

## 基于高光谱的砂岩露头孔隙度估算方法

盛洁<sup>1,2,3</sup>, 刘展<sup>1,3\*</sup>, 曾齐红<sup>2</sup>, 张友焱<sup>2</sup>, 白永良<sup>1,3</sup>, 刘兰法<sup>4</sup>

(1. 中国石油大学(华东)地球科学与技术学院, 山东 青岛 266580;

2. 中国石油勘探开发研究院测井遥感所, 北京 100083;

3. 海洋国家实验室海洋矿产资源评价与探测技术功能实验室, 山东 青岛 266071;

4. Institute for Cartography, TU Dresden, Dresden 01062, Germany)

**摘要:**为快速获得宏观、定量的砂岩露头孔隙度,提出了基于高光谱的孔隙度估算新方法.采集野外露头砂岩样品并测得其孔隙度,利用岩石薄片鉴定资料分析砂岩孔隙度的影响因素;对岩样实测光谱预处理,探索砂岩孔隙度的光谱响应机理;考虑到光谱波段高维性和波段间多重相关性,采用偏最小二乘方法构建孔隙度估算模型;通过变量投影重要性分析模型中重要波段.研究结果表明:基于砂岩填隙物与孔隙度的相关性以及填隙物的光谱特征,可间接反演孔隙度;砂岩孔隙度具有良好的光谱响应;反射率能够定量估算砂岩孔隙度(全波段模型  $R^2 = 0.72$ ,  $RMSE = 2.28$ ,  $RPD = 1.94$ );重要波段帮助降低自变量维度,发现孔隙度敏感波谱响应.本研究为基于高光谱图像的野外露头孔隙度表征奠定了基础.

**关键词:**砂岩露头;孔隙度估算;高光谱;偏最小二乘

**中图分类号:**TP79 TE122.2 **文献标识码:**A

## Porosity estimation method in sandstone outcrop based on hyperspectrum

SHENG Jie<sup>1,2,3</sup>, LIU Zhan<sup>1,3\*</sup>, ZENG Qi-Hong<sup>2</sup>, ZHANG You-Yan<sup>2</sup>, BAI Yong-Liang<sup>1,3</sup>, LIU Lan-Fa<sup>4</sup>

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

2. Research Department of Well-logging and Remote Sensing Technology, Research Institute of Petroleum Exploration & Development, PetroChina, Beijing 100083, China;

3. Laboratory for Marine Mineral Resources, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266071, China;

4. Institute for Cartography, TU Dresden, Dresden 01062, Germany)

**Abstract:** In order to obtain macroscopic and quantitative porosity data in outcrop rapidly, a new Hyperspectrum based porosity estimation method was proposed in this paper. Sandstone samples were collected from field outcrop and measured for porosity data, whose influence factors were analyzed with rock thin section. After the preprocessing on the spectral data of rock samples, the spectral response mechanism of porosity was preliminarily explored. Considering the high dimensionality of spectral bands and the multiple correlation between bands, the porosity estimation models were constructed using the partial least squares method. The important bands in the model were indicated by the variable importance in the projection. The results show that: sandstone porosity can be indirectly retrieved based on the correlation between interstitial fillings and porosity and the spectral characteristics of interstitial fillings; sandstone porosity shows good spectral response; reflectance has the ability to estimate porosity quantitatively (porosity estimation model based on full wavelengths:  $R^2 = 0.72$ ,  $RMSE = 2.28$ ,  $RPD = 1.94$ ); the important bands help to reduce the independent variable dimension and find the porosity-sensitive spectral response. This study lays the foundation for porosity characterization in the outcrop

收稿日期:2018-04-25,修回日期:2018-09-21

Received date: 2018-04-25, revised date: 2018-09-21

基金项目:国家科技重大专项(2017ZX05001001), 中国石油股份重大专项(2016B-03)

**Foundation items:** Supported by National Science and Technology Major Project (2017ZX05001001), and Science and Technology Major Project of PetroChina (2016B-03)

**作者简介 (Biography):** 盛洁(1990-),女,山东菏泽人,博士研究生,主要研究领域为地球及地质资源的数字化理论及技术. E-mail: hellosheengjie@163.com

\* 通讯作者 (Corresponding author): E-mail: liuzhan5791@sina.com

based on hyperspectral images.

**Key words:** sandstone outcrop, porosity estimation, hyperspectrum, PLSR

**PACS:** 07. 07. Df, 89. 30. aj, 91. 65. -n

## 引言

露头是地下储层的真实刻画,砂岩露头物性参数的定量表征能够为地下砂岩储层地质模型建立提供依据<sup>[1]</sup>.孔隙度可以反映岩石储存流体的能力,是砂岩露头的一个重要物性参数.目前,用于露头孔隙度估算的传统方法是:在露头表面取若干岩芯,测得岩芯孔隙度,插值得到整个露头的孔隙度数据<sup>[2]</sup>.该方法存在岩芯采样位置受限、岩芯资料少的问题.此外,孔隙度大小受粒度、填隙物成分及含量等因素的影响<sup>[3-4]</sup>.马石玉<sup>[5]</sup>、李凤杰<sup>[6]</sup>等人认为鄂尔多斯盆地陇东地区延长组砂岩孔隙度随粒径的减小而降低.填隙物主要由粘土矿物(高岭石、蒙脱石等),碳酸盐矿物(方解石、白云石等)和杂基等组成,填隙物可以充填粒间孔隙,填隙物含量越高,孔隙度越低<sup>[7]</sup>.这些影响因素在空间上的不均匀性会造成插值结果不准确的问题.传统的露头孔隙度估算方法不能快速获得宏观、定量、精确的露头孔隙度数值.

高光谱技术以获取数据快、光谱分辨率高、波段连续性强、数据丰富的特点,在地质领域得到广泛关注<sup>[8-9]</sup>,使得地质应用实现了定性解译到定量反演的飞跃.光谱特征与孔隙度之间存在关联关系:孔隙度是岩石内部结构参数的一种,直接影响岩石光谱反射率、谱带位置以及吸收特征<sup>[10-12]</sup>,但其物理机制尚不明确.光谱与粒度之间存在相关性:在矿物成分基本相同时,大多数岩石的反射率与岩石粒度之间存在负相关关系,而暗色岩矿的反射率与粒度之间存在正相关关系<sup>[13]</sup>.光谱特征与矿物成分、含量之间也存在关联关系:研究发现高岭石的吸收谱带在2165 nm和2205 nm附近,蒙托石的吸收谱带在2205 nm和2215 nm附近<sup>[14]</sup>;方解石和白云石在2350 nm附近有吸收特征<sup>[10]</sup>.因此,可以利用光谱数据定量反演岩石填隙物的构成<sup>[15-16]</sup>.光谱与孔隙度、岩石粒度、填隙物之间存在的相关关系,为基于高光谱数据估算孔隙度提供了基础.据此,有学者研究了光谱特征与孔隙度之间的定量关系:粉砂岩在645 nm处的光谱反射率与孔隙率有较为明显的正相关关系<sup>[17]</sup>;尤金凤等人的研究发现油砂样品表征粘土矿物含量的粘土基因吸收深度与孔隙度值呈负

相关关系<sup>[18]</sup>.由于岩石光谱响应的复杂性,前人主要是对孔隙度与光谱特征之间的相关性进行了探讨,在孔隙度定量估算方面的研究尚且少见.

