

基于光谱 FOD 与优化指数的银川平原土壤有机质含量反演

张俊华^{1,2} 尚天浩³ 陈睿华⁴ 王怡婧⁴ 丁启东¹ 李小林¹

(1. 宁夏大学生态环境学院, 银川 750021; 2. 西北土地退化与生态恢复国家重点实验室培育基地, 银川 750021;
3. 西安煤航遥感信息有限公司, 西安 710199; 4. 宁夏大学地理科学与规划学院, 银川 750021)

摘要: 土壤有机质(SOM)是土壤肥力的重要组成部分,是作物生长的主要养分来源。为探究分数阶微分(FOD)联合优化光谱指数对低肥力地区 SOM 的反演效果,以银川平原为研究对象,对野外土壤高光谱反射率原始数据进行 0~2 阶 FOD 处理(间隔 0.2 阶),构建优化光谱指数 DI/RDI、DI/NDI、NDI/RDI、RDI/NDI、DI/GDI 和 RI/GDI,分析各指数与土壤有机质含量间的二维相关性,筛选出最佳优化光谱指数,并建立基于支持向量机(SVM)的 SOM 含量反演模型。结果表明:银川平原 SOM 含量整体偏低,其中 93.05% 处于四级到六级水平。土壤野外原始光谱反射率吸收特征差异明显,在 1 400、1 900 nm 处有明显吸收峰。随着分数阶的不断增加,光谱反射率不断趋近于 0。土壤 DI/NDI、DI/GDI、RI/GDI、NDI/RDI 和 RDI/NDI 在 0~2 阶最大相关系数绝对值(MACC)均小于 0.80,DI/RDI 在 0.2~2.0 阶范围内的 MACC 为 0.996 5~0.998 6,其敏感波段主要集中在 1 450~1 750 nm 和 2 100~2 400 nm 之间。基于 0.2 阶微分处理的 DI/RDI-SVM 模型对 SOM 的反演精度最佳,建模决定系数 R_c^2 和验证决定系数 R_p^2 分别为 0.98 和 0.99,相对分析误差(RPD)为 4.31。研究结果可为低肥力地区的 SOM 含量快速、准确反演及制图提供科学依据。

关键词: 低含量有机质; 光谱特征; 分数阶微分; 支持向量机; 反演; 克里格插值

中图分类号: S127; TP79 **文献标识码:** A **文章编号:** 1000-1298(2022)11-0379-09

OSID:



Inversion of Soil Organic Matter Content in Yinchuan Plain Using Field Spectral Fractional-order Derivatives Combined with Spectral Optimization Index

ZHANG Junhua^{1,2} SHANG Tianhao³ CHEN Ruihua⁴ WANG Yijing⁴ DING Qidong¹ LI Xiaolin¹

(1. School of Ecology and Environment, Ningxia University, Yinchuan 750021, China

2. Breeding Base for State Key Laboratory of Land Degradation and Ecological Restoration in Northwestern China, Yinchuan 750021, China

3. Xi'an Meihang Remote Sensing Information Co., Ltd., Xi'an 710199, China

4. College of Geographical Sciences and Planning, Ningxia University, Yinchuan 750021, China)

Abstract: Soil organic matter (SOM) is an important part of soil fertility and the main nutrient source for crop growth. In order to explore the inversion effect of fractional-order derivatives (FOD) combined with spectral optimization index on SOM in low fertility areas, taking Yinchuan Plain as the study object, the original data of hyperspectral reflectance of field were processed by 0~2 order FOD (with an interval of 0.2 order) after log reciprocal transformation, the spectral optimization indices DI/RDI, DI/NDI, NDI/RDI, RDI/NDI, DI/GDI and RI/GDI were constructed, the two-dimensional correlation between each index and SOM content was analyzed, the optimal spectral optimization index was selected, and a support vector machine (SVM) model was established to inverse the SOM content. The results showed that the content of SOM in Yinchuan Plain was generally low, of which 93.05% was at the level of 4~6 class. There were obvious differences in the absorption characteristics of the original spectral reflectance of soil in the field, with obvious absorption peaks at 1 400 nm and 1 900 nm. With the increasing fractional order, the spectral reflectance was approaching 0. The maximum absolute correlation coefficient (MACC) values of soil DI/NDI, DI/GDI, RI/GDI, NDI/RDI and RDI/NDI were all less than 0.80 in

收稿日期: 2022-07-15 修回日期: 2022-09-15

基金项目: 国家重点研发计划项目(2021YFD1900602)、国家自然科学基金项目(42067003)和清华大学-宁夏银川水联网数字治水联合研究院联合开放基金项目(SKLHSE-2022-LOW11)

作者简介: 张俊华(1977—),女,研究员,博士生导师,主要从事土壤信息快速获取和土壤质量提升研究,E-mail: zhangjunhua728@163.com

order 0 ~ 2. The MACC values of DI/RDI in order 0.2 ~ 2.0 were ranged from 0.996 5 to 0.998 6, and their sensitive bands were mainly concentrated in 1 450 ~ 1 750 nm and 2 100 ~ 2 400 nm. The model inversion accuracy based on DI/RDI - SVM model was the best at order 0.2, modeling determination coefficient (R_c^2) and verification determination coefficient (R_p^2) were 0.98 and 0.99, and residual predictive derivation (RPD) got 4.31. The results can provide scientific basis for rapid and accurate estimation and mapping of SOM in areas with low organic matter content.

