Study on Characteristic Spectrum and Multiple Classifier Fusion With Different Particle Size in Marine Sediments

XUEYING LI and PINGPING FAN

1 Institute of Oceanographic Instrumentation, Qilu University of Technology (Shandong Academy of Sciences), Qingdao 266061, China
2 School of Geosciences, China University of Petroleum (Huadong), Qingdao 266580, China
3 National Engineering and Technological Research Center of Marine Monitoring Equipment, Qingdao 266061, China
4 Shandong Provincial Key Laboratory of Ocean Environmental Monitoring Technology, Qingdao 266061, China

Corresponding author: Pingping Fan (fanpp_sdioi@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 3130041; and in part by the Natural Science Foundation of Shandong Province, China, under Grant ZR2018LD007, Grant ZR2017BB037, and Grant ZR2019PD004.

ABSTRACT Marine sediments record much information of the ecological process, which highly correlated with global and local environmental change. Particle size and its distributions of sediments indicate different ecological functions, thus are the key questions in marine ecology. The analysis method is tedious and laborious, which is not conducive for in-situ monitoring. Here, a spectral analysis was explored using surface sediments sampled in the intertidal zone of Dongdayang village, Qingdao, China. These samples were dried and sieved to pass through the mesh size of 0.3 mm, 0.2 mm, 0.1 mm, and 0.075 mm, respectively. Then, four types of subsamples were collected with the particle size of 0.3-0.2 mm, 0.2-0.1 mm, 0.1-0.075 mm, and < 0.075 mm, respectively. The visible and near infrared reflectance spectra (226-975nm) of these subsamples with different particle size were measured. Results showed that there was a negative correlation between the spectral reflectance and the particle size. And the characteristic spectra for particle size classification were 926-975nm. These particle size were classified by the support vector machine algorithm. The classification accuracy for the calibration set and validation set was 100% and 89.06%, respectively. Furthermore, the fusion classifier was compared with the single classifier. Three spectral bands were selected as the single particle size classifier, that is 226-325nm, 826-925nm and 226-975nm. These three single classifiers were fused by voting method, forming multiple classifier fusion. A fusion classifier was recommended whose validation set had the classification accuracy of 93.75%, better than any single classifier. Multiple classifier fusion is a good tool for searching the characteristic spectra of the chemical and physical parameter in sediments. This method provides a solution for the division of particle size in sediment.

INDEX TERMS Marine sediment, particle size, spectroscopy, classification.

I. INTRODUCTION

Marine sediments are the medium and carrier of interaction among hydrosphere, biosphere and lithosphere [1]. They record much information of the ecological process, which highly correlated with global and local environmental change [2], [3]. Particle size is an important structural feature of sediments and the basis of its classification, indicating the source and environment of sediments. It is the most widely used parameter to study the characteristics and environment of sediments [4]. And its distribution in sediments reveals migration and settlement information of sediments [5]. Particle size analysis is widely used in sediment research, which is to determine the particle size and distribution of sediment components. The granularity analysis of marine sediments is not only the basis of dividing the types of seabed sediments and compiling the type map of seabed sediments, but also the material source and sedimentation of seabed sediments [6], [7]. Therefore, the analysis of different particle sizes of marine sediments is of great significance to the study of sediments.

The associate editor coordinating the review of this manuscript and approving it for publication was Alessandro Pozzebon.
microscopy [8], [9]. The most commonly used methods are sieve analysis and sedimentation. The sieve analysis method is to sieve the samples from coarse to fine through different sieve sizes [10]. The basic principle of sedimentation method is to use particle sedimentation velocity to divide particle distribution [11]. With the development of analytical technology, laser particle analyzer has been developed and widely used [12], [13]. The analysis speed and measurement range have been improved greatly, but it still needs tedious and laborious pretreatment. Furthermore, results from laser particle analyzer are often different from the traditional sedimentation method and sieve analysis method. Cuven et al. [14] used laser particle size analysis to study the relationship between chemical profile and particle size of sediments at macro and micro scales. Menéndez-Aguado et al. [15] carried out power spectrum analysis using laser diffraction spectrum to accurately represent the mathematical model of sediment particle size distribution. Image method is also one of the current particle size analysis methods, which can express the size of each particle and contain particle shape information. However, the data processing is complex and their interpretation is difficult [16]. Yates et al. [17] used the landsat5 TM satellite data to classify the surface sediments. Results showed a more accurate classification for the muddy areas. Kim et al. [18] studied the classification of sediments in xx sites using the UAV Orthophoto Image combined with the field survey (232 samples). Results showed that the overall accuracy was 72.8%.

Visible near infrared spectroscopy (VNIRS) is a fast, simple, nondestructive and green analysis technology [19], [20], thus has been widely used in many fields [21], [22]. The absorbance and reflectance spectra contain a lot of material information, which mainly reflects the types and quantities of O-H, N-H, C-H and other organic functional groups. Varied achievements have been made in the rapid determining many chemical and physical parameters in soils and sediments by VNIRS [23], [24]. Sediment size classification was also explored by VNIRS in limited studies but needs a further research [25].

In this paper, VNIRS was applied to analyze the particle size in the surface sediments of the intertidal zone in Qingdao, China. Two questions would be resolved. Firstly, how would the reflectance spectra of the sediments vary with different particle size? We hypothesized that the reflectance of the sediments increased with the decreased particle size because smaller size particle could produce larger reflectance. Secondly, how could the particle size of the sediments be expressed by the spectra? As other chemical or physical parameters, the particle size of sediments could be modelled by stoichiometry, such as multiple classifier fusion combined the support vector machine algorithm. Here, through the introduction of spectrum technology, we can find the characteristic wavelength to characterize the particle size characteristics, and establish the classification model of different particle sizes, so as to achieve the rapid division among different particle sizes of sediment. It provides a new method for rapid particle size analysis of sediments.

