

多光谱融合的海洋沉积物碳含量检测

李雪莹^{1,3}, 李宗民^{2*}, 侯广利³, 邱慧敏³, 吕红敏³, 陈光源⁴, 范萍萍^{3*}

1. 中国石油大学(华东)地球科学与技术学院, 山东 青岛 266580
2. 中国石油大学(华东)计算机科学与技术学院, 山东 青岛 266580
3. 齐鲁工业大学(山东省科学院), 山东省科学院海洋仪器仪表研究所, 山东 青岛 266061
4. 山东科技大学海洋科学与工程学院, 山东 青岛 266590

摘要 海洋沉积物中碳的变化是衔接海洋生态系统的过去与未来的信息桥梁, 揭示了海洋生态过程变化规律。因此开展海洋沉积物碳含量的研究, 对掌握海洋生态系统碳循环规律, 研究全球碳循环, 研究对气候变化的响应和反馈有着重要的作用。光谱技术是一种快速、无损的测量方法, 在定量分析中已有很成熟的应用。多光谱融合通过将多个光谱数据结合在一起, 获得比单一光谱更丰富的信息, 有利于物质的分析研究。将多光谱融合应用于海洋沉积物碳含量的研究, 以青岛海洋潮间 161 份沉积物为样品, 分别采用海洋光学 QE65000 光谱仪(光谱仪 1)和 AVANTES 光纤光谱仪 AvaSpec-ULS2048(光谱仪 2)采集沉积物可见-近红外光谱。将两种光谱仪的光谱进行多光谱融合, 分别采用偏最小二乘回归算法(PLSR)和 BP 神经网络算法(BPNN)建立沉积物碳含量模型。在 PLSR 沉积物碳含量建模结果中, 多融合光谱结果优于光谱仪 2, 略低于光谱仪 1, RPD 值为 1.968; 在 BPNN 沉积物碳含量建模结果中, 多融合光谱结果优于两个单光谱仪, RPD 值为 2.235。将多光谱融合后的光谱划分多个波段, 分别建立沉积物碳含量模型, 寻找沉积物碳的特征波段。通过分析多光谱融合各波段模型结果, 560~790 nm 的建模效果最好, R_p^2 为 0.949, RMSEC 为 0.550, R_p^2 为 0.874, RMSEP 为 0.733, RPD 值为 2.823。预测效果相较于光谱仪 1、光谱仪 2、多光谱融合全波段都有了显著的提高。因此采用多融合光谱特征波段建立海洋沉积物碳含量模型, 能够提高海洋沉积物碳含量的预测结果, 建立准确度更高的沉积物碳模型, 为沉积物碳的快速测定打下基础。

关键词 海洋沉积物; 多光谱融合; 碳; 光谱技术

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)09-2898-06

引言

海洋沉积物是大气圈、水圈、生物圈和岩石圈相互作用的介质和载体, 记录了海陆变迁、沉积环境变化等过程的详细信息^[1]。海洋系统中碳的含量非常丰富, 经过一系列的碳循环, 碳信息最终记录在沉积物中。海洋沉积物中碳的变化是衔接海洋生态系统的过去与未来的信息桥梁, 揭示了海洋生态过程变化规律^[2]。因此开展海洋沉积物碳含量的研究, 对掌握海洋生态系统碳循环规律, 研究全球碳循环, 研究对气候变化的响应和反馈有着重要的作用。

传统的测试方法费时、费力。目前碳含量分析可以使用元素分析仪测定, 但仍需烘干、粉碎等前处理。光谱技术是

一种快速、无损的测量方法, 已经应用于业、化工、食品等多个领域^[3-4]。在土壤养分含量分析中, 也取得了一定的成果^[5-6]。可见-近红外光谱包含大量的物质信息, 主要反映了样品的 O—H, C—H 等官能团的种类和数量。通过算法和分析获得样品特征光谱信息并解译, 测定其含量。目前沉积物碳含量检测的研究多为针对湖泊沉积物, 而对海洋沉积物碳含量的研究较少, 海洋沉积物的成分更为复杂, 在分析上更有难度, Jacq 等利用短波红外高光谱影像建立法国布尔吉湖岩芯沉积物样本碳偏最小二乘回归模型, 预测相关系数 r 为 0.95^[7]。