Key words: low content of soil organic matter; spectral characteristics; fractional-order derivative; support vector machine; inversion; Kriging interpolation method

0 引言

土壤有机质(Soil organic matter, SOM)是植物营养物质的主要来源之一,在土壤肥力质量评价中具有重要地位^[1]。SOM 成分复杂,光谱特性差异大,但总体呈现出 SOM 含量升高,整个谱线反射率逐渐降低的趋势^[2]。但当 SOM 含量较低,光谱波段被吸收部分的能量过低时,光谱特性分析效果不佳^[3]。ABERGAZ 等^[4]指出, SOM 含量(质量比)小于 20 g/kg 时,光谱曲线易受其他成分的影响,不能准确反映由官能团引起的吸收特征。因此,如何提高低含量 SOM 的反演精度是相关领域的研究热点之一。

利用不同的反射率变换方式筛选 SOM 敏感波段和采用不同建模方法都是提高低含量 SOM 反演模型精度的有效手段。XIE 等^[5]研究指出,与反射率的一阶微分、倒数的一阶微分、平方根的一阶微分和对数的一阶微分变换方式相比,特征波段选择一阶微分可以显著提高低含量 SOM 的反演精度。LIU 等^[6]在一阶微分、二阶微分和吸光度对数、吸光度对数的一阶和二阶微分及连续统去除等变换方式的基础上,采用竞争自适应加权算法筛选出 15 ~ 40 个不等的特征波段,发现采用连续统去除-随机森林建立反演陕西省靖边县 SOM 模型精度达 0.96,相对分析误差(Residual predictive derivation, RPD)为 3.02。在建模方法中,最小二乘法(Least square, LS)^[7-8]、支持向量机(Support vector machine, SVM)^[7-8]、支持向量机分类(Support vector machine classification, SVMC)^[9]、遗传算法(Genetic algorithm, GA)^[10]、最小二乘回归(Least square regression, LSR)^[11]、反向传播神经网络(Back propagation neural network, BPNN)^[12-13]、随机森林(Random forest, RF)^[14-15]等一种或多种方法被用于巴西、以色列荒漠区以及我国河北、新疆、西藏、宁夏等地区低含量 SOM 的反演,都取得了较好的效果。此外,引入其他变量也可以提高 SOM 反演精度^[13,16-17]。虽然变量的引入可以显著提高 SOM 含量的反演精度,但这些变量的获取技术难度较大,所以其应用和推广有一定困难。

为提高模型运算效率,减少建模变量输入个数,有学者在敏感波段筛选的基础上,用两个或两个以上特征波段分别进行波段耦合运算,构建比值指数(Ratio index, RI)、差值指数(Difference index, DI)、土壤沙化指数(Soil desertification index, SDI1)、土壤退化指数(Soil degradation index, SDI2)、修正归一化差分指数(Corrected normalized difference index, CNDI)等^[18-20],也有学者尝试在二维空间建立光谱数据与 SOM 间的全波段相关性运算,以解决单一敏感波段在一维空间构建光谱指数所造成的光谱信息缺失问题^[21-22]。二维空间所建模型精度较一维空间所建模型精度得到显著提升,但因研究区域不同,导致所选最佳建模光谱指数存在较大差异。综合来看,已有研究或基于不同二维光谱指数分别建模,或以分数阶微分(Fractional-order derivatives, FOD)、整数阶微分联合单一光谱指数建模,将 FOD 联合优化指数进行较低 SOM 含量反演的报道相对较少。

尚天浩等^[23]对比分析了基于岭回归、偏最小二乘回归、SVM 和反向神经网络的银川平原 SOM 反演模型,发现 SVM 模型精度相对最高(R_c^2 为 0.74),但仍未达到理想水平,说明光谱常规变换并不能准确凸显该地区较低含量 SOM 在光谱信息中的细微变化。因此,本研究以分数阶微分进行光谱数据预处理,结合不同光谱指数进行二维波段间的算法优化,建立“FOD + 优化光谱指数”的 SVM 模型,为银川平原 SOM 的定量反演提供科学依据。

1 材料与方法

1.1 研究区概况

银川平原(37°50' ~ 39°23'N, 104°17' ~ 107°39'E)地处宁夏北部,是我国粮食主产区之一,属温带干旱气候,年均降水量约 190 mm,全年蒸发量约 1 750 mm。土壤类型多以灌淤土、灰钙土和盐碱土为主^[24], SOM 含量普遍偏低。

1.2 土壤样本采集

根据银川平原气候、水文、地质地貌和土壤等基础资料数据,主要以农田土壤为研究对象,利用 ArcGIS 10.4 预设 5 km × 5 km 网格。于 2019 年 5

月以梅花五点法采集裸露地表 0 ~ 20 cm 的土样 (以光谱测定点为中心, 半径为 1 m 距离 4 个方向处各选 1 个点, 共 5 个点), 混合均匀后保留 0.5 kg 作为该点土样。采样过程中受建筑、道路和水体的影响, 适当调整采样点 (图 1)。土样风干后采用外加热-重铬酸钾法测定 SOM 含量, 共获得有效土壤样本 187 个。

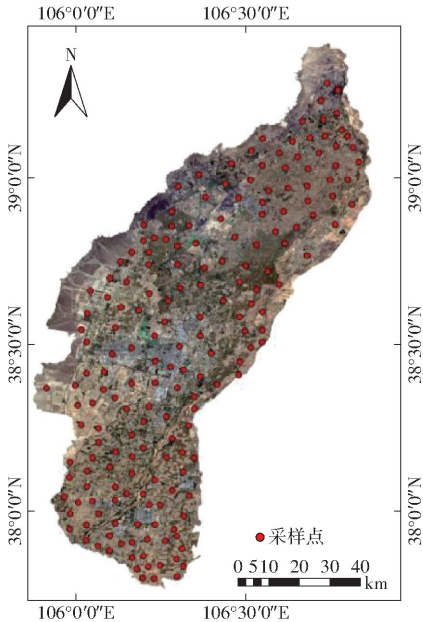


图 1 研究区及样点分布

Fig. 1 Location of study area and sample points

1.3 野外光谱采集及预处理

采用 ASD FieldSpec4 型地物光谱仪进行野外原位土壤的高光谱测定, 波段为 350 ~ 2 500 nm, 每次测定前均进行白板校正, 光谱探头始终保持在距地面约 80 cm, 每一样点重复测定 5 次, 取平均值作为

该样点的光谱反射率。测定时间为 10:00—14:00, 天气状况良好, 晴朗无风。

为消除仪器噪声和环境背景干扰, 去除噪声过大的边缘波段 (350 ~ 399 nm 和 2 401 ~ 2 500 nm)。在 Matlab 2018b 软件中对 400 ~ 2 400 nm 范围内的高光谱数据作 S - G 平滑去噪处理。FOD 作为一种光谱信号处理技术^[22], 通过细化阶数改变光谱曲线中的斜率和曲率, 使不同分数阶下光谱曲线间的特征波段处微弱差异得到明显放大^[25]。本研究以整数阶微分定义推广出的 G - L 算法进行分数阶微分计算, 计算式为

$$\frac{d^v f(\lambda)}{d\lambda^v} \approx f(\lambda) - v f(\lambda - 1) - \frac{v(-v+1)}{2} f(\lambda - 2) + \dots + \frac{I'(-v+1)}{n! I'(-v+n+1)} f(\lambda - n) \tag{1}$$