本文的研究目标是利用高光谱数据对砂岩露头孔隙度进行定量估算.尝试以砂岩样品的实测光谱数据和孔隙度数据为基础,首先探讨砂岩样品孔隙度的影响因素,通过分析不同孔隙度砂岩样品的光谱指标(反射率和连续统去除)响应特征,初步探索砂岩孔隙度的光谱响应机理.采用偏最小二乘方法构建孔隙度估算模型,并讨论模型中的重要波段.本研究为快速、准确获得露头孔隙度奠定基础.

## 1 数据采集与测定

基于砂岩样品开展孔隙度估算研究,因此需要对野外砂岩露头进行样品采集,测定砂岩样品的孔隙度数据和光谱数据.同时对研究区砂岩样品孔隙度因素进行分析,为探索光谱对砂岩孔隙度的响应机理奠定基础.

### 1.1 野外露头岩石样品采样与孔隙度测定

对陕西省彬县水北村东、芦河村北、芦河村南,延长县杨家沟共四个砂岩露头剖面进行岩石采样,这四个剖面上分别发育延长组长8<sub>1</sub>、长9<sub>2</sub>、长10<sub>1</sub>和长3油层组.延长组具有非常丰富的油气资源,油藏具有低孔、低渗、低压的特征,为典型的致密砂岩储层<sup>[19-20]</sup>.所采的55块岩石样品是典型块状致密砂岩,岩性以细砂岩为主.

为测得块状岩样的孔隙度数据,对块状岩样钻取柱样,所取柱状样直径约2.5 cm,长度在1~6 cm范围内.通过PHI-220型氦气孔隙度测量仪测得柱状样的孔隙度,测量的相对误差在0.5%以内.统计发现样品孔隙度的分布范围是0.83%~16.33%,其基本符合致密砂岩孔隙度的分布规律.各露头剖面上岩石样品采样点的空间分布如图1所示,孔隙度数值分布较为离散,砂体非均质性较强.

### 1.2 岩石样品孔隙度影响因素分析

为分析研究区砂岩孔隙度的影响因素,对岩石薄片镜下特征进行分析.延长组致密砂岩粒度以细砂、极细砂为主,填隙物含量(12%~17.2%)较高,其中粘土矿物含量较高<sup>[21]</sup>.杨家沟长3露头剖面17块砂岩样品薄片鉴定结果如表1所示,表1表明砂

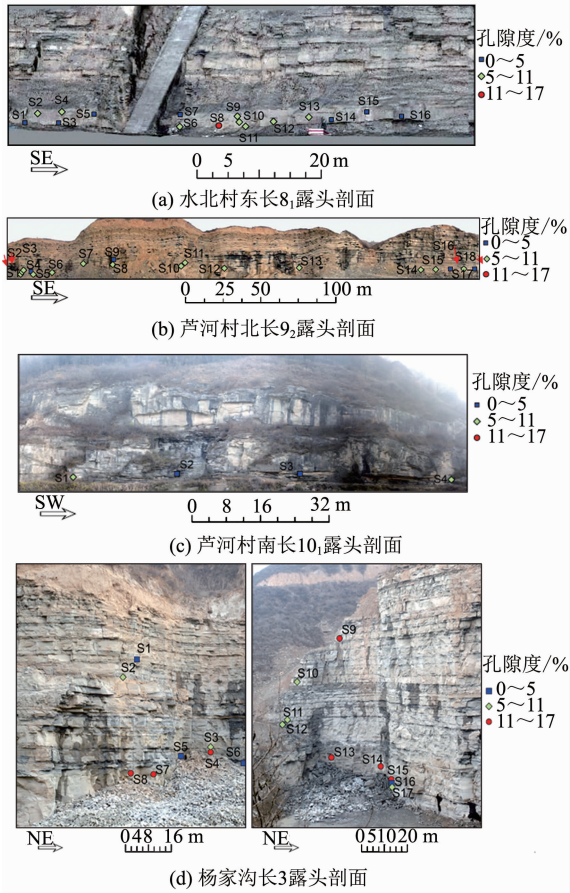


图 1 各露头剖面上岩石采样点的空间分布 (a) 水北村东, (b) 芦河村北, (c) 芦河村南, (d) 杨家沟  
 Fig. 1 The spatial distribution of sampling points on each outcrop section (a) Shuibeiandong, (b) Luhecunbei, (c) Luhecunnan, (d) Yangjiagou

岩样品的岩石结构以粉砂质结构和细粒碎屑结构为

主;胶结方式包括基底式胶结、接触式胶结和孔隙式胶结. 胶结类型以泥质胶结为主.

根据岩石样品的空间分布特征和镜下鉴定资料,将 17 块砂岩样品按孔隙度范围分为三组:低孔隙度 0% ~ 5% 组、中孔隙度 5% ~ 11% 组和高孔隙度 11% ~ 17% 组. 0% ~ 5% 组内岩样为泥质粉砂质结构和泥质细粒碎屑结构,胶结方式为基底式胶结,填隙物较多. 5% ~ 11% 组岩样为含细砂粉砂质结构和含粉砂质细粒碎屑结构,胶结方式有基底式胶结和接触式胶结,填隙物含量降低. 11% ~ 17% 组岩样以细粒碎屑结构为主,胶结方式有接触式胶结和孔隙式胶结,填隙物较少. 三组内典型岩样 S5、S17 及 S13 的显微特征如图 2 所示.

综上,研究区砂岩孔隙度受主要碎屑颗粒(石英、长石)和填隙物(绿泥石、绿帘石、碳酸盐矿物、黑云母等)的相对含量影响,碎屑颗粒含量较高、填隙物较少的岩样孔隙度较大,而孔隙碎屑颗粒含量较少、填隙物较多的岩样孔隙度较小.

