Journal of Environmental Science and Engineering B 14 (2025) 11-18
doi:10.17265/2162-5263/2025.01.002

Multiple Linear Regression Analysis of Vertical Distribution of Near-Shore Suspended Sediment

Mengmeng Wei, Wenjin Zhu, Xiaotian Dong and Xingyuan Chen
Jiangsu Ocean University, Lianyungang 222000, China

Abstract: According to some main assumptions in the Rouse Formula, it analyzes the applicability of Rouse distribution in the coastal region. Based on the classical Rouse Formula, the linear form of Rouse Formula and the transport characteristics of offshore sediment were used to take $\ln\frac{z}{h}$, $\ln c_a$, c_a, u, $\ln u$ and $\frac{z}{h}$ as the independent variables. The multiple liner regression method was used to analyze the influence of the independent variables on the vertical distribution of sediment concentration. By using the method of significance test, the factors ($\ln u$) that have less influence on sediment concentration among 6 variables were eliminated. The correlation coefficient between the calculated sediment concentration and the measured sediment concentration indicates that the adopted variables can reflect the characteristics of vertical distribution of concentration of fine sediment near shore under complex dynamic conditions.

Key words: Rouse Formula, multiple linear regression, vertical distribution of suspended sediment, Hai'an Bay.

1. Introduction

Suspended sediment concentration is an important parameter in the study of ocean water and sediment dynamics [1, 2], and it is an important factor to understand the transport and deposition of marine materials. So far, the theories used to study the distribution of suspended sediment concentration mainly include diffusion theory, mixing theory, energy theory, similarity theory and random theory [3-6]. Various theories have obtained many results from different perspectives. Through the comprehensive analysis of various theories, it is found that although the theories are different, the results are back to or close to the form of the diffusion equation, and the difference is only the diffusion coefficient. This understanding makes the study of fine-grained sediment suspension in estuaries and coasts often directly adopt the diffusion theory [7-10].

Generally speaking, the use of Rouse Formula [11] in the estuary and coastal waters of unsteady flow has very strict restrictions. Before the vertical distribution of flow velocity is not clear, their use can only be limited to the conditions of small wind and wave, no density stratification area, and the flow is close to stability [12, 13]. The non-stationarity and multi-factor of the dynamic field in the nearshore waters make the vertical concentration of sediment particularly complex. At present, there is no method to accurately describe the vertical distribution of sediment under complex dynamic conditions in the nearshore waters. In this paper, based on the traditional Rouse distribution, combined with the method of multiple linear regression, multiple independent variables affecting the distribution of suspended sediment vertical lines are introduced, and the characteristics of the distribution of suspended sediment vertical lines in the near shore are analyzed by linear regression.

2. Study Area

Hai'an Bay is located in the middle of the southern tip of the Leizhou Peninsula, China (Fig. 1). The geographical coordinates are 110°13′12″ E and 20°16′12″ N (Fig. 2). The sample data are as follows:

Corresponding author: Xiaotian Dong, Ph.D., research fields: port, coastal and offshore engineering.

From May 13, 2008 to May 28, 2008, SC1, SC2, SC3, SC4 and SC5 were observed for 15 days. Tidal current velocity data: The neap tide lasted 52 h from 10:00 on 14 May 2008 to 14:00 on 16 May 2008. The spring tide lasted 53 h from 17:00 on 20 May 2008 to 22:30 on 22 May 2008. Sediment data: The measurement time is synchronized with the flow velocity observation, sampling once every 2 h.