The outline of the paper is as follows. The model method and multiple classifiers fusion explained in Section II. Section III is the acquisition of sediments samples and data. Section IV and V presents and discusses the experimental studies that demonstrate the performances of the method. Finally, conclusion is given in Section VI.

II. THEORETICAL ANALYSIS
A. MODELING METHOD
In this paper, support vector machine algorithm (SVM) was used as the modeling method used. The core idea of SVM is to construct an optimal classification hyperplane by mapping vectors into higher dimensional space through kernel function [26], [27]. Two parallel hyperplanes are found on both sides of the hyperplane, which maximizes the distance between two hyperplanes. By constructing a hyperplane $f(x) = ax + b = 0$, where $a$ is the normal vector of the classification plane and $B$ is the offset of the classification plane. The classification function constructed is $f(x) = ax + b$. The larger the distance between the parallel hyperplanes is, the smaller the error of the classifier is. The classification function of SVM is as follows:

$$f(x) = \text{sign} \left( (\omega^T \varphi(x) + b^*) \right)$$

where $x_i$ is the support vector, $x$ is the unknown vector, $\text{sign}$ is the sign function.

SVM algorithm is usually divided into the following steps. Firstly, the kernel function is used to transform the sample space into a space that can be linearly separable. Then, the partition line with the largest interval is obtained by maximizing the interval, and the support vector is obtained. Finally, the new samples can be classified and predicted by using the segmentation lines and support vectors.

Radial basis function (RBF) is one of the kernel function of SVM algorithm. In the algorithm, C and g are two important parameters [28], [29]. When RBF function is selected as the kernel function, the parameter $g$ determines the distribution of the data mapped a new feature space implicitly. The function relationship is as follows:

$$K(x_i, x_j) = \exp \left( - \frac{|x_i - x_j|^2}{2 \cdot \delta^2} \right)$$

$$= \exp \left( -g |x_i - x_j|^2 \right)$$

The larger the $g$ is, the finer the classification will be. It attempts to distinguish each sample from other samples. If the $g$ is too large, it will lead to over fitting. The smaller the $g$ is, the rougher the classification will be. If the $g$ is too small, it is easy to cause the data undistinguished and appear under fitting.
The C is the penalty factor. From the perspective of risk, C weighs the empirical risk (the fitting ability to the sample) and the structural risk (the prediction ability to the test sample). After the introduction of C, the objective functions and constraints of SVM optimization are as follows:

\[
\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i \\
\text{subject to } y_i [(\omega x_i) + b] \geq 1 - \xi_i (i = 1, 2, \ldots, n) \\
\xi_i > 0
\]

(3)

\(\xi_i\) is the relaxation factor, which is the classification loss for the sample \(i\). If the classification is correct, it is 0. If the classification has some deviation, it corresponds to a linear value. \(\sum_{i=1}^{n} \xi_i\) is the total error. The goal of optimization is that the smaller the total error is, the better the effect will be. It means that training set classification is more accurate. The optimization direction of \(\|w\|^2\) minimization is to make the interval size \(2/\|w\|\) maximum. The value of penalty factor C weighs the empirical risk and structural risk. The greater the C is, the less empirical risk is, the greater structural risk is, and over fitting is likely to occur. The smaller the C is, the lower the model complexity is, and under fitting is likely to occur.

B. MULTIPLE CLASSIFIERS FUSION

The process of multiple classifier fusion is shown in Figure 1. Different models are established by using different classifiers for samples. Then appropriate fusion methods are selected according to the output data of different models. Finally, the end decision of fusion methods are obtained by rules [30]–[32].

![FIGURE 1. The multiple classifier fusion process flow diagram.](image)

During the process of multi classifier fusion, the input samples of each single classifier are the same, but the feature information of the samples can be different. Different feature information can be used for multi classifier fusion, and the same type of classifier can be used. And also, different classifiers are used for both the same and different feature information. The output of each classifier represents the decision for the same sample. Different types of single classifier determine different classification results. Single classifier includes decision tree, support vector machine, neural network, and other classification algorithms.

The fusion method can be divided into soft decision and hard decision. For the soft decision, the output of each single classifier is a probability value or range value, of which the probability value is processed to determine the sample category. Soft decision has many common fusion algorithms, such as product method, sum method, and other algorithms. The product method multiplies the posterior probability values of all single classifiers corresponding to the same category, and takes the largest one as the final classification result. The sum method is to add the posterior probability of the corresponding category to get the maximum value, which is the final classification result.

For the hard decision, the output of each single classifier is a unique category number. The common fusion method is voting method (MV). For each single classifier, the decision is taken as one vote, and then the most voted class is counted as the final output class. The formula of voting method is as follows:

\[
\tilde{f}_i = \begin{cases} 
1, & T(x_j) = \text{max} T(x_j) \\
0, & T(x_j) \neq \text{max} T(x_j) \quad (j = 1, 2, \ldots, m) 
\end{cases} \\
F_x = \sum_{i=1}^{N} \tilde{f}_i
\]

(4)

where \(T(x_j)\) is the single classifier classification result of the \(j\)th sample. The weighted voting method is based on the voting method. According to the different classification effect of each single classifier, the weight parameters of each classifier are determined by the classification accuracy of the calibration set. And then the multiple classifiers are fused according to the weight parameters. In addition, according to the empirical value, each single classifier is weighted based on the prediction probability of classification. It makes that the single classifier with good classification effect has more decision-making power.

