

连续小波变换的土壤有机质含量高光谱估测

玉米提·买明¹, 王雪梅^{1, 2*}

1. 新疆师范大学地理科学与旅游学院, 新疆 乌鲁木齐 830054

2. 新疆维吾尔自治区重点实验室“新疆干旱区湖泊环境与资源实验室”, 新疆 乌鲁木齐 830054

摘要 土壤有机质含量的高光谱估测可快速、准确监测土壤肥力,对现代化农业生产进行精准施肥提供科学依据。以新疆渭干河-库车河三角洲绿洲耕层土壤为研究对象,对采集的98个土壤样品的原始光谱反射率 R 分别进行传统倒数对数 $\lg(1/R)$ 、一阶微分 R' 和倒数对数一阶微分 $[\lg(1/R)]'$ 数学变换,以及基于小波母函数Bior1.3不同尺度分解的连续小波变换(CWT),并与实测土壤有机质含量进行相关分析,从而筛选出各类变换下与土壤有机质含量密切相关的特征波段和小波系数($p < 0.01$)。分别以原始光谱反射率(R)以及不同变换处理下的特征波段反射率和敏感小波系数作为自变量,土壤有机质含量作为因变量,采用偏最小二乘回归和支持向量机回归方法构建土壤有机质含量的估测模型。结果表明:(1)各类光谱变换方法有效提升光谱与土壤有机质含量之间的敏感性,其中经CWT变换后的土壤光谱反射率与有机质含量的相关性得到显著提高,相关系数由0.39提高到0.54($p < 0.01$)。(2)传统的 $[\lg(1/R)]'$ 变换构建的支持向量机回归模型,其决定系数(R^2)高于 $\lg(1/R)$ 和 R' 变换构建的模型,说明倒数对数一阶微分变换可有助于提高估测模型的精度,且支持向量机回归模型的精度和稳定性高于偏最小二乘回归模型。(3)经过CWT分解后,以原始光谱反射率在不同尺度上的敏感小波系数作为自变量建立的模型,估测精度和稳定性均有明显的提高,构建的R-CWT-2³-SVMR模型的决定系数(R^2)为0.84,均方根误差(RMSE)为1.48,相对分析误差(RPD)等于2.11,模型精度达到最高并拥有极好的预测能力。高光谱数据经多种变换处理后可有效去除白噪声,而连续小波变换处理比传统的数学变换方法更适合于挖掘土壤有效信息,实现光谱信号的近似特征和细节特征的有效分离,建立的反演模型可更加精准估测土壤有机质含量。

关键词 连续小波变换; 分解尺度; 高光谱估测; 土壤有机质; 渭干河-库车河三角洲绿洲

中图分类号: TP79 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)04-1278-07

引言

土壤有机质(soil organic matter, SOM)是衡量土壤肥力的重要指标,快速、准确获取土壤有机质含量,已成为现代农业发展的必然需要^[1]。目前,测量土壤有机质含量的方法已由传统的化学分析转变为现在的光谱分析方法。众多学者通过对土壤光谱反射率与有机质含量进行相关分析发现,土壤光谱反射率与有机质含量之间存在着显著的相关性^[2-3]。但由于土壤高光谱数据在采集过程中,易受外界环境、光源、仪器以及土壤样品的质量等诸多因素的影响,反射光谱中不可避免存在干扰噪声。因此,积极探索降低光谱噪声的有效方法对提高土壤有机质估测精度有着重要的现实意义。

研究者们尝试使用对土壤原始光谱进行倒数对数 $\lg(1/R)$ 、倒数对数一阶微分 $[\lg(1/R)]'$ 、一阶微分 R' 、二阶微分 R'' 、以及包络线去除CR等多种数学变换方法,从而消除样品中各类因素引起的噪声,增强土壤光谱反射率中的有效信号^[4]。Feng等^[5]对原始光谱进行一阶微分转换,并从中选出特征波段光谱值作为自变量建立了最优估测模型;Shen等^[6]采用一阶微分、倒数一阶微分等多种变换方法处理原始光谱,建立偏最小二乘回归模型用于估测土壤有机质。张锐等^[7]研究发现土壤原始光谱反射率经包络线去除虽在一定程度上提高了预测模型的精度,却难以有效去除白噪声。随着前人研究的不断深入,经过连续小波变换(continuous wavelet transformation, CWT)后的土壤光谱噪声可被有效去除,构建的估测模型具有更高的精度,能更全面而稳定地反演土