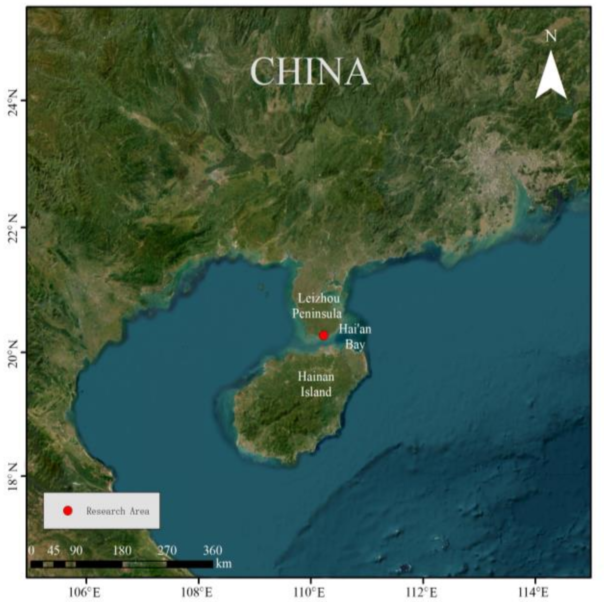


Fig. 1 Qiongzhou Strait Region.

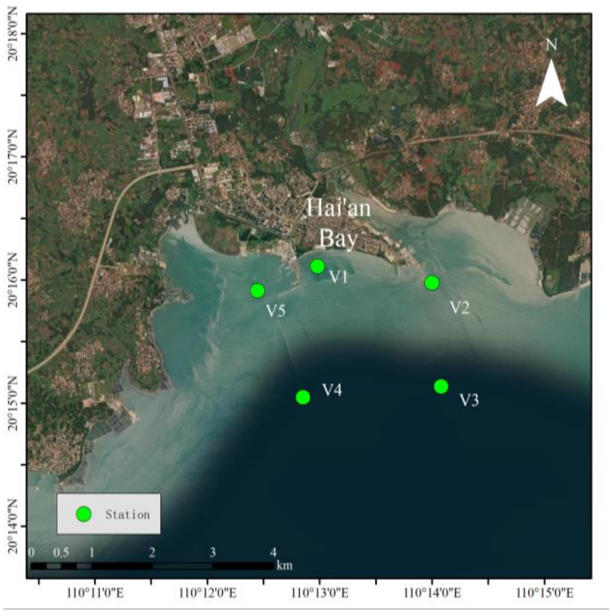


Fig. 2 Overview of Hai'an Bay and location of hydrological station.

3. Methods

3.1 The Linear Form of Rouse Formula

The basic idea of the diffusion theory is that the reason why the sediment can be suspended is the result of the combination of turbulent diffusion and sediment gravity [14-18]. The conservation equation of sediment movement reflects the continuous equation that the suspended sediment should satisfy when it is in an equilibrium state in the water flow. The form of the Rouse Formula describing the vertical distribution of suspended sediment is as follows:

$$\frac{c(z)}{c_a} = \left[\frac{a(h-z)}{z(h-a)}\right]^{w_s/\kappa u_*} \tag{1}$$

where, c_a is the reference point concentration.

In order to better calculate the vertical distribution curve of suspended sediment concentration, the logarithm of the Eq. (1) is written into the form of linear equation:

$$\ln c = a_0 + a_1 \ln \frac{z}{h} + a_3 \ln c_a \tag{2}$$

where a_0, a_1 is the undetermined coefficient, which is obtained by regression analysis of measured data; the undetermined parameters a_3 are used to improve the relationship between the sediment concentration of each layer and the sediment concentration of the reference point. Eq. (2) reflects the relationship between the concentration of each point on the vertical line and the relative water depth of the variable and the concentration of the reference point.

3.2 Multivariate Linear Regression Analysis Method Based on Rouse Formula

Multivariate linear regression analysis is a classical method in multivariate statistical analysis. It predicts one or more response variables (dependent variables) through a set of predictors (independent variables). After using multivariate linear regression analysis method to predict, the predicted value and change trend of dependent variables can provide managers with theoretical basis for decision support. Assuming that the dependent variable Y satisfies or approximately satisfies the linear relationship with multiple independent variables $X_1, X_2, \cdots, X_k$ the following multiple linear regression model can be established :

$$\begin{cases} Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \\ \varepsilon \sim N(0, \sigma^2) \end{cases} \tag{3}$$

where Y is an observable random variable, ε is an unobservable random error, $\beta_0, \beta_1, \cdots, \beta_k, \sigma^2$ are unknowns independent of $X_1, X_2, \cdots, X_k$.

