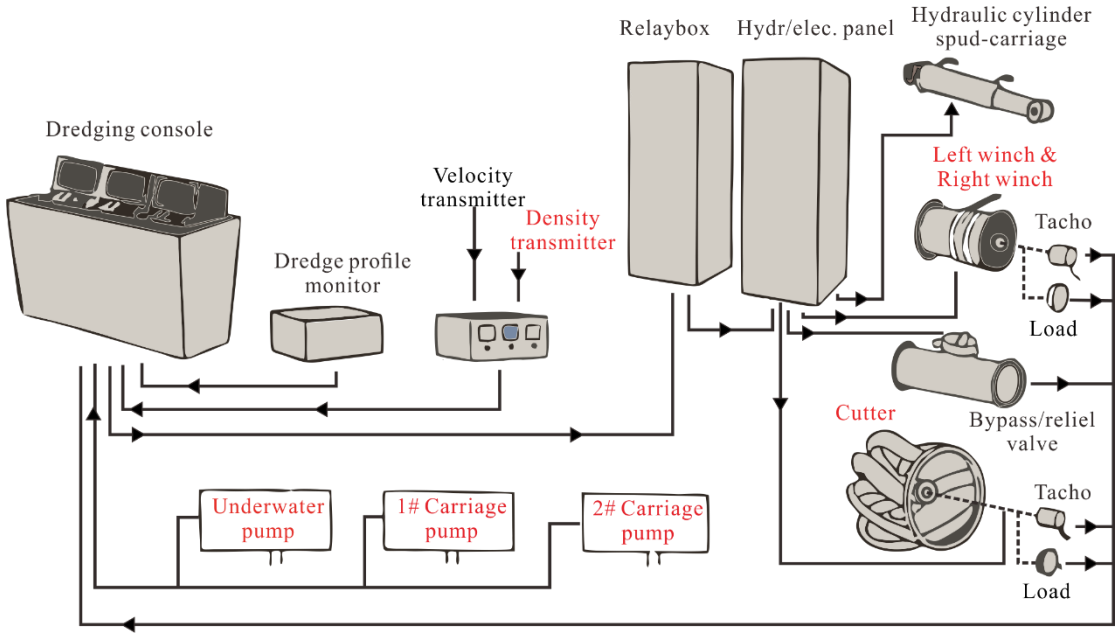


Figure 13. Control system of the CSD.

5.2. Experimental results

The time lag effect of the slurry concentration was first eliminated using the method of Section 4.1.1, as shown in Figure 14. For a further explanation, the vacuum data are also plotted in Figure 14. It can be seen that the real slurry concentration changes in sync with the vacuum. When the vacuum decrease, more mud will be sucked into the pipeline, so the slurry concentration will increase – and vice versa.