收稿日期: 2020-11-26, 修订日期: 2021-09-15

基金项目: 国家自然科学基金项目(41561051)和新疆维吾尔自治区自然科学基金项目(2020D01A79)资助

作者简介: 玉米提·买明, 1992年生, 新疆师范大学地理科学与旅游学院硕士研究生 e-mail: 812814409@qq.com

* 通讯作者 e-mail: 502529672@qq.com

壤有机质含量^[8]。叶红云等^[9-10]研究发现,土壤原始光谱反射率经CWT处理后可实现光谱信号的近似特征和细节差异的有效分离,构建的预测模型精度较单纯采用传统数学变换方法高,预测模型具有更好的稳定性。本研究通过对野外采集的98个土壤样品进行光谱数据的获取和有机质含量的测定,对原始光谱反射率进行传统数学变换和连续小波变换处理,选择Bior1.3为小波变换最佳母函数,通过不同分解尺度提取与土壤有机质含量密切相关的特征光谱波段和小波系数,采用偏最小二乘回归(partial least squares regression, PLSR)和支持向量机回归(super vector machine regression, SVMR)方法构建渭干河-库车河三角洲绿洲耕层土壤有机质含量估测模型,旨在确定土壤有机质含量的最佳小波分解尺度和最优估测模型,从而提高反演模型的精度和预测能力。

1 实验部分

1.1 土样采集

渭干河-库车河三角洲绿洲位于塔里木盆地北缘,为典型的干旱区扇形平原绿洲,辖区包括新疆维吾尔自治区阿克苏地区的库车市、沙雅县和新和县,北纬 $39^{\circ}30' - 42^{\circ}40'$,东经 $81^{\circ}27' - 84^{\circ}07'$ 。土壤类型以潮土、棕漠土、灌淤土,草甸土,沼泽土和盐土等为主^[11]。2019年7月中下旬以遥感影像及地形图作为参考底图在研究区进行合理布点(图1)。利用GPS系统对样点进行精准定位和现场调查。每个采样点采集0~20 cm土层的土壤500 g左右,共采集98个土壤样品。将土壤样品带回实验室后经自然风干,挑出杂物,研磨并过筛后送至中国科学院新疆生态与地理研究所测试分析中心,由工作人员通过重铬酸钾容量—外加加热法对土壤有机质含量进

行测定。

1.2 光谱数据预处理

采用ASDFieldSpec3光谱仪对土壤样品进行光谱采集,其有效光谱范围为350~2500 nm。首先进行光谱仪的白板校正,并尽量排除干扰土壤光谱的物体。测量时,按照标签顺序将土壤样品均匀摊开并完全覆盖在50 cm×50 cm的牛皮纸上,探头位于采样点的垂直上方,距离为15 cm。每个样品重复测量10条光谱曲线,取其平均值作为该样品的光谱反射率。为了避免所测光谱受背景干扰以及仪器等因素的影响而产生较大噪声,剔除光谱曲线的两端数据,同时考虑到水分对光谱反射率的影响,又剔除了1341~1400和1811~1950 nm的水分吸收波段,最终每份土样选取400~2450 nm范围内的1901个光谱数据用于后续处理和分析。进一步对原始光谱反射率(R)进行Savitzky-Golay(S-G)5点平滑处理,再采用倒数对数 $\lg(1/R)$ 、一阶微分 R' 和倒数对数一阶微分 $[\lg(1/R)]'$ 以及连续小波变换(CWT)处理,可有效减少外界因素对原始光谱反射率的干扰,旨在快速寻找对土壤有机质含量敏感的有效波段。

1.3 连续小波变换

连续小波变换(CWT)作为一种线性变换方法,常应用于各类噪声信号的处理上,具有一定程度的高分辨性和适应性,其变换公式如式(1)和式(2)

$$\begin{cases} W_f(a,b) \leq f \\ \Psi_{a,b} \geq \int_{-\infty}^{+\infty} f(\lambda) \Psi_{a,b}(\lambda) d\lambda \end{cases} \quad (1)$$

$$\Psi_{a,b} = \frac{1}{\sqrt{a}} \Psi \frac{\lambda - b}{a} \quad (2)$$

式中: λ 为高光谱波段数; a 为尺度因子; b 为平移因子; $f(\lambda)$ 为土壤光谱反射率; $\Psi_{a,b}$ 为小波基函数; $W_f(a,b)$ 为小波系数,包含二维,分别为波长(400~2450 nm)与分解尺度(1, 2, ..., 10)。

1.4 模型的建立与检验

通过The Unscrambler X软件使用偏最小二乘和支持向量机回归方法构建土壤有机质含量的估测模型,其中偏最小二乘模型的最佳潜在变量依入选波段和小波系数的数量而定;支持向量机回归函数可通过多次构建模型和模型调试后最终确定,其中Cost惩罚参数为1, Gamma参数为0.01。采用决定系数(R^2)、均方根(RMSE)和相对分析误差(RPD)检验模型的拟合效果和估测精度。 R^2 越大, RMSE越小,则模型的精度越高, RPD越大则说明模型的预测能力越好^[12]。