以青岛鳌山湾海洋潮间 161 份沉积物为样品, 分别采用海洋光学 QE65000 光谱仪和 AVANTES 光纤光谱仪 AvaSpec-ULS2048 采集沉积物可见-近红外光谱。将两种光

收稿日期: 2020-06-05, 修订日期: 2020-10-04

基金项目: 国家自然科学基金项目(3130041)和山东省自然科学基金项目(ZR2018LD007, ZR2017BB037, ZR2019PD004, ZR2018ZB0523)资助

作者简介: 李雪莹, 女, 1990 年生, 中国石油大学(华东)地球科学与技术学院博士研究生 e-mail: 412973984@qq.com

* 通讯作者 e-mail: lizongmin@upc.edu.cn; fanpp_sdioi@126.com

谱仪测得的光谱进行多光谱融合,分别采用偏最小二乘回归和BP神经网络建立沉积物碳含量模型。多光谱融合通过将多个光谱数据结合在一起,获得比单一光谱更丰富的信息,有利于对物质进行分析。将多光谱融合模型预测结果与单一光谱沉积物碳含量预测结果进行对比分析。对多光谱融合后的光谱进行特征波长筛选,寻找沉积物碳的特征波段,建立准确度更高的沉积物碳模型,实现对沉积物碳的快速测定。

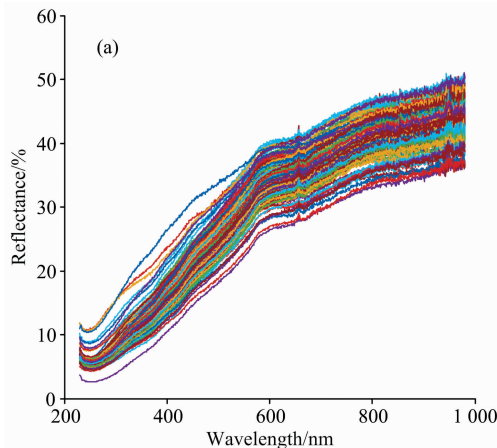
1 实验部分

1.1 材料

采样地点为青岛鳌山湾海洋潮间带两个地点的沉积物样品S1和S2,共计161个沉积物样品。采样深度为0~10 cm,每个采样点间隔10 m。潮间带沉积物组成以黏土和粉砂为主。将采集的样品风干研磨后,过60目筛,分成两部分,分别用于碳含量分析和光谱采集。沉积物碳含量采用重铬酸钾氧化法测定,S1和S2的碳含量见表1。

1.2 光谱数据采集

沉积物光谱1采用海洋光学QE65000光谱仪搭配DH-



2000-BAL型光谱采集样品的反射光谱,光谱采样间隔为1 nm,积分时间600 ms,谱区范围200~1 000 nm。光谱仪和光源通过Y型光纤连接,Y型光纤探头依靠支架固定,将样品放在自制样品盒中,探测样品的反射光谱。每个样品测定5次光谱反射率,取平均值作为这个样品的反射光谱。由于光谱的前段和后端受噪声影响,因此沉积物光谱1反射取220~980 nm,如图1(a)所示。

表1 沉积物碳($\text{g} \cdot \text{kg}^{-1}$)含量统计表
Table 1 The statistical results of carbon ($\text{g} \cdot \text{kg}^{-1}$) content in soil samples

沉积物样品	样品个数	最大值	最小值	平均值	标准差
S1	120	7.404	0.860	2.077	0.892
S2	41	9.487	5.177	7.001	0.853
S1+S2	161	9.487	0.860	3.331	2.320

