

岩性不同的含水岩石近红外光谱特征对比分析

张芳^{1,2}, 卢佐乐^{1,2}, 王东升^{1,2}, 刘雨濛^{1,2}, 谢运鑫^{1,2}, 卓慧慧², 何满潮^{1*}

1. 中国矿业大学(北京)深部岩土力学与地下工程国家重点实验室, 北京 100083

2. 中国矿业大学(北京)力学与建筑工程学院, 北京 100083

摘要 对不同岩性的岩石构建近红外光谱含水量反演模型过程中,其特征集是否会因岩性不同而变化,是建模时必须面对的关键问题。针对该问题,首先利用深部软岩水力作用智能测试系统进行了岩石吸水室内实验,分别采集了砾岩、粉砂岩、夯土不同时刻3种不同岩性的含水岩石的近红外光谱各为51条、106条和149条;然后,采用一阶导数法对原始光谱进行预处理,消除环境干扰等对光谱的影响;接着,采用几何特征法提取光谱特征,进行归一化处理,消除量纲和变化幅度不同带来的影响;分析了初始特征变量之间以及初始特征变量与含水量之间的相关程度,并通过抑制阈值大小,剔除冗余特征,得到了两个强相关谱段处的特征变量;最后利用最大信息系数(MIC)值作为度量标准,对比分析了不同岩性的含水岩石近红外光谱的特征选择结果,以期评价岩性对含水岩石光谱特征的影响。结果表明:(1)砾岩、粉砂岩、夯土三种含水岩石的近红外光谱在1400和1900 nm附近都有着明显的吸收峰,且随着含水量的变化,吸收强度越来越强,与含水量大小有明显的相关性;(2)对比砾岩、粉砂岩、夯土近红外光谱特征变量与其含水量的最大相关系数MIC值,表明夯土的近红外光谱与含水量之间的相关性最强;(3)不同岩性的近红外光谱各个特征值与含水量的相关程度不同,具体表现为在1400 nm附近峰高、右肩宽与含水量都具有较高相关系数,只是相关性大小会因岩性不同而不同,1900 nm附近的右肩宽和峰面积与含水量都具有较高相关系数,且右肩宽的相关程度高于峰面积;(4)不同岩性的含水岩石介质近红外光谱的特征变量与含水量相关性具有相似规律:峰高、右肩宽、峰面积是相关程度最高的三个特性。

关键词 岩石;光谱特征;含水量;最大信息系数;相关性

中图分类号: P574.1 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)03-0971-09

引言

近红外光谱分析(near infrared spectroscopy technique, NIRS)是近几十年来发展最快,最引人注目的光谱分析技术之一,被广泛应用于食品、化工、农业等领域^[1]。在岩土工程领域应用近红外光谱分析检测、监测含水岩石水分状态,是近些年发展起来的新思路,其原理是根据岩土介质O—H基团的吸收强度,建立含水量与近红外光谱特征之间的定量关系,实现预测岩土介质中含水量的目的。与传统分析方法相比,该方法具有无损、实时、定量的优势。

对于含水量与光谱特征之间建模研究,诸多学者进行了大量的研究工作,尤其在土壤方面研究成果颇多:金慧凝等^[2]通过提取反射光谱特征指标,定量分析土壤含水量与反

射光谱特征之间关系,建立了土壤水分含量光谱预测模型。包青岭等^[3]利用包络线消除法提取反射光谱水分吸收特征,并与土壤含水量进行相关性分析,通过随机森林方法对光谱水分吸收特征参数进行分类,获取各参数对土壤含水量的重要性,运用多元逐步回归方法,建立土壤水分含量反演模型。娄径等^[4]通过对光谱数据进行倒数、对数、均方根及一阶导数微分等光谱变换,分析光谱特征,并与土壤含水量进行相关分析,利用多元线性回归分析建立土壤含水量监测模型。

可见,光谱特征选择是构建准确、稳健的定量模型的关键。国内学者对此进行了很多研究,如:孔清清等^[5]基于随机森林,结合博弈论进行了近红外光谱特征选择,利用互信息选择出无冗余的特征子集,此法应用于近红外光谱分类中有较高的分类识别率。

收稿日期:2019-01-09,修订日期:2019-04-18

基金项目:国家自然科学基金青年基金项目(51604276)资助

作者简介:张芳,1976年生,中国矿业大学(北京)高级工程师

* 通讯联系人 e-mail: 201203@cumtb.edu.cn

e-mail: zhangf76@163.com

利用信息度量法进行波谱的特征选择具有无参、非线性的优势,能有效度量两随机变量之间相关性,能较好地解决特征变量选择问题,该方法在特征选择算法中得到广泛应用。

但是,上述研究成果大多集中于研究土壤含水量与光谱特征之间的相关性,针对岩石的研究工作很少,更没有涉及岩性对含水岩石光谱特征选择的影响,即不同岩性,能否选择相同特征集,或是需要具体问题具体分析,针对不同现场选择对应的特征集合,这个问题对生产实践中构建模型时选择特征变量具有非常重要的指导意义。

故此,利用互信息作为相关程度的度量标准,对比分析不同岩性的含水岩石近红外光谱的特征选择结果,以期评价岩性对含水岩石光谱的影响。

1 基于互信息的特征选择

采用最大信息系数(maximal information coefficient, MIC)^[6]进行特征选择,其计算思想是:设任意的行数 x_i 列数 y_i 下的含水量 C 与特征参数 f 的散点图网格 $G_{x_i y_i}(j)$,计算最大互信息和最大信息系数

$$I_{\max}(C, f, x_i, y_i) = \max I((C, f) |_{G_{x_i y_i}(j)}), \quad j = 1, 2, \dots \quad (1)$$

$$\text{MIC} = \max \left\{ \frac{I_{\max}(C, f, x_i, y_i)}{\log \min(x_i, y_i)} \right\} \quad (2)$$