### 1.3 室内光谱数据测量及预处理

采用美国 ASD 公司生产的 ASD FieldSpec 3 非成像光谱仪测量岩样光谱,其波段范围是 350 ~ 2500 nm,光谱重采样间隔为 1 nm. 测量前严格按照操作规范去除暗电流影响,并进行标准白板定标. 为了保证光谱测定结果的可靠性,对每一个块状砂岩样品新鲜面测量 20 次并取其算术平均值作为该砂岩样品的反射光谱数据.

预处理过程包括平滑和光谱指标计算. 实测光谱可能会受到测试环境、仪器本身、杂散光等因素的

表 1 杨家沟剖面 17 块砂岩样品的镜下鉴定结果

Table 1 Microscopic identification results of 17 sandstone samples from Yangjiagou section

岩石样品	孔隙度	结构描述	胶结方式	组别
S5	3.345	泥质粉砂质结构	基底式胶结	低孔隙度 0% ~ 5%
S1	3.507	泥质细粒碎屑结构	基底式胶结	低孔隙度 0% ~ 5%
S16	4.734	泥质粉砂质结构	基底式胶结	低孔隙度 0% ~ 5%
S6	4.748	泥质细粒碎屑结构	基底式胶结	低孔隙度 0% ~ 5%
S10	6.600	含细砂粉砂质结构	接触式胶结	中孔隙度 5% ~ 11%
S12	6.700	含粉砂质细粒碎屑结构	基底式胶结	中孔隙度 5% ~ 11%
S11	7.061	含粉砂质细粒碎屑结构	基底式胶结	中孔隙度 5% ~ 11%
S3	7.540	含细砂粉砂质结构	基底式胶结	中孔隙度 5% ~ 11%
S17	10.835	含粉砂质细粒碎屑结构	接触式胶结	中孔隙度 5% ~ 11%
S2	10.860	含粉砂质细粒碎屑结构	接触式胶结	中孔隙度 5% ~ 11%
S15	11.030	细粒碎屑结构	接触式胶结	高孔隙度 11% ~ 17%
S14	11.611	细粒碎屑结构	孔隙式胶结	高孔隙度 11% ~ 17%
S4	13.865	中细粒碎屑结构	接触式胶结	高孔隙度 11% ~ 17%
S8	14.635	中粒碎屑结构	孔隙式胶结	高孔隙度 11% ~ 17%
S9	14.950	细粒碎屑结构	孔隙式胶结	高孔隙度 11% ~ 17%
S13	15.562	细中粒碎屑结构	孔隙式胶结	高孔隙度 11% ~ 17%
S7	16.329	细粒碎屑结构	接触式胶结	高孔隙度 11% ~ 17%

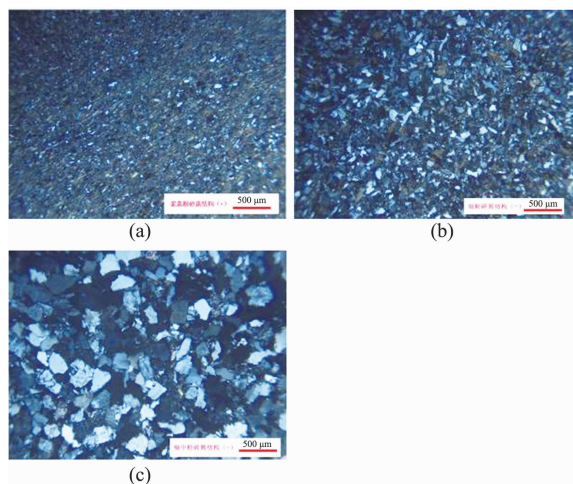


图2 典型岩样的显微特征(a)0% ~ 5%组内S5, (b)5% ~ 11%组内S17, (c)11% ~ 17%组内S13

Fig.2 Microscopic characteristics of typical rock samples, (a) S5 in group 0% ~ 5%, (b) S17 in group 5% ~ 11%, (c) S13 in group 11% ~ 17%

影响而产生噪音,为了消除上述干扰,需要首先对实测光谱数据进行平滑处理,本文采用的平滑方法是9点移动平均法.在此基础上,对光谱曲线进行数学变换,计算得到用于建模的光谱指标.连续统去除法可以抑制背景信息,突出弱吸收特征<sup>[22]</sup>.因此,本文对反射率数据进行连续统去除变换,最终得到反射率及连续统去除值两个光谱指标.

## 2 砂岩孔隙度的光谱响应

砂岩孔隙度的光谱响应特征分析是利用高光谱进行孔隙度估算的基础.目前,岩石和矿物介质与辐射之间相互作用的机理研究还比较滞后,光谱对砂岩孔隙度的响应机理尚不明确,本文初步探索了致密砂岩孔隙度的光谱响应机理.

根据上述岩样分组方式,将55块砂岩样品分为三组:低孔隙度0% ~ 5%组(17块样品)、中孔隙度5% ~ 11%组(30块样品)和高孔隙度11% ~ 17%组(8块样品).在可见-近红外光谱范围内(350 ~ 2500 nm),统计光谱反射率和连续统去除值的组内平均值,两个光谱指标平均值曲线如图3(a)和图3(b)所示.本文进一步分析了光谱指标与孔隙度之间定量的相关关系.计算得到二者之间的相关系数,如图3所示,并且对相关系数进行显著性检验(本文中采用的显著性水平为0.01),通过显著性检验的波段为显著性波段.相关系数和显著性波段为光谱指标的优选提供依据,优选出的光谱指标最终用

于孔隙度估算模型建立.

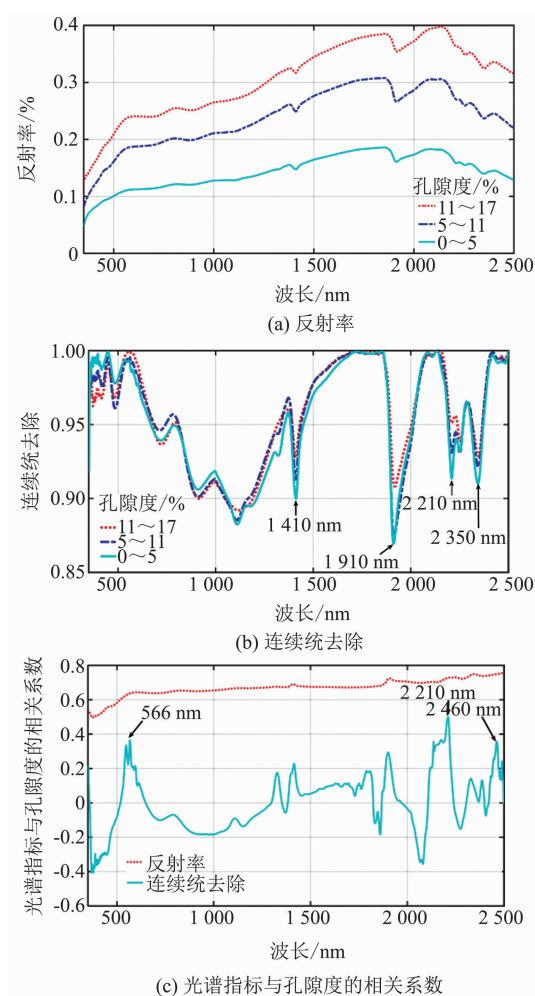


图3 不同孔隙度分组的岩样光谱响应 (a)反射率, (b)连续统去除, (c)光谱指标与孔隙度的相关系数  
Fig.3 Spectral response of sandstone samples with different porosity groups (a)reflectance, (b) continuum removal reflectance, (c) the correlation coefficient between two spectral indexes and porosity