C. EVALUATION STANDARD

Accuracy is a common evaluation standard in classification. The correct number of samples divided by all the samples is the accuracy of classification. Generally, the higher the accuracy is, the better the classifier is.

III. MATERIALS AND EXPERIMENT

A. SAMPLES ACQUISITION

The sampling site is located in the intertidal zone of Dongdayang village, Qingdao, Shandong Province, China. In August 2019, 59 sediment samples were collected with the help of bamboo rafts. The collected sediment samples were put in the laboratory for air drying, crushing, low-temperature drying and grinding. All the samples were passed the 0.3mm sieve. And then these samples were screened through...
three layers of sieves of 0.2mm, 0.1mm and 0.075mm, respectively (Sieve method). Subsamples were obtained with different particle size of 0.3-0.2mm, 0.2-0.1mm, 0.1-0.075mm and < 0.075mm, respectively (Table 1). These subsamples were classified as four particle size type according to the classification of Udden and Wentworth [33], [34]. In these types, only the fine sand type (0.2-0.1mm) was pure, others were all mixed (Table 1). Our classification was rough, but increased the interpretation difficulties.

TABLE 1. Particle size type of each parcel size sediment.

| Particle size | Particle size type                  |
|---------------|-------------------------------------|
| 0.3-0.2mm     | Medium sand and fine sand           |
| 0.2-0.1mm     | Fine sand                           |
| 0.1-0.075mm   | Fine sand and very fine sand        |
| <0.075mm      | very fine sand, silt and clay       |

B. DATA ACQUISITION

The reflection spectra of the sediment samples was collected by QE65000 spectrometer and DH-2000 light source (Ocean Optics, USA). The spectral sampling interval was 1nm, the integration time was 600ms, and the spectral range was from 200nm to 1100nm. The spectrometer and the light source were connected by the Y-type optical fiber (QR400-7-UV-VIS). The Y-type optical fiber probe was fixed by the bracket (RPH-1) with the angle of 45°. The samples were placed in a self-made sample box. The size and length of the self-made sample box were the same as that of the bracket. There were two spherical sample chambers on the sample box. The position and diameter of the sample chamber coincided with the probe hole of the bracket. More details see Li et al. [37].

Before collecting the spectrum, turned on the light source and the spectrometer for 0.5-1 hour to make them work normally. After calibration by the white board, the sample was put into the self-made sample box and gently levelled off. Then the bracket were put on the sample box tightly and the reflectance of the sample was collected. Each sample was measured 5 times and the average of its reflectance was used. The schematic diagram of sediment samples spectral measurement was shown in Figure 2. Finally, 128 reflectance spectra were collected. Since the front and back segments of the spectrum were affected by noise, only the spectral band of 226-975 nm was taken.

The concentrations of total carbon (TC) and total nitrogen (TN) in sediments with different particle size was determined using carbon and nitrogen analyzer (PE 2400II, Perkin Elmer, USA). Results shown in Table 2, the concentrations of TN and TC in sediments with different particle size was different.

C. DATA ANALYSIS

SVM algorithm was used to establish the spectral model using the RBF as the kernel function. The model based on different spectral band was used as the single classifier. The voting method was used to fuse different single classifier to make the final decision of classification. If the classification of each single classifier is not always consistent, the best classification for a single classifier model is the decision value.

Because of the complexity of sediment composition, it was difficult to find its characteristic wavelength. Therefore, the full spectral band was roughly divided. The spectral bands of the sediments were divided into 8 groups per 100nm. That is, 226-325nm, 326-425nm, 426-525nm, 526-625nm, 626-725nm, 726-825nm, 826-925nm and 926-975nm. The group 9 was full spectrum band, 226-976nm. The spectral data of these 9 different bands were normalized first, and then SVM was used to build the classification models for the four particles size types. 32 sediment samples were randomly selected as calibration set and validation set, which were used to build and validate the classification model. And other 27 sediment samples were test set, which were used to predict the classification of particle size. The proportion of the calibration set and validation set was 1:1 (64 calibration set samples, 64 validation set samples) and 2:1 (86 calibration set samples, 42 validation set samples), respectively. The sampling order among particle size was sequentially 0.3-0.2mm, 0.2-0.1mm, 0.1-0.075mm, and < 0.075mm. There were 108 samples in the test set. All data analysis were completed in MATLAB R2017a. The SVM algorithm is implemented by the libsvm toolbox on MATLAB platform.

IV. RESULTS

A. THE VISIBLE AND NEAR INFRARED SPECTRA OF SEDIMENTS WITH DIFFERENT PARTICLE SIZE

The average spectral reflection of 32 sediment samples with different particle size of 0.3-0.2mm, 0.2-0.1mm,
0.1-0.075mm and < 0.075mm were calculated. Four average spectra of different particle size are obtained, as shown in Figure 3.

The visible and near infrared spectral reflectance was significantly different among the four particle size types of sediments. But all types had the same spectral characteristics. The reflectance increased with the wave length for all types. All types had a peak reflectance at about 580nm.

The reflectance of sediments increased with the decreased particle size (Figure 2). The reflectance was largest for the particle size <0.075mm, followed by the particle size of 0.1-0.075mm and 0.2-0.1mm, and smallest for the particle size of 0.3-0.2mm.