其中, $I_{\max}(C, f, x_i, y_i)$ 表示在固定行列数 x_i 与 y_i 情况下,不同的网格划分方式下的互信息最大值(最大互信息); $I((C, f) |_{G_{x_i y_i}(j)})$ 表示含水量 C 与特征参数 f 的散点图在网格 $G_{x_i y_i}(j)$ 下的互信息, $x_i, y_i < B(n)$, $B(n)$ 为网格分割细度^[7]。

最后选择满足预先给定的阈值的特征参数,组成近红外光谱特征集 S ,完成特征的选择。

2 实验部分

2.1 样品制备

砾岩、粉砂岩采自敦煌莫高窟北区的崖壁,样品信息如表 1、表 2 所示。因为该砾岩呈半胶结状态,不易加工成标准试件,实验时将砾岩加工成尺寸约为 80 mm×90 mm×60 mm 的不规则形状(图 1),使之能够满足实验仪器的放置要求。

表 1 样品基本信息

Table 1 Sample basic information

岩性	取样地点	岩性描述	颜色	干燥后质量/g	吸水后质量/g	尺寸/mm
砾岩	莫高窟北区	疏松多孔,块状构造	青灰色	922.60	968.46	约为 80×90×60
粉砂岩	莫高窟北区	断面粗糙,较硬	灰白色	371.94	383.79	φ50×85
夯土	莫高窟 108 洞室	泥质结构,疏松多孔	黄褐色	400.516	544.416	φ50×100

表 2 矿物成分信息

Table 2 Mineral composition information

岩性	石英/%	钾长石/%	钠长石/%	方解石/%	白云石/%	角闪石/%	石盐/%	黏土矿物总量/%
砾岩	30.4	2.2	23.5	11.8	9.7	2.0	6.9	13.5
粉砂岩	24.3	0.5	6.3	22.8	16.4	0.0	0.0	29.7
夯土	30.0	2.9	9.1	15.8	7.6	1.2	0.0	33.4

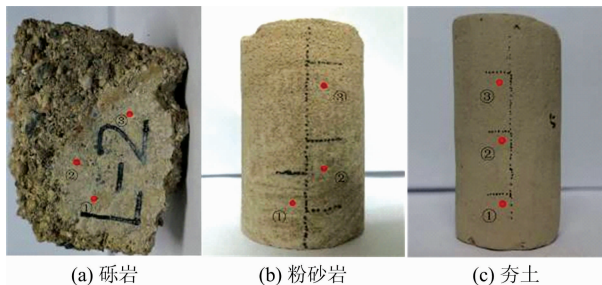


图 1 岩样及其测试点位置(红点)

(a): 砾岩; (b): 粉砂岩; (c): 夯土

Fig. 1 Rock sample and its test point location (red dot)

(a): Conglomerate; (b): Siltstone; (c): Rammed soil

夯土试样(表 1 和表 2)采自敦煌莫高窟 108 洞室,该夯土强度较低,无法直接用钻机取出土样,所以先取土块,然

后加工成长方体土样,之后再采用打磨的方法制成 φ50mm 的标准样品(图 1)。

实验中,试样不断吸水,水分自下而上运移,故沿着试样高度方向,选取 3 个近红外光谱测试点,并尽量避开签字笔标记区域,三种岩性样本的测试点位置如图 1 中红点位置。

1.2 原理及设备

将岩样放置真空干燥箱内,设置箱内温度 105~110 ℃,干燥 24 h,取出干燥后岩样冷却 12 h 称重,利用中国矿业大学(北京)深部岩土力学与地下工程国家重点实验室何满潮^[8]自主研发的“深部软岩水理作用智能测试系统”,进行岩样吸水模拟室内实验。该系统主要由主体实验箱、称重系统和数据采集系统三部分组成,如图 2 所示。

1.3 光谱采集

在试样吸水过程中,采用瑞士万通的 XDS SmartProbe

近红外光谱分析仪采集不同位置不同时刻的近红外光谱。测试时将光纤探头分别接触试样的 3 个测量点(图 1 红点)，自下往上依次测量，测量的频率随试样的吸水速率适时调整。实验参数如表 3，实验装置如图 3。



图 2 岩石吸水过程中的近红外光谱采集实验设备

Fig. 2 Near-infrared spectrum acquisition experimental equipment for determining water absorption of rock

表 3 近红外光谱分析仪的实验参数^[9]

Table 3 Experimental parameters of the near infrared spectroscopy analyzer^[9]

探头类型	反射探头
采样方式	漫反射-固体
光谱范围	400~2 500 nm
测样点个数	4
测样角度	90°
测样方式	直接与样品接触



图 3 XDS SmartProbe 近红外光谱分析仪

Fig. 3 XDS SmartProbe near infrared spectrum analyzer

整个实验，在砾岩从干燥到饱和的吸水全过程中，共采集了 51 条近红外光谱信息，分别为：1 号点 17 条，2 号点 18 条，3 号点 16 条。

在粉砂岩吸水全过程中，共采集了 106 条近红外光谱信息，分别为：1 号点 51 条，2 号点 34 条，3 号点 21 条。

在夯土吸水全过程中，共采集了 149 条近红外光谱信息，分别为：1 号点 24 条，2 号点 59 条，3 号点 66 条。

1.4 近红外光谱预处理

利用 XDS SmartProbe 近红外光谱分析仪配套软件提供

的一阶导数法对采集的光谱进行预处理，消除背景的常数平移对近红外光谱的影响，使数据具有更好的连续性，处理前后的光谱如图 4—图 6 所示，限于篇幅，仅列出砾岩、粉砂岩、夯土 1 号点的光谱图。