式中 n ——上下限微分阶数之差
 $f(\lambda)$ ——以波长 λ 处反射率为自变量的函数
 v ——阶数

本研究分数阶设置为 0 ~ 2 阶, 以 0.20 阶为间隔, 共 11 阶。

1.4 优化光谱指数

基于高光谱遥感信息进行表层土壤属性快速反演, 其中光谱信息多以连续窄波段为主。本文在全波段 400 ~ 2 400 nm 范围, 通过对差值指数、比值指数、归一化指数 (Normalized difference index, NDI)、再归一化差值指数 (Renormalized difference index, RDI) 和广义差值指数 (Generalized difference index, GDI) 进行两两算法优化 (表 1), 建立全波段 (400 ~ 2 400 nm) 与 SOM 含量间的二维相关性图。

表 1 优化光谱指数及公式

Tab. 1 Optimize spectral index and formula

优化光谱指数	DI/RDI	DI/NDI	NDI/RDI	RDI/NDI	DI/GDI	RI/GDI
公式	$\sqrt{R_i + R_j}$	$R_i + R_j$	$\frac{\sqrt{R_i + R_j}}{R_i - R_j}$	$\frac{R_i + R_j}{\sqrt{R_i + R_j}}$	$\frac{R_i - R_j}{\sqrt{R_i^2 + R_j^2}}$	$\frac{R_i}{R_j \sqrt{R_i^2 + R_j^2}}$

注: R_i 和 R_j 分别表示 400 ~ 2 400 nm 中的任意两个波段, 且 $R_i \neq R_j$ 。

1.5 建模方法与精度评价

SVM 模型基于结构风险最小化处理, 能够很好地解决模型中存在的局部最优问题^[26]。本研究选用决定系数 R^2 、均方根误差 (Root mean square error, RMSE) 和 RPD 作为所建模型的精度和预测效果评价。一般将 RPD 分为 3 类来评价模型的可靠性: RPD 小于等于 1.4, 表明模型预测能力极差; RPD 大于 1.4、小于等于 2.0, 表明模型能够具有较为可靠预测能力; RPD 大于 2.0, 表明模型预测能力极好^[23]。

2 结果与分析

2.1 银川平原土壤有机质含量特征

根据全国第二次土壤普查推荐的土壤肥力分级标准, 将银川平原 187 个土样按 SOM 含量共分为 6 个等级 (表 2): SOM 含量为 2.81 ~ 56.14 g/kg, 均值 12.48 g/kg、变异系数 57.26%, 其中处于四级 ~ 六级 (SOM 含量小于 20 g/kg) 的土样有 174 个, 占总数的 93.05%, 该范围 SOM 含量平均值为 11.06 g/kg, 变异系数为 39.30%, 本文选取这 174 个土样为研究对象。

表 2 银川平原土壤 SOM 统计特征

Tab. 2 Basic statistical characteristics of soil organic matter content of Yinchuan Plain

分级	样本数/个	有机质含量/ (g·kg ⁻¹)	最大值/ (g·kg ⁻¹)	最小值/ (g·kg ⁻¹)	平均值/ (g·kg ⁻¹)	标准差/ (g·kg ⁻¹)	变异系数/%
一级	2	>40	56.14	52.05	54.09	2.89	5.34
二级	3	30~40	37.08	30.20	32.74	12.08	36.90
三级	8	20~30	27.20	21.99	24.20	3.60	41.00
四级	107	10~20	19.31	10.07	13.99	4.35	31.10
五级	38	6~10	9.92	6.04	8.19	2.75	33.53
六级	29	0~6	6.00	2.81	4.33	2.38	55.03

利用 K-S 算法^[27] 中的欧氏距离划分总样本集,其中建模和验证样本分别为 117 和 57 (表 3)^[23],所划分的建模数据和验证数据较总样本 SOM 含量区间值和标准差较为一致,变异系数说明 SOM 均属中度变异。

表 3 建模与验证土壤样本有机质含量统计特征

Tab. 3 Statistical characteristics of soil organic matter content of model calibration and validation set

数据集	样本数/个	最大值/ (g·kg ⁻¹)	最小值/ (g·kg ⁻¹)	均值/ (g·kg ⁻¹)	标准差/ (g·kg ⁻¹)	变异系数/%
总样本	174	19.31	2.81	11.02	4.36	39.6
建模样本	117	19.25	2.81	11.01	4.36	39.7
验证样本	57	19.31	3.01	11.16	4.31	38.6

2.2 基于原始光谱反射率的分数阶微分光谱特征

由图 2 可知,原始光谱反射率吸收带整体较宽,吸收特征差异明显,其中 1 400、1 900 nm 处存在明显吸收峰。在 400~1 800 nm 范围内,光谱反射率随波长的增加而增大;在 1 900~2 200 nm 范围内,光谱重叠严重。随着分数阶的不断 增加,重叠峰和基线漂移逐渐消除,光谱强度的幅度下降,光谱反射率不断趋近于 0,其中 1 900 nm 处的吸收谷逐渐由正峰变为负峰;近红外波段玫红色区域的光谱标准差差异明显大于可见光区域。当 FOD 从 0 阶增加到 1.6 阶时,土壤光谱反射率逐渐降低并最终稳定在 -0.07~0.07;当从 1.6 阶增加到 2.0 阶时,土壤光谱反射率在 -0.06~0.06 稳定不变,且高分数阶微分的吸收谷明显小于低分数阶微分。

2.3 不同分数阶下的光谱指数与土壤有机质含量相关性 & 敏感波段优选

比较土壤 NDI/RDI、DI/NDI、DI/RDI、RDI/NDI、DI/GDI 和 RI/GDI 在不同分数阶微分下的最大相关系数绝对值 (Maximum absolute correlation coefficient, MACC) 可以看出 (图 3),优化指数 DI/RDI 与 SOM 的 MACC 显著高于其他指数,平均值高达 0.997 6,其次为 NDI/RDI 和 RDI/NDI,均值分别为 0.740 6 和 0.730 2,DI/NDI 最低,MACC 均