2 结果与讨论

2.1 土壤有机质含量特征统计

依据全国第二次土壤有机质分级标准,对研究区98个土壤样品按有机质含量的高低分为3个等级(<6 g·kg⁻¹为极缺乏、6~10 g·kg⁻¹为很缺乏、10~20 g·kg⁻¹为缺乏)。通过分析有机质含量的基本特征(表1),可以看出研究区总体样品土壤有机质含量在1.150~17.582 g·kg⁻¹范围内变化,即研究区土壤有机质含量在极缺乏和缺乏类别之间,土

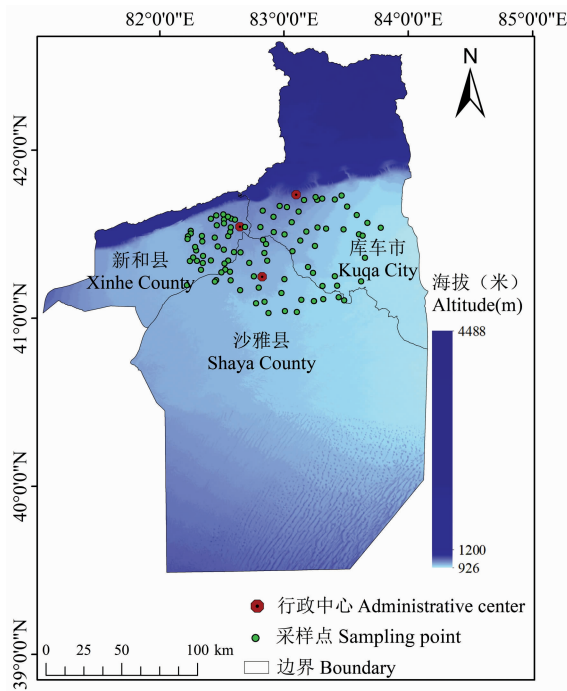


图1 采样点分布图

Fig. 1 Distribution of sampling points

壤有机质平均含量为 $8.574 \text{ g} \cdot \text{kg}^{-1}$ ，属于很缺乏水平。分析认为研究区气候干旱、降水稀少等因素导致土壤中微生物活性较低，从而使有机质在土壤中的积累较慢，土壤养分匮乏。为了对研究区耕层土壤有机质含量的离散程度进行深入分析，经计算土壤有机质的变异系数为 41.86% ，属于中等程度的空间变异，说明研究区土壤有机质含量的空间变异程度较高，构建模型可具有代表性。

表 1 土壤样品的基本统计特征

Table 1 Basic statistical characteristics of soil samples

样品类型	样品数 / 个	有机质含量 / ($\text{g} \cdot \text{kg}^{-1}$)				变异系数 / %
		最大值	最小值	平均值	标准差	
总体样品	98	17.582	1.150	8.574	3.589	41.86

2.2 土壤光谱特征分析

由于影响土壤光谱数据的因素较多，在野外采集光谱数据后剔除较明显的受干扰波段，并求出每个等级的光谱反射率平均值，按照每个等级的平均光谱反射率做出不同等级的土壤光谱反射率曲线(图 2)。不同等级有机质含量的光谱曲线具有以下特征：(1)三个不同肥力级别的土壤光谱反射率曲线大致走向类似，在 $400 \sim 1\,000 \text{ nm}$ 区间，随着波长的增加，反射率呈现上升趋势；在 $1\,000 \text{ nm}$ 以后，除了水分吸收谷外，曲线整体上较为平稳。(2)在 $1\,400, 1\,900$ 和 $2\,200 \text{ nm}$ 附近存在水分吸收谷。(3)土壤有机质含量与土壤的光谱反射率 R 呈负相关，有机质含量越高，土壤反射率越低。本研究样品的平均值为 $8.57 \text{ g} \cdot \text{kg}^{-1}$ 属于很缺乏级别，故土壤样品平均值的光谱反射率与很缺乏曲线基本吻合，说明土壤的光谱反射率与有机质含量具有明显的相关性，土壤光谱反射率可准确反映有机质含量信息。