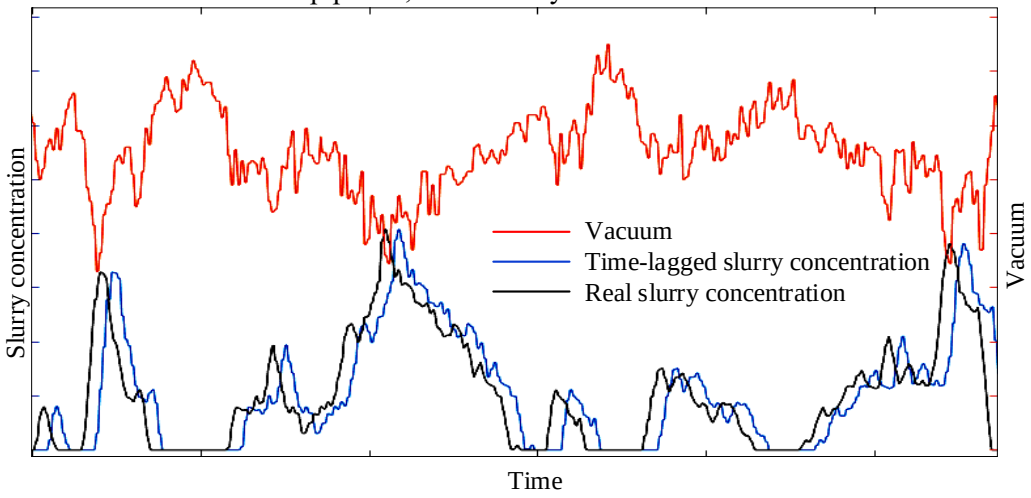
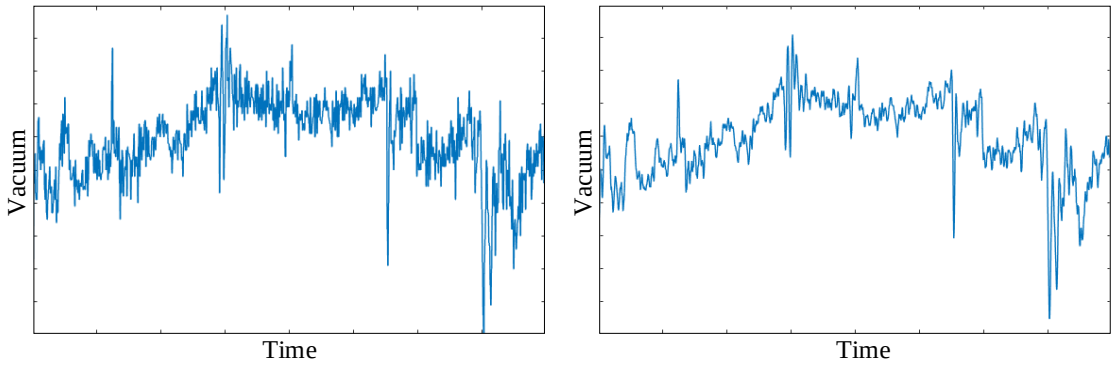


Figure 14. Comparison between the time-lagged slurry concentration and the real values.

All the normal construction data were then selected using the method proposed in Section 4.1.2. After that, the data of seven indices were filtered using the method proposed in Section 4.1.3, including the vacuum of the underwater pump, the drive power of the shaft of the underwater pump, the rotate speed of the underwater pump, the motor speed of the underwater pump, the motor power of the underwater pump, the cut angle, and the time-lagged slurry concentration. Figure 15 shows the raw data and the filtered data of the vacuum of the underwater pump.



(a) Raw vacuum data (b) Filtered vacuum data

Figure 15. Example of data filtering.