值为 0.309 3。同时,DI/RDI、NDI/RDI 和 RDI/NDI 与 SOM 含量间的 MACC 在 0.8 阶时达到峰值后逐渐下降,此时 MACC 分别为 0.998 6、0.762 1 和 0.799 3。

为筛选最佳建模变量,本研究以 MACC 大于等于 0.800 0 为阈值进行优选。土壤 DI/NDI、DI/GDI、RI/GDI、NDI/RDI 和 RDI/NDI 在 0~2 阶范围内并无 MACC 大于等于 0.800 0 的敏感波段组合存在,故在后续建模中不作考虑。DI/RDI 在 0.2~2.0 阶范围内的 MACC 为 0.996 5~0.998 6,全部高于 0.800 0,故在后续建模中采用 DI/RDI 指数。

2.4 优化光谱指数与土壤有机质含量二维相关性分析

图 4 为 SOM 含量与优化光谱指数 DI/RDI 的二维相关系数分布图,该图能够清晰展示相关系数较高特征指数的波段分布区间。与优化指数 NDI/RDI 和 RDI/NDI 的最大 MACC 相比 (分布图未展示),DI/RDI 最大相关系数绝对值整体提高 0.25,其敏感波段主要集中在 1 450~1 750 nm 和 2 100~2 400 nm 之间 (图 4),这也进一步说明分数阶微分联合优化指数在较低含量 SOM 的敏感波段筛选中,能够有效消除土壤水分敏感波段在 1 400 nm 和 1 900 nm 处的干扰^[9]。

2.5 基于支持向量机的 SOM 含量反演模型构建

以通过阈值筛选的 DI/RDI 指数为建模变量,SOM 含量为因变量,分别基于不同 FOD 下的优化指数建立 SVM 模型 (表 4)。在优化指数 DI/RDVI 中,0.2~2.0 阶范围内 R_c^2 和 R_p^2 整体为 0.81~0.99, RMSE 为 0.84~1.38 g/kg,其中 0.2 阶和 0.4 阶时模型的 R_c^2 和 R_p^2 均高于 0.95,RPD 分别为 4.31 和 4.27,故 0.2 阶下的 DI/RDI-SVM 模型对较低含量 SOM 具有最佳反演精度。

2.6 克里格模型与银川平原 SOM 含量分布反演

在 ArcGIS 10.2 中,利用模型残差结合普通克里格对实测和 0.2 阶下 DI/RDI-SVM 模型的结果分别作反演插值 (图 5),可以看出,银川平原南部