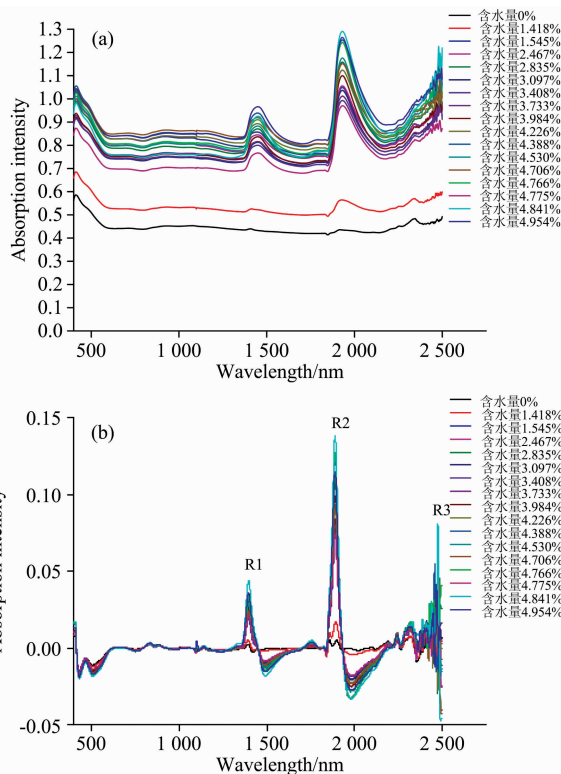


图 4 砾岩 1 号点近红外光谱

(a): 原始光谱; (b): 一阶导数预处理后光谱

Fig. 4 Near-infrared spectra of conglomerate at point 1

(a): Original; (b): First derivative pre-processed

1.5 谱段选择和初始特征变量

分析图 4—图 6 可知，在 400~2 500 nm 波长范围内有 3 个明显的吸收峰，分别在 1 400，1 900 和 2 300 nm 附近，其光谱反射率随试样含水量变化而变化，依次将其命名为峰 R₁、峰 R₂、峰 R₃。随着含水量的不断增大，R₁ 和 R₂ 两个吸收峰的波峰越来越高，峰 R₁ 中心点位置最终停留在 1 400 nm 左右，峰 R₂ 中心点位置最终停留在 1 900 nm 左右，而 R₃ 吸收峰的波峰随含水量增加逐渐减小，信号特征逐渐减弱，因 2 400 nm 之后的噪音干扰强烈，故峰 R₃ 不适合作为含水量信息的特征谱段。因此，选择峰 R₁、峰 R₂ 所在的 1 400 和 1 900 nm 谱段进行含水试样光谱特征分析，具体提取的特征变量如图 7 所示，分别为峰面积 (Area)、峰高 (Height)、半高宽 (FWHM)、左肩宽 (left half width, LHW)、右肩宽 (right half width, RHW)、左右肩宽比 (LHW/RHW) 共计 6 个初始特征参数，设定初始特征集 F 为 F = {f₁, f₂, f₃, f₄, f₅, f₆} = {Area, Height, FWHM, LHW, RHW, LHW/RHW}，各含水试样近红外光谱的初始特征数值如表 4 所示。