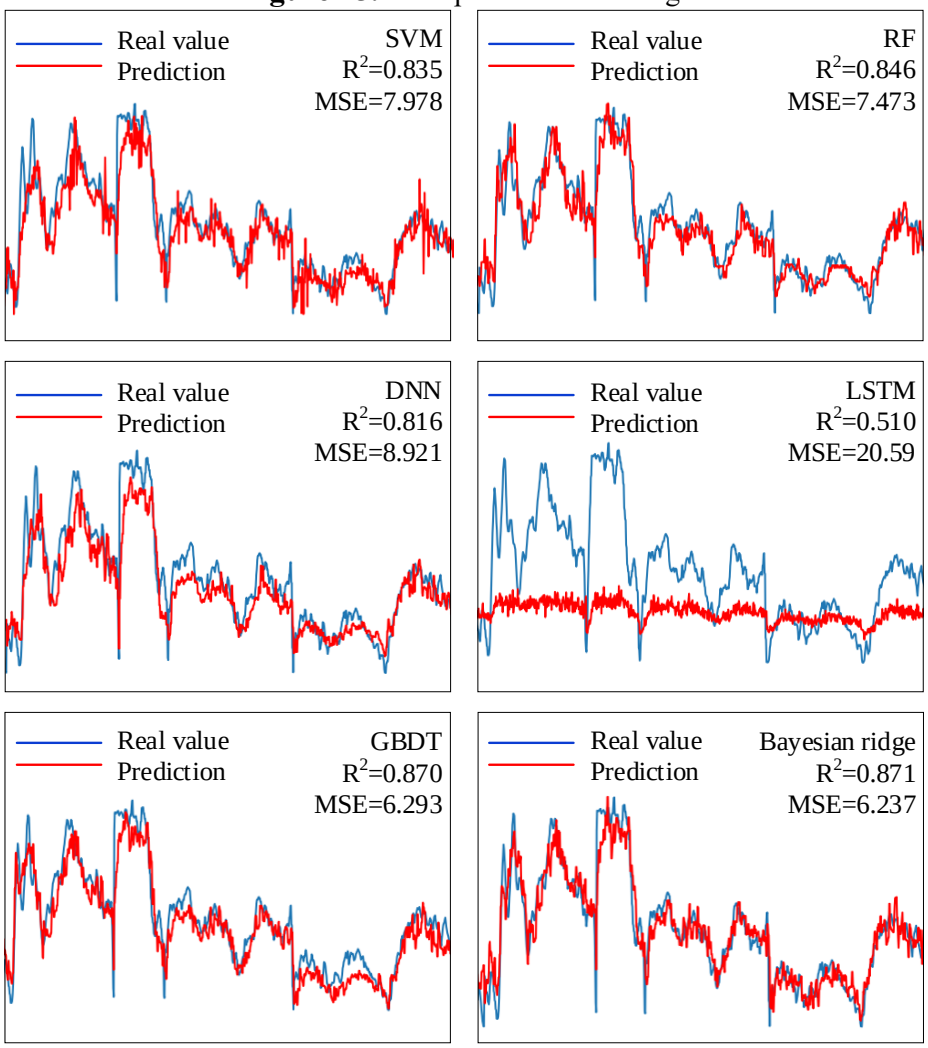


Figure 16. Results of SVM, RF, DNN, LSTM, GBDT, and BR (the plots just shows a part of the result).

After data preprocessing, we used the proposed ensemble learning algorithm and another six algorithms to establish the relationship between the slurry concentration and the selected 23 indices. Three days' data were used to train the algorithms, and the last four days' data were used to test them. A part of the test results is shown in Figure 16-17. The details of the predictions results of all the tested algorithms are list in Table 3.

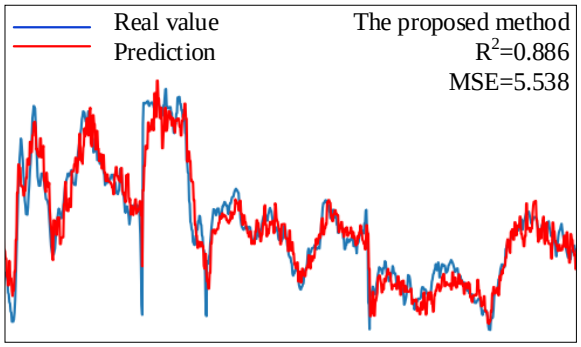


Figure 17. Result of the proposed ensemble learning algorithm (the plots just shows a part of the result).

Table 3. Structure of the logging file of the CSD monitoring system

	SVM	RF	DNN	LSTM	GBDT	BR	The proposed method
R^2	0.835	0.846	0.816	0.510	0.870	0.871	0.886
MSE	7.987	7.473	8.921	20.56	6.293	6.237	5.538

5.3. Discussion

5.3.1. About the selection of meta-learner

Although it seems from the plots that the RF, GBDT, and Bayesian ridge can also fit the data well, the R^2 and MSE show that the proposed ensemble learning algorithm is significantly better than the other algorithms. DNN and LSTM have the worst performances. Actually, at the beginning, DNN was the first choice for the study because the problem was complex and the number of data was large enough, and is thought to have the best performance as long as the network was well-designed and the parameters were suitable. However, numerous trials showed that DNN was not suitable for this task, and its performance was even not as good as SVM and RF. Meanwhile, the linear regression algorithms, such as Bayesian ridge and logistic regression, were found to have good effects, although their principles were not complicated at all.

This discovery made us realize that the relationship between the slurry concentration and the other selected indices should be closer to a linear relationship than a non-linear relationship. Therefore, when designing the ensemble learning algorithm, half of the meta-learners were set as the Bayesian ridge algorithms to specially learn the linear part of the relationship, and finally, it produced the best results.

