doi:10.6041/j.issn.1000-1298.2018.11.018

基于优化光谱指数的土壤有机质含量估算

尼加提·卡斯木[1,2]　茹克亚·萨吾提[2,3]　师庆东[2,3]　买合木提·巴拉提[3]
米热阿地力·库尔班[3]　苏比努尔·居来提[3]
(1. 新疆大学干旱生态环境研究所, 乌鲁木齐 830046; 2. 新疆大学绿洲生态教育部重点实验室, 乌鲁木齐 830046;
3. 新疆大学资源与环境科学学院, 乌鲁木齐 830046)

摘要: 为了寻求估算土壤有机质含量的最佳光谱参数,实现土壤养分无损监测,使用 ASD Field－Spec3 型高光谱仪对野外采集的土壤样品进行室内光谱测定,并通过重铬酸钾氧化容量法测定土壤样品有机质质量比;利用两波段优化算法对构建的新算法(SOMCI/ND)进行波段优化,筛选基于不同光谱数据(原始光谱反射率及其对应的 4 种数学变换)运算下的最敏感波段组合,从而建立土壤有机质质量比高光谱估算模型。结果表明:通过归一化光谱指数(I_{ND})和概念指数(I_{CI})比值构建的新算法(SOMCI/ND)优化后与土壤有机质质量比之间的相关性显著提高,在光谱原始数据及其平方根、倒数变换形式下,相关系数绝对值达到 0.82,且敏感的组合波段集中在 2 220～2 240 nm 和 2 160～2 195 nm。基于平方根波段优化的估算模型效果最佳,估算精度 R_p^2 为 0.84,RMSEP 为 2.24 g/kg,RPD 为 2.89。对光谱数据的适当数学变换有利于优化光谱指数,更好地估算土壤有机质质量比,进一步实现土壤有机质质量比的高精度动态监测。

关键词: 土壤有机质; 光谱分析; 高光谱; 优化光谱指数
中图分类号: O433.4; S27　　**文献标识码**: A　　**文章编号**: 1000-1298(2018)11-0155-09

Estimation of Soil Organic Matter Content Based on Optimized Spectral Index

NIJAT Kasim[1,2]　RUKEYA Sawut[2,3]　SHI Qingdong[2,3]　MAIHEMUTI Balati[3]　MIREADILI Kuerban[3]
SUBINUER Julaiti[3]
(1. *Institute of Arid Ecology Environment*, *Xinjiang University*, *Urumqi* 830046, *China*
2. *Key Laboratory of Oasis Ecology*, *Ministry of Education*, *Xinjiang University*, *Urumqi* 830046, *China*
3. *College of Resources and Environmental Sciences*, *Xinjiang University*, *Urumqi* 830046, *China*)

Abstract: The rapid monitoring of soil organic matter content based on hyperspectral data is of great significance for evaluating soil fertility. The best spectral parameters for predicting soil organic matter content were tried to find and non-destructive monitoring of soil nutrients was achieved. ASD Field－Spec3 spectrometer was used to measure the indoor spectra of soil samples collected in the field, and the organic matter content of soil samples was measured by the potassium dichromate oxidation capacity method; the nitrogen planar component index (SOMCI/ND) was optimized by two-band optimization algorithm. Band optimization, screening the most sensitive spectral parameters of different spectral data (the original spectral reflectance and its corresponding four mathematical transformations), thus establishing a hyperspectral estimation model of soil organic matter content. The results showed that the correlations between soil organic matter content and the new algorithm (SOMCI/ND) optimized by the normalized spectral index (I_{ND}) and conceptual index (I_{CI}) ratios were significantly improved. The raw data in the spectrum and its square root and reciprocal transformation form, the absolute value of correlation coefficient reached 0.82, and the sensitive combination bands were concentrated in 2 220～

收稿日期: 2018－06－08　修回日期: 2018－07－11
基金项目: 国家自然科学基金项目(41762019、U1703237)、新疆大学博士科研启动基金项目(BS160232)和新疆大学大学生创新训练项目(201710755099)
作者简介: 尼加提·卡斯木(1991—),男,博士生,主要从事生态规划与管理研究,E-mail: NejatKasim@126.com
通信作者: 买合木提·巴拉提(1981—),男,副教授,主要从事水文-水环境数值模拟与地质灾害研究,E-mail: bmaihemuti@xju.edu.cn

2 240 nm and 2 160 ~ 2 195 nm. The prediction model based on the square root band optimization had the best effect. The prediction accuracy was R_P^2 of 0.84, RMSEP of 2.24 g/kg and RPD of 2.89. Therefore, the appropriate mathematical transformation of the spectral data was conducive to optimizing the spectral index to better estimate the soil organic matter content, and further achieve high-precision dynamic monitoring of soil organic matter.

Key words: soil organic matter; spectrum analysis; hyperspectra; optimized spectral index

0 引言

土壤有机质(Soil organic matter, SOM)是土壤中固态物质的重要组成部分,具有改良土壤结构、疏松土壤以及提供植物所需养分等作用,是评价土壤肥沃程度的一个重要衡量指标[1-2]。高效监测区域土壤有机质含量可为及时改善区域土壤状况、优化作物施肥管理,以及农业发展提供科学依据。传统的土壤有机质含量测定方法是野外采集土壤样品后,通过化学方法测定与分析,该方法时效低、成本高、代表性不足,无法快速无损地监测区域土壤有机质含量[3-4]。土壤光谱反射率与有机质含量之间存在一定的光谱响应关系,这为现阶段高光谱遥感在区域尺度土壤有机质含量的快速、准确监测提供了可能,从而成为定量遥感的研究热点之一[5]。对高光谱数据进行不同的预处理及建模方法的选择是构建估算模型的重要步骤,对模型的精度和稳定性具有重要意义[6]。KRISHNAN 等[7]选取 4 种不同的土壤为研究对象,发现土壤光谱反射率与其有机质含量呈现显著负相关关系,探测出土壤有机质含量最佳波段是 0.62 μm 和 0.56 μm,可见光波段优于近红外波段,在近红外区域没有由有机质引起的反射峰和吸收谷。王延仓等[8]对光谱数据进行6 种分解,并与土壤有机质含量进行相关性分析,发现光谱反射率经小波变换后,参与建模的特征波段集中在近红外波段范围,即近红外波段估算有机质含量优于可见光波段。徐明星等[9]研究了原始反射率、微分光谱、吸光度和波段深度对不同成陆年代土壤有机质含量的估算精度,发现微分光谱是估算成陆年代土壤有机质含量的最佳光谱指标,以 1 400 nm 为中心及 1 900 ~ 2 450 nm 的近红外波段是估算土壤有机质含量的最佳波段。HUMMEL 等[10]在室内分析了土壤有机质和光谱曲线之间的关系,采用光谱反射率吸光度建立多元逐步回归模型。CONFORTI 等[11]对来自不同地区的 215 个土样进行偏最小二乘回归(PLSR)建模,取得了较好的估算效果。NOWKANDEH 等[12]以高光谱遥感为数据源,利用 PLSR、PCR、MinR 和 SWR 方法进行建模,取得显著的估算效果。栾福明等[13]选取 40 个土壤样本,采用多元线性逐步回归和人工神经网络两种方法建立有机质含量估算模型,得出人工神经网络(ANNs)集成模型表现最优。叶勤等[14]将原始光谱反射率、一阶导数光谱分别与土壤有机质含量进行相关性分析,直接建立基于一阶导数光谱的多元线性逐步回归分析(MLSR)模型。