需要特别强调的是，在整个实验历程中，粉砂岩 3 号点

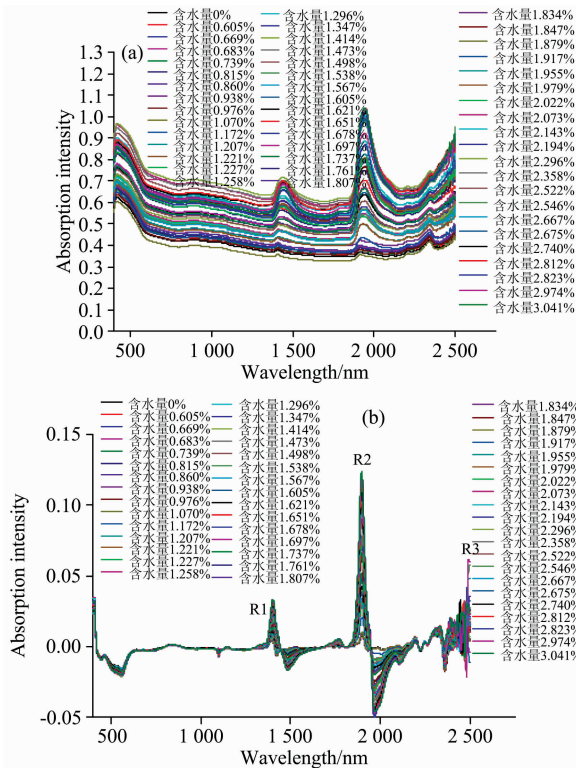


图 5 粉砂岩 1 号点近红外光谱

(a): 原始光谱; (b): 一阶导数预处理后光谱

Fig. 5 Near-infrared spectra of siltstone at point 1

(a): Original; (b): First derivative pre-processed

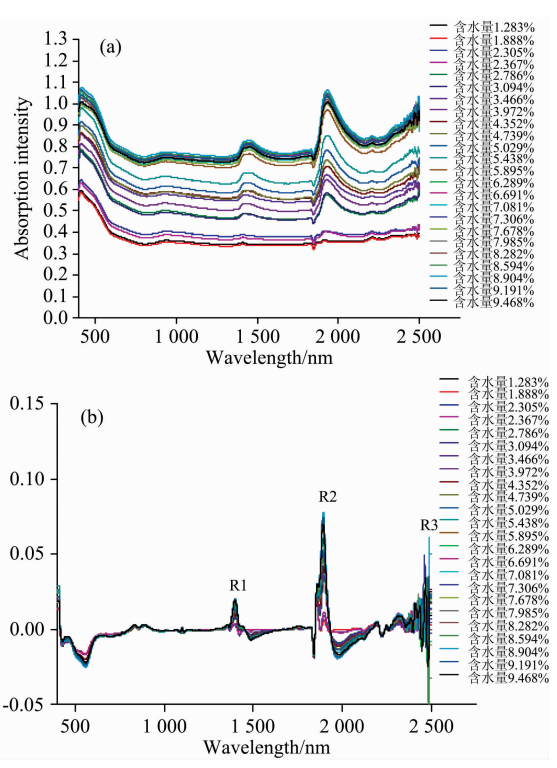


图 6 夯土 1 号点近红外光谱

(a): 原始光谱; (b): 一阶导数预处理后光谱

Fig. 6 Near-infrared spectra of rammed soil at point 1

(a): Original; (b): First derivative pre-processed

处采集的 21 条近红外光谱没有明显的吸收峰, 分析其原因。3 号点位于粉砂岩顶端, 当水分没有达到这个位置并浸润到它时, 该点始终处于干燥状态, 所以没有吸收峰。3 号点采集的 21 条光谱都没有采集到含水情况下的光谱, 为失效光谱。故在后续分析中将该组实验数据去掉。

1.6 特征变量归一化

分析表 4 可知, 由于 6 个初始特征变量的量纲不同, 且特征变量之间的变化幅度不同, 可能导致在分析计算过程中, 一些数量级较小的特征变量的作用无法体现, 因此对上述表中的原始数据进行归一化变换, 将所有变量转换成 0-1 内的数值, 消除量纲和变化幅度不同带来的影响。

归一化的方法是将原始数据矩阵的各元素减去该元素所在列的最小值后再除以该列元素的极差, 公式如下

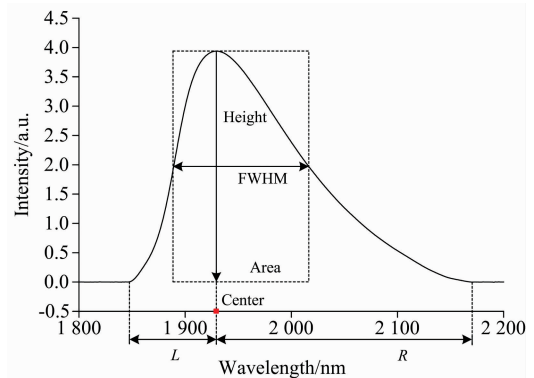


图 7 近红外光谱的初始特征变量几何意义示意图^[9]

Fig. 7 Schematic diagram of geometric meaning of initial characteristic variables of near-infrared spectroscopy^[9]

表 4 砾岩在峰 R_1 处的初始特征变量(只列出部分)

Table 4 Initial characteristic variables of conglomerate at the peak R_1 (only some data shown in the table)

近红外光谱	特征值 F						含水量 $C/\%$
	f_1/nm	f_2	f_3/nm	f_4/nm	f_5/nm	f_6	
O_{1-1}	0.075 06	0.003 12	24.115 56	12.650 07	11.465 49	1.103 32	0
O_{1-2}	0.154 18	0.005 53	27.745 35	14.025 93	13.719 42	1.022 34	1.418
O_{1-3}	0.702 56	0.019 60	32.274 25	15.517 77	16.756 48	0.926 08	1.545
O_{1-4}	0.756 08	0.021 26	32.283 11	15.474 91	16.808 20	0.920 68	2.467
O_{1-5}	1.085 27	0.028 36	34.134 17	16.244 76	17.889 41	0.908 07	2.835

续表 4

O ₁₋₆	0.858 17	0.023 00	33.442 99	15.644 15	17.798 84	0.878 94	3.097
O ₁₋₇	0.707 91	0.019 52	32.949 73	15.486 05	17.463 68	0.886 76	3.408
O ₁₋₈	0.757 88	0.020 17	33.442 24	16.000 75	17.441 49	0.917 40	3.733
O ₁₋₉	0.985 46	0.025 94	34.016 37	16.231 61	17.784 76	0.912 67	3.984
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
O ₃₋₁₆	1.492 56	0.039 00	34.051 89	16.336 99	17.714 90	0.922 22	4.971

$$x'_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)}$$

归一化结果如表 5 所示。

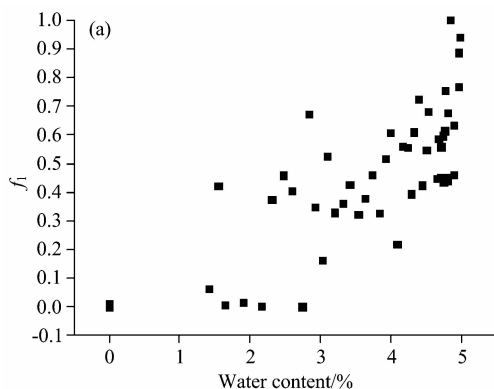
表 5 砾岩在峰 R₁ 处初始特征变量归一化值(只列出部分)