沉积物光谱2采用AVANTES光纤光谱仪AvaSpec-ULS2048,谱曲范围160~1 100 nm,采用同样的方法测得光谱能量值,见图1(b)。

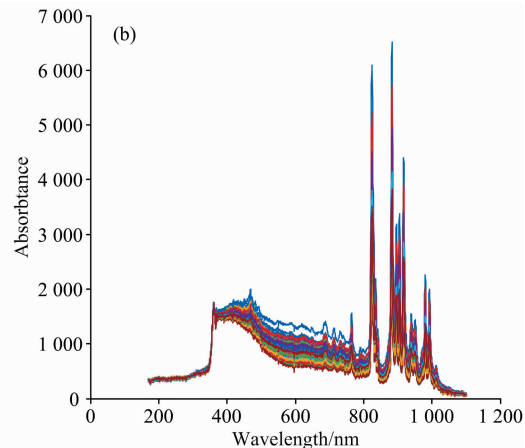


图1 沉积物光谱1和沉积物光谱2

Fig. 1 The spectra 1 and spectra 2 of sediment samples

1.3 多光谱融合

多光谱图像融合是将在空间上存在冗余或互补的图像数据,按照一定的规则进行处理,从而得到比单一图像更准确的信息^[8]。多光谱融合通过将多个光谱数据结合在一起,获得比单一光谱更丰富的信息,有利于对物质进行分析。两种光谱仪的波段均为200~1 000 nm,但包含的具体信息有所差异。将这两个光谱仪的光谱数据融合处理,获取更多更全的光谱信息,能够更加完整的表征样品的信息。

由于光谱仪1获取的是光谱反射率,而光谱仪2获取的是光谱能量值,因此先将两个光谱仪的光谱数据归一化处理,然后再将光谱仪1和光谱仪2的光谱混合在一起,得到多光谱融合光谱。融合光谱具备光谱仪1和光谱仪2的所有光谱值,主要集中在220~980 nm。因此,多融合光谱包括了更多的光谱信息。

1.4 碳含量建模方法

偏最小二乘回归(PLSR)是将相关分析、多元线性回归和主成分的优点集合在一起,在计算过程中同时考虑自变量(光谱数据)和因变量矩阵(化学参考值)对建模效果的影响,能够较好地处理数据多重共线性、因子结果不确定性和数据非正态分布等问题^[9]。通过降维运算得到潜在变量,达到消除无用信息的目的。在已知光谱数据和化学值数据的基础上,通过偏最小二乘回归建立相关模型,实现对未知样品的预测。该方法是目前最常用的化学计量学建模方法之一。

BP神经网络(BPNN)是一种基于误差反向传播算法的多层前向神经网络,通过反向传播不断调整网络的权值和阈值,使网络的误差平方和最小^[10]。反向传播神经网络运行过程分为两个部分:一是正向计算过程,从样本中选取信息,从输入层通过隐含层计算出各单元的输出值;二是误差反向过程,误差由输出层计算,隐含层各元素的误差逐层计算,并修改前一层的权重值。

1.5 模型评价标准

为保证每次划分建模集和检验集均为一致,采用顺序分类按 2:1 比例划分(即按顺序取两个沉积物样品为建模集,再取一个沉积物样品为检验集,依次类推),分别采用 PLSR 和 BPNN 建立土壤碳含量校正模型,并预测检验集沉积物碳含量。通过建模集绝对系数 R_c^2 、校正均方根误差 RMSEC、检验集绝对系数 R_p^2 、预测均方根误差 RMSEP 和相对分析误差 RPD 对模型进行评价, R_c^2 , R_p^2 , RPD 越大,且 RMSEC, RMSEP 越小,则模型效果越好^[11]。

2 结果与讨论

2.1 两种光谱仪及多光谱融合后建模结果-PLSR

采用顺序分类划分建模集和检验集,以 PLSR 建立光谱仪 1、光谱仪 2、多光谱融合的海洋沉积物碳含量光谱模型,模型评价结果见表 2。

表 2 海洋沉积物碳含量光谱模型评价结果-PLSR

Table 2 Model evaluation of carbon content detection in marine sediments-PLSR

光谱数据	建模集		检验集		
	R_c^2	RMSEC	R_p^2	RMSEP	RPD
光谱仪 1	1.000	0.039	0.786	1.018	2.032
光谱仪 2	0.688	1.363	0.575	1.433	1.444
多融合光谱	0.880	0.845	0.753	1.052	1.968