上述对于土壤有机质含量的研究[7-14]是通过全波段反射率及其对应的数学变换,在一维层面上选取单个敏感波段或多个敏感波段,并利用多种回归方法建立估算模型,而对于两波段优化组合算法进行光谱指数的波段二维层面优化,并在土壤有机质含量估算中的应用研究相对较少。基于此,本文以新疆乌鲁木齐安宁渠为研究靶区,进行野外土壤样品采集和室内土壤光谱测定及理化分析,尝试对土壤原始光谱反射率及其对应的 4 种数学变换进行波段优化运算,分析基于光谱数据不同变换形式的波段优化组合与土壤有机质质量比的相关性,通过相关系数的显著性检验,筛选最适宜敏感光谱指数,并利用偏最小二乘回归(Partial least square regression, PLSR)分析法建立土壤有机质质量比优化光谱指数的估算模型,得出反演土壤有机质的最佳光谱参数。

1 材料与方法

1.1 研究区概况

研究区位于乌鲁木齐市北郊的冲积平原上,地势平缓,地理坐标为 87°25′ ~ 87°30′E,43°57′ ~ 44°06′N[15]。地处中温带干旱气候区,光热充足,土壤肥沃,适合一般的农作物、蔬菜和果树生长[16]。该区域水资源丰富,既有水质良好的地下水资源,又有纵贯南北的和平渠提供的地表径流,形成井渠双灌的地方特色,为农牧业的丰收奠定了良好的基础[17-18]。

1.2 土壤样品采集

土壤样品的采集时间为 2017 年 10 月 15 日,针对乌鲁木齐市安宁渠蔬菜基地研究区典型景观特征,设置 5 m × 5 m 样方(图 1)。利用 GPS 记录每个样方位置,采用 5 点混合法进行样品采集,土壤样品采样深度为 0 ~ 10 cm,共计 57 个土壤样品。土壤样品带回实验室并进行自然风干、研磨后,过 2 mm 孔筛分为两部分,分别用于室内高光谱数据的测定及土壤有机质的化学分析。其中,土壤有机质质量比

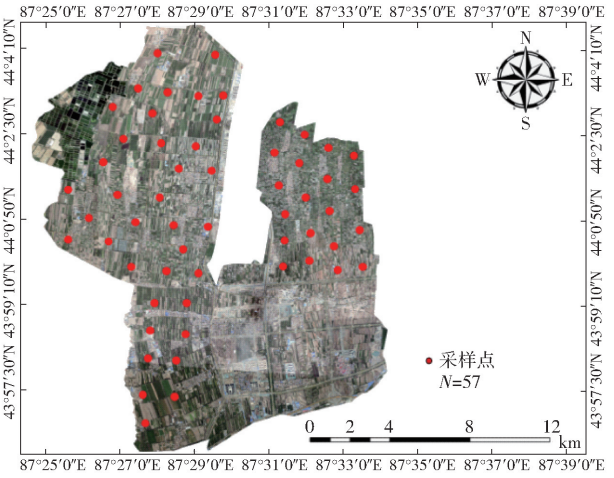


图 1　研究区位置和采样区分布

Fig. 1　Location of study area and distribution of sampling sites

采取重铬酸钾容量-硫酸氧化法进行测定[19]。

1.3　光谱测定及预处理

采用美国 ASD(Analytical Spectral Devices)公司生产的 Field - Spec3 型光谱仪(波段范围 350 ~ 2 500 nm)。光谱在 350 ~ 1 000 nm 与 1 000 ~ 2 500 nm 区间的采样间隔分别为 1.4 nm 与 2 nm,重采样间隔为 1 nm。将准备的 57 个土壤样品分别装入黑色盛样皿(直径 12 cm,深 1.8 cm)中,装满土壤样品后,对表面进行平滑处理。光谱测定的光源为 50 W 卤素灯,测量时距土壤样品表面 50 cm,光源的天顶角为 15°,探头至待测样品表面的距离为 5 cm[20]。高光谱数据测定之前进行白板校正,每个土壤样品重复测量 5 次,取 5 条光谱曲线的算术平均值作为该样品的实际光谱反射率。

1.4　研究方法

近几年的研究结果[21]表明,为了估算不同区域土壤有机质含量最佳普适性光谱,参数的选择具有一定的难度,同时对土壤有机质含量选择敏感光谱指数的波段组合呈现多样化。

连续窄波段光谱包含更多的光谱信息,通过两波段算法计算所有可能两两组合的光谱指数能够充分提取与土壤有机质含量相关性最强的相关波段[22]。常用的优化光谱指数包括:比值光谱指数 $I_{RI}(R_{\lambda1}/R_{\lambda2})$、归一化光谱指数 $I_{ND}((R_{\lambda1}-R_{\lambda2})/(R_{\lambda1}+R_{\lambda2}))$ 和概念指数 $I_{CI}((R_{\lambda2}^{-1}-R_{\lambda1}^{-1})R_{\lambda2})$。其中,$R_\lambda$ 为波长 λ 对应的反射率[23-25]。

其中,张娟娟等[22]利用 350 ~ 2 500 nm 光谱反射率,系统构造了原始光谱任意两波段组合而成的比值、归一化光谱指数,建立了特征光谱指数的土壤全氮含量估算模型。本研究在已有成果的基础上,通过利用原始光谱数据系统地构建任意两波段组合的优化光谱指数,并分析光谱指数与土壤有机质质量比的相关性等势图(图 2)。