(1) 岩样孔隙度越小,其填隙物含量越高,在全波段范围内对应的光谱反射率越低,原因是绿帘石、绿泥石、黑云母等填隙物的光谱反射率低于主要碎屑颗粒(石英、长石).反射率与孔隙度在全波段范围内显著正相关,相关系数曲线较为平缓,相关系数绝对值的最大值 $|R|_{\max}$ 达到0.76,说明不同孔隙度的砂岩样品具有较好的光谱响应差异性.

(2) 连续统去除值可以突出岩石组成成分的吸收特征,很多学者研究了绿帘石、绿泥石、高岭石、蒙脱石、方解石和白云石的吸收谱带,以及利用光谱吸收特征来反演填隙物矿物含量<sup>[14-16]</sup>.连续统去除曲线表明在1410、1910、2210、2350 nm等附近存在明显的吸收谱带.1410、1910 nm处的

吸收谱带可能由矿物中的水分子引起,2 210、2 350 nm 附近的吸收谱带可能由绿泥石和绿帘石引起<sup>[10]</sup>. 连续统去除值与孔隙度的相关系数曲线有较大起伏,出现多个谷值和峰值,  $|R|_{\max}$  为 0.50. 连续统去除值共有 95 个显著波段,主要在分布在 566(具有矿物阳离子的吸收谱带)、2 210 和 2 460 nm(具有粘土矿物的吸收谱带)附近. 粘土矿物在 2 210 nm 附近具有吸收谱带,该谱带处的连续统去除值与孔隙度显著正相关. 吸收深度值等于 1 减去连续统去除值,因此粘土基团响应谱带的吸收深度与孔隙度存在显著负相关的关系,即粘土矿物含量越多,该谱带的吸收深度值越大,孔隙度越小,这与前人研究结论一致<sup>[23]</sup>.

通过砂岩填隙物与孔隙度的相关性,基于砂岩填隙物的光谱特征,可以间接反演砂岩孔隙度. 砂岩孔隙度具有良好的光谱响应,反射率与孔隙度的正相关性较好,显著性波段较多,是较优的光谱指标. 因此,本文选择了反射率来构建孔隙度估算模型. 本文岩石样品来自于四个野外露头剖面,沉积环境及后期埋藏成岩作用存在一定的差异性. 这种混合效应影响了光谱指标与孔隙度的相关性,连续统去除值受到的影响较大.

### 3 孔隙度估算模型

本节研究孔隙度估算模型的构建方法,并对模型中波段重要性的计算方法和模型精度评价方法进行论述.

#### 3.1 模型构建方法

构建孔隙度估算模型,即通过回归分析方法来确定孔隙度和光谱指标之间的关系模式. 偏最小二乘方法(Partial Least Squares Regression, PLSR)是一种经典的多元回归方法,其综合了主成分分析,典型相关分析和多元线性回归方法<sup>[24]</sup>. 在近红外、红外等光谱分析中,偏最小二乘方法已被广泛用于构建线性定量模型<sup>[25-29]</sup>,因其在解决光谱波段的高维性和波段之间的多重相关性等问题上具有优势,本文采用该方法建立孔隙度估算模型.

利用偏最小二乘方法构建孔隙度估算模型的步骤如下:

首先,对建模集的反射率(自变量  $X_{s \times n}$ ,  $s$  为样本数,  $n$  为光谱波段数)、孔隙度(因变量  $y_s$ ,  $s$  为样本数)进行标准化,标准化后的数据矩阵分别记为  $E_0$  和  $F_0$ .

随后,分别提取  $E_0$  和  $F_0$  的第一对主成分  $t_1$

和  $u_1$ :

$$\begin{cases} t_1 = E_0 w_1 \\ u_1 = F_0 c_1 \end{cases}, \quad (1)$$

其中,  $w_1$  是  $E_0$  的第一个轴,  $\|w_1\| = 1$ ,  $c_1$  是  $F_0$  的第一个轴,  $\|c_1\| = 1$ . 在偏最小二乘回归中,要求  $t_1$  与  $u_1$  的协方差达到最大,采用拉格朗日算法求解该最优化问题.  $w_1$  是对应于矩阵最大特征值的单位特征向量,  $c_1$  是对应于矩阵  $F_0' E_0 E_0' F_0$  最大特征值的单位特征向量. 计算出  $w_1$  和  $c_1$  后,可求得  $t_1$  和  $u_1$ .

再分别求  $E_0$  和  $F_0$  对  $t_1$  的回归方程:

$$\begin{cases} E_0 = t_1 p_1' + E_1 \\ F_0 = t_1 r_1' + F_1 \end{cases}, \quad (2)$$

其中,  $E_1$  和  $F_1$  记为残差矩阵,回归系数向量为:

$$\begin{cases} p_1 = \frac{E_0' t_1}{\|t_1\|^2} \\ r_1 = \frac{F_0' t_1}{\|t_1\|^2} \end{cases}, \quad (3)$$

用残差矩阵  $E_1$  和  $F_1$  取代  $E_0$  和  $F_0$ ,提取第二个主成分. 重复该步骤,直到主成分个数  $H$  满足要求. 推导可得  $F_0$  关于  $t_h$  ( $1 \leq h \leq H$ ) 的回归模型,由于  $t_h$  为  $E_0$  的线性组合,有:

$$t_h = E_{h-1} w_h = E_0 w_h^*, \quad (4)$$

其中,  $w_h^* = \prod_{j=1}^{h-1} (I - w_j p_j')$ .