5.3.2. About the index class

At the end of the research, we studied the importance of the four index class proposed. Figure 18 shows the MSEs when the indices of one of a certain index class were not used. It can be seen that the time-lagged slurry concentration has the most significant importance, and the MSE will be three times as large as the best result if this index is not used. The Pump-related indices also have a significant effect on the regression.

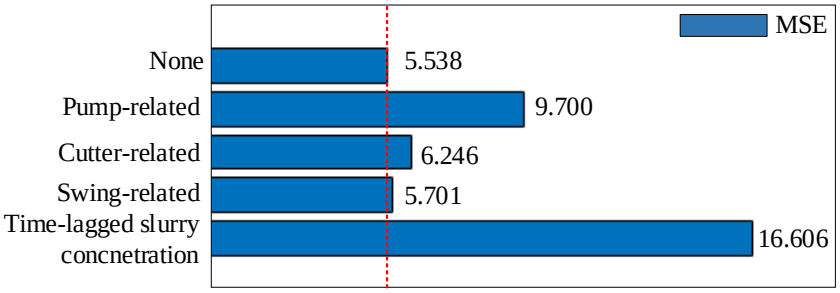


Figure 18. Important analysis of index classes.

Furtherly, the variance inflation factor (VIF) was used to check the multicollinearity among the indices, and the indices with low VIF values (<10) are listed in Table 4. After that, we just used the six

listed indices to test the predictive effect, and the R^2 is 0.839, 5% lower than the best result, 0.886; the MSE is 7.80, 40% higher than the best result, 5.538. Such a result indicates: (1) there is at least one key index in every index class, i.e., swing direction in the swing–related class, motor power in the cutter-related class, vacuum in the pump-related class, diesel speed in the pump-related class, and time-lagged concentration; (2) low VIF indices may not have bigger impacts on the prediction than high VIF indices, e.g., the swing-direction has the lowest VIF value (Table 3), while the swing-related index class contributes the least among the four classes (Figure 18); (3) according to the R^2 and MSE, although the other indices are with high VIF values, they are important to improve the predictive effect, meaning that the proposed algorithm is effective in extracting valuable information from different indices even if the indices have high multicollinearity with others.

Table 4. Indices with low VIF values (<10)

Index	Class	VIF
Swing direction	Swing-related	1.06
Motor power	Cutter-related	5.33
Vacuum (underwater pump)	Pump-related	2.31
Diesel speed (1# Carriage pump)	Pump-related	2.41
Diesel speed (2# Carriage pump)	Pump-related	4.48
Time-lagged concentration	Time-lagged concentration	7.75

Index selection is one of the most challenging problems in this study because of the large number of the indices. Moreover, different CSDs have different indices, so the indices that appropriate for the slurry prediction of one CSD may not exist in another CSD. Therefore, we suggest taking the proposed "index class" as a reference to select indices. Actually, another case study was also carried out in the experiment to test the proposed method further. The data were collected from another CSD, and the indices of this CSD were less than the CSD in the first case. Totally 11 indices were then selected according to Table 2. Figure 19 shows that the proposed method works well, as shown in. However, the result is worse than in the first case. It mainly because the number of indices was less, and the number of data was not enough (less than one days' data). Overall, this method are with general applicability.

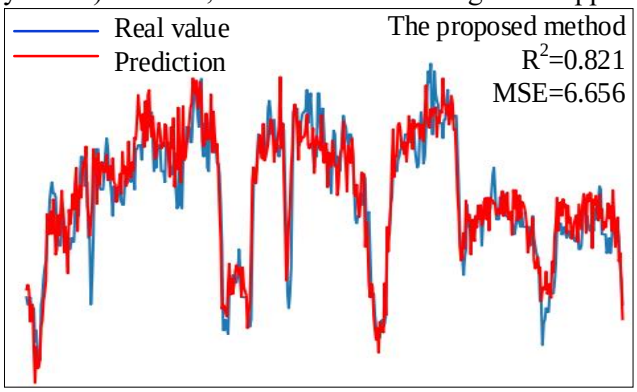


Figure 19. Results of another case study.