为进一步研究优化算法在土壤有机质质量比估算中的应用,分析发现波段组合在算法 $(R_{\lambda1}-R_{\lambda2})/(R_{\lambda1}+R_{\lambda2})$、$(R_{\lambda2}^{-1}-R_{\lambda1}^{-1})R_{\lambda2}$ 运算后,对土壤有机质质量比具有显著相关,相关性绝对值均超过 0.6。因此,选择对土壤有机质贡献突出的两种算法,通过比值关系建立新的优化算法,进一步研究该算法对土壤有机质的敏感性。计算过程为

$$R_{SOMCI/ND}=\frac{I_{CI}}{I_{ND}}=\frac{(R_{\lambda2}^{-1}-R_{\lambda1}^{-1})R_{\lambda2}}{\frac{R_{\lambda1}-R_{\lambda2}}{R_{\lambda1}+R_{\lambda2}}}=\frac{R_{\lambda1}+R_{\lambda2}}{R_{\lambda2}} \quad (1)$$

式中　$R_{SOMCI/ND}$——两波段优化算法应用在全波段范围的优化光谱指数

通过该算法计算 57 个土壤样品的原始光谱反射率 R 以及对应的 4 种光谱变换(倒数 $1/R$、对数 $\lg R$、对数倒数 $1/\lg R$、平方根 $\sqrt{R}$)在波段 400 ~ 2 400 nm 之间进行所有可能两波段组合的优化光谱指数。为了选取模型最佳变量参数,本文将变量投影重要性(Variable importance in projection, VIP)准则应用到其中,VIP 值代表自变量对模型拟合的程度,自变量对因变量的解释作用相同,则自变量的 VIP 均接近于 1。WOLD[26]建议 VIP 值小于 1 的自变量对因变量的贡献较小,可以考虑剔除。因此,本研究自变量 VIP 的阈值为大于等于 1。

为了使建模集和验证集可以充分反映研究区土壤有机质质量比的实际情况,将 57 个样本按有机质质量比从高到低进行排序,等间隔抽取 38 个建模数据和 19 个验证数据样本,分别作为模型的建立与精度验证数据。将通过显著性检验及 VIP(≥1.0)的优化光谱指数作为自变量,利用 PLSR 建立土壤有

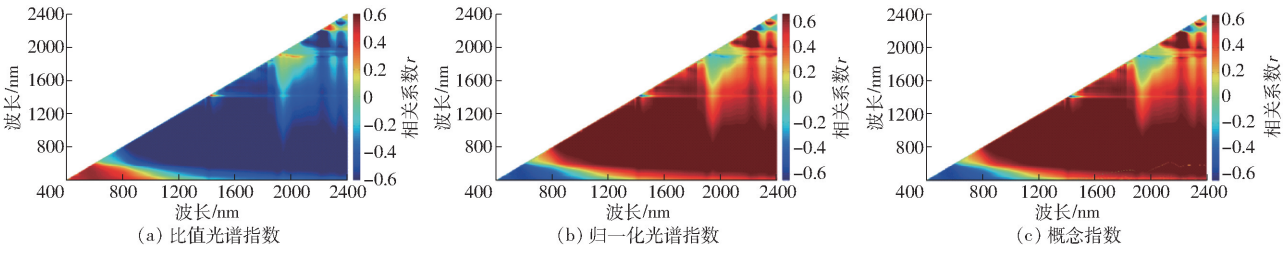


图 2　优化光谱指数与土壤有机质质量比之间的相关性等势图

Fig. 2　Equipotential patterns of correlation between optimized spectral index and soil organic matter content

机质质量比的估算模型。结合估算模型的建模决定系数(Calibration coefficient of determination, R_C^2)、预测决定系数(Predictive coefficient of determination, R_P^2)、校正均方根误差 r_{RMSEC}(Root mean square error of calibration, RMSEC)、预测均方根误差 r_{RMSEP}(Root mean square error of prediction, RMSEP)、相对分析误差 r_{RPD}(Relative prediction deviation, RPD)筛选最佳模型,精度参数如表 1 所示[23]。

表 1 模型精度参数

Tab. 1 Model precision parameters

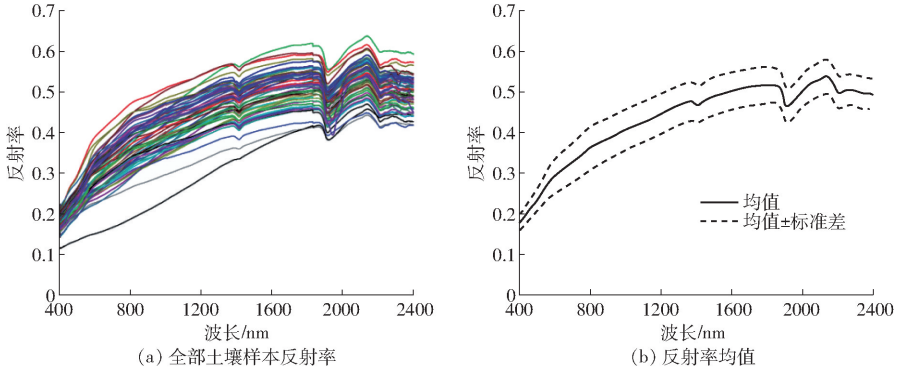
参数	简称	公式	描述
建模决定系数	R_C^2	$\dfrac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2+\sum_{i=1}^{N}(y_i-\bar{y})^2}}$	判定模型的稳定性,其值越接近于 1,模型的稳定性越好
预测决定系数	R_P^2		
校正均方根误差	RMSEC	$\sqrt{\dfrac{\sum_{i=1}^{N}(y_i-x_i)^2}{N}}$	表征模型的准确性,其值越接近于 0,表明模型的精度越高
预测均方根误差	RMSEP		
相对分析误差	RPD	$\dfrac{S_{SD}}{r_{RMSEP}}$	0 ~ 1.4,差 1.4 ~ 2.0,中 大于 2.0,好

注:x_i 和 y_i 分别为实测值与估算值,$\bar{x}$ 和 $\bar{y}$ 分别为实测与估算的平均值,N 为样本数量。

2 结果与分析

2.1 数据处理与分析

对获取的土壤样品高光谱数据利用 ViewSpec Pro 软件进行数据算术平均并导出,为减少噪声影响,剔除信噪比较低的边缘波段(350 ~ 399 nm 及 2 401 ~ 2 500 nm)后,利用 Savitzky - Golay 滤波对 57 个土壤样品 400 ~ 2 400 nm 的高光谱数据进行平滑去噪处理,去噪后的光谱曲线如图 3 所示。