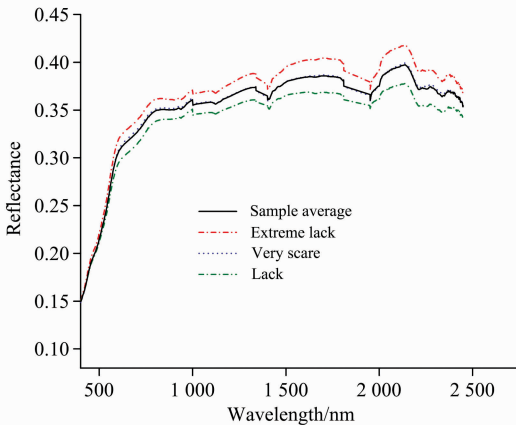


图 2 不同等级有机质含量的光谱曲线

Fig. 2 Spectral curves of organic matter content in different grades

2.3 基于传统变换方法的相关性分析

分别对原始光谱反射率 R 以及经 $\lg(1/R)$ 、 R' 和 $[\lg(1/R)]'$ 转换的光谱反射率与土壤有机质含量进行相关性分析(图 3)。从图中可看出， R 和 $\lg(1/R)$ 在可见光和近红外波段范围内有较多光谱波段可通过 $p < 0.01$ 显著性检验，其中大

部分显著波段分布在可见光范围内。通过对 R 和 $\lg(1/R)$ 与土壤有机质含量进行相关分析发现，这两种光谱反射率与有机质含量之间的相关性并不高。究其原因可能是土壤光谱数据在收集过程中无法避免会受到周围环境中的植被、云层、风力以及光照强度等因素的影响，使得光谱信息被不同程度的干扰。通过对原始光谱进行 R' 、 $[\lg(1/R)]'$ 变换处理后，其光谱反射率与土壤有机质含量之间的相关性有所提高，通过 $p < 0.01$ 显著性检验的波段零散分布在 $430 \sim 2\,100 \text{ nm}$ 范围内(表 2)。通过一阶微分变换可有效进行混合光谱的分解，降低背景噪声的干扰或消除基线漂移，有助于扩大样品之间的微小光谱特征差异，提升光谱的识别能力以及有效波段的获取。以 4 种光谱的特征波段反射率作为自变量，土壤有机

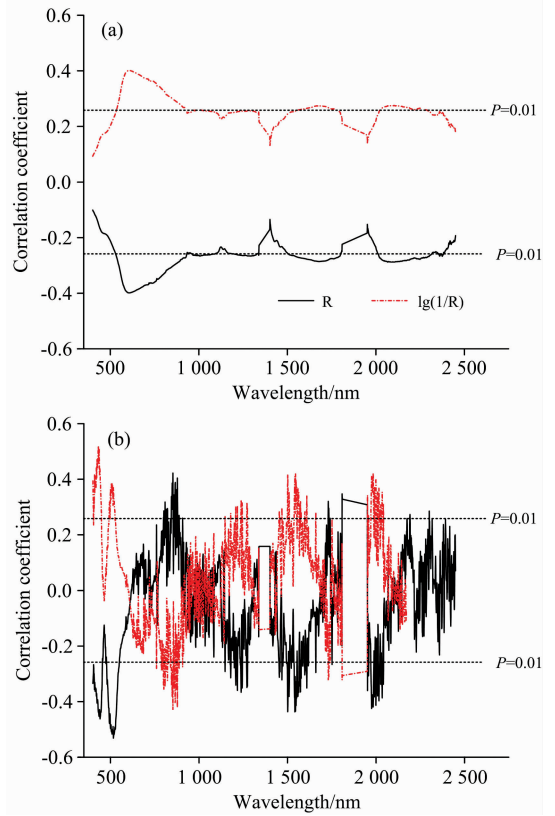


图 3 土壤有机质含量与光谱反射率及其变换的相关性分析
Fig. 3 Correlation analysis of soil organic matter content with reflectance spectra and spectral transformations

表 2 不同变换下的特征波段

Table 2 The characteristic bands selected by different transformations

光谱变换	特征波段数量	对应波段范围/nm	显著性水平 (p 值)
R	1 135	529~933, 2 351~2 382	< 0.01
$\lg(1/R)$	1 137	526~934, 2 349~2 378	< 0.01
R'	318	431~538, 800~889, 1 451~1 660, 1 974~2 035	< 0.01
$[\lg(1/R)]'$	293	431~538, 795~881, 1 731~2 042	< 0.01

质含量作为因变量，构建传统数学变换下的土壤有机质含量的偏最小二乘回归 PLSR 和支持向量机回归 SVMR 估测模型。

2.4 基于 CWT 的相关性分析

为了进一步探索土壤光谱与有机质含量之间的相关性，对原始光谱反射率 R 进行连续小波变换，光谱数据可分解成不同尺度，有效去除了土壤高光谱数据中的白噪音。由于小波基函数的选择对于原始光谱进行 CWT 较为重要，相关学者选择不同小波母函数作为基函数对原始光谱反射率进行