According to the turbulent diffusion theory, the variation law of sediment vertical concentration is consistent with the variation law of flow turbulence intensity, and there is also an internal relationship between suspended sediment concentration and flow velocity and velocity gradient. The measured velocity of each layer of the vertical line in the nearshore waters is the result of the combined action of various dynamic factors. The measured velocity can not only reflect the strength of the tidal current power, but also reflect the stage of the water flow. As long as the measured velocity data are fully used, the influence of waves and tidal currents can be indirectly reflected, which is of great significance for further correcting the distribution of sediment vertical concentration.

According to the above analysis, it can be seen that the three physical quantities of relative water depth $\frac{z}{h}$, reference point sediment concentration c_a and velocity u are very important for the calculation of the vertical distribution of sediment concentration. Therefore, in this paper, the sediment concentration c is regarded as a random variable, and the three physical quantities of relative water depth $\frac{z}{h}$, reference point sediment concentration c_a and flow velocity u are regarded as independent variables. At the same time, in order to clarify the form of the three independent variables in the formula, referring to the linear form of the Rouse Formula, six variables such as $\frac{z}{h}$, $\ln\frac{z}{h}$, c_a, $\ln c_a$, u, and $\ln u$ are used as independent variables to construct Eq. (4) (list of variables is shown in Table 1).

$$\ln c = a_0 + a_1 \ln\frac{z}{h} + a_2 c_a + a_3 \ln c_a + a_4 u + a_5 \ln u + a_6 \frac{z}{h} \tag{4}$$

Table 1 List of variables.

Dependent variable	Constant a_0	Independent variable a_1	Independent variable a_2	Independent variable a_3	Independent variable a_4	Independent variable a_5	Independent variable a_6
$\ln c$	_	$\ln\frac{z}{h}$	c_a	$\ln c_a$	u	$\ln u$	$\frac{z}{h}$

4. Result and Discussion

4.1 Multivariate Linear Regression Analysis Method Based on Rouse Formula

The derivation of the Rouse Formula has the following main assumptions: (1) The vertical average flow velocity is 0; (2) Some assumptions in the turbulent mixing length theory (including the logarithmic distribution of velocity); (3) The sediment settling velocity does not change with the water depth as a constant. These assumptions are more stringent in coastal waters.

Firstly, the settling velocity ω_s is no longer a constant, because the settling velocity of fine sediment in fresh water and salt water is different [19]. In the range of small salinity, the average settling velocity of flocs increases rapidly with the increase of salinity. After the salinity exceeds a certain value, the further increase of salinity has no great influence on the average settling velocity. In addition, the sediment concentration at the bottom of the water body is large, and the sedimentation process will be hindered by other sediment particles, and the sedimentation rate is no longer a constant.

Secondly, the horizontal time-averaged velocity in the nearshore waters of the estuary is affected by factors such as accelerated flow and decelerated flow. The vertical line is no longer a logarithmic distribution (Fig. 3). The distribution of vertical diffusion coefficient of sediment is more complicated when the density is stratified. At present, there are relatively few studies on the vertical mixing problem of stratified estuaries. It is generally believed that the vertical distribution of vertical diffusion coefficient in density stratification is more complex, which is related to the vertical gradient of density and velocity, rather than a simple parabolic distribution.

Therefore, before the vertical distribution of the velocity is not clear, the use of the Rouse Formula can only be limited to the area where the wind wave is small, the acceleration of the water flow is small and there is no density stratification, otherwise it will produce large errors.