Table 5 Normalized values of initial characteristic variables of conglomerate at the peak R₁ (partial listing)

近红外光谱	特征值 F					
	f ₁	f ₂	f ₃	f ₄	f ₅	f ₆
O ₁₋₁	0.013 46	0.006 32	0.190 05	0.305 06	0.170 43	0.669 85
O ₁₋₂	0.065 19	0.029 74	0.452 66	0.564 39	0.409 94	0.469 75
O ₁₋₃	0.423 74	0.166 49	0.780 30	0.845 59	0.732 68	0.231 88
O ₁₋₄	0.458 73	0.182 62	0.780 94	0.837 51	0.738 18	0.218 54
O ₁₋₅	0.673 97	0.251 63	0.914 86	0.982 62	0.853 07	0.187 38
O ₁₋₆	0.525 48	0.199 53	0.864 86	0.869 41	0.843 45	0.115 41
O ₁₋₇	0.427 24	0.165 71	0.829 17	0.839 61	0.807 83	0.134 73
O ₁₋₈	0.459 91	0.172 03	0.864 80	0.936 62	0.805 47	0.210 43
O ₁₋₉	0.608 71	0.228 11	0.906 34	0.980 14	0.841 95	0.198 75
⋮	⋮	⋮	⋮	⋮	⋮	⋮
O ₃₋₁₆	0.940 27	0.355 04	0.908 91	1.000 00	0.834 53	0.222 35
均值	0.443 99	0.189 43	0.700 87	0.749 84	0.669 98	0.267 86
标准方差	0.249 14	0.148 66	0.269 65	0.251 80	0.257 85	0.216 78
最大值	1.000 00	1.000 00	1.000 00	1.000 00	1.000 00	1.000 00
最小值	0.000 00	0.000 00	0.000 00	0.000 00	0.000 00	0.000 00

2 结果与讨论

在进行特征选择之前,需要对初始特征集各特征变量之间、特征变量与含水量之间的相关性进行筛选,以便去掉冗余特征,本文参考文献[13]中采用的阈值及结论,取初始特征变量之间的相关系数的阈值为 0.95,初始特征变量与含水



量之间的相关系数的阈值为 0.5。则利用相关系数评价上述变量间的相关程度,选取的特征变量如表 6。

表 6 各试样在峰 R₁, 峰 R₂ 处满足相关系数阈值要求的特征变量

Table 6 Characteristic variables satisfying the correlation coefficient threshold at peaks R₁ and peaks R₂ for each sample

岩性	特征变量	
	R ₁	R ₂
砾岩	f ₁ , f ₅	f ₁ , f ₄
粉砂岩	f ₂ , f ₄ , f ₅	f ₁ , f ₅
夯土	f ₂ , f ₅	f ₁ , f ₄ , f ₅

2.1 砾岩的特征选择分析

将砾岩特征变量与含水量数据做成散点图,如图 8 所示。

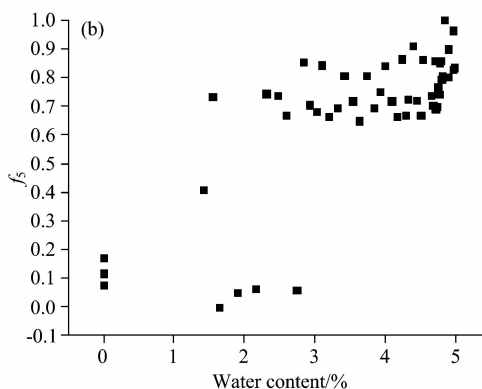
利用第 1 节中的公式,分别计算峰 R₁ 和 R₂ 处的 f₁, f₅, f₁, f₄ 与含水量 C 之间的 MIC 值,如表 7。

表 7 砾岩特征变量与含水量间的 MIC 值

Table 7 MIC values between characteristic variables and water content of conglomerate

吸收峰	MIC(C, f ₁)	MIC(C, f ₄)	MIC(C, f ₅)
峰 R ₁	0.549 43	/	0.452 73
峰 R ₂	0.536 65	0.400 06	/

由表 7 可知,砾岩在峰 R₁ 处有 MIC(C, f₁) > MIC(C, f₅),说明峰面积与含水量相关关系最强,右肩宽次之;对于峰 R₂ 有 MIC(C, f₁) > MIC(C, f₄),说明峰面积与含水量相关关系最强,左肩宽次之。表 7 中的 MIC 值位于 0.4 ~



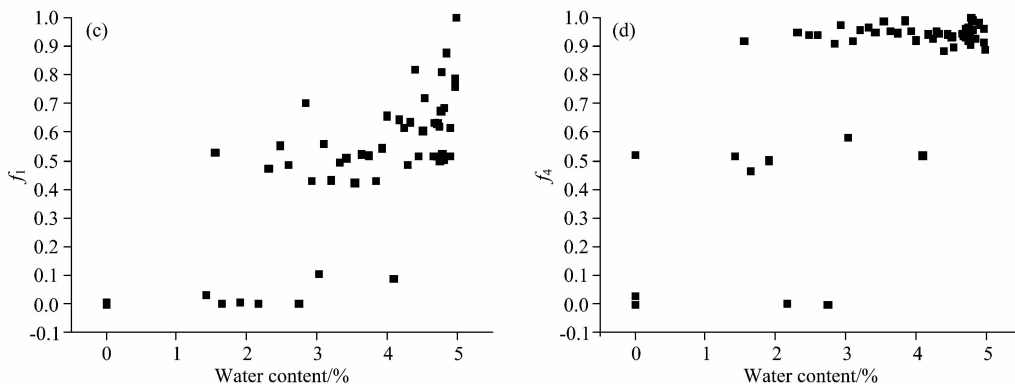


图 8 砾岩特征变量与含水量散点图(归一化)