(a) 全部土壤样本反射率 (b) 反射率均值

图 3 平滑处理后的土壤样本室内光谱反射率

Fig. 3 Spectral reflectance of soil sampling after smoothing

由表 2 知,建模集和验证集对应的土壤有机质质量比最大值分别为 31.78、15.79 g/kg,最小值分别为 4.39、2.70 g/kg,均值分别为 11.84、8.73 g/kg,变异系数分别为 53.82% 和 36.54%;研究区所有采样点土壤有机质质量比平均值为 10.89 g/kg,变异系数为 53.17%,介于建模集与验证集之间,数据离散程度一般,属于中等变异强度(15% ~100%)[24]。

表 2 土壤有机质质量比的统计分析

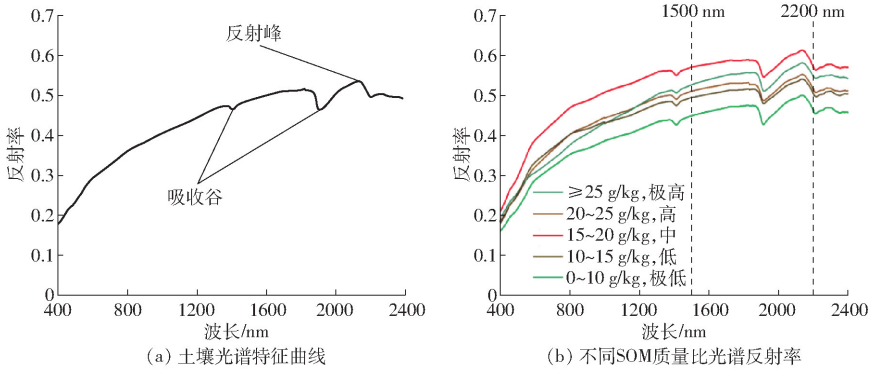
Tab. 2 Statistics analysis of soil organic content

样本集	样本数	有机质质量比/(g·kg^{-1})			标准差/(g·kg^{-1})	变异系数/%
		最大值	最小值	均值		
全部	57	31.78	2.70	10.89	5.79	53.17
建模集	38	31.78	4.39	11.84	6.37	53.82
验证集	19	15.79	2.70	8.73	3.19	36.54

2.2 土壤光谱曲线特征分析

为研究土壤有机质质量比对土壤光谱反射率的影响,对土壤样本反射率进行分级标准归类、求平均,绘制不同土壤有机质质量比的土壤光谱曲线(图 4)。土壤有机质为土壤重要的赋色成分,土壤光谱反射对其质量比的高低产生一定响应[25]。研究区不同有机质质量比的土壤光谱反射曲线的变化趋势基本一致。土壤样品的光谱反射率与其有机质质量比并没有呈现明显正负相关关系,从图 4b 中可以看出,在可见光及近红外波段(400 ~ 1 500 nm)范围内光谱曲线呈快速上升和平缓上升趋势,并在1 400 ~ 1 500 nm 处有 1 个比较浅的水分吸收谱段,而在 1 800 ~ 2 100 nm 波段范围内出

现了 1 个较为明显的水分吸收谱段(图 4a)。整体光谱曲线而言,除 400 ~ 1 500 nm 波段范围,不同土壤有机质质量比的光谱反射率曲线较为容易区分。



(a) 土壤光谱特征曲线　　(b) 不同SOM质量比光谱反射率

图 4　不同有机质质量比的土壤样本光谱发射率

Fig. 4　Spectral reflectance of soil samples with different organic content

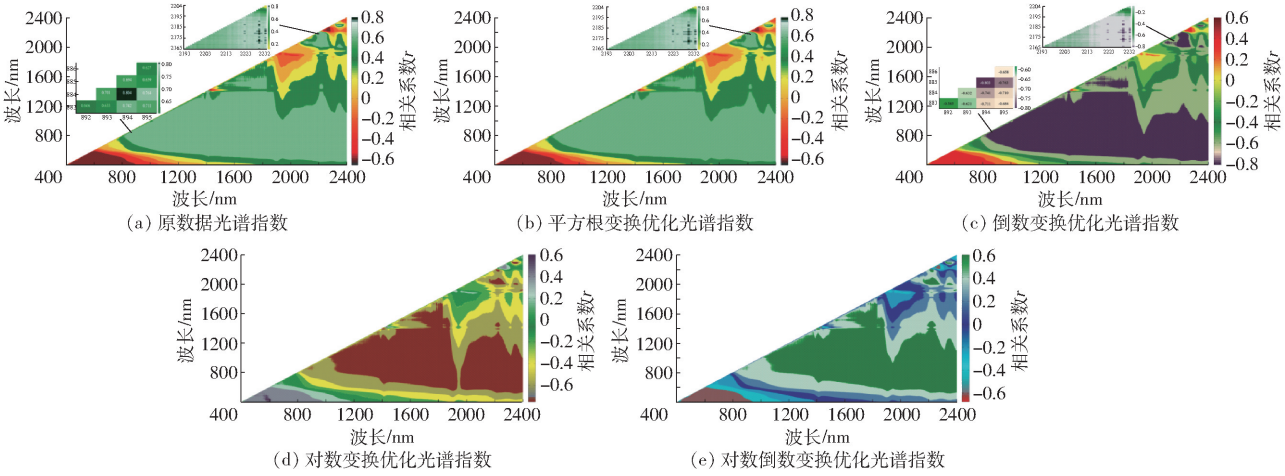
2.3　优化光谱指数与有机质质量比相关性分析

该研究利用两波段优化算法,通过最优两波段组合计算构建的新算法 SOMCI/ND,选取与土壤有机质质量比相关性最高并通过显著性检验($p < 0.01$)及 VIP 值大于 1 的优化指数进行偏最小二乘回归建模。