可以得到  $F_0$  关于  $E_0$  的回归模型:

$$\begin{aligned} \hat{F}_0 &= E_0 w_1^* r_1' + E_0 w_2^* r_2' + \cdots + E_0 w_H^* r_H' + F_H \\ &= E_0 (w_1^* r_1' + w_2^* r_2' + \cdots + w_H^* r_H') + F_H \end{aligned} \quad (5)$$

若记  $x_j^* = E_0 w_j^*$ ,  $y^* = F_0$ ,  $\alpha_j^* = \sum_{h=1}^H r_h w_{hj}^*$ ,  $j = 1, 2, \cdots, n$ , 其中  $w_{hj}^*$  是  $w_h^*$  的第  $j$  个分量. 能得到基于标准化光谱反射率的估算模型:

$$\hat{y}^* = \alpha_1^* x_1^* + \alpha_2^* x_2^* + \cdots + \alpha_n^* x_n^* \quad (6)$$

最终得到  $H$  个主成分下,基于原始光谱反射率的孔隙度估算模型:

$$\hat{y} = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_{n-1} X_{n-1} + \alpha_n X_n \quad (7)$$

其中,  $\hat{y}$  为孔隙度估算值.  $n$  为所用波段数.  $X_1 \sim X_n$  为波段处的光谱反射率.  $\alpha_0 \sim \alpha_n$  为孔隙度估算模型系数:

$$\begin{cases} \alpha_0 = E(y) - \sum_{i=1}^n \alpha_i^* \frac{S_y}{S_{x_i}} E(X_i) \\ \alpha_i = \alpha_i^* \frac{S_y}{S_{x_i}} X_i \end{cases} \quad i = 1, 2, \cdots, n \quad (8)$$

其中,  $E(y)$ ,  $E(X_i)$  分别为孔隙度  $y$ , 光谱反射率  $X$  的均值,  $S_y$ ,  $S_{X_i}$  为  $y$ ,  $X$  的方差.

### 3.2 波段重要性分析

偏最小二乘方法可以评价每个光谱波段对于孔隙度估算模型的重要性<sup>[30]</sup>,其指示的重要光谱波段能够用于波段优选,为孔隙度估算模型的建立、基于高光谱图像的孔隙度反演提供依据,因此本文对孔隙度估算模型中的波段重要性进行计算.变量投影重要性(Variable importance in the projection, VIP)被广泛应用于变量选择<sup>[31]</sup>,VIP可以评价自变量对偏最小二乘模型的贡献度,本文采用VIP评价孔隙度定量估算模型中的重要光谱波段.第*i*个光谱波段VIP值的计算公式为:

$$VIP_i = \sqrt{\frac{\sum_{h=1}^H w_{ih}^2 \cdot SSY_h \cdot N}{SSY_{total} \cdot H}}, \quad (9)$$

其中,*H*为模型中主成分的总数, $w_{ih}$ 为第*i*个波段在*h*主成分上的权重, $w_{ih}^2$ 说明了第*i*个波段在*h*主成分上的重要性, $SSY_h$ 为第*h*个主成分的解释方差平方和,*N*为自变量总数, $SSY_{total}$ 为自变量解释方差平方总和.波段的VIP值越大,说明其对因变量解释方差贡献越大,说明该波段越重要,也是后续孔隙度定量估算中需要重点考察的波段.通常认为VIP值大于1的波段为重要波段<sup>[31]</sup>.

### 3.3 模型精度评价

为评价各个孔隙度模型的应用效果,需要采用一定的手段对模型的精度进行评价.具体地,采用决定系数 $R^2$ 、均方根误差RMSE及相对分析误差RPD<sup>[32]</sup>三个指标来评价模型精度.

$$R^2 = \frac{\sum_{i=1}^s (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^s (y_i - \bar{y})^2}, \quad (10)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^s (\hat{y}_i - y_i)^2}, \quad (11)$$

$$RPD = \frac{SD}{RMSE}, \quad (12)$$

其中,*s*为样本数,*y*是实测值, $\bar{y}$ 为实测值均值, $\hat{y}$ 为估算值,SD为标准差. $R^2$ 能反映模型建立和预测的稳定性, $R^2 \leq 1$ , $R^2$ 越大说明模型的稳定性越好、拟合程度越高.RMSE用来检验模型的预测能力,RMSE越小表明模型预测能力越好.RPD常被用于分析模型的定量估算能力<sup>[33]</sup>,当 $RPD \leq 1.00$ 时,说明模型精度很低,不具备估算能力;当 $1.00 < RPD \leq 1.40$ 时,说明模型精度低,具有区别高值和低值的能力;当 $1.40 < RPD \leq 1.80$ 时,说明模型精度一般,具有一般的估算能力;当 $1.80 < RPD \leq 2.00$ 时,说明模型精度较高,具有定量估算能力;当 $2.00 <$

$RPD \leq 2.50$ 时,说明模型精度很高,具有很好的定量估算能力;当 $RPD > 2.50$ 时,说明模型具有极好的定量估算能力.

## 4 结果与讨论

### 4.1 模型建立与应用结果分析

本文利用反射率构建孔隙度估算模型,并对构建的模型进行应用效果分析.对于采集的55块砂岩样品,本文将水北村东、芦河村北和芦河村南三个露头剖面的38块砂岩样品用于模型训练,将杨家沟露头剖面的17块砂岩样品用于模型预测.为了保证孔隙度估算模型的精度,将全波段(350~2500 nm)数据用于模型建立.利用反射率构建的孔隙度估算模型如表2所示,主成分数为9, $X_1 \sim X_{2151}$ 为波段350~2500 nm处的光谱反射率.从建模结果看,模型 $R^2$ 为0.91,RSME为0.86,建模精度较高.

表2 利用反射率构建的孔隙度估算模型

Table 2 Porosity estimation model based on reflectance

光谱指标	孔隙度估算模型	$R^2$	RSME
反射率	$y = -0.879 - 0.985X_1 - 7.123X_2 + \dots + 1.323X_{2150} + 1.779X_{2151}$	0.91	0.86

为检验孔隙度估算模型的应用效果,对杨家沟露头剖面的17块砂岩样品进行孔隙度反演.基于上述孔隙度估算模型,利用预测集砂岩样品全波段的反射率数据,反演得到预测集的孔隙度数据.根据预测孔隙度,按照上述分组标准对砂岩样品进行分组.孔隙度估算模型的应用效果如图4所示.该模型应用评价指标为 $R^2 = 0.72$ ,RMSE = 2.28,RPD = 1.94 > 1.80,模型具有定量估算孔隙度的能力.在17个砂岩样品中有13个砂岩样品的预测组别与实际组别一致,S11、S2、S15及S9的预测组别与实际组别不一致.对这四个岩石样品的镜下特征进行分析,S11、S2及S15碳酸盐填隙物较多,S9填隙物中出现了锆石、磷灰石,四者填隙物成分及含量与其他岩石样品存在明显差异,导致预测组别与实际组别不符.孔隙度估算模型的应用效果验证了本文方法的可行性,本文的研究为野外露头上孔隙度的估算奠定了基础.砂岩样品来自于四个野外露头剖面,其填隙物存在一定差异性,该差异性影响了孔隙度估算模型的应用效果.