(a): 峰 R_1 ; (b): 峰 R_2

Fig. 8 Scatter plot of characteristic variables and water content (normalized) for conglomerate

(a): Peak R_1 ; (b): Peak R_2

0.55 之间,说明特征变量与含水量之间相关性偏弱,这与图 8 的散点图的规律相一致。究其原因,砾岩形状不规则,因而导致体积计算不准确,含水量计算误差较大,数据规律分散,表现出特征变量之间的相关程度不强,这组数据真正反

映的近红外光谱特征与含水量的相关性不具有代表性意义。

2.2 粉砂岩的特征选择分析

将粉砂岩特征变量与含水量数据做成散点图,如图 9 所示。

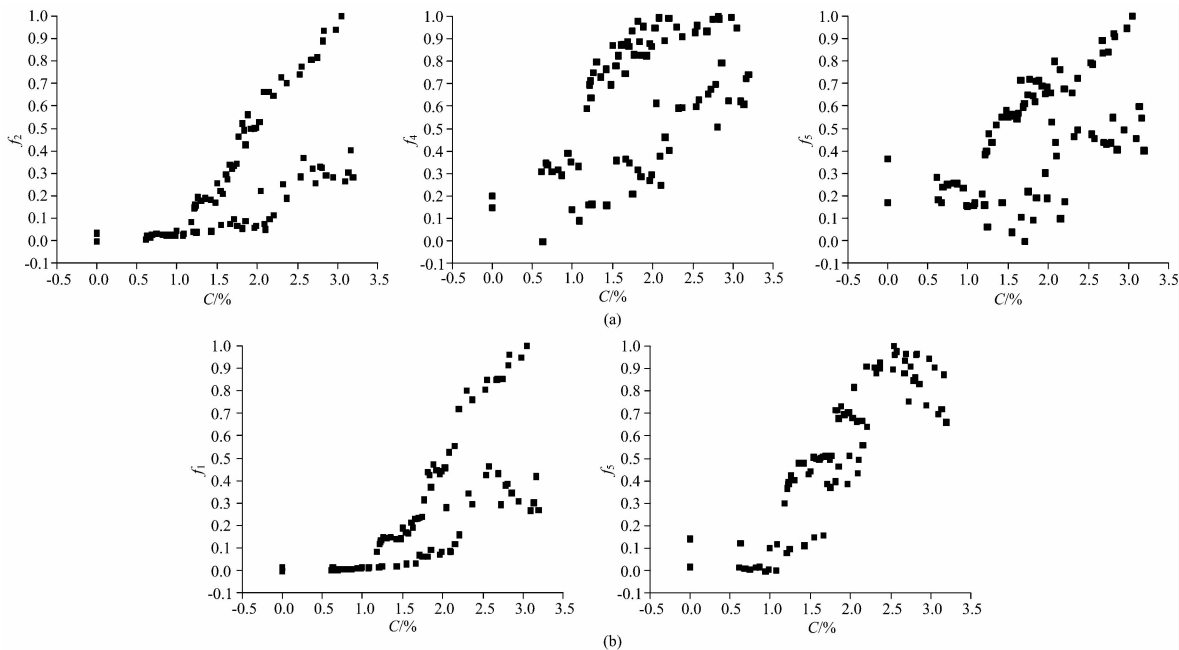


图 9 粉砂岩特征变量与含水量散点图(归一化)

(a): 峰 R_1 ; (b): 峰 R_2

Fig. 9 Scatter plot of characteristic variables and water content (normalized) for siltstone

(a): Peak R_1 ; (b): Peak R_2

利用第 1 节中的公式,分别计算峰 R_1 和 R_2 处的 f_2 , f_4 , f_5 , f_1 , f_3 与含水量之间的 MIC 值,如表 8。

由表 8 可知,粉砂岩在峰 R_1 处有 $MIC(C, f_2) > MIC(C, f_5) > MIC(C, f_4)$,说明峰高与含水量相关性最强,其次是右肩宽,最后是左肩宽。对于峰 R_2 有 $MIC(C, f_5) > MIC(C, f_1)$,说明右肩宽与含水量相关关系最强,峰面积次之。表 8 中的 MIC 值位于 0.48~0.90 之间,说明特征变量与含水量

表 8 粉砂岩特征变量与含水量间的 MIC 值

Table 8 MIC values between characteristic variables and water content of siltstone

吸收峰	$MIC(C, f_1)$	$MIC(C, f_2)$	$MIC(C, f_4)$	$MIC(C, f_5)$
峰 R_1	/	0.657 21	0.482 03	0.482 30
峰 R_2	0.845 27	/	/	0.896 22