### B. CHARACTERISTIC BAND OF REFLECTION SPECTRUM OF SEDIMENT WITH DIFFERENT PARTICLE SIZE

The proportion of calibration set and validation set was 1:1 and 2:1 by equal distance, respectively. The classification results of 9 groups’ spectral band with different particle size were shown in Table 3 and Table 4.

#### TABLE 3. Classification results of calibration set and validation set in 1:1 scale for each spectral band.

| Number | Spectral band | Accuracy of calibration set | Accuracy of validation set |
|--------|---------------|----------------------------|---------------------------|
| 1      | 226-325nm     | 100.0% (64/64)             | 73.44% (47/64)            |
| 2      | 326-425nm     | 64.06% (41/64)             | 48.44% (31/64)            |
| 3      | 426-525nm     | 67.19% (43/64)             | 50.00% (32/64)            |
| 4      | 526-625nm     | 85.94% (55/64)             | 53.13% (34/64)            |
| 5      | 626-725nm     | 89.06% (57/64)             | 53.13% (34/64)            |
| 6      | 726-825nm     | 85.94% (55/64)             | 59.38% (38/64)            |
| 7      | 826-925nm     | 100% (64/64)               | 73.44% (47/64)            |
| 8      | 926-975nm     | 100% (64/64)               | 89.06% (57/64)            |
| 9      | 226-975nm     | 95.31% (61/64)             | 70.31% (45/64)            |

According to the Table 3, under the condition of 1:1 ratio of calibration set and validation set, the classification results under different wavebands were quite different. The calibration set and validation set accuracy of full spectrum band 226-976nm with different particle size were 95.31% and 70.31%. The classification accuracy of 326-425nm, 426-525nm, 526-625nm, 626-725nm and 726-825nm spectral band were lower than that of the full band. While the classification accuracy of 226-325nm, 826-925nm and 926-975nm spectral band were higher than that of the full band. The classification accuracy of 226-325nm and 826-925nm spectral band were the same, the accuracy of the calibration set and the validation set were both 100.0% and 73.44%. 926-975nm spectral band had the best classification result, and the classification accuracy of calibration set and validation set were 100% and 89.06%, respectively. Table 4 under the condition of 2:1 ratio of calibration set and validation set also showed that the validation set accuracy of 226-325nm, 826-925nm and 926-975nm spectral band were better than that of the full band. The calibration set accuracy of 226-325nm and 826-925nm were slightly lower than that of the full band. The classification accuracy of 926-975nm spectral band was the highest, and the accuracy of calibration set and validation set were 100% and 88.10%. Therefore, 926-975nm spectral band could best represent the reflection spectrum of different particle size sediments. Choosing 926-975nm spectral band for different particle size classification could not only improve the overall classification results, but also improved the classification accuracy of each particle size. In the 526-625nm band, the calibration set R² was basically the same under the different proportion of the model set and validation set. The R² of the 2:1 proportion increased from 53.13% to 71.43%. It showed that the 580nm characteristic band could improve the final classification results when the proportion of calibration sets was increased. However, the classification results were still not ideal, and the 580nm band did not have the ability to classify different particle size sediments.

Furthermore, whether 926-975 nm was the characteristics spectral band was tested by the randomly sampling method (Table 5-6).

Table 5 was the classification results of randomly selected calibration set and validation set in the scale of 1:1, the calibration set accuracy of the two spectral band was not much
different. The accuracy of 926-975nm spectral band was slightly better than that of 226-975nm, and the average accuracy of random 10 times was 98.28% and 93.13%, respectively. There was a big difference between the validation set accuracy of the two spectral band. The classification accuracy of 926-975nm spectral band was significantly higher than that of 226-975nm spectral band, and the average accuracy of random 10 times was 87.62% and 73.00%, respectively. The validation set accuracy of 226-975nm spectral band was between 60% and 70%. While the validation set accuracy of 926-975nm spectral band was between 80% and 90%, and the best classification accuracy was 90.63%.

Table 6 showed the same regularity in the classification results of randomly selected calibration set and validation set in the scale of 2:1. The calibration set accuracy of 926-975nm spectral band was slightly better than that of 226-975nm spectral band, and the average accuracy of random 10 times was 98.28% and 93.13%, respectively. There was a big difference between the validation set accuracy of the two spectral band. The classification accuracy of 926-975nm spectral band was significantly higher than that of 226-975nm spectral band, and the average accuracy of random 10 times was 87.62% and 73.00%, respectively. The validation set accuracy of 226-975nm spectral band was between 60% and 70%. While the validation set accuracy of 926-975nm spectral band was between 80% and 90%, and the best classification accuracy was 90.63%.