在波段 400 ~ 2 400 nm 范围内,通过两波段优化算法计算土壤样品的原始光谱反射率 R 及其对应的 4 种光谱变换下所有可能两波段组合的优化光谱指数,并与土壤有机质质量比进行 Pearson 相关性分析(图 5)。由图 5a、5b、5c 可知,原数据及其对应的平方根、倒数变换后优化的光谱指数与有机质质量比的相关性呈现最高。其中原数据和平方根变换后得出的优化光谱指数与有机质质量比呈现正相关关系,均为出现在 2 230 nm 和 2 215 nm 波段附近两波段组合的光谱指数;而倒数变换后的优化指数与有机质质量比为负相关关系,主要是 2 230 nm 和 2 215 nm波段组合得出的光谱指数;原数据和平方根变换、倒数变换后计算的新算法与土壤有机质质量比相关系数绝对值均大于 0.80;其余的 2 种光谱变换($1/\lg R$、$\lg R$)后得出的优化光谱指数与有机质质量比的相关系数绝对值最高达到 0.75。为了分析光谱数据 4 种变换对优化光谱的影响,统计了不同数据变换下与土壤有机质质量比相关系数在($|r| \geqslant 0.80, \rho < 0.01$)范围内的数量,统计结果如表 3 所示。其中,$1/\lg R$、$\lg R$ 变换后与有机质质量比的相关系数绝对值均低于 0.80,而以原数据及 2 种变换($\sqrt{R}$、$1/R$)的优化光谱指数与有机质质量比的相关系数绝对值超出 0.80 的数量分别达到 48、47、42;不同光谱数据变换中,敏感指数(SOMCI/ND,$|r| \geqslant 0.80$)的数量从大到小依次为 R、$\sqrt{R}$、$1/R$、$\lg R$、$1/\lg R$。

2.4　土壤有机质质量比 PLSR 估算模型建立及精度分析

该研究以室内土壤光谱反射率及 4 种变换数据,对新算法进行波段组合优化;结合变量投影重要性(VIP)算法筛选最优两波段组合光谱指数并作为自变量,利用偏最小二乘回归法(PLSR)建立土壤有



(a) 原数据光谱指数　(b) 平方根变换优化光谱指数　(c) 倒数变换优化光谱指数

(d) 对数变换优化光谱指数　(e) 对数倒数变换优化光谱指数

图 5　400 ~ 2 400 nm 波段范围内所有可能波段组合与土壤有机质质量比的相关性

Fig. 5　Correlation coefficient between organic content and narrow band calculated all possible two-band combinations in range of 400 ~ 2 400 nm

表 3 不同光谱数据变换下优化光谱指数与有机质质量比相关性统计分析

Tab. 3 Statistic analysis between organic content and optimized spectral index based on different spectral data transform

光谱数据变换	相关系数绝对值超出的数量（$\|r\|\geq 0.80, \rho<0.01$）	相关系数绝对值（$\|r\|$）的最值
原数据 R	48	0.82
对数 $\lg R$	0	0.75
对数倒数 $1/\lg R$	0	0.75
倒数 $1/R$	42	0.82
平方根 $\sqrt{R}$	47	0.82

表 4 偏最小二乘回归建模精度分析

Tab. 4 Accuracy analysis for predictive models by partial least squares regression

光谱数据数学变换	自变量数量	建模集		验证集		RPD
		R_C^2	RMSEC/（g·kg^{-1}）	R_P^2	RMSEP/（g·kg^{-1}）	
原数据 R	18	0.73	3.14	0.66	3.21	2.05
对数 $\lg R$	6	0.71	3.22	0.54	5.36	1.11
对数倒数 $1/\lg R$	8	0.71	3.15	0.48	5.29	1.12
倒数 $1/R$	19	0.80	2.47	0.77	2.51	2.19
平方根 $\sqrt{R}$	19	0.86	2.43	0.84	2.24	2.89

机质质量比估算模型，估算模型的精度参数如表 4 所示。为 PLSR 更深入地分析数据，利用 VIP 技术进一步对变量进行筛选，基于 PLSR 建模方法的优点，VIP 技术主要用于样本较少且几个自变量间相关性较强的情形，在一定程度上弥补了传统线性回归的不足[26]。由图 6 可知，基于光谱反射率及其

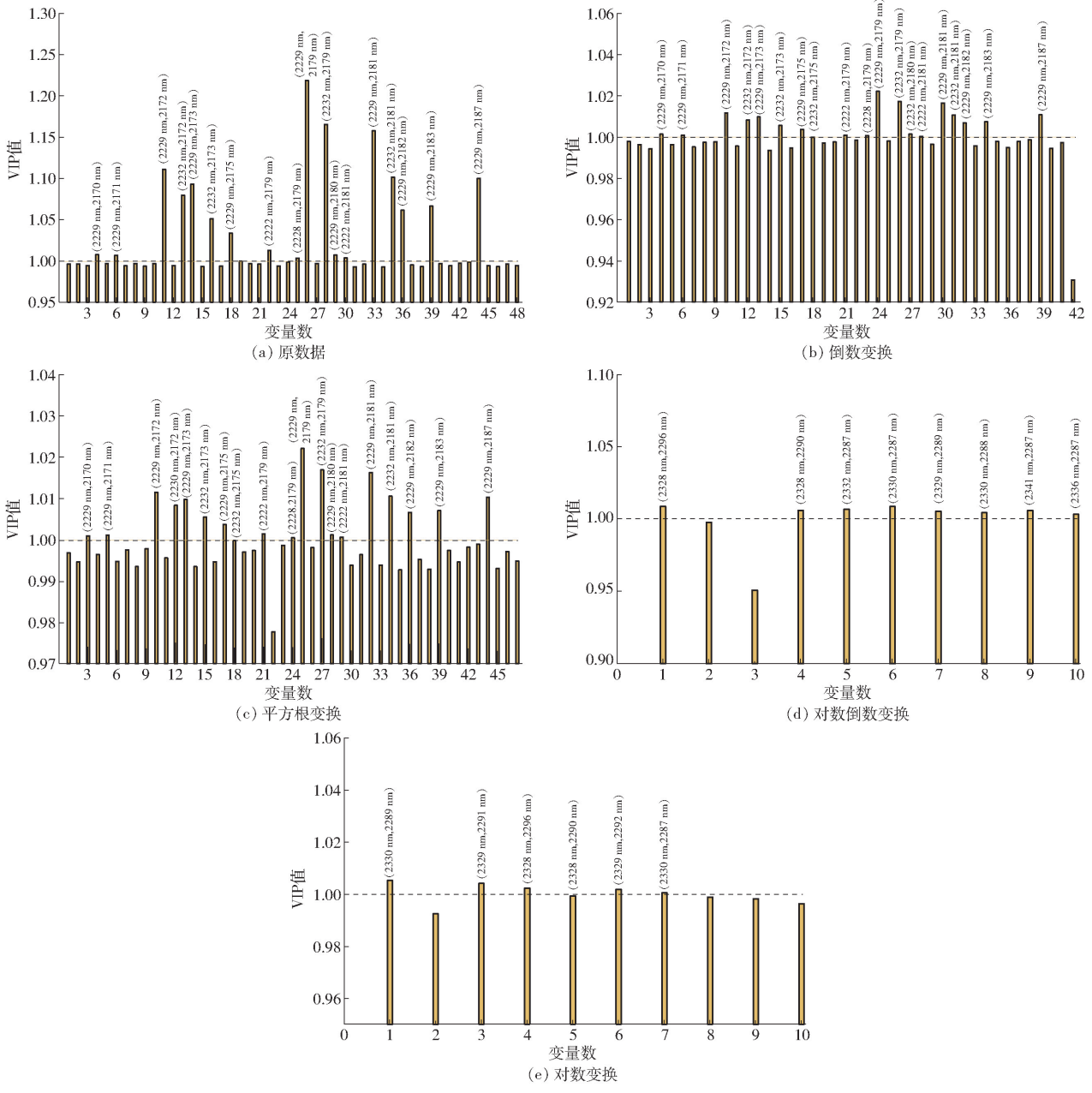


图 6 基于变量 VIP 值筛选最佳模型自变量

Fig. 6 VIP values of variables to select the best model based on independent variables