之间相关程度较强，这与图 9 的散点图表现出来的规律相一致。

2.3 夯土的特征选择分析

将夯土特征变量与含水量数据做成散点图，如图 10 所

示。

利用第 1 节中的公式，分别计算峰 R_1 和 R_2 处的 f_2 , f_5 , f_1 , f_4 , f_5 与含水量之间的 MIC 值，如表 9。

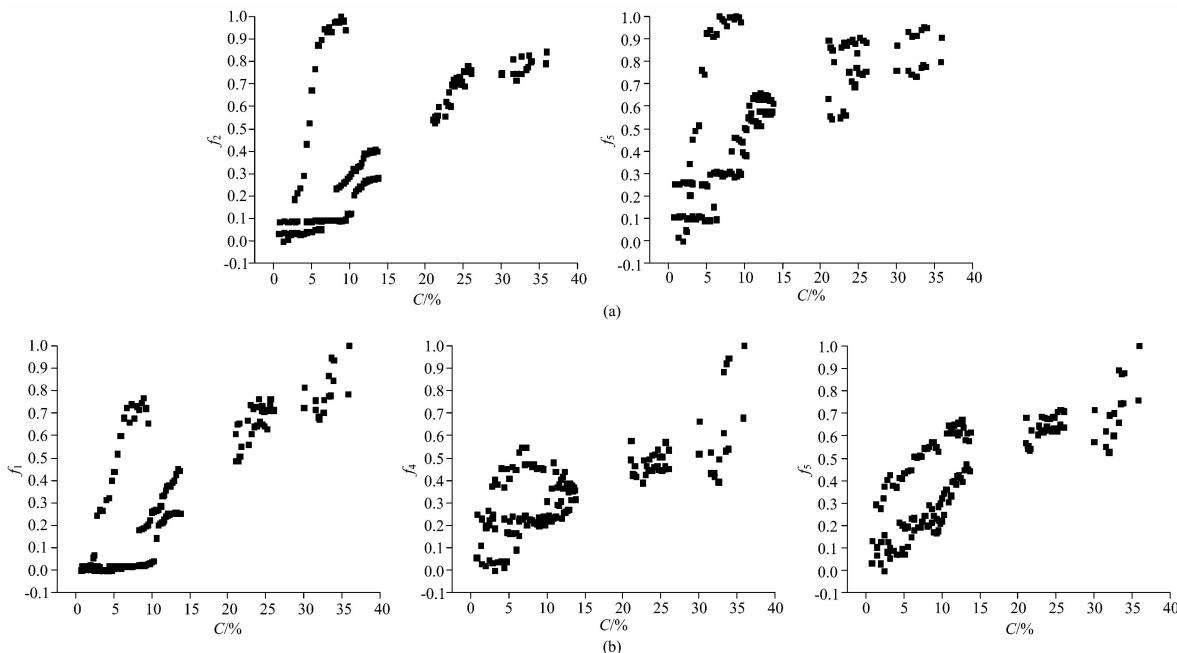


图 10 夯土特征变量与含水量散点图(归一化)

(a): 峰 R_1 ; (b): 峰 R_2

Fig. 10 Scatter plot of characteristic variables and water content (normalized) for rammed soil

(a): Peak R_1 ; (b): Peak R_2

表 9 夯土特征变量与含水量间的 MIC 值

Table 9 MIC values between characteristic variables and water content of rammed soil

吸收峰	MIC(C, f_1)	MIC(C, f_2)	MIC(C, f_4)	MIC(C, f_5)
峰 R_1	/	0.825 54	/	0.940 26
峰 R_2	0.681 53	/	0.622 47	0.818 99

由表 9 可知，夯土在峰 R_1 处有 $MIC(C, f_5) > MIC(C, f_2)$ ，说明右肩宽与含水量相关性最强，峰高次之。对于峰 R_2 有 $MIC(C, f_5) > MIC(C, f_1) > MIC(C, f_4)$ ，说明右肩宽与含水量相关性最强，峰面积次之，最后是左肩宽。表 9 中的 MIC 值位于 0.62~0.95 之间，说明特征变量与含水量之间相关程度较强，这与图 10 的散点图表现出来的规律相一致。

2.4 不同岩性近红外光谱特征选择对比分析

以最大相关系数 MIC 值作为指标，评价岩石近红外光谱的特征变量与其含水量之间的相关性，对于砾岩、粉砂岩、夯土各特征变量按照相关性大小排序结果具体见表 10。

由表 10 中可知，对于粉砂岩和夯土，若只取两个特征变量，峰 R_2 处只需要关注右肩宽和峰面积即可。峰 R_1 处只需要关注峰高和右肩宽即可，二者的特征变量选择一致。因砾岩形状不规则，计算体积时产生误差较大，不予以考虑。

综上所述，对于含水岩土介质的近红外光谱的特征选

择，在 1 400 nm 附近可选择峰高、右肩宽作为特征变量，参与到含水量预测模型的构建中；在 1 900 nm 附近可选择峰面积和右肩宽作为特征变量，参与含水量预测模型的构建。

表 10 近红外光谱特征变量相关程度排序

Table 10 Near-infrared spectral feature variables correlation degree ranking

岩性	相关程度	
	R_1	R_2
砾岩	峰面积 > 右肩宽	峰面积 > 左肩宽
粉砂岩	峰高 > 右肩宽 > 左肩宽	右肩宽 > 峰面积
夯土	右肩宽 > 峰高	右肩宽 > 峰面积 > 左肩宽

3 结 论

(1) 砾岩、粉砂岩、夯土三种含水岩石的近红外光谱在 1 400 和 1 900 nm 附近都有着明显的吸收峰，且随着含水量的变化，吸收强度越来越强，与含水量大小有明显的相关性，故波段 1 400 和 1 900 nm 附近可作为分析光谱特征的基本波段。