CWT 处理，从而得到了较好的预测结果^[13-14]。研究采用 Bior1.3, db4, Gaus4 和 mexh 等基函数，对原始光谱反射率 R 进行多次小波变换实验，其中以 Bior1.3 为基函数的小波变换结果较为理想，故选择小波变换基函数为 Bior1.3 函数。通过 MATLAB 编程，对原始光谱反射率 R 进行 CWT 处理，为了减少数据冗余，尺度分别选择 $2^1, 2^2, 2^3, \dots, 2^{10}$ ，并对生成的小波变换后的光谱数据（即小波系数）与土壤有机质含量进行相关性分析（见图 4）。

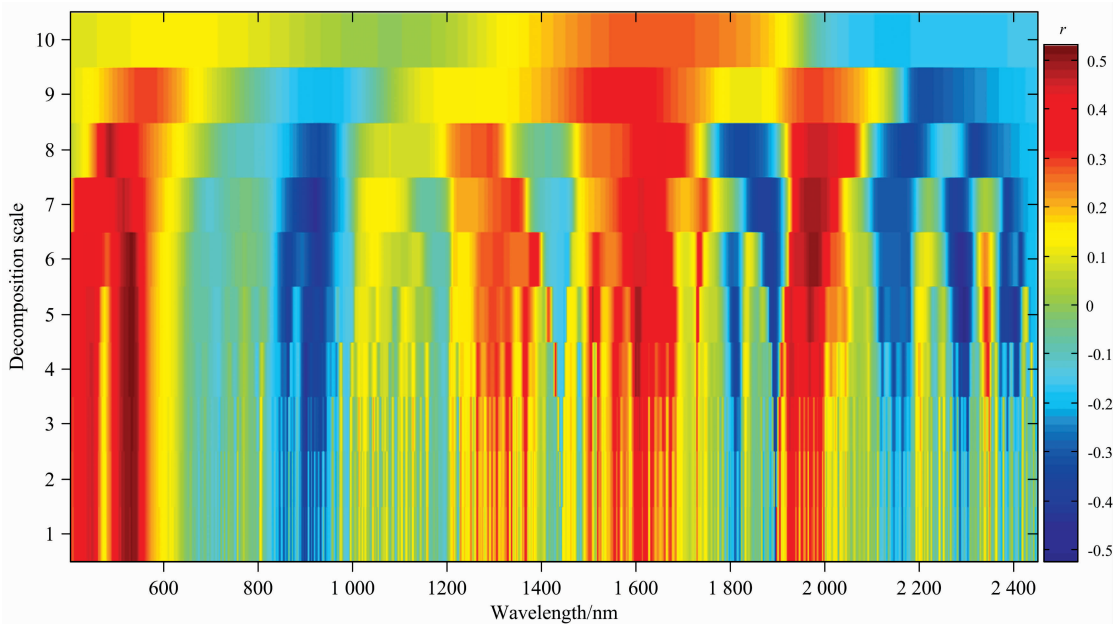


图 4 土壤有机质含量与小波系数的相关性

Fig. 4 The correlation between soil organic matter content and wavelet coefficients

通过对 CWT 分解后的 10 种尺度的小波系数与土壤有机质含量进行相关分析，发现各尺度的小波系数与土壤有机质含量之间的相关性较传统数学变换处理有所提高，相关性较高的区域主要位于 410~525, 790~890, 1 244~1 274 和 2 008~2 038 nm 处。从不同尺度来看，CWT 分解后的光谱反射率与有机质含量之间的相关性较高的区域出现在 1~4 尺度的可见光波段和 5~7 尺度的近红外波段，第 8~10 尺度下的 CWT 处理对提升光谱与有机质含量之间的关联性无明显作用。经各分解尺度变换后，通过 $p < 0.01$ 显著性水平检验的小波系数总数为 2 719 个。同时，从图 4 中可以看出原来与土壤有机质含量只存在负相关关系的原始光谱反射率 R ，通过 CWT 处理后各尺度都存在一定的正负相关性。分析说明，CWT 处理能有效放大光谱中的微妙信号，更有利于获取土壤中有有机质含量信息。

2.5 土壤有机质高光谱反演模型的建立

分别以特征波段的原始光谱反射率 R 和三种传统数学变换下的光谱反射率以及经过 CWT 分解后选取的敏感小波系数作为自变量，土壤有机质含量作为因变量，通过偏最小二乘回归 (PLSR) 和支持向量机回归 (SVMR) 方法构建土壤有机质含量估测模型。为了找出原始光谱反射率 R 经 CWT