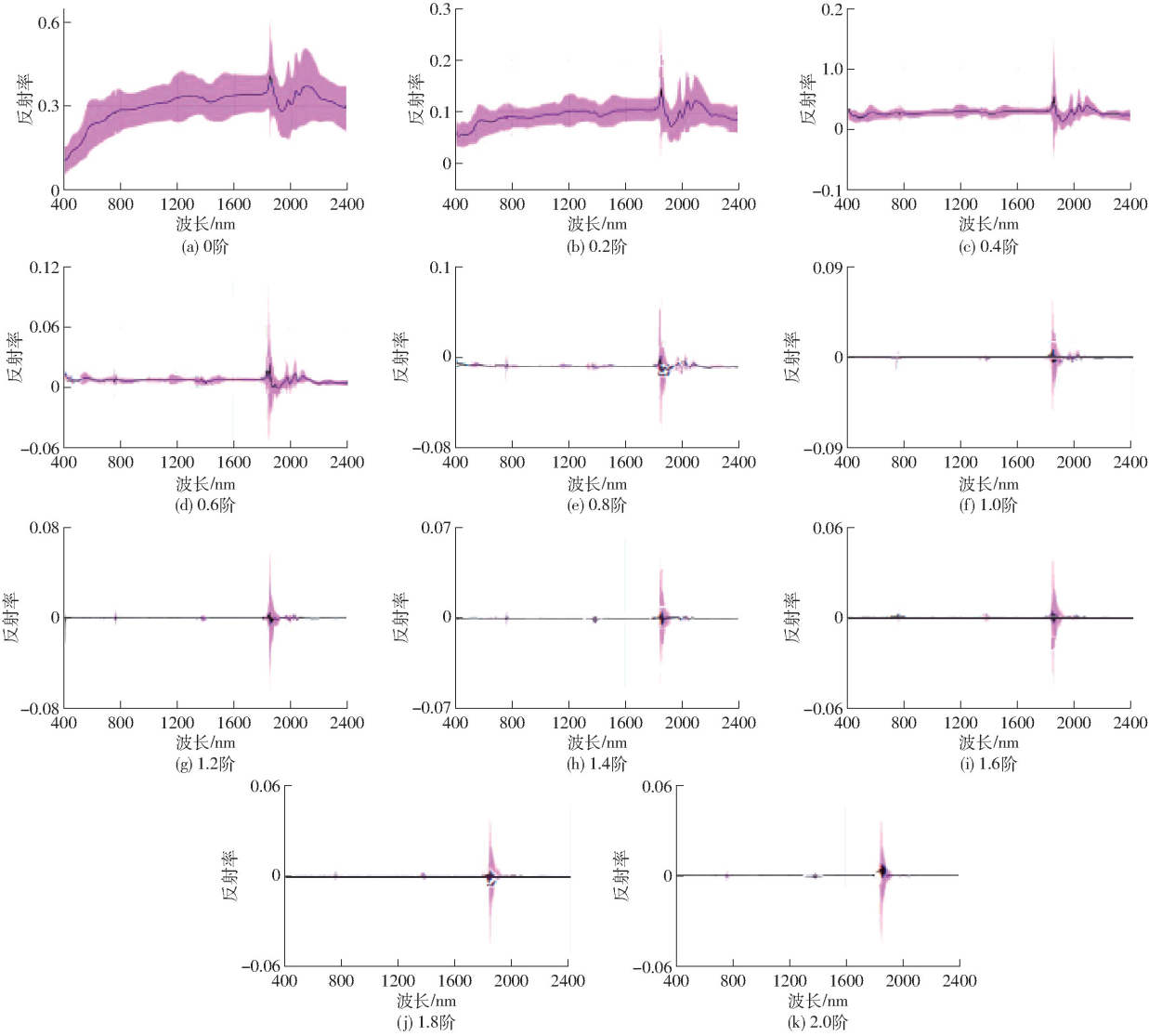


图 2 土壤反射率分数阶微分曲线

Fig. 2 Spectral reflectance fractional differential curves of soil

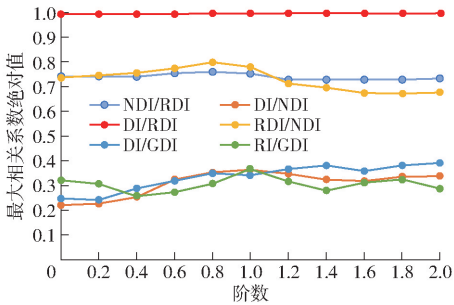


图 3 不同分数阶微分下光谱优化指数最大
相关系数绝对值

Fig. 3 MACC value of spectral optimization index
under different FOD

的吴忠市利通区、青铜峡市和巴音陶亥镇黄河以西地区 SOM 含量明显高于其他地区。贺兰县整体及惠农区东北部 SOM 含量较高,而银川市西夏区西南部、永宁县东西两侧及石嘴山市大武口区 SOM 含量整体较低,这是由于该区域地势低洼、蒸发强烈,加之多年引黄灌溉,导致该区域土壤盐渍

化较重,土壤结构差,保水、保肥性能下降,故 SOM 含量低下。

与实测值相比,四级水平 (SOM 含量 10 ~ 20 g/kg) SOM 的反演值所占面积比实测值高 2.80% (表 5),而五级 (SOM 含量 6 ~ 10 g/kg) 和六级 (SOM 含量小于 6 g/kg) 水平面积低了 2.07% 和 0.73%,说明该方法反演的银川平原 SOM 在相对较高水平反演值略高于实测值,而低水平反演值则略低于实测值。

3 讨论

光谱技术在快速监测土壤属性方面具有较大优势,其中高光谱以光谱信息丰富、波段连续性强而用于土壤属性的快速反演^[28],但同时也存在数据冗余、共线性强等特点。光谱微分变换则在消除基线漂移或平滑背景的条件下,提供较原始光谱精度更高的分辨率和更清晰的光谱轮廓,增强反射率与土

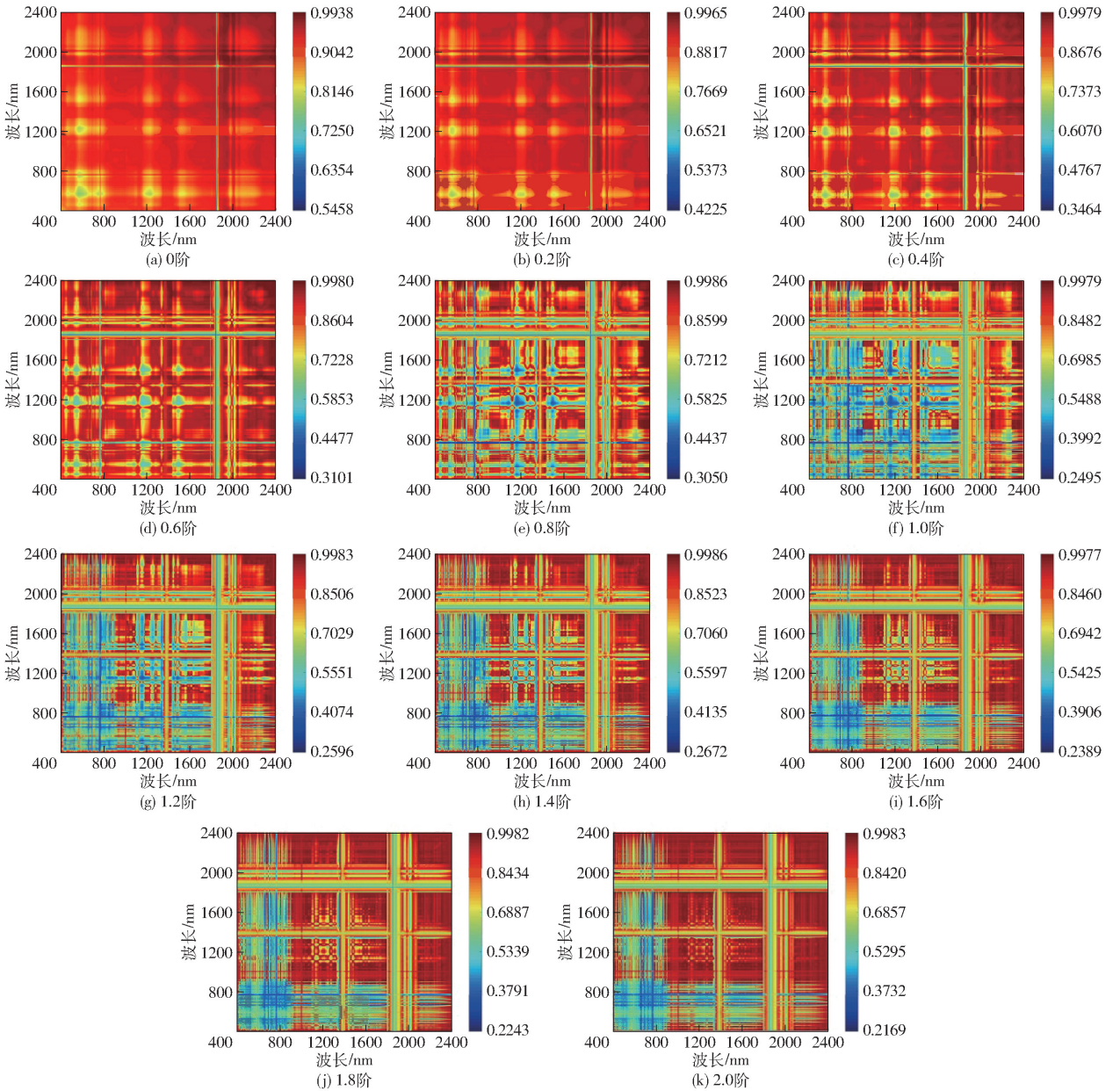


图 4 最优 DI/RDI 指数与 SOM 含量的二维相关系数

Fig. 4 Two-dimensional correlation coefficient between optimal DI/RDI index and SOM content

表 4 基于 DI/RDI - SVM 的 SOM 建模与预测结果
Tab. 4 Modeling and validation results of SOM
based on DI/RDI - SVM