(2) 对比砾岩、粉砂岩、夯土近红外光谱特征变量与其含水量的最大相关系数 MIC 值，表明夯土的近红外光谱与

含水量之间的相关性最强。

(3)不同岩性的近红外光谱各个特征值与含水量的相关程度不同,具体表现为在 1 400 nm 附近峰高、右肩宽与含水量都具有较高的相关系数,只是相关性大小会因岩性不同而不同;1 900 nm 附近的右肩宽和峰面积与含水量都具有较高

相关系数,且右肩宽的相关程度高于峰面积。

(4)不同岩性的含水岩土介质近红外光谱的特征变量与含水量相关性具有相似规律:峰高、右肩宽、峰面积是相关程度最高的三个特性。

References

- [1] FENG Fang(冯 放). Life Science Instruments(生命科学仪器), 2007, 5(10): 9.
- [2] JIN Hui-ning, ZHANG Xin-le, LIU Huan-jun, et al(金慧凝, 张新乐, 刘焕军, 等). Journal of Soil Science(土壤学报), 2016, 53(3): 627.
- [3] BAO Qing-ling, DING Jian-li, WANG Jing-zhe(包青岭, 丁建丽, 王敬哲). Progress in Laser and Optoelectronics(激光与光电子进展), 2018, 55(11): 470.
- [4] LOU Jing, XU Jian-hui, ZHOU Yan-lin, et al(娄 径, 徐建辉, 周燕林, 等). Anhui Agricultural Sciences(安徽农业科学), 2016, 44(13): 99.
- [5] KONG Qing-qing, DING Xiang-qian, GONG Hui-li, et al(孔清清, 丁香乾, 宫会丽, 等). Journal of Analytical Testing(分析测试学报), 2017, (10): 1203.
- [6] Reshef D N, Reshef Y A, Finucane H K, et al. Science, 2011, 334(6062): 1518.
- [7] LIANG Ji-ye, FENG Chen-jiao, SONG Peng(梁吉业, 冯晨娇, 宋 鹏). Journal of Computers(计算机学报), 2016, (1): 1.
- [8] HE Man-chao, ZHANG Guo-feng, ZHAO Jian(何满潮, 张国锋, 赵 健). Chinese Patent(中国专利), CN102253181A, 2011.
- [9] ZHANG Fang, HU Zuo-le, HOU Xin-li, et al(张 芳, 户佐乐, 侯欣莉, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2019, 39(11): 3395.

Comparative Analysis of Near-Infrared Spectral Characteristics of Water-Bearing Rocks with Different Lithologies

ZHANG Fang^{1,2}, HU Zuo-le^{1,2}, WANG Dong-sheng^{1,2}, LIU Yu-meng^{1,2}, XIE Yun-xin^{1,2}, ZHUO Hui-hui², HE Man-chao^{1*}

1. State Key Laboratory for Geomechanics & Deep Underground Engineering, China University of Mining & Technology, Beijing 100083, China
2. School of Mechanics and Civil Engineering, China University of Mining & Technology, Beijing 100083, China

Abstract During the process of building inversion model of water content based on NIR spectrum, it is the key issue to figure out whether the feature set of rock will change along with the differences of lithology or not. Aiming at this problem, firstly conducting a laboratory experiment on water absorbability of rock by intelligent test system of hydrologic action on deep soft rock, this paper measures near infrared spectrum of water-bearing rock in three different lithology at different times. For conglomerate, siltite and rammed soil, we collect 51 pieces, 106 pieces and 149 pieces correspondingly. Afterwards, first-order derivative method is adopted in dealing with pretreatment of original spectrum, avoiding the influence of environmental interference on the spectrum. Next, the geometric feature method is used to extract the spectral features and normalize it to eliminate the influence caused by different dimensions and change amplitudes. Then, the correlation degree between initial characteristic variables and that between initial characteristic variables and water content are analyzed, and the characteristic variables at two strong correlation bands are obtained by referring suppression threshold size and eliminating redundant features. At last, Maximal Information Coefficient (MIC) is used as a metric to compare and analyze the feature selection results on near infrared spectrum of water-bearing rock with different lithology, in order to evaluate the influence of lithology on the spectral characteristics of water-bearing rocks. The results are as shown below: (1)The near infrared spectra of conglomerate, siltstone and rammed soil have obvious absorption peaks near 1 400 and 1 900 nm, and with the change of water content, the absorption intensity becomes much more stronger, which shows a significant correlation with the size of the water content. (2)The maximum correlation coefficient between the characteristic variables of near infrared spectra of conglomerate, siltstone and rammed soil and their water content value shows that the correlation between the near infrared spectrum of rammed soil and water content is the strongest. (3)Each characteristic value of near infrared spectrum of different lithology has different correlation degree with water content, which

shows that the peak height near 1400nm, the right shoulder width and the water content all have high correlation. However, the correlation will vary because of different lithology, with the right shoulder width and peak area in the vicinity of 1900nm having higher correlation coefficient with water content, and the correlation degree of right shoulder width being higher than that of peak area. (4)The characteristic variables of near infrared spectrum of water-bearing rock with different lithology are similar to the correlation of water content—peak height, right shoulder width and peak area are the three characteristics with the highest correlation degree.

Keywords Rock; Spectral characteristics; Water content; Maximal information coefficient; Correlation

(Received Jan. 9, 2019; accepted Apr. 18, 2019)

* Corresponding author