4 种变换优化在变量选择上呈现不同，原数据优化指数中筛选的变量数为 18 个，对应的波段组合见图 6a，光谱数据 4 种变换优化指数中筛选变量数分别为 19、19、8、6 个，对应波段组合见图 6b ~ 6e。

根据 PLSR 估算模型精度参数，对建立的 5 种模型进行分析发现，RPD 大于 2 的模型有 3 个，分别为基于原数据、倒数、平方根变换下优化建立的估算模型。基于对数及对数倒数预处理的光谱数据优化对土壤有机质质量比估算能力较低，估算精度 R_p^2 分别为 0.54、0.48，RMSEP 分别为 5.36、5.29 g/kg。而基于原数据及 2 种数学变换（$1/R$、$\sqrt{R}$）优化建立的估算模型具有一定的定量反演能力，模型精度 R_p^2 分别为 0.66、0.77、0.84，RMSEP 分别为 3.21、2.51、2.24 g/kg，RPD 分别为 2.05、2.19、2.89。其中，基于平方根（$\sqrt{R}$）变换预处理光谱反射率并对光谱指数（SOMCI/ND）进行波段优化运算后，通过 VIP 技术筛选的优化指数（（2 229 nm，2 170 nm）、（2 229 nm，2 171 nm）、（2 229 nm，2 172 nm）、（2 230 nm，2 172 nm）、（2 229 nm，2 173 nm）、（2 232 nm，2 173 nm）、（2 229 nm，2 175 nm）、（2 232 nm，2 175 nm）、（2 222 nm，2 179 nm）、（2 228 nm，2 179 nm）、（2 229 nm，2 179 nm）、（2 232 nm，2 179 nm）、（2 229 nm，2 180 nm）、（2 222 nm，2 181 nm）、（2 229 nm，2 181 nm）、（2 232 nm，2 181 nm）、（2 229 nm，2 182 nm）、（2 229 nm，2 183 nm）、（2 229 nm，2 187 nm））为自变量建立的 PLSR 估算模型最佳，估算精度为 R_p^2 为 0.84，RMSEP 为 2.24 g/kg，RPD 为 2.89。利用土壤有机质质量比验证数据对模型进行估算能力的检验，结果如图 7 所示。与其余模型相比较，该模型在土壤有机质质量比估算中表现了最优的估算能力。

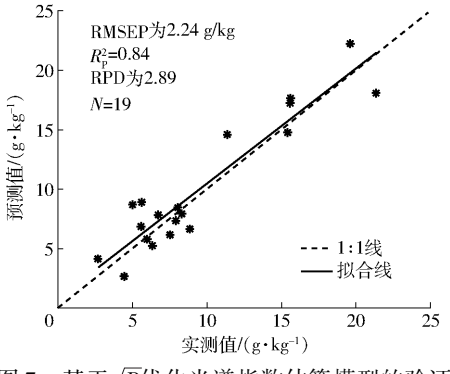


图 7　基于$\sqrt{R}$优化光谱指数估算模型的验证

Fig. 7　Validation of model based on $\sqrt{R}$ optimized spectral index

3　讨论

遥感估算土壤有机质含量都是以土壤光谱理论为指导，选择计算在土壤有机质特有的特征波段的反射率和基于相同位置的光谱指数来达到土壤有机质光谱特性研究的目的，适合土壤有机质含量估算的模型必须具有理论依据和应用能力[27]。本文利用野外土壤样品采样，室内土壤光谱采集及理化分析，进行土壤原始光谱反射率及对应的 4 种数学变换，对构建的新算法（SOMCI/ND）进行波段优化运算，分析基于光谱数据不同变换形式的优化光谱指数与土壤有机质质量比的相关性，通过相关系数的显著性检验（$p = 0.01$，阈值为 ±0.70）筛选对应最适波段组合的敏感光谱指数，并利用 PLSR 建立土壤有机质质量比优化光谱指数的估算模型，得出反演土壤有机质的最适光谱参数。光谱反射率及其 4 种变换预处理后，通过两波段优化算法运算的光谱指数与土壤有机质质量比的相关性显著提高。其中，主要位于 2 220 ~ 2 240 nm 和 2 160 ~ 2 195 nm 的组合波段对应光谱指数具有很强的相关性，相关系数绝对值最高达到 0.82。除此之外，原数据和倒数变换在波长 894 nm 与 884 nm 组合计算的新算法与有机质质量比的相关系数绝对值也超出了 0.80。此说明优化算法的组合及波段组合对土壤有机质质量比的敏感组合波段筛选具有很重要的意义。

研究指出对有机质质量分数低于 2% 的荒漠土壤进行有机质高光谱定量反演时，敏感波段的选择相对困难[28]。BAUMGARDNER 等[29]在研究印度土壤时发现，当土壤有机质质量分数大于 2% 时，有机质就有可能掩盖其他因素对土壤光谱特征的影响；当低于 2% 时，随着有机质含量降低，其掩盖其他成分的能力变弱，对土壤光谱的影响就越小。该研究区土壤有机质质量分数最值分别为 3.178%、0.270%，平均值为 1.089%，土壤有机质含量属于偏低。武彦清等[30]通过对土壤光谱数据进行预处理，比较逐步回归和 PLSR 模型，确定了对土壤有机质含量敏感波段为 590、2 010、2 270、2 280 nm。何挺等[31]运用不同的数学变换及统计学的方法研究表明，在原数据变换前敏感波段区域为 675 nm 附近，数学变换后近红外波段的 2 187 nm 处与有机质的相关性达到最强，参与最佳模型的敏感波段为 849、1 681、2 187 nm。上述学者提出的敏感波段范围与本研究结果相似，说明构建的新算法对土壤有机质敏感波段范围是可利用的。与光谱数据的对数、对数倒数变换结果相比，基于平方根（$\sqrt{R}$）变换并对光谱指数进行波段组合优化的 PLSR 估算模型为最佳，估算精度 R_p^2 分别提高了 0.30、0.36，估算平方根误差 RMSEP 分别降低了 3.12、3.05 g/kg，相对分析误差 RPD 分别提高了 1.78、1.77。结合模型精