### 4.2 孔隙度定量估算模型中的重要波段

分析孔隙度定量估算模型中的重要波段对后续波段优选、高光谱图像反演有重要意义.反射率模型

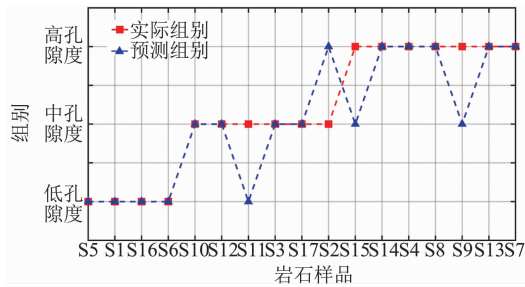


图 4 孔隙度估算模型应用效果  
Fig. 4 Prediction results of porosity estimation model

在全波段上的 VIP 值曲线如图 5 所示,重要性临界值为 1,即 VIP 值大于 1 的波段为重要波段.

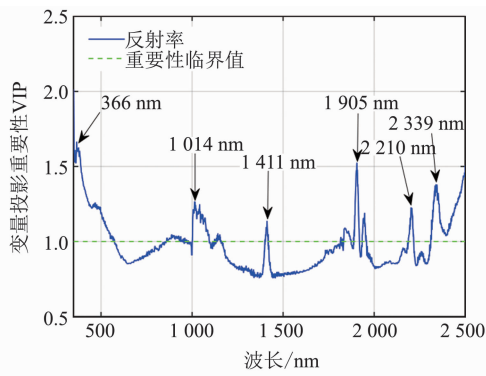


图 5 孔隙度定量估算模型中的重要波段  
Fig. 5 Important variables in the porosity quantitative estimation model

对反射率模型中的重要波段进行统计分析. 结果表明,模型中的重要波段共有 785 个,具有离散分布特征,在 366 nm、1 905 nm、2 339 nm、1 014 nm、2 210 nm 和 1 411 nm 附近有 6 个明显的峰值. 其中,1 014 nm 附近可能存在铁质胶结物等的吸收谱带,1 411、1 905 nm 附近处可能存在矿物中水分子的吸收谱带,2 210、2 339 nm 附近可能存在绿泥石和绿帘石等的吸收谱带. 因此,光谱对孔隙度的响应特征应主要是由砂岩填隙物成分引起的.

为了验证重要波段的提取效果,需要基于重要波段重新构建孔隙度模型,并分析其应用效果. 本文提取了重要波段处的反射率作为新数据集. 在此基础上,仍以水北村东、芦河北和芦河南三个露头剖面的 38 块砂岩样品作为建模集,以杨家沟露头剖面的 17 块砂岩样品作为预测集,主成分数为 6,最终构建了基于重要波段的孔隙度估算模型. 利用预测集砂岩样品重要波段的反射率数据,反演得到预测集砂岩样品的孔隙度数据. 同样根据预测孔隙度

对砂岩样品进行分组. 该模型应用效果如图 6 所示,  $R^2 = 0.72$ ,  $RMSE = 2.26$ ,  $RPD = 1.96$ . 在 17 个砂岩样品中有 14 个砂岩样品的预测组别与实际组别一致, S11、S2 及 S5 的预测组别与实际组别不一致. 模型的预测精度比全波段模型略有提高,说明采用重要波段代替全波段来进行孔隙度预测具有可行性.

重要波段的筛选降低了波段维度,一些机器学习方法在全波段范围预测效果欠佳的情况会得到一定程度的改善<sup>[34]</sup>,同时会提升大样本情况下的计算速度.

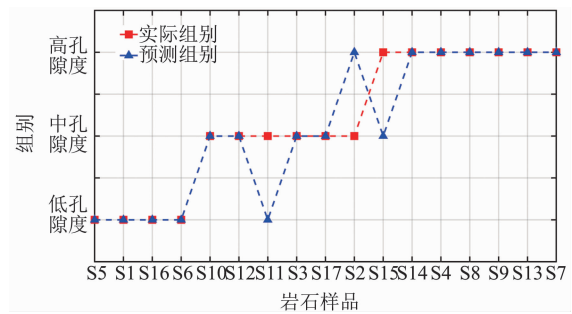


图 6 基于重要波段的孔隙度估算模型预测效果  
Fig. 6 Prediction results of porosity estimation model based on important bands

### 4.3 露头剖面尺度上的孔隙度反演

为了估算露头尺度上的孔隙度数据,需要结合高光谱图像进行反演,露头的高光谱图像通常由地面高光谱成像仪获得. 地面高光谱成像仪的光谱范围一般在 380 ~ 1 000 nm、600 ~ 1 600 nm、1 000 ~ 2 500 nm 等范围内,其光谱分辨率一般高于 10 nm. 根据图 5 中 VIP 值指示的重要波段,本文以光谱范围为 1 000 ~ 2 500 nm、光谱分辨率为 10 nm 的地面高光谱成像仪为例,论述露头剖面尺度上的孔隙度反演流程. 利用建模集样品实测光谱反射率构建孔隙度估算模型,对露头高光谱图像数据进行预处理,并通过迁移方法使该孔隙度估算模型适用于高光谱图像<sup>[35]</sup>,再利用该模型对高光谱图像进行孔隙度反演,最终得到露头剖面尺度上的孔隙度数据.

首先,利用建模集样品实测光谱反射率构建孔隙度估算模型. 对于光谱范围为 1 000 ~ 2 500 nm、光谱分辨率为 10 nm 的地面高光谱成像仪,本文提取了 151 个波长. 依据上文建模集和预测集的划分,利用建模集砂岩样品 151 个波长处的反射率数据构建孔隙度估算模型,主成分数为 8,所建模型如表 3 所示. 该模型在预测集上预测效果如图 7 所示,应用评价指标为  $R^2 = 0.71$ ,  $RMSE = 2.31$ ,  $RPD = 1.91 >$

1.80,模型具有定量估算孔隙度的能力.在17个砂岩样品中有13个砂岩样品的预测组别与实际组别一致,S10、S2、S15及S9的预测组别与实际组别不一致.模型预测效果与利用全波段、重要波段构建的孔隙度估算模型效果相近.由于用于建模的波长较少,本文未为该模型中的重要波段进行提取.该模型应用效果表明利用高光谱图像在露头剖面尺度上的孔隙度反演研究中具有应用潜力.