C. MULTIPLE CLASSIFIERS FUSION WITH SINGLE CLASSIFIER IN EACH BAND

It can be seen from the table 2 and table 3 that the calibration set and validation set accuracy of 226-325nm, 826-925nm and 926-975nm spectral band were all higher than that of the full spectral band 226-975nm. The first four spectral bands with classification accuracy ranging from large to small were 926-975nm, 226-325nm, 826-925nm and 226-975nm spectral band. The classification results of 226-325nm and 826-925nm spectral band were consistent. Therefore, the four band classification model were selected as single classifier. Using the four single classifiers of 926-975nm, 226-325nm, 826-925nm and 226-975nm, five recommendation results were got. Recommendation result 1 were fused by the three single classifiers of 926-975nm, 226-325nm and 826-925nm. Recommendation result 2 were fused by the four single classifiers of 926-975nm, 226-325nm, 826-925nm and 226-975nm. Recommendation result 3 were fused by the three single classifiers of 926-975nm, 826-925nm and 226-975nm. Recommendation result 4 were fused by the three single classifiers of 926-975nm, 226-325nm and 226-975nm. Recommendation result 5 were fused by the four single classifiers of 226-325nm, 826-925nm and 226-975nm. The classification accuracy and the number of correct samples with each four particle sizes sediment samples, in 926-975nm, 226-325nm, 826-925nm and 226-975nm spectral band, were shown in Table 7. The overall classification accuracy were shown in Figure 4.
The classification accuracy of each four particle size was more than 80%, and the accuracy of particle size. The classification accuracy of each four particle size in the 926-975nm spectral band was the same as that of 0.1-0.075mm particle size. Only one of the four particle size classification results reached more than 80%, and the other three particle size was less than 70%. The accuracy of 0.2-0.1mm particle size and 0.1-0.075mm particle size were the same. The classification results of 0.3-0.2mm and 0.2-0.1mm particle size in 826-925nm spectral band were still the best, followed by 826-925nm and 226-975nm spectral bands, which was 100%. The other three particle sizes did not reach the optimal classification accuracy of four single classifier. The overall classification accuracy was better than the single classifier of 226-325nm, 826-925nm and 226-975nm spectral bands, but lower than that of 926-975nm.

Recommendation result 3, recommendation result 4 and recommendation result 5 were all lower than the recommendation result 1. This three recommendation results did not choose the best three single classifier, which could not cause the better results. Therefore, the single classifiers chosen were very important, which decided the final recommendation results.

108 sediment samples of test set were predicted in the classification model based on 926-975nm, 226-325nm, 826-925nm and 226-975nm spectral band. The classification accuracy and the number of correct samples of four spectral bands and recommendation 1-5 with overall classification and each particle sizes were shown in table 8.

According to table 8, in the classification results of four spectral bands, the overall classification results of 926-975nm spectral band were still the best, followed by 826-925nm, 226-975nm and 226-325nm spectral band. The overall accuracy of the first three spectral bands were more than 80%. The classification results of 0.3-0.2mm and < 0.075mm were better, which were more than 90%. Among the classification results of recommendation 1-5 based on multiple classifiers fusion, recommendation 2 (the classification results fused by four single classifiers) had the highest accuracy of overall classification and each particle sizes, and the overall accuracy was 90.74%. The second was recommendation 1, recommendation 3 and recommendation 4, and the overall
The characteristic wavelength is the key to spectral analysis. Li [36-38] also pointed out that the characteristic wavelengths of the single classifiers have similar classification accuracy, which affected the full band classification accuracy. The correct extraction of characteristic wavelengths is very important in spectral analysis, which can characterize the key information of sediment particle size. In this paper, the characteristic wavelength of sediments is a rough band range, and the optimal characteristic characteristic band range can produce larger reflectance (Figure 2). The pattern had been rarely reported, but was the basis of spectral analysis. And 926-975 nm was the characteristic spectra of the particle size of sediments. As hypothesis 2, the particle size of sediments could be spectrally analyzed by stoichiometry. In this study, we found that the particle size of sediments could be quantitatively analyzed by the multi spectral classifier fusion. We found that the classification accuracy of fusion classifier was determined by the compositions of the single classifiers. Therefore, getting the characteristic spectra was the most important step.

Whether for chemical or physical parameters, the characteristic wavelength is the key to spectral analysis. Li et al. [38] extracted the characteristic wavelengths of soil total nitrogen (TN) by non-information variable elimination (UVE) and successive projection algorithm (SPA), and obtained better prediction results of total nitrogen content. Jiang et al. [39] selected two characteristic bands of PCs by principal component analysis (PCA), and obtained higher accuracy for detection of pear bruise. Zhou et al. [40] used wavelet transform to extract the characteristic wavelength and established the water content model of lettuce samples, which improved the accuracy of water content detection. In this paper, the characteristic wavelengths of sediment particle size were selected manually, and the classification results were improved to a certain extent. By analyzing the classification results with different particle size in each band, the optimal characteristic band was 926-975nm spectral band, followed by 226-325nm and 826-925nm spectral band. Compared with the full band 226-975nm spectral band, the classification accuracy had been improved. It showed that the 226-975nm contained the spectral band which was not related to the particle size characteristics, that is useless information. 426-525nm, 626-725nm, 726-825nm and other bands had low classification accuracy, which affected the full band classification results. The correct extraction of characteristic wavelengths is very important in spectral analysis, which can characterize the key information of sediment particle size. In this paper, the characteristic wavelength of sediments is a rough band range, furthermore we will study the band range which can better characterize its characteristics. In the future, the algorithm that can automatically select the characteristic wavelength of sediment will be studied, to further improve the classification result of sediment particle size.

In the spectral model of marine sediment with different particle sizes, different proportion of model set and validation set, different division methods of model set and validation set and different characteristic bands of marine sediment will lead to different models. In particular, the selection of the characteristic spectral band which can best characterize the sediment particles has the greatest impact on the model results. Different spectral bands contain different information, and different information can represent different characteristics of marine sediment. In order to improve the accuracy of the model, the characteristic spectral band which can best represent the sediment particles is selected to build the model. Therefore, it is very important to choose the appropriate model set method and characteristic spectral band for the establishment of stable and accurate model.