度参数分析,基于平方根优化算法所建立的土壤有机质质量比估算模型具有一定的估算性能。

综上所述,对土壤有机质质量比定量估算主要以计算高光谱数据的敏感波段、对原始数据的不同数学变换、计算对应光谱指数等方法,而本研究在已有研究基础上,进一步优化算法构建新的优化算法,并进行波段组合与研究分析。高光谱数据提供了精细的光谱波段信息,可以解决光谱指数与土壤有机质质量比存在的饱和性问题[32]。利用两波段优化算法对光谱指数的波段组合进行优化可以提高土壤有机质质量比估算模型的精度。本研究通过高光谱的精细光谱范围对 SOMCI/ND 的波段进行优化,建立了适宜该研究区的土壤有机质估算模型,并表现出了很好的估算能力。今后将通过高光谱数据探讨光谱波段优化理论,在不同地理位置、不同土壤类型中进一步应用两波段优化算法并得到更好的验证,为不同尺度、实时性的土壤有机质动态监测提供理论基础。

4 结论

(1)与优化光谱指数(I_{RI}、I_{ND}、I_{CI})的算法相比,构建的新算法显著提高了与土壤有机质之间的相关性,通过光谱数据的 4 种数学变换形式发现,除了在原数据波段优化外,$\sqrt{R}$和 $1/R$ 形式下的优化光谱指数均与土壤有机质质量比呈现很高的相关性,相关系数绝对值最高达到 0.82。

(2)原数据及其不同数学变换发现,位于 2 220 ~ 2 240 nm和 2 160 ~ 2 195 nm 波段范围组合的光谱指数与土壤有机质质量比之间相关性显著提高。另外,原数据和倒数变换在波长 894 nm 与 884 nm 组合计算 SOMCI/ND 对土壤有机质质量比,也具有显著的敏感程度。

(3)对土壤有机质高光谱 PLSR 建模发现,基于平方根($\sqrt{R}$)变换预处理光谱反射率,并对构建的新算法进行波段优化运算后,通过 VIP 技术筛选的优化指数为自变量建立的 PLSR 估算模型为最佳,估算精度 R_P^2 为 0.84,RMSEP 为 2.24 g/kg,RPD 为 2.89。这为研究区蔬菜基地土壤有机质动态监测提供了基础理论。

参 考 文 献

1 方少文, 杨梅花, 赵小敏,等. 红壤区土壤有机质光谱特征与定量估算:以江西省吉安县为例[J]. 土壤学报, 2014, 51(5):1003 - 1010.
FANG Shaowen, YANG Meihua, ZHAO Xiaomin, et al. Spectral characteristics and quantitative estimation of SOM in red soil typical of Ji'an county, Jiangxi province[J]. Acta Pedologica Sinica, 2014,51(5):1003 - 1010. (in Chinese)

2 王海峰, 张智韬, ARNON K,等. 基于灰度关联-岭回归的荒漠土壤有机质含量高光谱估算[J]. 农业工程学报, 2018, 34(14):124 - 131.
WANG Haifeng, ZHANG Zhitao, ARNON K, et al. Estimation model of desert soil organic matter content using hyperspectral data [J]. Transactions of the CSAE, 2018, 34(14):124 - 131. (in Chinese)

3 田永超, 张娟娟, 姚霞,等. 基于近红外光声光谱的土壤有机质含量定量建模方法[J]. 农业工程学报, 2012, 28(1):145 - 152.
TIAN Yongchao, ZHANG Juanjuan, YAO Xia, et al. Quantitative modeling method of soil organic matter content based on nearinfrared photoacoustic spectroscopy[J]. Transactions of the CSAE, 2012, 28(1):145 - 152. (in Chinese)

4 宋海燕, 秦刚, 韩小平,等. 基于可见光谱的不同质地土壤有机质快速测定[J/OL]. 农业机械学报, 2012, 43(7):69 - 72. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no=20120712&flag=1. DOI: 10.6041/j.issn.1000-1298.2012.07.012.
SONG Haiyan, QIN Gang, HAN Xiaoping, et al. Rapid prediction of soil organic matter by using visible infrared spectral technology[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(7):69 - 72. (in Chinese)

5 李美清, 李晋阳. 基于反射光谱特性的土壤有机质含量测定仪改进设计[J]. 农业机械学报, 2008, 39(10):148 - 151.
LI Meiqing, LI Jinyang. Improved research on intellectual system for measuring soil organic matter content with spectrum method [J]. Transactions of the Chinese Society for Agricultural Machinery, 2008, 39(10):148 - 151. (in Chinese)

6 KOOISTRA L, WANDERS J, EPEMA G F, et al. The potential of field spectroscopy for the assessment of sediment properties in river floodplains[J]. Analytica Chimica Acta, 2003, 484(2): 189 - 200.

7 KRISHNAN P, ALEXANDER J D, BUTLER B J, et al. Reflectance technique for predicting soil organic matter[J]. Soil Science Society of America Journal, 1980, 44(6): 1282 - 1285.

8 王延仓, 杨贵军, 朱金山,等. 基于小波变换与偏最小二乘耦合模型估测北方潮土有机质含量[J]. 光谱学与光谱分析, 2014, 34(7):1922 - 1926.
WANG Yancang, YANG Guijun, ZHU Jinshan, et al. Estimation of organic matter content of north fluvo-aquic soil based on the coupling model of wavelet transform and partial least squares[J]. Spectroscopy and Spectral Analysis, 2014, 34(7):1922 - 1926. (in Chinese)

9 徐明星, 周生路, 丁卫,等. 苏北沿海滩涂地区土壤有机质含量的高光谱估算[J]. 农业工程学报, 2011, 27(2):219 - 223.
XU Mingxing, ZHOU Shenglu, DING Wei, et al. Hyperspectral reflectance models for predicting soil organic matter content in coastal tidal land area, northern Jiangsu[J]. Transactions of the CSAE, 2011, 27(2):219 - 223. (in Chinese)