Based on the above reasons, this paper verifies Eq. (2) linearized by Eq. (1). The correlation between the independent variables $\ln\frac{z}{h}$ and $\ln c_a$ was analyzed by taking the sediment vertical concentration of five measured points in Hai'an Bay of Qiongzhou Strait in May 2008. The analysis results are shown in Fig. 4. It can be seen that if the correlation coefficient between the sediment concentration calculated by the correlation analysis of these two independent variables and the measured value is between 0.5 and 0.6, it shows that for the fine-grained sediment in Hai'an Bay, under the action of complex dynamic conditions, neither Eq. (1) nor Eq. (2) can accurately describe the vertical distribution of sediment concentration.

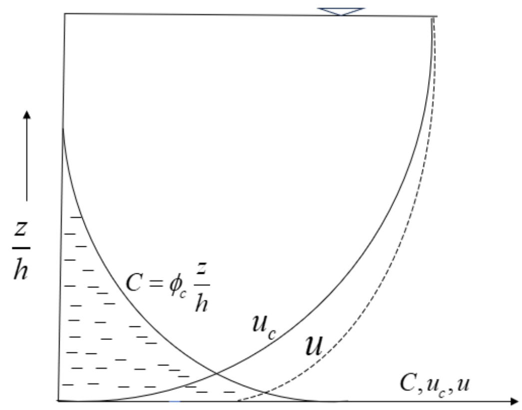


Fig. 3 Vertical distribution of concentration and velocity (u_c is the velocity of sediment-laden flow).

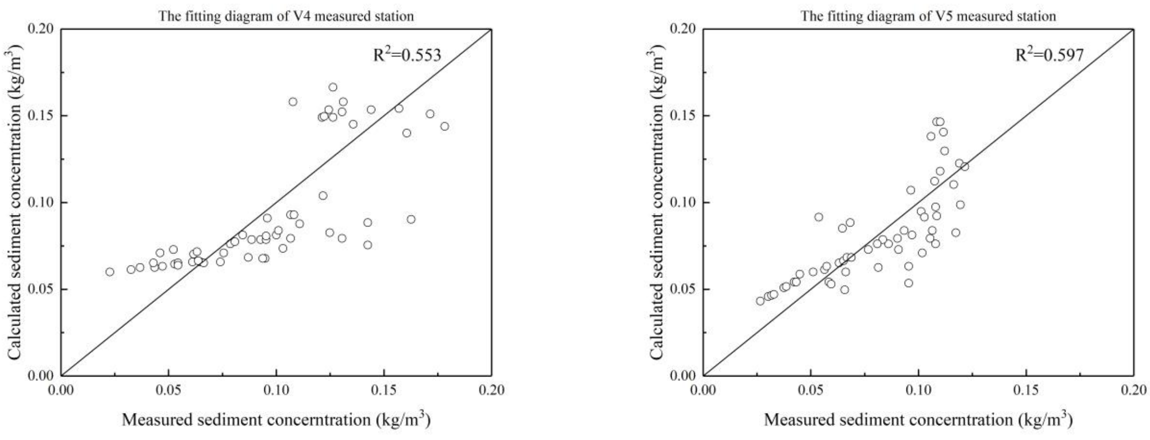


Fig. 4 The correlation between calculated and measured values of vertical sediment concentration when $\ln\frac{z}{h}$ and $\ln c_a$ are considered.

Table 2 Correlation of sediment concentration between layers.

Station		Spring tide				Neap tide			
	Relative depth	0.0	0.6	1.0	Average	0.0	0.6	1.0	Average
V5	0.0	1	0.548	0.196	0.550	1	0.732	0.103	0.506
	0.6	-	1	0.699	0.915	-	1	0.394	0.718
	1.0	-	-	1	0.767	-	-	1	0.663
	Average	-	-	-	1	-	-	-	1
V4	0.0	1	0.055	0.013	0.427	1	0.715	0.624	0.821
	0.6	-	1	0.442	0.545	-	1	0.927	0.949
	1.0	-	-	1	0.525	-	-	1	0.934
	Average	-	-	-	1	-	-	-	1
V3	0.0	1	0.521	0.378	0.708	1	0.944	0.861	0.961
	0.6	-	1	0.719	0.896	-	1	0.958	0.994
	1.0	-	-	1	0.848	-	-	1	0.963
	Average	-	-	-	1	-	-	-	1
V2	0.0	1	0.799	0.604	0.876	1	0.394	0.170	0.603
	0.6	-	1	0.816	0.952	-	1	0.562	0.763
	1.0	-	-	1	0.856	-	-	1	0.665
	Average	-	-	-	1	-	-	-	1
V1	0.0	1	0.676	0.426	0.862	1	0.871	0.705	0.913
	0.6	-	1	0.542	0.882	-	1	0.855	0.972
	1.0	-	-	1	0.717	-	-	1	0.908
	Average	-	-	-	1	-	-	-	1