In the research of the fusion with the single classifier, the classification accuracy in order from large to small was recommendation result 1 > the optimal single classifier > recommendation result 2, recommendation result 3, recommendation result 4 and recommendation result 5. The reason was that the classification results of four single classifiers were uneven. Poor single classifier had certain influence on final particle size classification. The useless information or noise information would be introduced into the fusion results, which leads to unsatisfactory final classification result. The classification accuracy of single classifier (926-975nm spectral band) was much higher than the other three single classifiers (226-325nm, 826-925nm and 226-975nm spectral band). The classification accuracy of the three single classifiers would affect the results of the single classifier (926-975nm spectral band). In the test set, the four single classifiers have similar classification accuracy, which can complement other’s classification results, so that the classification accuracy of all recommended results is higher than that of each single classifier.

The physical properties of the sediment can be estimated with the spectra baseline that is affected by the grain size in the visible and near-infrared ranges [35]. Jacq et al. [36]

### TABLE 8. The test set classification results of four spectral bands and five recommendation results with 0.3-0.2 mm, 0.2-0.1 mm, 0.1-0.075 mm, <0.075 mm particle sizes.

| Spectral band | Overall accuracy | 0.3-0.2 mm | 0.2-0.1 mm | 0.1-0.075 mm | <0.075 mm |
|---------------|-----------------|------------|------------|--------------|-----------|
| 926-975nm     | 87.04%          | 96.30%     | 85.19%     | 74.07%       | 92.99%    |
| (26/27)       | (25/27)         | (25/27)    | (21/27)    | (21/27)      |
| 826-925nm     | 85.19%          | 92.59%     | 74.07%     | 77.78%       | 96.30%    |
| (92/108)      | (25/27)         | (25/27)    | (21/27)    | (21/27)      |
| 226-325nm     | 76.85%          | 88.89%     | 70.37%     | 62.96%       | 85.19%    |
| (83/108)      | (24/27)         | (19/27)    | (17/27)    | (23/27)      |
| 226-975nm     | 83.33%          | 92.59%     | 77.78%     | 66.67%       | 96.30%    |
| (90/108)      | (25/27)         | (21/27)    | (18/27)    | (26/27)      |
| Recomm 1      | 89.81%          | 96.30%     | 88.89%     | 77.78%       | 96.30%    |
| dation 1      | (97/108)        | (26/27)    | (24/27)    | (21/27)      | (26/27)   |
| Recomm 2      | 90.74%          | 96.30%     | 88.89%     | 81.48%       | 96.30%    |
| dation 2      | (98/108)        | (26/27)    | (24/27)    | (22/27)      | (26/27)   |
| Recomm 3      | 89.81%          | 96.30%     | 85.19%     | 81.48%       | 96.30%    |
| dation 3      | (97/108)        | (26/27)    | (23/27)    | (22/27)      | (26/27)   |
| Recomm 4      | 89.81%          | 96.30%     | 88.89%     | 77.78%       | 96.30%    |
| dation 4      | (97/108)        | (26/27)    | (24/27)    | (21/27)      | (26/27)   |
| Recomm 5      | 87.96%          | 96.30%     | 85.19%     | 77.78%       | 92.99%    |
| dation 5      | (95/108)        | (26/27)    | (23/27)    | (21/27)      | (25/27)   |

The contents in the brackets were the number of samples with correct classification / the total number of samples.
used visible and near-infrared hyperspectral imaging to predict particle size fractions and distribution at a resolution of 200 μm on a previously well-studied sediment core taken from Lake Bourge. In this paper, different particle sizes of intertidal sediments were classified by visible near infrared spectroscopy. Compared to using the UAV red, green and blue (RGB) orthophotometer, the accuracy assessment of the surface sediment classification based on these six types indicated an overall accuracy of 72.8% [18]. The best classification accuracy of this paper is up to 93.75%. The former was based on the sediment in the field, while this paper was aimed at the simple treated sediment samples. And this method will be applied to the analysis of the sediment in the field later. The result of the sieve method was used as the criteria. Compared with the sieve method, the classification accuracy was higher and the classification speed was much faster than that of sieve method. Visible - near infrared spectroscopy will be a new technique for rapid particle size analysis of Marine sediments.

VI. CONCLUSION

The particle size of sediment has a close relation to the identification of sediment source, the influence of organic matter in sediments, and the formation of sediment hydrate. Therefore, it is of great significance to analyze the different particle size of marine sediments, and the formation of sediment hydrate. In this paper, the particle size of marine sediments in Qingdao, China by the sieve method was used as the criteria. Compared to using the sieve method, the classification accuracy was higher and the classification speed was much faster than that of sieve method. Visible - near infrared spectroscopy will be a new technique for rapid particle size analysis of Marine sediments.

REFERENCES

[1] T. Luo, Y. Song, Y. Zhu, W. Liu, Y. Liu, Y. Li, and Z. Wu, “Triaxial experiments on the mechanical properties of hydrate-bearing marine sediments of south China sea,” Mar. Petroleum Geol., vol. 77, pp. 507–514, Nov. 2016.

[2] X. Gu, R. Tenzer, and V. Gladikikh, “Empirical models of the ocean-sediment and marine sediment-bedrock density contrasts,” Geosci. J., vol. 18, no. 4, pp. 439–447, Dec. 2014.

[3] R. Lin, X. Lin, T. Guo, L. Wu, W. Zhang, and W. Lin, “Metaproteomic analysis of bacterial communities in marine mudflat aquaculture sediments,” World J. Microbiol. Biotechnol., vol. 31, no. 9, pp. 1397–1408, Sep. 2015.