10 HUMMEL J W, SUDDUTH K A, HOLLINGER S E. Soil moisture and organic matter prediction of surface and subsurface soils using an NIR soil sensor[J]. Computers and Electronics in Agriculture, 2001, 32(2): 149 - 165.
11 CONFORTI M, BUTTAFUOCO G, LEONE A P, et al. Studying the relationship between water-induced soil erosion and soil organic matter using vis-NIR spectroscopy and geomorphological analysis: a case study in southern Italy[J]. Catena, 2013, 110: 44 - 58.
12 NOWKANDEH S M, HOMAEE M, NOROOZI A A. Mapping soil organic matter using hyperion images[J]. International Journal of Agronomy & Plant Production,2013,4:1753 - 1759.
13 栾福明, 张小雷, 熊黑钢,等. 基于不同模型的土壤有机质含量高光谱反演比较分析[J]. 光谱学与光谱分析, 2013, 33(1):196 - 200.
LUAN Fuming, ZHANG Xiaolei, XIONG Heigang, et al. Comparative analysis of soil organic matter content based on different hyperspectral inversion models[J]. Spectroscopy and Spectral Analysis, 2013, 33(1):196 - 200. (in Chinese)
14 叶勤, 姜雪芹, 李西灿,等. 基于高光谱数据的土壤有机质含量反演模型比较[J/OL]. 农业机械学报, 2017, 48(3): 164 - 172. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no=20170321&flag=1&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2017.03.021.
YE Qin, JIANG Xueqin, LI Xican, et al. Comparison on inversion model of soil organic matter content based on hyperspectral data[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(3): 164 - 172. (in Chinese)
15 张小禹. 乌鲁木齐市蔬菜基地土壤有效态镍的空间变异特征[J]. 新疆环境保护, 2012, 34(3): 41 - 45.
ZHANG Xiaoyu. Spatial variation characteristic of soil active nickel in Urumqi Vegetable Base[J]. Environmental Protection of Xinjiang, 2012, 34(3): 41 - 45. (in Chinese)
16 王灵, 朱建雯, 钱翌,等. 乌鲁木齐市菜地土壤和蔬菜铅含量及风险评价[J]. 新疆农业大学学报, 2008, 31(1):46 - 50.
WANG Ling, ZHU Jianwen, QINA Yi, et al. Lead content in vegetable and soil of vegetable plot in Urumqi City and its risk evaluation[J]. Journal of Xinjiang Agricultural University, 2008, 31(1):46 - 50. (in Chinese)
17 ZHENG C. Analysis of Cd pollution & spatial variability characteristics of typical agricultural soils in suburbs of Urumqi City, Xinjiang[J]. Spectroscopy and Spectral Analysis, 2012, 32(2): 537 - 540.
18 ZHENG Jiang, WANG Ling. Spatial variation characteristics of soil available lead in vegetable base soil in Urumqi City[J]. Agricultural Sciences & Technology, 2012, 13(1): 155 - 157.
19 鲍士旦. 土壤农化分析[M]. 北京:中国农业出版社,2000.
20 尼加提·卡斯木, 师庆东,王敬哲,等. 基于高光谱特征和偏最小二乘法的春小麦叶绿素含量估算[J]. 农业工程学报, 2017, 33(22):208 - 216.
NIJIATI Kasimu, SHI Qingdong, WANG Jingzhe, et al. Estimation of spring wheat chlorophyll content based on hyperspectral features and PLSR model[J]. Transactions of the CSAE, 2017, 33(22): 208 - 216. (in Chinese)
21 李丹. 基于优化光谱指数的小麦玉米冠层氮素含量估算[D]. 呼和浩特: 内蒙古农业大学, 2015.
LI Dan. Optimaized spectral indices based on estimation of canopy nitrogen content in wheat and maize[D]. Huhhot: Inner Monolia Agricultural University, 2015. (in Chinese)
22 张娟娟, 田永超, 姚霞,等. 基于高光谱的土壤全氮含量估测[J]. 自然资源学报, 2011, 26(5):881 - 890.
ZHANG Juanjuan, TIAN Yongchao, YAO Xia, et al. Estimating soil total nitrogen content based on hyperspectral analysis technology[J]. Journal of Natural Resources, 2011, 26(5):881 - 890. (in Chinese)
23 JORDAN C F. Derivation of leaf-area index from quality of light on the forest floor[J]. Ecology, 1969, 50(4): 663 - 666.
24 ROUSE JR J W, HAAS R H, SCHELL J A, et al. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation[C]//Prog Rep RSC1978 - 1, Remote Sensing Center, Texas A&M University, College Station,1973:93.
25 GITELSON A A, VINA A, ARKEBAUER T J, et al. Remote estimation of leaf area index and green leaf biomass in maize canopies[J]. Geophysical Research Letters, 2003, 30(30):335 - 343.
26 WOLD S. PLS for multivariate linear modeling[J]. Chemometric Methods in Molecular Design, 1995, 2: 195 - 218.
27 MAIMAITIYIMING M, GHULAM A, BOZZOLO A, et al. Early detection of plant physiological responses to different levels of water stress using reflectance spectroscopy[J]. Remote Sensing, 2017, 9(7): 745.
28 BO S, ROSSEL R A V, MOUAZEN A M, et al. Chapter five-visible and near infrared spectroscopy in soil science[M]// Advances in Agronomy. Elsevier Science & Technology, 2010:163 - 215.
29 BAUMGARDNER M F, SILVA L F, BIEHL L L, et al. Reflectance properties of soils[J]. Advances in Agronomy, 1985, 38:1 - 44.
30 武彦清, 张柏, 宋开山,等. 松嫩平原土壤有机质含量高光谱反演研究[J]. 中国科学院大学学报, 2011, 28(2):187 - 194.
WU Yanqing, ZHANG Bai, SONG Kaishan, et al. Retrieval of soil organic matter content from hyper-spectra in Songnen Plain [J]. Journal of the Graduate School of the Chinese Academy of Sciences, 2011, 28(2):187 - 194. (in Chinese)
31 何挺, 王静, 林宗坚,等. 土壤有机质光谱特征研究[J]. 地球空间信息科学学, 2009, 12(1):975 - 979.
HE Ting, WANG Jing, LIN Zongjian, et al. Spectral features of soil organic matter[J]. Geospatial Information Science, 2009, 12(1):975 - 979. (in Chinese)
32 安海波. 基于优化光谱指数的草地生物量估算[D]. 呼和浩特: 内蒙古农业大学, 2015.
ZHANG Haibo. Estimating grassland biomass based on optimized spectral indices[D]. Huhhot: Inner Monolia Agricultural University, 2015. (in Chinese)