4.2 Multiple Linear Regression Analysis of Nearshore Suspended Sediment Vertical Distribution

The sediment concentration of each layer has a good correlation with the sediment concentration of the adjacent layer under the influence of the sedimentation of the upper layer and the turbulent diffusion of the lower layer. The correlation analysis of the sediment concentration of each layer during the spring and neap tides in Hai'an Bay also has a good correlation. Table 2 shows the correlation coefficients between the water surface, the bottom layer, the 0.6 layer and the average concentration.

The variables considered in the regression analysis are four more variables (c_a, u, $\ln u$, $\frac{z}{h}$) than the two variables ($\ln\frac{z}{h}$ and $\ln c_a$) in the original Rouse Formula. It can be seen from Table 3 that the concentration of the reference point has the most significant effect on the sediment concentration. The correlation between the sediment concentration calculated only by the relative water depth and the measured value is only about 0.3. The variable 1 is combined with the variable 2, the variable 3, the variable 4, the variable 5 and the variable 6 respectively. The calculation results are shown in Table 3. The correlation coefficient calculated by "variable 1 + variable 2" is significantly higher than that of other two combinations, that is "variable 1+ variable 2" is very important for the vertical distribution of sediment in Hai'an sea area. Therefore, "variable 1 + variable 2" was used as the basic combination to combine with other variables in turn, and the correlation coefficients under 16 different combination conditions were calculated. Compared with the combination of the 13th group and the 14th group, the combination effect with the "variable 4" and the "variable 5" was slightly lower than that with the "independent variable 6".

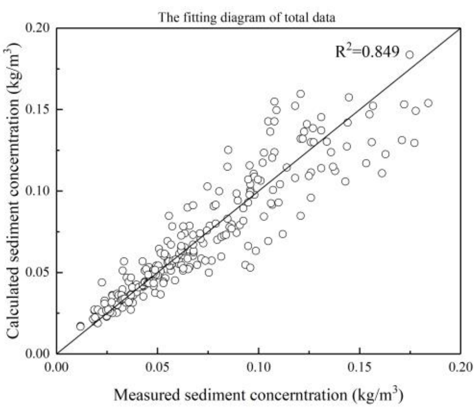
For fine sediment in Hai'an Bay, relative water depth, reference point concentration and flow velocity all have a certain influence on the vertical distribution. According to the calculation results of Tables 2 and 3, the influence degree of the three is the reference point concentration, the relative water depth and the velocity. The combination of the 14th group, the 15th group and the 16th group shows that the influence of the "variable 5" ($\ln u$) is small. Therefore, the final calculation result of this paper adopts the combination of "variable 1 + variable 2 + variable 3 + variable 4 + variable 6" to calculate the sediment vertical distribution in Hai'an Bay. The multiple linear regression analysis method based on the Rouse equation is as follows:

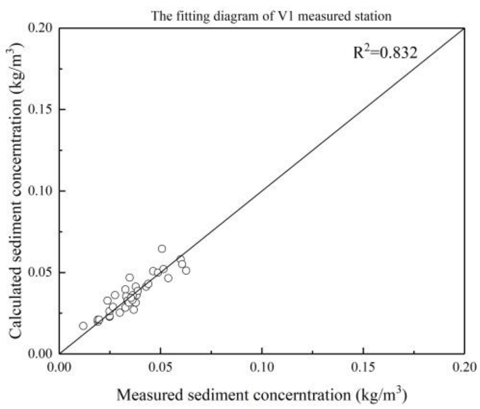
$$\ln c = a_0 + a_1 \ln \frac{z}{h} + a_2 c_a + a_3 \ln c_a + a_4 u + a_6 \frac{z}{h} \quad (5)$$