阶数	建模集		验证集		
	R_c^2	RMSE/ ($\text{g}\cdot\text{kg}^{-1}$)	R_p^2	RMSE/ ($\text{g}\cdot\text{kg}^{-1}$)	RPD
0.2	0.98	0.92	0.99	0.84	4.31
0.4	0.99	0.87	0.99	0.84	4.27
0.6	0.81	1.38	0.82	1.32	2.03
0.8	0.82	1.12	0.84	1.10	2.20
1.0	0.81	1.13	0.91	0.89	2.36
1.2	0.82	1.16	0.86	1.08	2.19
1.4	0.82	1.16	0.86	1.06	2.22
1.6	0.90	1.21	0.84	1.06	2.04
1.8	0.92	1.18	0.87	1.01	2.27
2.0	0.97	0.88	0.95	1.21	2.85

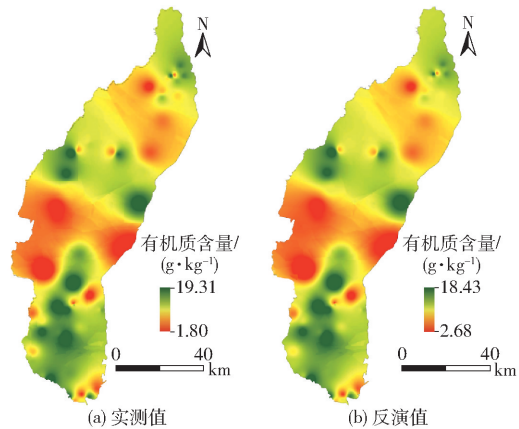


图 5 银川平原土壤有机质含量实测值和 0.2 阶下
DI/RDI - SVM 反演值插值图

Fig. 5 Measured and inversion values of 0.2 - DI/RDI - SVM
for soil organic matter of Yinchuan Plain

表 5 银川平原不同等级有机质实测值与反演值所占面积

Tab.5 Area occupied by measured and inversion values of different levels of soil organic matter in Yinchuan Plain

类型	面积/km ²				面积比例/%		
	四级	五级	六级	总数	四级	五级	六级
实测	5 204.54	1 804.90	98.56	7 108.00	73.22	25.39	1.39
反演	5 403.58	1 657.66	46.76	7 108.00	76.02	23.32	0.66

壤属性间的相关性^[29],本研究结果也证实了这一观点。当分数阶增加时,重叠峰和基线漂移逐渐消除,但光谱强度的幅度也同时下降,这与 HONG 等^[15]对美国爱荷华州东南部农场裸露表层、ZHANG 等^[20]对新疆艾比湖湿地国家自然保护区裸露表土 SOM 反演中表现一致。在分数阶选择方面,国佳欣等^[30]基于 FOD 进行江西省红壤光谱的敏感波段优选,指出 1.5 阶所建反向神经网络模型预测精度最佳,RPD 为 2.75;张智韬等^[9]利用 FOD 结合 NDI 进行以色列南部 SOM 含量估算,指出黏壤土在 0.6 阶和沙质土在 1.2 阶时达到 RF 模型的最佳反演精度,RPD 分别为 4.32 和 3.69。HONG 等^[15]基于 0.75 阶反射率和最佳频段组合的射频模型,获得了美国爱荷华州东南部农场裸露表层的 SOC 估计最高精度。本研究以 0.2 阶为间隔进行土壤原始高光谱反射率的 FOD 处理,对构建的不同光谱指数进行全波段优化运算,对比分析发现基于建立的模型 RPD 高达 4.31,说明光谱数据分数阶微分处理可以扩大二维光谱的空间,为获取高相关性敏感指数提供了更大可能,有利于提高 SOM 估测模型精度^[28]。WANG 等^[31]对 2015 年 5 月中下旬新疆艾比湖湿地国家自然保护区裸露表土 SOM 的估测中指出,基于反射率的 1.8 阶微分模型最稳健,然后采用偏最小二乘回归进行模型校准,模型 R_c^2 达 0.92,RPD 为 3.42。但 WANG 等^[32]对 2016 年 5 月中下旬和 8 月中上旬相同地区的表土 SOM 含量进行了反演,指出 SOM 含量与最佳遥感指数(RI、DI、NDI)的相关性在 1.2 阶达到峰值,利用灰色关联分析反向神经网络,最优模型为 1.2 阶微分模型。此后,ZHANG 等^[20]对 2017 年 10 月相同地区表土 SOM 的估测中指出,FOD 光谱和 SOM 含量之间的相关性在可见光区域显著增强,且在 1.05 ~ 1.45 阶范围内影响最为明显。由此可以看出,不同区域、相同区域不同质地,甚至相同区域相同质地但不同时期的土壤反射率进行分数阶转换时最佳阶数都有可能不同。

光谱指数能够从二维光谱空间凸显 SOM 的响应特征,充分利用高光谱信息,减少其他土壤信息对 SOM 的影响,降低模型的复杂性、去除冗余信息变量^[29]。郭燕等^[21]分析了 NDI、DI、RI 与 SOM 的二

维相关性,最大相关系数在 0.5 ~ 0.7 之间;HONG 等^[22]基于不同光谱变换的光谱指数与 SOM 的二维相关性,得到最大相关系数为 0.810 0,比一维光谱与 SOM 的相关性有了显著提升。尚天浩等^[23]采用单一光谱指数对银川平原 SOM 估算,发现 RDI 的 MACC 最高,MACC 可达 0.801 0,RDI-SVM 模型估测精度最高,RPD 为 2.32。目前大多应用 RI、DI 和 NDI 等单一光谱指数来反演土壤属性,而两两组合很少使用^[15]。本研究基于 5 种光谱指数的两两算法优化,发现 NDI/RDI 和 RDI/NDI 的 MACC 最大值分别为 0.762 1 和 0.769 3,说明光谱指数 NDI 和 RDI 无论作何算法,其比值优化波段指数间的 MACC 范围都差异不大;与此相比,DI/RDI 的 MACC 显著提升 20% 以上,最高达 0.998 6,表明在土壤低有机质含量条件下,DI/RDI 指数作算法优化时的作用优于单一 NDI 或 RDI 指数。这也说明 SOC 对优化光谱指数的灵敏度明显高于使用单一光谱指数时的灵敏度,从而突出了使用光谱指数组合的优势^[7]。在 SOM 估测建模方法中,SVM 作为一种有效监督机器学习的方法,通过对数据的智能化处理和数据价值的充分挖掘,在一定程度上解决了数据处理过程中的“过学习”和“离散值多”问题^[23]。本研究利用分数阶微分联合不同优化指数所建 SVM 模型,预测 RPD 普遍高于 2.00,其中 DI/RDI-SVM 模型 RPD 高达 4.31,这是因为 SVM 为机器算法,其模型在数据驱动中具有极强的自学习能力,能够准确获取土壤光谱中极为复杂的非线性特征^[33],SVM 模型依靠核函数,经过多次训练后,将输入数据绘制到新的超空间中,在该超空间中执行分离,最终使用 ϵ -不敏感损失函数获得用于数据拟合和预测的最优超空间^[17],该超空间可以容忍小于常数 ϵ 集作为阈值的误差^[34],使得模型拟合精度不断接近 100%^[35]。ZHANG 等^[20]还基于波段优化算法和两波段指数形式构建了一个新的三波段指数——修正归一化差分指数,并指出该三维指数在利用可见-近红外光谱估计土壤其他生化参数方面具有很强的应用潜力。本研究后期也可以尝试采用三维光谱指数建立 SOM 估算模型。此外,还需考虑银川平原土壤属性中的水分、盐分、pH 值和养分等因素对 SOM 光谱反演的影响。