由表 2 可知,采用 PLSR 建立模型的建模集的评价结果由好到坏依次为光谱仪 1、多融合光谱、光谱仪 2,检验集的评价结果与建模集一致。光谱仪 1 的 R_c^2 和 RPD 值均略高于多融合光谱, RMSEP 略低于多融合光谱。

2.2 两种光谱仪及多光谱融合后建模结果-BPNN

同样采用顺序分类划分建模集和检验集,以 BPNN 分别建立光谱仪 1、光谱仪 2、多光谱融合的海洋沉积物碳含量光谱模型,模型评价结果见表 3。

表 3 海洋沉积物碳含量光谱模型评价结果-BPNN

Table 3 Model evaluation of carbon content detection in marine sediments-BPNN

光谱数据	建模集		检验集		
	R_c^2	RMSEC	R_p^2	RMSEP	RPD
光谱仪 1	0.961	0.492	0.809	0.936	2.212
光谱仪 2	0.894	0.814	0.690	1.176	1.760
多融合光谱	0.944	0.596	0.814	0.926	2.235

由表 3 可知,采用 BPNN 建立碳含量模型的建模集的评价结果与 PLSR 建模结果一致,依次为光谱仪 1、多融合光谱、光谱仪 2。而检验集的评价结果有所差异,评价效果最好的为多融合光谱, R^2 和 RPD 值均略高于光谱仪 1 的值,分别为 0.814 和 2.235,可用于沉积物碳含量的定量分析。

2.3 多融合光谱寻找沉积物碳的特征波段

采用 PLSR 建模来寻找沉积物碳的特征波段。多融合光谱的波段范围取 180~1100 nm,每间隔 100 nm 作为一个光谱波段,以 PLSR 建立各光谱波段沉积物碳模型评价结果见表 4。

表 4 海洋沉积物各光谱波段碳含量光谱模型评价结果

Table 4 Model evaluation of carbon content detections in marine sediments for various spectral bands

光谱波段 /nm	建模集		检验集		
	R_c^2	RMSEC	R_p^2	RMSEP	RPD
180~280	0.111	2.302	0.190	1.901	1.088
280~380	0.221	2.155	0.312	1.755	1.179
380~480	0.444	1.820	0.442	1.534	1.349
480~580	0.555	1.628	0.486	1.510	1.370
580~680	0.806	1.075	0.718	1.101	1.879
680~780	0.827	1.014	0.781	0.972	2.129
780~880	1.000	0.011	0.683	1.184	1.748
880~980	0.546	1.646	0.512	1.440	1.437
980~1100	0.017	2.420	0.058	1.992	1.039

由表 4, 580~680 和 680~780 nm 波段的建模集和检验集评价结果最好, R_c^2 均在 0.8 以上, R_p^2 在 0.7 以上,其他波段的结果较差。因此,可判定海洋沉积物碳的特征波段在 580~780 nm 附近。在 580~780 nm 附近寻找沉积物碳的特征波段,将 780 nm 固定,以 580 nm 为中点,上下各取 50 nm 波段,间隔为 10 nm,寻找特征波段的下限,得到模型评价结果,见表 5。

表 5 海洋沉积物碳含量光谱模型评价结果

(530~780 nm 到 630~780 nm)

Table 5 Model evaluation of carbon content detection in marine sediments (from 530~780 nm to 630~780 nm)

光谱波段 /nm	建模集		检验集		
	R_c^2	RMSEC	R_p^2	RMSEP	RPD
530~780	1.000	0.046	0.833	0.860	2.407
540~780	0.948	0.556	0.862	0.774	2.672
550~780	0.950	0.547	0.867	0.756	2.736
560~780	0.949	0.552	0.868	0.751	2.754
570~780	0.944	0.579	0.855	0.794	2.607
580~780	0.943	0.585	0.854	0.796	2.601
590~780	0.942	0.590	0.852	0.802	2.580
600~780	0.938	0.607	0.851	0.804	2.573
610~780	0.937	0.614	0.845	0.819	2.527
620~780	0.934	0.627	0.835	0.845	2.448
630~780	0.930	0.648	0.828	0.868	2.385