At the same time, the correlation coefficient of each station calculated by the combination of these five variables can reach 0.8, and the V5 station can reach 0.92 (Fig. 5). Therefore, Eq. (5) is used to analyze the vertical distribution of fine sediment concentration in Hai'an Bay, and satisfactory accuracy can be achieved.

Table 3 The correlation between the calculated vertical sediment concentration and the measured value under the average combination of variables.

	Combination	V1	V2	V3	V4	V5
1	variable 1	0.225	0.180	0.370	0.555	0.360
2	variable 1 + variable 2	0.786	0.782	0.713	0.562	0.783
3	variable 1 + variable 3	0.756	0.800	0.384	0.565	0.815
4	variable 1 + variable 4	0.394	0.227	0.424	0.572	0.417
5	variable 1 + variable 5	0.380	0.235	0.383	0.560	0.408
6	variable 1 + variable 6	0.240	0.197	0.738	0.741	0.445
7	variable 1 + variable 2 + variable 3	0.787	0.810	0.740	0.567	0.816
8	variable 1 + variable 2 + variable 4	0.808	0.785	0.713	0.578	0.824
9	variable 1 + variable 2 + variable 5	0.801	0.790	0.767	0.567	0.821
10	variable 1 + variable 2 + variable 6	0.801	0.808	0.713	0.748	0.868
11	variable 1 + variable 2 + variable 3 + variable 4	0.810	0.827	0.740	0.580	0.831
12	variable 1 + variable 2 + variable 3 + variable 5	0.802	0.846	0.740	0.570	0.828
13	variable 1 + variable 2 + variable 3 + variable 6	0.802	0.838	0.794	0.763	0.901
14	variable 1 + variable 2 + variable 3 + variable 4 + variable 5	0.815	0.865	0.740	0.619	0.831
15	variable 1 + variable 2 + variable 3 + variable 4 + variable 6	0.832	0.849	0.797	0.769	0.924
16	variable 1 + variable 2 + variable 3 + variable 4 + variable 5 + variable 6	0.836	0.889	0.799	0.798	0.913