变换后建模的最佳尺度，把各个尺度上的敏感小波系数单独作为自变量建立估测模型。通过建模软件选取约 30% 的样品作为交叉验证样品集，用于检验模型的估测精度。因 CWT 处理后，以 $2^8, 2^9, 2^{10}$ 尺度下的小波系数作为自变量建立的反演模型精度低且不稳定，决定系数小于 0.4，故在最终结果中没有展示。分别以不同变换下的特征波段与不同尺度下的敏感小波系数作为自变量建立的反演模型，其建模和验证结果如表 3 所示。

表 3 土壤有机质含量反演模型的建模集与验证集结果
Table 3 Inversion results of modeling set and validation set for soil organic matter content

模型	建模集		验证集		RPD
	R^2	RMSE	R^2	RMSE	
R-PLSR	0.59	2.29	0.46	2.96	1.55
R-SVMR	0.61	2.90	0.32	2.92	1.23
$\lg(1/R)$ -PLSR	0.58	2.31	0.46	2.64	1.54
$\lg(1/R)$ -SVMR	0.63	2.23	0.31	2.99	1.60
R' -PLSR	0.61	2.24	0.46	2.64	1.59
R' -SVMR	0.69	2.05	0.48	2.59	1.74
$[\lg(1/R)]'$ -PLSR	0.61	2.23	0.46	2.62	1.60

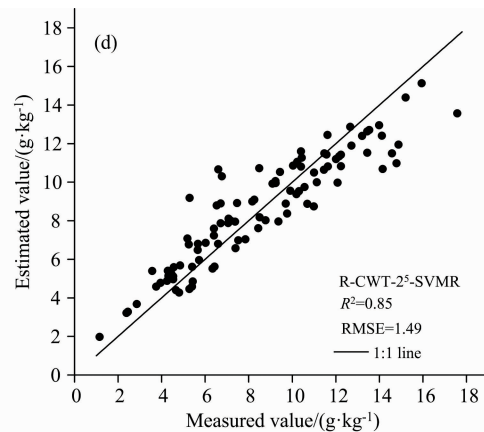
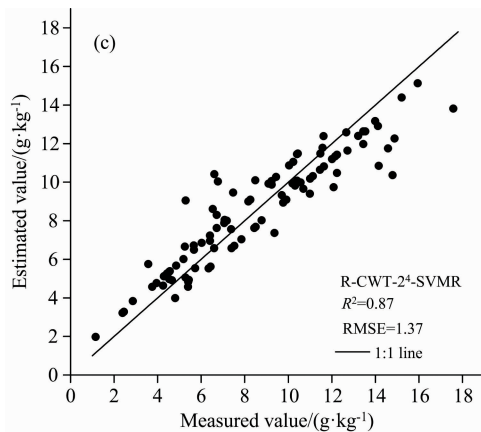
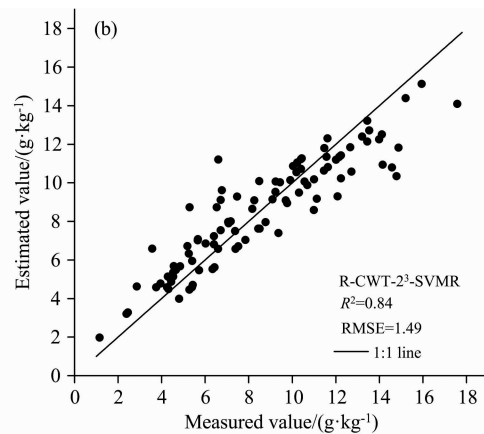
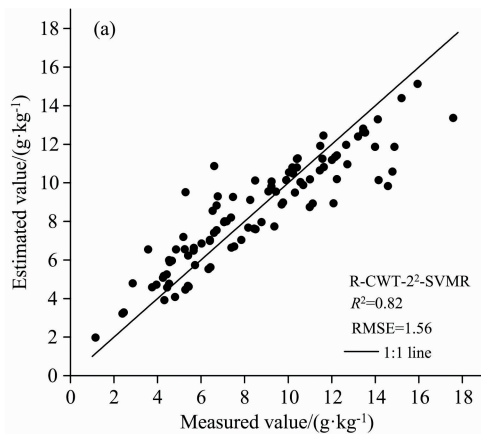
续表 3