然后,基于构建的孔隙度估算模型,对高光谱图像进行反演得到露头剖面的孔隙度数据.对地面高光谱成像仪获取的露头高光谱图像进行几何校正、去噪等预处理,通过迁移方法使孔隙度估算模型适用于高光谱图像<sup>[37]</sup>,再利用该模型对高光谱图像进行孔隙度反演,最终得到露头剖面尺度上的孔隙度数据.

表3 利用地面高光谱成像仪波长的孔隙度估算模型  
Table 3 porosity estimation model based on the wavelengths of ground hyperspectral imager

光谱指标	孔隙度估算模型	$R^2$	RSME
反射率	$y = -1.393 - 88.717X_1 + 2.133X_2 + \dots - 10.849X_{150} + 81.028X_{151}$	0.82	1.25

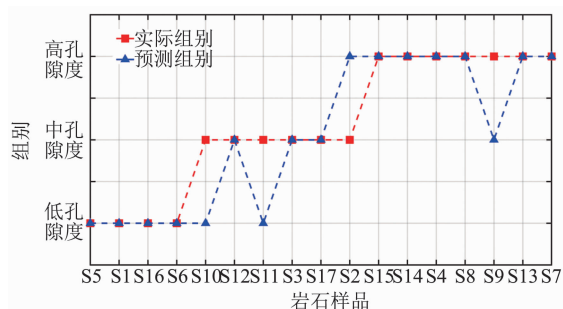


图7 基于地面高光谱成像仪波长的孔隙度估算模型预测效果

Fig. 7 Prediction results of porosity estimation model based on the wavelengths of ground hyperspectral imager

## 5 结论

孔隙度是评价砂岩露头的重要物性参数,并与岩石光谱特征有明显的关联关系.本文提出了基于高光谱的孔隙度估算新思路.本文采集了4个野外露头剖面的55块砂岩样品孔隙度数据,并分析了其影响因素;通过数据预处理得到反射率和连续统去除两个光谱指标,初步探索了孔隙度的光谱响应机理;引入偏最小二乘方法构建了孔隙度估算模型;通过变量投影重要性分析了模型中的重要波段.

本文的研究表明,砂岩孔隙度受填隙物影响,其

成分及含量影响了砂岩的光谱特征.砂岩孔隙度具有良好的光谱响应,反射率与孔隙度在全波段范围内存在显著的正相关关系, $|R|_{\max}$ 为0.76,连续统去除值 $|R|_{\max}$ 为0.50,具有离散分布的显著性波段.本文优选出反射率用于孔隙度估算模型建立,反射率具有定量估算孔隙度的能力(全波段模型 $R^2 = 0.72$ , $RMSE = 2.28$ , $RPD = 1.94$ ).VIP值指示了孔隙度估算模型中的重要波段,帮助找到地面高光谱成像仪的孔隙度敏感响应波段、降低自变量维度.

本文提出的基于高光谱的砂岩露头孔隙度估算方法具有可行性,为基于高光谱图像反演孔隙度、野外露头孔隙度表征提供方法支持.下一步的研究重点是深入分析高光谱对孔隙度的研究机理,为提高孔隙度估算精度奠定基础.

## References

- [1] ZHU Ru-Kai, BAI Bin, YUAN Xuan-Jun, *et al.* A new approach for outcrop characterization and geostatistical analysis of meandering channels sandbodies within a delta plain setting using digital outcrop models: Upper Triassic Yanchang tight sandstone Formation, Yanhe outcrop, Ordos Basin [J]. *Acta Sedimentologica Sinica* (朱如凯, 白斌, 袁选俊等. 利用数字露头模型技术对曲流河三角洲沉积储层特征的研究. *沉积学报*), 2013, **31**(5):867-877.
- [2] LI Xiao-Chun. Research on sedimentary facies and reservoir characteristics of Xujiahe Formation in Anyue-Jianyang area [D]. Chengdu University of Technology (李晓春. 安岳~简阳地区须家河组沉积相及储层特征研究. 成都理工大学), 2015.
- [3] ZHONG Da-Kang, ZHU Xiao-Min, ZHANG Zhi-Huang, *et al.* Controlling factors of sandstone reservoir of the Paleogene in Dongying Sag [J]. *Petroleum Exploration and Development* (钟大康, 朱筱敏, 张枝焕, 等. 东营凹陷古近系砂岩储集层物性控制因素评价. *石油勘探与开发*), 2003, **30**(3):95-98.
- [4] GAO Zhi-Yong, FENG Jia-Rui, CUI Jing-Gang, *et al.* Physical simulation and quantitative calculation of increased feldspar dissolution pores in deep reservoirs [J]. *Petroleum Exploration and Development* (高志勇, 冯佳睿, 崔京钢, 等. 深层储集层长石溶蚀增孔的物理模拟与定量计算. *石油勘探与开发*), 2017, **44**(3):359-369.
- [5] MA Shi-Yu. Research on physical property feature and controlling factor of Yanchang Formation in Longdong area of Ordos Basin [D]. China University of Petroleum (Beijing) (马石玉. 鄂尔多斯盆地陇东地区延长组储层物性特征与控制因素. 中国石油大学(北京)), 2011.
- [6] LI Feng-Jie, WANG Duo-Yun, XU Xu-Hui. The influential factors and characteristic of triassic Yanchang Formation reservoir in Longdong area, Ordos Basin [J]. *Petroleum Geology & Experiment* (李凤杰, 王多云, 徐旭辉. 鄂尔多斯盆地陇东地区三叠系延长组储层特征及影响因素分析. *石油实验地质*), 2005, **27**(4):365-370.
- [7] ZHANG Ji-Zhi, CHEN Shi-Jia, XIAO Yan, *et al.* Characteristics of the Chang 8 tight sandstone reservoirs and their genesis in Huaqing area, Ordos Basin [J]. *Oil & Gas Geolo-*