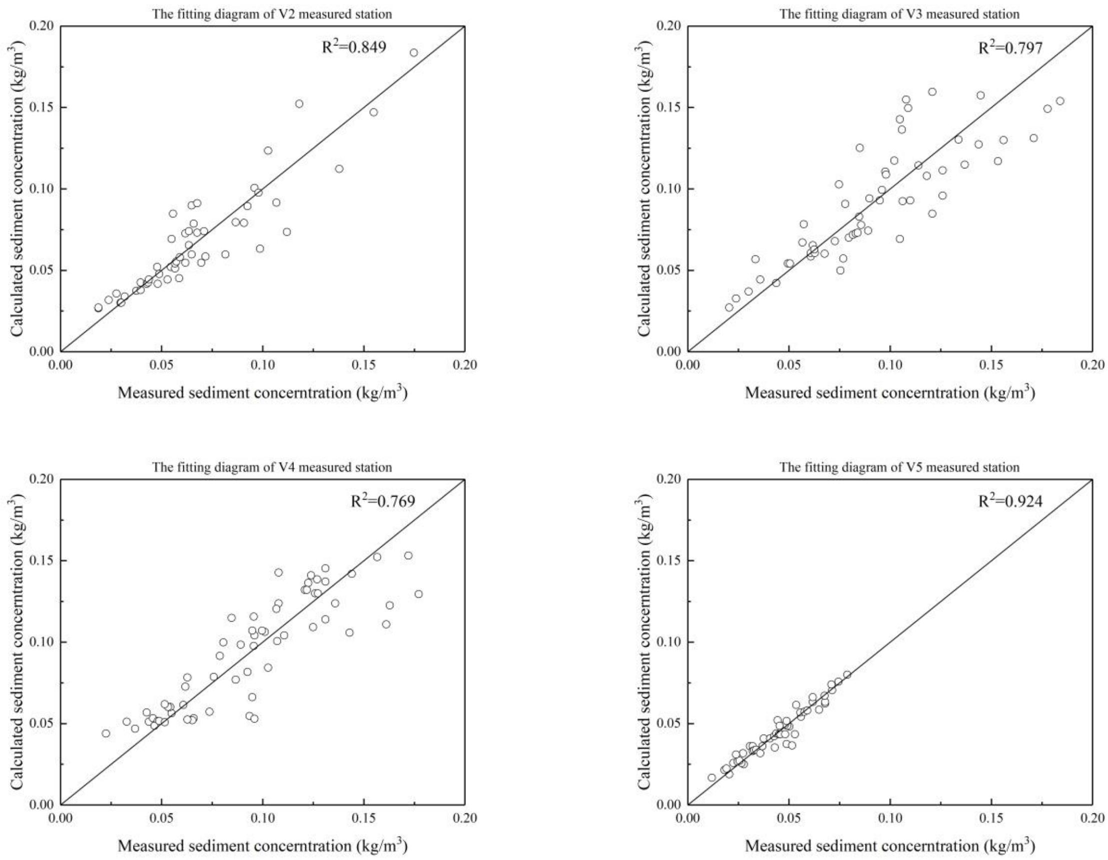


Fig. 5 The correlation between calculated and measured sediment vertical concentration.

5. Conclusions

The use of the Rouse Formula under complex dynamics in the estuary and coastal areas has very strict restrictions. There are not many basic theories that can be relied on to determine the vertical distribution of fine-grained sediment concentration in the estuary and coastal areas. In this paper, the sediment concentration is regarded as a "random variable" by means of sediment movement mechanics theory and multiple linear regression method. According to the linear form of Rouse equation and the transport characteristics of nearshore sediment, $\ln\frac{z}{h}$, $\ln c_a$, c_a, u, $\ln u$, $\frac{z}{h}$ are selected as independent variables. The multiple linear regression method was used to analyze the influence of each variable on the sediment concentration. According to the significance test method, the smaller influence factor ($\ln u$) of the six independent variables on the sediment concentration was eliminated. The correlation coefficient between the calculated sediment concentration and the measured sediment concentration shows that the independent variables used can reflect the characteristics of the vertical distribution of fine sediment concentration in the nearshore under complex dynamic conditions.

Acknowledgements

Hanbo Chen, Jia Ding and Zhujun Lin, all postgraduate students of Jiangsu Ocean university, are thanked for their help during data analysis activities.

References

[1] Wiechen, V. P., Vries, D. S., and Reniers, A. 2024. "Field Observations of Wave-Averaged Suspended Sediment Concentrations in the Inner Surf Zone with Varying Storm Conditions." *Marine Geology* 473: 107302.

[2] Ji, H. Y., Chen, S. L., Li, P., Pan, S. Q., Gong, X. L., and Jiang, C. 2024. "Spatiotemporal Variability of Suspended Sediment Concentration in the Coastal Waters of Yellow River Delta: Driving Mechanism and Geomorphic Implications." *Marine Geology* 470: 107266.

[3] Zhang, W. H., and Ran, J. 2023. "Tracking and Monitoring of Suspended Sediment Diffusion Generated during the Trenching of Submarine Pipeline in Northern Beibu Gulf." *Ships and Offshore Structures* 18 (11): 1528-34.

[4] Snehasis, K., and Ranjan, R. S. 2022. "A Unified Model of Suspension Concentration Distribution in Sediment Mixed Turbulent Flows Using Generalized Fractional Advection-Diffusion Equation." *Fluid Dynamics Research* 54 (6): 065505.