4 结论

- (1) 银川平原土壤样本 SOM 含量平均值为 12.48 g/kg,属于高度变异,其中 93.05% 的土样 SOM 处于四级~六级水平。
- (2) 土壤野外原始光谱反射率吸收带整体较

宽,吸收特征差异明显,在 1 400 nm 和 1 900 nm 处有明显吸收峰。随着分数阶的不断增加,光谱反射率不断趋近于 0;高分数阶微分的吸收谷明显小于低分数阶微分。

(3)优化光谱指标 MACC 从大到小依次为 DI/RDI、NDI/RDI、RDI/NDI、DI/GDI、DI/NDI 和 RI/

GDI,敏感波段主要集中在 1 450 ~ 1 750 nm 和 2 100 ~ 2 400 nm。

(4)基于 0.2 阶微分联合 DI/RDI - SVM 建立的银川平原 SOM 含量反演模型精度最高, R_c^2 和 R_p^2 为 0.98 和 0.99, RMSE 分别为 0.92、0.84 g/kg, RPD 达到 4.31。

参 考 文 献

- [1] YADAV V, MALANSON G. Progress in soil organic matter research: litter decomposition, modeling, monitoring and sequestration[J]. Progress in Physical Geography, 2007, 31(2): 131 - 154.
- [2] SHARMA V, CHAUHAN R, KUMAR R. Spectral characteristics of organic soil matter: a comprehensive review[J]. Microchemical Journal, 2021, 171: 106836.
- [3] VISCARRA R R, HICKS W. Soil organic carbon and its fractions estimated by visible-near infrared transfer functions[J]. Eur. J. Soil Sci., 2015, 66: 438 - 450.
- [4] ABEGAZ A, WINOWIECKI L A, VAGEN T G, et al. Spatial and temporal dynamics of soil organic carbon in landscapes of the upper Blue Nile Basin of the Ethiopian Highlands[J]. Agriculture, Ecosystems & Environment, 2016, 218: 190 - 208.
- [5] XIE S G, DIN F F J, CHEN S G, et al. Prediction of soil organic matter content based on characteristic band selection method[J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2022, 273: 120949.
- [6] LIU J B, DONG Z Y, XIA J S, et al. Estimation of soil organic matter content based on CARS algorithm coupled with random forest[J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2021, 258: 119823.
- [7] ZHANG Z P, DING J L, ZHU C M, et al. Strategies for the efficient estimation of soil organic matter in salt-affected soils through Vis - NIR spectroscopy: optimal band combination algorithm and spectral degradation[J]. Geoderma, 2021, 382: 114729.
- [8] DE SANTANA F B, OTANI S K, DE SOUZA A M, et al. Comparison of PLS and SVM models for soil organic matter and particle size using Vis - NIR spectral libraries[J]. Geoderma Regional, 2021, 27: e00436.
- [9] 张智韬, 劳聪聪, 王海峰, 等. 基于 FOD 和 SVMDA - RF 的土壤有机质含量高光谱预测[J]. 农业机械学报, 2020, 51(1): 156 - 167.
ZHANG Zhitao, LAO Congcong, WANG Haifeng, et al. Estimation of desert soil organic matter through hyperspectral based on fractional-order derivatives and SVMDA - RF[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(1): 156 - 167. (in Chinese)
- [10] SUN W C, LIU S, ZHANG X, et al. Estimation of soil organic matter content using selected spectral subset of hyperspectral data[J]. Geoderma, 2022, 409: 115653.
- [11] 玉米提·买明, 王雪梅. 连续小波变换的土壤有机质含量高光谱估测[J]. 光谱学与光谱分析, 2022, 42(4): 1278 - 1284.
YUMITI Maiming, WANG Xuemei. Hyperspectral estimation of soil organic matter content based on continuous wavelet transformation[J]. Spectroscopy and Spectral Analysis, 2022, 42(4): 1278 - 1284. (in Chinese)
- [12] YU Q, YAO T C, LU H W, et al. Improving estimation of soil organic matter content by combining Landsat 8 OLI images and environmental data: a case study in the river valley of the southern Qinghai - Tibet Plateau[J]. Computers and Electronics in Agriculture, 2021, 185: 106144.
- [13] MORAIS T G, TUFIK C, RATO A E, et al. Estimating soil organic carbon of sown biodiverse permanent pastures in Portugal using near infrared spectral data and artificial neural networks[J]. Geoderma, 2021, 404: 115387.
- [14] MENG X T, BAO Y L, ZHANG X L, et al. Prediction of soil organic matter using different soil classification hierarchical level stratification strategies and spectral characteristic parameters[J]. Geoderma, 2022, 411: 115696.
- [15] HONG Y S, GUO L, CHEN S C, et al. Exploring the potential of airborne hyperspectral image for estimating topsoil organic carbon: effects of fractional-order derivative and optimal band combination algorithm[J]. Geoderma, 2020, 365: 114228.
- [16] 郑文博. 基于多光谱影像的有机质含量反演探究[J]. 科学技术创新, 2021(9): 46 - 47.
ZHENG Wenbo. Retrieval of organic matter content based on multispectral images[J]. Scientific and Technological Innovation, 2021(9): 46 - 47. (in Chinese)
- [17] FORKUOR G, HLUNKPATIN O K L, WELP G, et al. High resolution mapping of soil properties using remote sensing variables in south-western Burkina Faso: a comparison of machine learning and multiple linear regression models[J]. PLoS ONE, 2017, 12(1): e0170478.
- [18] 赵瑞, 崔希民, 刘超. GF - 5 高光谱遥感影像的土壤有机质含量反演估算研究[J]. 中国环境科学, 2020, 40(8): 3539 - 3545.
ZHAO Rui, CUI Ximin, LIU Chao. Inversion estimation of soil organic matter content based on GF - 5 hyperspectral remote