由表 5, 560~780 nm 的模型效果最好, R_p^2 为 0.868, RPD 值为 2.754,因此将 560 nm 作为海洋沉积物碳的特征波段上限。将 560 nm 固定,以 780 nm 为中点,上下各取 50 nm 波段,间隔为 10 nm,寻找特征波段的下限,得到模型评

价结果, 见表 6。

由表 6 可知, 560~790 nm 的模型效果最好, R_p^2 为 0.874, RPD 值为 2.823, 将 790 nm 作为海洋沉积物碳的特征波段下限。因此海洋沉积物碳的特征波段为 560~790 nm, 以 PLSR, BPNN 建模的多光谱融合 560~790 nm 拟合结果见图 2(a)和(b)。

由图 2, PLSR, BPNN 建模的多光谱融合 560~790 nm 拟合结果, 两种建模方法得到评价结果基本相同, PLSR 略好一些, R_c^2 在 0.95 左右, R_p^2 在 0.85 以上, 均很好的预测沉积物碳含量。

2.4 讨论

在光谱仪 1、光谱仪 2、多光谱融合全波段建立模型中, BPNN 的模型效果均优于 PLSR 结果。由于 BPNN 是一种非线性的学习机制, 数据越多, 学习能力越强, 建立的模型越好; 而 PLSR 是一种线性的建模方法, 数据量越大, 出现的数据冗余会对模型结果又一定的影响。在多融合光谱全波段 BPNN 模型中, 其结果优于两种单光谱仪, 多融合光谱数据

表 6 海洋沉积物碳含量光谱模型评价结果 (560~730 nm 到 660~830 nm)

Table 6 Model evaluation of carbon content detection in marine sediments (from 560~730 nm to 660~830 nm)

光谱波段 /nm	建模集		检验集		
	R_c^2	RMSEC	R_p^2	RMSEP	RPD
560~730	0.940	0.599	0.862	0.778	2.660
560~740	0.940	0.597	0.874	0.739	2.799
560~750	0.993	0.203	0.870	0.742	2.789
560~760	0.995	0.179	0.855	0.788	2.625
560~770	0.948	0.558	0.862	0.767	2.697
560~780	0.949	0.552	0.868	0.751	2.754
560~790	0.949	0.550	0.874	0.733	2.823
560~800	0.951	0.542	0.869	0.751	2.757
560~810	0.955	0.518	0.860	0.778	2.661
560~820	0.959	0.494	0.855	0.785	2.637
560~830	0.936	0.616	0.834	0.846	2.447