5.3.3. About data filtering

Filtering is a standard method in signal analysis. In most instances, data filtering is helpful for analysis. However, the poignant matter needs to be aware of is whether or not data filtering is appropriate. In this research, we think that it depends on the roles of the data: the training set can be filtered, while the test set cannot be filtered. Training data are used to establish the algorithm, and they are the given data. However, the test set is the unknown data, and they should be regarded as the real-time data of a real construction process. For the S-G filter, the value of a point is modified according to the points before and after it. But during a construction process, it is impossible to get a series of future value to filter the current value. In our research, all the data of the test set were not preprocessed, as shown in Figure 20.

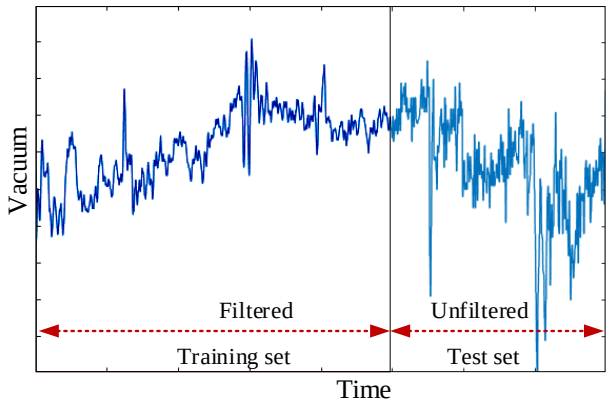


Figure 20. Illustration of which data can be filtered.

There is also a kind of method named real-time filterings, such as the Forward Linear Prediction (FLP) and the Kalman Filter. However, the FLP needs a control function, and the Kalman Filter needs a covariance matrix, both of which are not easy to be determined, and more works are needed to test the applicability of them on the research.

6. CONCLUSION

In this research, an ensemble learning-based method is proposed to predict the real-time slurry concentration of CSD dredging construction. There are three main innovations: at first, the preprocessing method for the CSD monitoring data is proposed; then we put forward the concept of "index class" to select the important indices; after that, an ensemble learning algorithm is presented to fit the relationship between the slurry concentration and the indices of the index classes. The results show that the proposed method is effective. The R^2 and MSE of the ensemble learning algorithm are 0.886 and 5.538, respectively. Such a fitting effect is better than almost all the commonly used regression algorithms, even including the DNN and LSTM. Besides, our method is with general applicability. We also test our method with the data of another CSD, as discussed in Section 5.3.2. It shows that the R^2 can reach to 0.82 and the MSE is less than 7, even though there are only several hours' data for training. Some important conclusions are also drawn from the research: (1) linear regression meta-models play an important role in the proposed algorithm; (2) the pump-related indices are the most significant indices except for the time-lagged slurry concentration; (3) the proposed algorithm is universal for different types of CSDs.

In conclusion, the proposed method can help the CSD operators to obtain the slurry concentration immediately; thus, it can help to improve the stationarity and production efficiency of dredging construction.

ACKNOWLEDGEMENTS

This work was supported by the Tianjin Science Foundation for Distinguished Young Scientists of China [Grant no. 17JCJQJC44000] and the National Natural Science Foundation of China [Grant no. 51879185].

REFERENCES

[1] R.E. Turner, "Fundamental of Hydraulic Dredging", 2d ed. ASCE, New York, 1996.
[2] W.J. Vlasblom, "Designing dredging equipment, Delft: Delft University of Technology", 2003.
[3] R.M.C. De Keyser, L. de Coen, P. Verdiere, "Multi-Microprocessor simulation of a Cutter Suction Dredging Ship", in: Digital Computer Applications to Process Control. Pergamon, pp. 307-312, 1986.
[4] Y. Jiang, X. Chen, J. Tan, "Comparison and analysis on ship types of small and medium size trailing suction hopper dredger", in: 2011 International Conference on Remote Sensing, Environment and Transportation Engineering, IEEE, pp. 181-184, 2011.
[5] J.Z. Tang, Q.F. Wang, "Online fault diagnosis and prevention expert system for dredgers, Expert Systems with Applications", vol. 34, no. 1, pp. 511-521, 2008.
[6] F.S. Ni, L.J. Zhao, L. Gu, S. Jiang, L.N. Qian, L.Q. Xu, K.J. He, "Simulation of dredging processes of a cutter suction dredger", in: Audio Language and Image Processing (ICALIP), 2010 International