[4] C. Booth, J. Walden, A. Neal, and J. Smith, “Use of mineral magnetic concentration data as a particle size proxy: A case study using marine, estuarine and fluvial sediments in the Carmarthen Bay area, South Wales, UK,” Sci. Total Environ., vol. 347, nos. 1–3, pp. 241–253, Jul. 2005.

[5] S. Walker, “Ecological functions and biogenic element cycling roles of marine sediment/particles,” CCAMLR Sci., vol. 24, nos. 1–2, pp. 47–62, 2017.

[6] M. Boudreault, A. J. Koster, D. A. Lobb, K. Liu, G. Benoy, P. N. Owens, S. Daniesescu, and S. Li, “Using colour, shape and radionuclide fingerprints to identify sources of sediment in an agricultural watershed in Atlantic Canada,” Can. Water Resour. J., vol. 43, no. 3, pp. 347–356, 2018.

[7] Y. Wang, M. Zhang, D. Zhang, and Z. Shen, “The influence of sediment particle size on the properties of adsorbed dissolved organic matter in the Yangtze estuary and its interactions with As/Se,” Mar. Pollut. Bull., vol. 105, no. 1, pp. 351–358, Apr. 2016.

[8] Y. Zhang, X.-S. Li, Y. Wang, Z.-Y. Chen, and K.-F. Yan, “Decomposition conditions of methane hydrate in marine sediments from South China Sea,” Fluid Phase Equilibria, vol. 413, pp. 110–115, Apr. 2016.

[9] P. D. Thorne and D. Hurter, “An overview on the use of backscattered sound for measuring suspended particle size and concentration profiles in non-cohesive inorganic sediment transport studies,” Continental Shelf Res., vol. 73, pp. 97–118, Feb. 2014.

[10] S. Hunt and H. F. E. Jones, “Sediment grain size measurements are affected by site-specific sediment characteristics and analysis methods: Implications for environmental monitoring,” New Zealand J. Mar. Freshwater Res., vol. 53, no. 2, pp. 244–257, Apr. 2019.

[11] S. Haba, K. Fukuda, Y. Ohta, Y. Kobuchi, and T. Katouda, “Fumed silica slurry stabilizing methods for chemical mechanical polishing,” Jpn. J. Appl. Phys., vol. 42, no. 2A, pp. 418–423, Feb. 2003.

[12] O. S. Aginemlen, A. J. Mulholland, and J. Sefick, “Modelling of artefacts in estimations of particle size of needle-like particles from laser diffraction measurements,” Chem. Eng. Sci., vol. 158, pp. 445–452, Feb. 2017.

[13] L. Pan, B. Ge, and F. Zhang, “Indetermination of particle sizing by laser diffraction in the anomalous size ranges,” J. Quant. Spectrosc. Radiat. Transfer, vol. 199, pp. 20–25, Sep. 2017.

[14] S. Cuvan, P. Francus, and S. F. Lamoureux, “Estimation of grain size variability with micro X-ray fluorescence in laminated Lacustrine sediments, cape bounty, Canadian high arctic,” J. Paleolimnology, vol. 44, no. 3, pp. 803–817, Oct. 2010.

[15] J. M. Menéndez-Aguado, E. Peña-Carpio, and C. Sierra, “Particle size distribution fitting of surface detrital sediment using the Swerebec function,” J. Soils Sediments, vol. 15, no. 9, pp. 2004–2011, Sep. 2015.

[16] B. L. DeCost and E. A. Holm, “Characterizing powder materials using keypoint-based computer vision methods,” Comput. Mater. Sci., vol. 126, pp. 438–445, Jan. 2017.

[17] M. G. Yates, A. R. Jones, S. McGrorty, and J. D. Goss-Custard, “The use of satellite imagery to determine the distribution of intertidal surface sediments of the wash, England,” Estuarine, Coastal Shelf Sci., vol. 36, no. 3, pp. 333–344, Jul. 1993.

[18] K.-L. Kim, B.-J. Kim, Y.-K. Lee, and J.-H. Ryu, “Generation of a large-scale surface sediment classification map using unmanned aerial vehicle (UAV) data: A case study at the Hwang-Do tidal flat, Korea,” Remote Sens., vol. 11, no. 3, p. 229, Feb. 2019.

[19] N. Naseer, F. M. Noori, N. K. Qureshi, and K.-S. Hong, “Determining optimal feature-combination for LDA classification of functional near-infrared spectroscopy signals in brain-computer interface application,” Frontiers Hum. Neurosci., vol. 10, p. 237, May 2016.

[20] K.-S. Hong and H. Santosa, “Decoding four different sound-categories in the auditory cortex using functional near-infrared spectroscopy,” Hearing Res., vol. 333, pp. 157–166, Mar. 2016.

[21] J. Peng, K. Song, H. Zhu, W. Kong, F. Liu, T. Shen, and Y. He, “Fast detection of tobacco mosaic virus infected tobacco using laser-induced breakdown spectroscopy,” Sci. Rep., vol. 7, no. 1, Mar. 2017, Art. no. 44551.

[22] W. Ng, B. P. Malone, and B. Minasny, “Rapid assessment of petroleum-contaminated soils with infrared spectroscopy,” Geoderma, vol. 289, pp. 150–160, Mar. 2017.