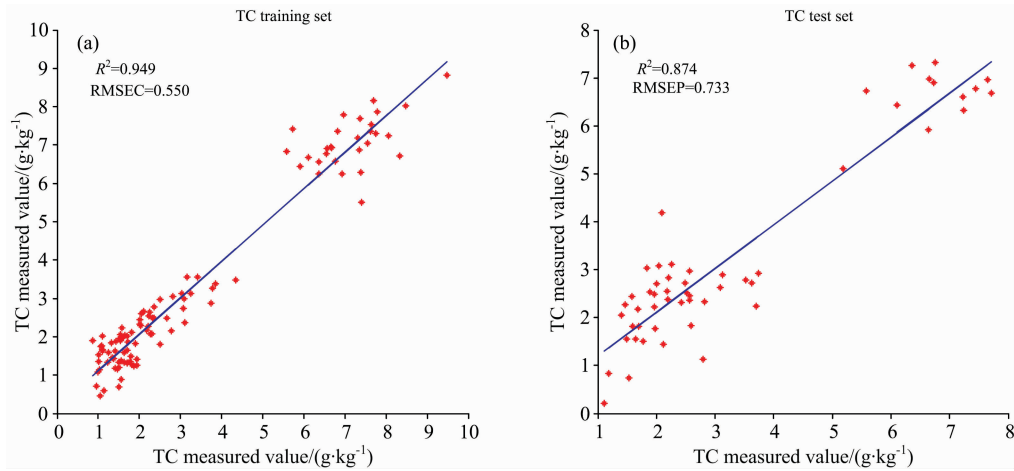


图 2(a) 多光谱融合 560~790 nm 拟合结果-PLSR

Fig. 2(a) Fitting results of multispectral fusion in 560~790 nm range-PLSR

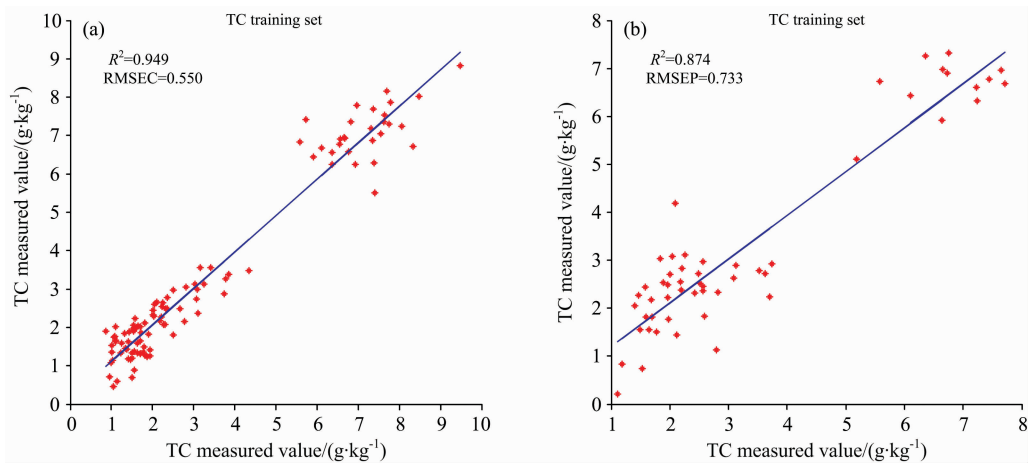


图 2(b) 多光谱融合 560~790 nm 拟合结果-BPNN

Fig. 2(b) Fitting results of multispectral fusion in 560~790 nm range-BPNN

量更大, 所含有信息更多, 更有利于 BPNN 学习和训练, 因此得到更好的预测结果。

在寻找海洋沉积物碳最优波段中, 通过尝试不同波段的评价结果, 获得了沉积物碳的特征波段, 在 560~790 nm。多融合光谱 560~790 nm 波段中包含了两种单光谱仪该波段的全部信息, 含有的光谱更加丰富, 能够表征更多的沉积物碳的信息。采用多融合光谱特征波段建立的海洋沉积物碳模型结果优于各单光谱仪全波段和多融合光谱全波段的模型结果。因此采用多融合光谱特征波段建立海洋沉积物碳含量模型, 能够提高海洋沉积物碳含量的预测结果。

3 结 论

以青岛鳌山湾海洋潮间 161 份沉积物为样品, 分别采用海洋光学 QE65000 光谱仪(光谱仪 1)和 AVANTES 光纤光

谱仪 AvaSpec-ULS2048(光谱仪 2)采集沉积物可见-近红外光谱。将两种光谱仪的光谱进行多光谱融合, 分别采用 PLSR 和 BPNN 建立沉积物碳含量模型。PLSR 建模结果, 多融合光谱结果优于光谱仪 2, 略低于光谱仪 1, RPD 值为 1.968; BPNN 建模结果, 多融合光谱结果优于两个单光谱仪, RPD 值为 2.235。将多光谱融合后的光谱划分多个波段, 寻找沉积物碳的特征波段。通过分析多光谱融合各波段模型结果, 560~790 nm 的建模效果最好, R_c^2 为 0.949, RMSEC 为 0.550, R_p^2 为 0.874, RMSEP 为 0.733, RPD 值为 2.823。预测效果相较于光谱仪 1、光谱仪 2、多光谱融合全波段都有了显著的提高。因此采用多融合光谱特征波段建立海洋沉积物碳含量模型, 能够提高海洋沉积物碳含量预测结果的准确度, 建立准确度更高的沉积物碳模型, 对沉积物碳的快速测定打下基础。