- sensing image[J]. *China Environmental Science*, 2020, 40(8): 3539–3545. (in Chinese)
- [19] 张娟娟. 土壤养分信息的光谱估测研究[D]. 南京: 南京农业大学, 2009.
ZHANG Juanjuan. Estimation soil nutrient information based on spectral analysis technology [D]. Nanjing: Nanjing Agricultural University, 2009. (in Chinese)
- [20] ZHANG Z P, DING J L, WANG J Z, et al. Prediction of soil organic matter in northwestern China using fractional order derivative spectroscopy and modified normalized difference indices[J]. *Catena*, 2020, 185: 104257.
- [21] 郭燕, 纪文君, 吴宏海, 等. 基于野外 Vis – NIR 光谱的土壤有机质预测与制图[J]. *光谱与光谱分析*, 2013, 33(4): 1135–1140.
GUO Yan, JI Wenjun, WU Honghai, et al. Estimation and mapping of soil organic matter based on Vis – NIR reflectance spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2013, 33(4): 1135–1140. (in Chinese)
- [22] HONG Y S, CHEN S C, ZHANG Y, et al. Rapid identification of soil organic matter level via visible and near-infrared spectroscopy: effects of two-dimensional correlation coefficient and extreme learning machine[J]. *Science of the Total Environment*, 2018, 644: 1232–1243.
- [23] 尚天浩, 毛鸿欣, 张俊华, 等. 基于 PCA 敏感波段筛选与 SVM 建模的银川平原土壤有机质高光谱估算[J]. *生态学杂志*, 2021, 40(12): 4128–4136.
SHANG Tianhao, MAO Hongxin, ZHANG Junhua, et al. Hyperspectral estimation of soil organic matter content in Yinchuan Plain, China based on PCA sensitive band screening and SVM modeling[J]. *Chinese Journal of Ecology*, 2021, 40(12): 4128–4136. (in Chinese)
- [24] 马玉兰, 陈润邑, 张学俭, 等. 宁夏耕地土壤与肥力[M]. 银川: 黄河出版传媒集团阳光出版社, 2020.
- [25] HONG Y S, SHEN R L, CHEN H, et al. Estimating lead and zinc concentrations in periurban agricultural soils through reflectance spectroscopy: effects of fractional order derivative and random forest[J]. *Science of the Total Environment*, 2019, 165(2): 1969–1982.
- [26] PENG X T, SHI T Z, SONG A H, et al. Estimating soil organic carbon using VIS/NIR spectroscopy with SVM and SPA methods[J]. *Remote Sensing*, 2014, 6(4): 2699–2717.
- [27] KENNARD R W, STONE L A. Computer aided design of experiments[J]. *Technometrics*, 1969, 11: 137–148.
- [28] MOURA-BUENO J M, DALMOLIN R S D, TEN C A, et al. Stratification of a local VIS – NIR – SWIR spectral library by homogeneity criteria yields more accurate soil organic carbon predictions[J]. *Geoderma*, 2019, 337: 565–581.
- [29] 张俊华, 孙媛, 贾科利, 等. 不同盐结皮光谱特征及其盐渍化信息预测研究[J]. *农业机械学报*, 2018, 49(12): 325–333.
ZHANG Junhua, SUN Yuan, JIA Keli, et al. Spectral characteristics and salinization information prediction of different soil salt crusts[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2018, 49(12): 325–333. (in Chinese)
- [30] 国佳欣, 朱青, 赵小敏, 等. 不同土地利用类型下土壤有机碳含量的高光谱反演[J]. *应用生态学报*, 2020, 31(3): 863–871.
GUO Jiaxin, ZHU Qing, ZHAO Xiaomin, et al. Hyper-spectral inversion of soil organic carbon content under different land use types[J]. *Chinese Journal of Applied Ecology*, 2020, 31(3): 863–871. (in Chinese)
- [31] WANG J Z, TIYIP T, DING J L, et al. Quantitative estimation of organic matter content in arid soil using Vis – NIR spectroscopy preprocessed by fractional derivative[J]. *Journal of Spectroscopy*, 2017, 9: 1375158.
- [32] WANG X P, ZHANG F, KUNG H, et al. New methods for improving the remote sensing estimation of soil organic matter content (SOMC) in the Ebinur Lake Wetland National Nature Reserve (ELWNNR) in northwest China[J]. *Remote Sensing of Environment*, 2018, 218: 104–118.
- [33] 朱亚星, 于雷, 洪永胜, 等. 土壤有机质高光谱特征与波长变量优选方法[J]. *中国农业科学*, 2017, 50(22): 4325–4337.
ZHU Yaxing, YU Lei, HONG Yongsheng, et al. Hyperspectral features and wavelength variables selection methods of soil organic matter[J]. *Scientia Agricultura Sinica*, 2017, 50(22): 4325–4337. (in Chinese)
- [34] SHRESTHA N K, SHUKLA S. Support vector machine based modeling of evapotranspiration using hydro-climatic variables in a sub-tropical environment[J]. *Agricultural and Forest Meteorology*, 2015, 200: 172–184.
- [35] 刁淑娟, 刘春玲, 张涛, 等. 基于 SVM 的湖泊咸度等级遥感信息提取方法——以内蒙古巴丹吉林沙漠为例[J]. *国土资源遥感*, 2016, 28(4): 114–118.
DIAO Shujuan, LIU Chunling, ZHANG Tao, et al. Extraction of remote sensing information for lake salinity level based on SVM: a case from Badain Jaran desert in Inner Mongolia[J]. *Remote Sensing for Land & Resources*, 2016, 28(4): 114–118. (in Chinese)