- gy (张纪智, 陈世加, 肖艳, 等. 鄂尔多斯盆地华庆地区长 8 致密砂岩储层特征及其成因. *石油与天然气地质*), 2013, **34**(5):679-684.
- [8] Bioucas-Dias J M, Plaza A, Camps-Valls G, *et al.* Hyperspectral remote sensing data analysis and future challenges [J]. *IEEE Geoscience & Remote Sensing Magazine*, 2013, **1**(2):6-36.
- [9] MA Xiu-Qiang, PENG Ling, XU Su-Ning, *et al.* Application of GF-2 satellite data to mine geological environment investigation in Daye, Hubei [J]. *Remote Sensing for Land & Resources* (马秀强, 彭令, 徐素宁, 等. 高分二号数据在湖北大冶矿山地质环境调查中的应用. *国土资源遥感*), 2017, **29**(s1):127-131.
- [10] ZHANG Zong-Gui. The research of method techniques on mineral and rock identification using imaging spectrometer remote sensing data and analyzing the influence factors [D]. China University of Petroleum (Beijing) (张宗贵. 成像光谱岩矿识别方法技术研究和影响因素分析. 中国地质大学(北京)), 2004.
- [11] WANG Run-Sheng, XIONG Sheng-Qing, NIE Hong-Feng, *et al.* Remote sensing technology and its application in geological exploration [J]. *Acta Geologica Sinica* (王润生, 熊盛青, 聂洪峰, 等. 遥感地质勘查技术与应用研究. *地质学报*), 2011, **85**(11):1699-1743.
- [12] Hapke B. Bidirectional reflectance spectroscopy: 6. Effects of porosity [J]. *Icarus*, 2008, **195**(2):918-926.
- [13] TONG Qing-Xi, ZHANG Bin, ZHENG Lan-Fen. *Hyperspectral remote sensing* [M]. Higher Education Press (童庆禧, 张兵, 郑兰芬. 高光谱遥感:原理、技术与应用. 高等教育出版社), 2006.
- [14] Cloutis E A, Gaffey M J, Moslow T F. Characterization of minerals in oil sands by reflectance spectroscopy [J]. *Fuel*, 1995, **74**(6):874-879.
- [15] Mielke C, Rogass C, Boesche N, *et al.* EnGeoMAP 2.0—automated hyperspectral mineral identification for the German EnMAP space mission [J]. *Remote Sensing*, 2016, **8**(2):127-153.
- [16] Brossard M, Marion R, Carrère V. Deconvolution of SWIR reflectance spectra for automatic mineral identification in hyperspectral imaging [J]. *Remote Sensing Letters*, 2016, **7**(6):581-590.
- [17] YANG Chang-Bao, LIU Na, ZHOU Zhen-Chao, *et al.* Research on the relationship between main rock metal elements content, physical parameters and spectral features in Tahe area [J]. *Spectroscopy and Spectral Analysis* (杨长保, 刘娜, 周振超, 等. 塔河地区主要岩石金属元素含量、物性参数与光谱特征间的关系研究. *光谱学与光谱分析*), 2017, **37**(8):2569-2574.
- [18] YOU Jin-Feng, XIN Li-Xin, PAN Jun, *et al.* Characteristics of spectral response of oil sand components and their significance [J]. *Chinese Journal of Geophysics* (尤金凤, 邢立新, 潘军, 等. 油砂组分遥感光谱响应特征及意义. *地球物理学报*), 2016, **59**(8):3068-3078.
- [19] YAO Jing-Li, DENG Xiu-Qin, ZHAO Yan-De, *et al.* Characteristics of tight oil in Triassic Yanchang Formation, Ordos Basin [J]. *Petroleum Exploration and Development* (姚泾利, 邓秀芹, 赵彦德, 等. 鄂尔多斯盆地延长组致密油特征. *石油勘探与开发*), 2013, **40**(2):150-158.
- [20] ZHOU Jin-Gao, YAO Gen-Shun, DENG Hong-Ying, *et al.* Exploration potential of Chang 9 member, Yanchang Formation, Ordos Basin [J]. *Petroleum Exploration and Development* (周进高, 姚根顺, 邓红婴, 等. 鄂尔多斯盆地延长组长 9 油层组勘探潜力分析. *石油勘探与开发*), 2008, **35**(3):289-293.
- [21] YANG Hua, LIU Xian-Yang, Zhang Cai-Li, *et al.* The main controlling factors and distribution of low permeability lithologic reservoirs of Triassic Yanchang Formation in Ordos Basin [J]. *Lithology Reservoir* (杨华, 刘显阳, 张才利, 等. 鄂尔多斯盆地三叠系延长组低渗透岩性油藏主控因素及其分布规律. *岩性油气藏*), 2007, **19**(3):1-6.
- [22] PENG Jie, CHI Chun-Ming, XIANG Hong-Ying, *et al.* Inversion of soil salt content based on continuum-removal method [J]. *Acta Pedologica Sinica* (彭杰, 迟春明, 向红英, 等. 基于连续统去除法的土壤盐分含量反演研究. *土壤学报*), 2014, **51**(3):459-469.
- [23] YOU Jin-Feng. Research on spectral response mechanisms of oil sands components and ore prospecting by remote sensing technology [D]. Jilin University (尤金凤. 油砂组分光谱响应机理及遥感找矿研究. 吉林大学), 2015.
- [24] WANG Hui-Wen. *Partial least-squares regression-method and application* [M]. Beijing: National Defense Industry Press (王惠文. 偏最小二乘回归方法及其应用. 北京: 国防工业出版社), 1999.
- [25] ZHANG Fang, XIONG Hei-Gang, LUAN Fu-Ming, *et al.* Characteristics of field-measured spectral response to alkalization soil [J]. *J. Infrared Millim. Waves* (张芳, 熊黑钢, 栾福明, 等. 土壤碱化的实测光谱响应特征. *红外与毫米波学报*), 2011, **30**(1):55-60.
- [26] SHEN Zhang-Quan, WANG Ke, HUANG Xuewen. Estimating the content of soil carbon by using near-infrared spectra [J]. *J. Infrared Millim. Waves* (沈掌泉, 王珂, Xuewen HUANG. 用近红外光谱预测土壤碳含量的研究. *红外与毫米波学报*), 2010, **29**(1):32-37.
- [27] Nanni M R, Cezar E, Junior C A D S, *et al.* Partial least squares regression (PLSR) associated with spectral response to predict soil attributes in transitional lithologies [J]. *Archives of Agronomy & Soil Science*, 2017, **64**(5):682-695.
- [28] Buondonno A, Amenta P, Viscarra-Rossel R A, *et al.* Prediction of Soil Properties with PLSR and vis-NIR Spectroscopy: Application to Mediterranean Soils from Southern Italy [J]. *Current Analytical Chemistry*, 2012, **8**(2):283-299.
- [29] YU Lei, HONG Yong-Sheng, GENG Lei, *et al.* Hyperspectral estimation of soil organic matter content based on partial least squares regression [J]. *Transactions of the Chinese Society of Agricultural Engineering* (*Transactions of the CSAE*) (于雷, 洪永胜, 耿雷, 等. 基于偏最小二乘回归的土壤有机质含量高光谱估算. *农业工程学报*), 2015, **31**(14):103-109.
- [30] Chong I G, Jun C H. Performance of some variable selection methods when multicollinearity is present [J]. *Chemometrics & Intelligent Laboratory Systems*, 2005, **78**(1):103-112.
- [31] Farrés M, Platikanov S, Tsakovski S, *et al.* Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation [J]. *Journal of Chemometrics*, 2015, **29**(10):528-536.