$[\lg(1/R)]^t$ -SVMR	0.72	1.93	0.53	2.44	1.85
R-CWT-2 ¹ -PLSR	0.68	2.02	0.46	2.63	1.77
R-CWT-2 ¹ -SVMR	0.79	1.68	0.52	2.48	1.92
R-CWT-2 ² -PLSR	0.61	2.23	0.48	2.64	1.72
R-CWT-2 ² -SVMR	0.82	1.56	0.56	2.38	1.98
R-CWT-2 ³ -PLSR	0.63	2.18	0.46	2.62	1.73
R-CWT-2 ³ -SVMR	0.84	1.49	0.53	2.38	2.11
R-CWT-2 ⁴ -PLSR	0.65	2.11	0.52	2.57	1.86
R-CWT-2 ⁴ -SVMR	0.87	1.37	0.50	2.44	1.91
R-CWT-2 ⁵ -PLSR	0.65	2.11	0.55	2.40	1.89
R-CWT-2 ⁵ -SVMR	0.85	1.44	0.47	2.53	1.99
R-CWT-2 ⁶ -PLSR	0.65	2.12	0.56	2.39	1.88
R-CWT-2 ⁶ -SVMR	0.85	1.45	0.50	2.53	1.86
R-CWT-2 ⁷ -PLSR	0.60	2.25	0.51	2.50	1.58
R-CWT-2 ⁷ -SVMR	0.80	1.64	0.41	2.74	1.97

通过分析表 3 可知,以原始光谱反射率 R 与倒数对数 $\lg(1/R)$ 为自变量的 PLSR 和 SVMR 模型的确定系数 R^2 较小,而通过一阶微分变换后模型的精度显著高于前者,其中 $[\lg(1/R)]^t$ -SVMR 模型的建模集和验证集确定系数 R^2 分别为 0.72 和 0.53,均方根误差 RMSE 分别为 1.93 和 2.44,相对分析误差 RPD 为 1.85,说明以倒数对数一阶微分变换处理后构建的模型拥有较高的估测精度,但是模型的预测能力一般。以 CWT 处理下的不同尺度(2¹, 2², 2³, ..., 2⁷)敏感小波系数为自变量的估测模型,其建模集和验证集的估测精度较传统数学变换有所提高, R^2 和 RPD 增大, RMSE 减小,

预测能力提升明显。且经 CWT 处理后的 SVMR 模型在第 2, 3, 4, 5, 6 和 7 分解尺度下的建模集确定系数均可达 0.8 以上,相对分析误差 RPD 大于 1.8,说明这些模型均具有较高的估测能力和稳定性。其中,第 3 尺度 R-CWT-2³-SVMR 模型的估测精度最高,稳定性最好。综合分析各模型估测结果认为,经 CWT 处理后的模型精度明显优于传统数学变换后的模型, SVMR 模型的精度高于 PLSR 模型。

为了更好地展示模型的估测结果,选择第 2, 3, 4, 5, 6 和 7 分解尺度下估测效果较好的 SVMR 模型,分别以样本实测值与估测值作为横纵坐标做出散点图(见图 5)。从图中可清晰地看到模型的估测值和实测值基本分布在 1:1 拟合线附近,散点分布较为集中,模型估测精度较高。综合分析认为,利用 CWT 分解后的高光谱数据结合 SVMR 方法建立的模型估测精度较高,可有效反演土壤有机质含量信息。在以第 3 尺度敏感小波系数为自变量建立的 R-CWT-2³-SVMR 模型中,确定系数 R^2 相较于原始光谱建模结果提升了 0.23,均方根误差 RMSE 降低了 1.41,模型的相对分析误差 RPD 达到了最大(值为 2.11),预测能力很好。分析说明,经 CWT 第 3 尺度分解处理后的 SVMR 模型可作为研究区土壤有机质含量的有效估测模型。通过多种变换处理,在不同程度上可有效提高光谱反射率与土壤有机质含量的相关性以及反演模型的估测精度;相较于传统变换方法,多种分解尺度下的连续小波变换处理能明显消除不确定性因素对光谱信息的干扰,从而快速提升反演模型的预测能力和适用性^[15-16]。



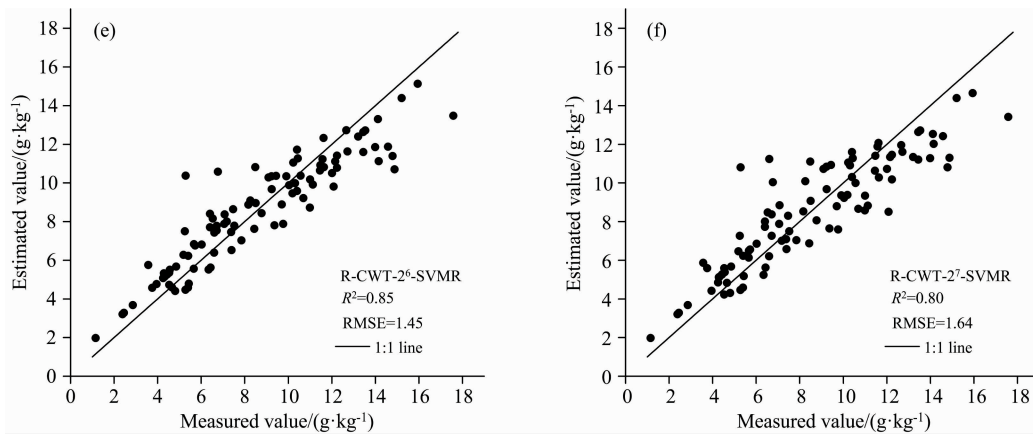


图5 CWT下的土壤有机质含量实测值与估测值比较

Fig. 5 Comparison of measured and estimated values of soil organic matter content using CWT