Conference on IEEE, pp. 628-632, 2010.
[7] J. Henriksen, R. Randall, S. Socolofsky, “Near-field resuspension model for a cutter suction dredge”, Journal of Waterway, Port, Coastal, and Ocean Engineering, vol. 138, no. 3, pp. 181-191, 2011.
[8] M. Zhang, S. Fan, H. Zhua, S. Han, “Numerical Simulation of Solid-Fluid 2-Phase-Flow of Cutting System for Cutter Suction Dredgers”, Polish Maritime Research, vol. 25, no. s2, pp. 117-124, 2018.
[9] M.C. Li, R. Kong, S. Han, G.P. Tian, L. Qin, “Novel method of construction-efficiency evaluation of cutter suction dredger based on real-time monitoring data”, Journal of Waterway Port Coastal and Ocean Engineering, vol. 144, no. 6, pp. 05018007, 2018.
[10] P. Yue, D.H. Zhong, Z. Miao, J. Yu, “Prediction of dredging productivity using a rock and soil classification model”, Journal of Waterway, Port, Coastal, and Ocean Engineering, vol. 141, no. 4, pp. 06015001, 2015.
[11] R. Setiwan, “Parametric analysis on off-shore dredging process using cutter suction dredgers”, ASEAN Engineering Journal, vol. 6, no. 1, pp. 37-46, 2015.
[12] J. Yang, F. Ni, C. Wei, “A BP neural network model for predicting the production of a cutter suction dredger”, in: The 3rd International Conference on Material, Mechanical and Manufacturing Engineering (IC3ME 2015), Atlantis Pres, 2015.
[13] S.A. Miedema, “Automation of a Cutter Dredge, Applied to the Dynamic Behaviour of a Pump/Pipeline System”, in: Proc. WODCON VI, Kuala Lumpur, Malaysia, 2001.
[14] J.Z. Tang, Q.F. Wang, Z.Y. Bi, “Expert system for operation optimization and control of cutter suction dredger”, Expert Systems with Applications, vol. 34, no. 3, 2180-2192, 2008.
[15] J. Tang, Q. Wang, T. Zhong, “Automatic monitoring and control of cutter suction dredger”, Automation in Construction, vol. 18, no. 2, pp. 194-203, 2009.
[16] Y. Ye, M. Bai, Z. Zhang, W. Qiu, R. Li, “A design of dredger cutter motor synchronous speed control system based on ADRC”, in: Control and Decision Conference (CCDC), IEEE, pp. 1646-1650. 2016.
[17] S. Jiang, F. Ni, L. Qian, “Swing Process Model Design of a Cutter Suction Dredger Based on RBF-ARX Model”, in: IOP Conference Series: Materials Science and Engineering, IOP Publishing, vol. 428, no. 1, p. 012018, 2018.
[18] S. Bai, M.C. Li, R. Kong, S. Han, H. Li, L. Qin, “Data mining approach to construction productivity prediction for cutter suction dredgers”, Automation in Construction, vol, 105, pp. 102833, 2019.
[19] Q. Ma, S. Liu, X. Zhao, “PCA-NARX Time Series Prediction Model of Surface Settlement during Excavation of Deep Foundation Pit”, IOP Conference Series: Earth and Environmental Science, vol. 560, no. 1, IOP Publishing, 2020.
[20] Q. Liu, X. Wang, X. Huang, X. Yin, “Prediction model of rock mass class using classification and regression tree integrated AdaBoost algorithm based on TBM driving data”, Tunnelling and Underground Space Technology, vol. 106, p. 103595, 2020.
[21] H. Chen, C. Xiao, Z. Yao, H. Jiang, T. Zhang, Y. Guan, “Prediction of TBM Tunneling Parameters through an LSTM Neural Network”, in: 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO) (pp. 702-707). IEEE, 2019, Dec.
[22] P. Zhang, R.P. Chen, H.N. Wu, “Real-time analysis and regulation of EPB shield steering using Random Forest”, Automation in Construction, vol. 106, p. 102860, 2019.
[23] X. Gao, M. Shi, X. Song, C. Zhang, H. Zhang, “Recurrent neural networks for real-time prediction of TBM operating parameters”, Automation in Construction, vol. 98, pp. 225-235, 2019.
[24] L.J. Jing, J.B. Li, Y. Chen, S. Chen, N. Zhang, X.X. Peng, “A case study of TBM performance prediction using field tunnelling tests in limestone strata”, Tunnelling and Underground Space Technology, vol. 83, pp. 364-372, 2019.
[25] B. Gao, R. Wang, C. Lin, X. Guo, B. Liu, W. Zhang, “TBM penetration rate prediction based on the long short-term memory neural network”, Underground Space, in press, 2020.
[26] S. Leng, J.R. Lin, Z.Z. Hu, X. Shen. “A Hybrid Data Mining Method for Tunnel Engineering Based on Real-Time Monitoring Data From Tunnel Boring Machines”, IEEE Access, vol. 8, pp. 90430-90449, 2020.
[27] V. Matousek, “Solids transportation in a long pipeline connected with a dredge”, Terra et Aqua, vol. 62, pp. 3-11, 1996.
[28] I.G. Currie, “Fundamental mechanics of fluids”, McGraw-Hill, New York, 1974.
[29] A. Savitzky, M.J. Golay, “Smoothing and differentiation of data by simplified least squares procedures”, Analytical chemistry, vol. 36, no. 8, pp. 1627-1639, 1964.
[30] Q. Wang, J. Tang, “Research on expert system for dredging production optimization”, in: 2006 6th