References

- [1] Luo T, Song Y, Zhu Y, et al. *Marine and Petroleum Geology*, 2016, 77: 507.
- [2] Dang H Y, Jiao N. *Biogeosciences*, 2014, 11: 3887.
- [3] Naseer N, Noori F M, Qureshi N K, et al. *Frontiers in Human Neuroscience*, 2016, 10: 237.
- [4] Hong K S, Santosa H. *Hearing Research*, 2016, 333: 157.
- [5] Dinakaran J, Bidalia A, Kumar A, et al. *Communications in Soil Science and Plant Analysis*, 2016, 47(12): 1503.
- [6] Wetterlind J, Stenberg B, Söderström M. *Geoderma*, 2010, 156: 152.
- [7] Jacq K, Perrette Y, Fanget B, et al. *The Science of the Total Environment*, 2019, 663: 236.
- [8] Bai L, Xu C, Wang C. *Optik International Journal for Light & Electron Optics*, 2015, 126(24): 4804.
- [9] Wold S, Sjöström M, Eriksson L. *Chemometrics & Intelligent Laboratory Systems*, 2001, 58(2): 109.
- [10] Ghose D K, Panda S S, Swain P C. *Journal of Hydrology*, 2010, 394(3-4): 296.
- [11] Douglas R K, Nawar S, Alamar M C, et al. *Science of the Total Environment*, 2017, 616-617: 147.

Carbon Content Detection of Marine Sediments Based on Multispectral Fusion

LI Xue-ying^{1,3}, LI Zong-min^{2*}, HOU Guang-li³, QIU Hui-min³, LÜ Hong-min³, CHEN Guang-yuan⁴, FAN Ping-ping^{3*}

1. School of Geosciences, China University of Petroleum (East China), Qingdao 266580, China

2. College of Computer Science and Technology, China University of Petroleum (East China), Qingdao 266580, China

3. Institute of Oceanographic Instrumentation, Qilu University of Technology (Shandong Academy of Sciences), Qingdao 266061, China

4. College of Ocean Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China

Abstract The change of carbon in marine sediment is the information bridge between the past and the future of the marine ecosystem, which reveals the law of the marine ecological process. Therefore, the research on the carbon content of marine sediments plays an important role in mastering the carbon cycle law of the marine ecosystems, studying the global carbon cycle, and studying the response and feedback to climate change. Spectrum technology is a fast and non-destructive measurement method, which has been widely used in quantitative analysis. Multispectral fusion based on spectral technology, through the combination of multiple spectral data, we can get more information than a single spectrum, which is conducive to the analysis of substances. In this paper, multispectral fusion was applied to the study of carbon content in marine sediments. 161 samples of Qingdao, China intertidal sediments are taken as samples. The visible near infrared spectra of sediments were collected by QE65000 spectrometer (spectrometer 1) and AVANTES optical fiber spectrometer (AvaSpec-ULS2048) (spectrometer 2),

respectively. The spectra of the two spectrometers were fused, PLSR and BPNN were used to establish the carbon content model. The results of PLSR modeling showed that the results of multi fusion spectrum were better than that of spectrometer 2, slightly lower than that of spectrometer 1, and the RPD value was 1.968. The results of BPNN modeling showed that the results of multi fusion spectrum were better than those of two single spectrometers, and the RPD value was 2.235. The spectra after multispectral fusion were divided into several bands to find the characteristic bands of carbon in sediments. By analyzing the results of the multispectral fusion model, 560~790 nm was the best, the R_c^2 was 0.949, the RMSEC was 0.550, the R_p^2 was 0.874, the RMSEP was 0.733, and the RPD was 2.823. Compared with spectrometer 1, spectrometer 2 and multispectral fusion, the prediction effect was significantly improved. Therefore, using the multi fusion spectral characteristic band to establish the model of the carbon content of marine sediments can improve the prediction results of the carbon content in marine sediments. It can establish a more accurate model of the organic carbon content of sediments, which will lay a foundation for the rapid determination of the organic carbon content of sediments.

Keywords Marine sediments; Multispectral fusion; Carbon; spectroscopy

(Received Jun. 5, 2020; accepted Oct. 4, 2020)

* Corresponding authors