[23] D. F. Malley, L. Lockhart, P. Wilkinson, and B. Hauser, “Determination of carbon, carbonate, nitrogen, and phosphorus in freshwater sediments by near-infrared reflectance spectroscopy: Rapid analysis and a check on conventional analytical methods,” J. Paleolimn., vol. 24, no. 4, pp. 415–425, Nov. 2000.
X. Li, P. Fan: Study on Characteristic Spectrum and Multiple Classifier Fusion With Different Particle Size in Marine Sediments

[24] M. Z. Martin, N. Labbé, N. André, S. D. Wallschleger, R. D. Harris, and M. H. Ebinger, “Novel multivariate analysis for soil carbon measurements using laser-induced breakdown spectroscopy,” Soil Sci. Soc. Amer. J., vol. 74, no. 1, pp. 87–93, Jan./Feb. 2010.

[25] M. R. Nanni, E. Cezar, C. A. D. S. Junior, G. F. C. Silva, and A. A. da Silva Gualberto, “Partial least squares regression (PLSR) associated with spectral response to predict soil attributes in transitional lithologies,” Arch. Agronomy Soil Sci., vol. 64, no. 5, pp. 682–695, Apr. 2018.

[26] J. Gao, L. Xu, and F. Huang, “A spectral–textural kernel-based classification method of remotely sensed images,” Neural Comput. Appl., vol. 27, no. 2, pp. 431–446, Feb. 2016.

[27] J. Xu, Y. Y. Tang, B. Zou, Z. Xu, L. Li, Y. Lu, and B. Zhang, “The generalization ability of SVM classification based on Markov sampling,” IEEE Trans. Cybern., vol. 45, no. 6, pp. 1169–1179, Jun. 2015.

[28] X. Tan, F. Yu, and X. Zhao, “Support vector machine algorithm for artificial intelligence optimization,” Cluster Comput., vol. 22, no. S6, pp. 15015–15021, Nov. 2019.

[29] Q. Liu, C. Chen, Y. Zhang, and Z. Hu, “Feature selection for support vector machines with RBF kernel,” Artif. Intell. Rev., vol. 36, no. 2, pp. 99–115, Aug. 2011.

[30] L. I. Kuncheva, J. C. Bezdek, and R. P. W. Duin, “Decision templates for multiple classifier fusion: An experimental comparison,” Pattern Recognit., vol. 34, no. 2, pp. 299–314, Feb. 2001.

[31] S. Sun, “Local within-class accuracies for weighting individual outputs in multiple classifier systems,” Pattern Recognit. Lett., vol. 31, no. 2, pp. 119–124, Jan. 2010.

[32] H. Imani and H. Ghassemian, “An overview on spectral and spatial information fusion for hyperspectral image classification: Current trends and challenges,” Inf. Fusion, vol. 59, pp. 50–83, Jul. 2020.

[33] J. A. Udden, “Mechanical composition of clastic sediments,” Geol. Soc. Amer. Bull., vol. 25, no. 1, pp. 655–744, Jan. 1914.

[34] C. K. Wentworth, “A scale of grade and class terms for clastic sediments,” J. Geol., vol. 30, no. 5, pp. 377–392, Jul. 1922.

[35] D. L. Wetzel, “Near-infrared reflectance analysis,” Anal. Chem., vol. 55, no. 12, pp. 1165A–1176A, Oct. 1983.

[36] K. Jacq, C. Giguet-Covex, P. Sabatier, Y. Perrette, B. Fanget, D. Coquin, M. Dehret, and F. Arnaud, “High-resolution grain size distribution of sediment core with hyperspectral imaging,” Sedimentary Geol., vols. 393–394, Dec. 2019, Art. no. 105536.

[37] X.-Y. Li, P.-P. Fan, Y. Liu, G.-L. Hou, Q. Wang, and M.-R. Lv, “Prediction results of different modeling methods in soil nutrient concentrations based on spectral technology,” J. Appl. Spectrosc., vol. 86, no. 4, pp. 765–770, Sep. 2019.

[38] H. Li, S. Jia, and Z. Le, “Quantitative analysis of soil total nitrogen using hyperspectral imaging technology with extreme learning machine,” Sensors, vol. 19, no. 20, p. 4355, Oct. 2019.

[39] H. Jiang, C. Zhang, Y. He, X. Chen, F. Liu, and Y. Liu, “Wavelength selection for detection of slight bruises on pears based on hyperspectral imaging,” Appl. Sci., vol. 6, no. 12, p. 450, Dec. 2016.

[40] X. Zhou, J. Sun, H. Mao, X. Wu, X. Zhang, and N. Yang, “Visualization research of moisture content in leaf lettuce leaves based on WT-PLSR and hyperspectral imaging technology,” J. Food Process Eng., vol. 41, no. 2, Apr. 2018, Art. no. e12647.

XUEYING LI was born in Jilin, China, in 1990. She received the B.S. and M.Eng. degrees from China Agricultural University, Beijing, China, in 2013 and 2016, respectively. She is currently an Assistant Professor with the Institute of Oceanographic Instrumentation, Qilu University of Technology (Shandong Academy of Sciences). Her research interests include algorithm research and spectral analysis.

PINGPING FAN was born in Shandong, China, in 1980. She received the B.S. and M.Eng. degrees from China Agricultural University, Beijing, China, in 2013 and 2016, respectively. She is currently an Assistant Professor with the Institute of Oceanographic Instrumentation, Qilu University of Technology (Shandong Academy of Sciences). Her research interests include marine sediment monitoring and soil monitoring.

***