World Congress on Intelligent Control and Automation, IEEE, vol. 1, pp. 2526-2530, 2006.
[31] T.G. Dietterich, “Ensemble learning. The Handbook of Brain Theory and Neural Networks”. 2nd ed. Cambridge, MA: MIT Press, pp. 405–408, 2002.
[32] X. Wang, M. You, Z. Mao, P. Yuan, “Tree-structure ensemble general regression neural networks applied to predict the molten steel temperature in ladle furnace”, Advanced Engineering Informatics, vol. 30, no. 3, pp. 368-375, 2016.
[33] M. Zounemat-Kermani, D. Stephan, M. Barjenbruch, R. Hinkelmann, “Ensemble data mining modeling in corrosion of concrete sewer: A comparative study of network-based (MLPNN & RBFNN) and tree-based (RF, CHAID, & CART) models”, Advanced Engineering Informatics. 43 (2020) 101030.
[34] A.B. Chan, N. Vasconcelos, “Counting people with low-level features and Bayesian regression”, IEEE Transactions on image processing, vol. 21, no. 4, pp. 2160-2177, 2011.
[35] C. Ding, D. Wang, X. Ma, H. Li, “Predicting short-term subway ridership and prioritizing its influential factors using gradient boosting decision trees”, Sustainability, vol. 8, no. 11, pp. 1100, 2016.
[36] J. Schmidhuber, “Deep learning in neural networks: An overview”, Neural networks, vol. 61, pp. 85-117, 2015.
[37] Q. Fang, H. Li, X. Luo, L. Ding, T.M. Rose, W. An, Y. Yu, “A deep learning-based method for detecting non-certified work on construction sites”, Advanced Engineering Informatics, vol. 35, pp. 56-68, 2018.
[38] X. Ma, Z. Tao, Y. Wang, H. Yu, Y. Wang, “Long short-term memory neural network for traffic speed prediction using remote microwave sensor data”, Transportation Research Part C: Emerging Technologies, vol. 54, pp. 187-197, 2015.
[39] A.J. Smola, B. Schölkopf, “A tutorial on support vector regression”, Statistics and computing, vol, 14, no. 3, pp. 199-222, 2004.
[40] B. Zhong, X. Xing, H. Luo, Q. Zhou, H. Li, T. Rose, W. Fang, “Deep learning-based extraction of construction procedural constraints from construction regulations”, Advanced Engineering Informatics, vol. 43 pp. 101003, 2020.
[41] L. Breiman, “Random forests”, Machine learning, vol. 45, no. 1, pp. 5-32, 2001.