[5] Traykovski, P., Trowbridge, J., and Kineke, G. 2015. "Mechanisms of Surface Wave Energy Dissipation over a High-Concentration Sediment Suspension." *Journal of Geophysical Research: Oceans* 120 (3): 1638-81.

[6] Peng, Y., Yu, X. L., Yan, H. X., and Zhang, J. P. 2020. "Stochastic Simulation of Daily Suspended Sediment Concentration Using Multivariate Copulas." *Water Resources Management* 34 (12): 3913-32.

[7] Henderson, F. M. 1966. *Open Channel Flow*. New York: MocMillan Co., pp. 85-6.

[8] Li, R. J., Luo, F., and Zhu, W. J. 2009. "The Suspended Sediment Transport Equation and Its Near-Bed Sediment Flux." *Science in China Series E: Technological Sciences* 52 (2): 387-91.

[9] Maa, J., and Kwon, J. 2007. "Using ADV for Cohesive Sediment Settling Velocity Measurements." *Estuarine, Coastal and Shelf Science* 73 (1): 351-4.

[10] Quan, X. F., Li, R. J., Li, Y. T., Luo, F., Fu, X. Y., and Gou, H. 2022. "The Influence of Vertical Velocity Distribution on the Calculation of Suspended Sediment Concentration. Discrete Dynamics in Nature and Society." *Discrete Dynamics in Nature and Society* 2022: 1-15.

[11] Rouse, H. 1937. "Modern Conceptions of the Mechanics of Fluid Turbulence." *Transactions of the American Society of Civil Engineers* 102: 463-505.

[12] Eggenhuisen, T. J., Tilston, C. M., Leeuw, J., Pohl, F., and Cartigny, B. J. M. 2020. "Turbulent Diffusion Modelling of Sediment in Turbidity Currents: An Experimental Validation of the Rouse Approach." *The Depositional Record* 6 (1): 203-16.

[13] Snehasis, K. 2018. "Suspension Concentration Distribution in Turbulent Flows: An Analytical Study Using Fractional Advection-Diffusion Equation." *Physica A: Statistical Mechanics and Its Applications* 506: 135-55.

[14] Van Rijn. 1993. *Principles of Sediment Transport in Rivers, Estuaries and Coastal Seas.* Amsterdam, the Netherlands: Aqua Publications, p. 715.

[15] Yuan, Y., Wei, H., Zhao, L., and Jiang, W. 2008. "Observations of Sediment Resuspension and Settling Off the Mouth of Jiaozhou Bay, Yellow Sea." *Continental Shelf Research* 28 (19): 2630-43.

[16] Bouchez, J., Métivier, F., Lupker, M., Maurice, L., Perez, M., Gaillardet, J., and France-Lanord, C. 2011. "Prediction of Depth-Integrated Fluxes of Suspended Sediment in the Amazon River: Particle Aggregation as a Complicating Factor." *Hydrological Processes* 25 (5): 778-94. doi: 10.1002/hyp.7868.

[17] Lupker, M., France-Lanord, C., Lavé, J., Bouchez, J., Galy, V., Métivier, F., Gaillardet, J., Lartiges, B., and Mugnier, J. L. 2011. "A Rouse-Based Method to Integrate the Chemical Composition of River Sediments: Application to the Ganga Basin." *Journal of Geophysical Research* 116: 24.

[18] Chen, C., Yao, Z. S., Gang, Y. W., and Zhang, J. S. 2012. "An Approximation of the Improved Rouse Equation." *Applied Mechanics and Materials* 256-259: 2480-5.

[19] Koichi, Y., Shota, N., Masahiko, S., Tsuyoshi, I., Takaya, H., Ariyo, K., Takaharu, H., and Katsuhide, Y. 2013. "Dependence of the Settling Velocity Distributions of Suspended Coastal Sediment on Salinity." *Proceedings of Civil Engineering in the Ocean* 69 (2): 916-21.