3 结论

以连续小波变换方法对渭干河-库车河三角洲绿洲耕层土壤光谱反射率 R 进行 $2^1, 2^2, 2^3, \dots, 2^{10}$ 尺度分解, 将筛选出来的敏感小波系数和 $R, \lg(1/R), R', [\lg(1/R)]'$ 处理筛选出来的特征波段光谱值与 98 个土壤样品的土壤有机质含量进行了高光谱反演模型的构建, 并分析对比偏最小二乘回归 (PSLR) 和支持向量机回归 (SVMR) 模型的精度和预测能力, 可得出以下结论:

(1) 原始光谱反射率 R 进行各类变换后与土壤有机质含量的相关性系数明显提高, 由原光谱数据的 0.39 提高到 0.54。说明不管是传统变换方式还是连续小波变换, 都可以不同程度地放大一些细小的光谱吸收特征。

(2) 通过对 $R, \lg(1/R), R'$ 和 $[\lg(1/R)]'$ 构建的反演模

型估测结果比较认为, R' 和 $[\lg(1/R)]'$ 构建的模型精度较高, 其中 $[\lg(1/R)]'$ -SVMR 模型精度最高, 决定系数为 0.72, 说明一阶微分变换在一定程度上可以提高模型的估测精度, 且支持向量机回归方法更适合于进行本研究区域土壤有机质含量的估测, 模型精度和稳定性略高于偏最小二乘回归模型。

(3) 经过 CWT 分解后, 以原始光谱反射率 R 在不同尺度上的敏感小波系数作为自变量建立的模型, 估测精度较传统变换均有明显的提高, 其中 R -CWT- 2^3 -SVMR 模型精度最高, 且预测能力极好, 其建模 R^2 等于 0.84, RMSE 为 1.49, RPD 等于 2.11。CWT 处理后的模型精度优于传统变换下建立的模型, 说明 CWT 处理方法与传统的光谱变换方法相比更适合于挖掘土壤有效信息, 建立的模型可更加精准反演土壤有机质含量。

References

- [1] YANG Yue-chao, ZHAO Ying-jun, QIN Kai, et al(杨越超, 赵英俊, 秦凯, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2019, 35(20): 94.
- [2] YE Qin, JIANG Xue-qin, LI Xi-can, et al(叶勤, 姜雪芹, 李西灿, 等). Transactions of the Chinese Society for Agricultural Machinery(农业机械学报), 2017, 48(3): 164.
- [3] ZHOU Qian-qian, DING Jian-li, TANG Meng-ying, et al(周倩倩, 丁建丽, 唐梦迎, 等). Acta Pedologica Sinica(土壤学报), 2018, 55(2): 313.
- [4] ZHANG Dong-hui, ZHAO Ying-jun, QIN Kai, et al(张东辉, 赵英俊, 秦凯, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2018, 34(20): 141.
- [5] NAN Feng, ZHU Hong-fen, BI Ru-tian(南锋, 朱洪芬, 毕如田). Scientia Agricultura Sinica(中国农业科学), 2016, 49(11): 2126.
- [6] Shen L, Gao M, Yan J, et al. Remote Sensing, 2020, 12(7): 1206.
- [7] ZHANG Rui, LI Zhao-fu, PAN Jian-jun(张锐, 李兆富, 潘剑君). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2017, 33(1): 175.
- [8] ZHANG Tao, YU Lei, YI Jun, et al(章涛, 于雷, 易军, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2019, 39(10): 3217.
- [9] YE Hong-yun, XIONG Hei-gang, ZHANG Fang, et al(叶红云, 熊黑钢, 张芳, 等). Advances in Laser & Optoelectronics Progress(激光与光电子学进展), 2019, 56(5): 115.
- [10] WANG Yong-min, LI Xi-can, TIAN Lin-ya, et al(王永敏, 李西灿, 田林亚, 等). Remote Sensing for Land and Resources(国土资源遥感), 2019, 31(1): 110.

- [11] DENG Kai, DING Jian-li, YANG Ai-xia, et al(邓凯, 丁建丽, 杨爱霞, 等). Journal of Ecology(生态学报), 2016, 36(20): 6387.
- [12] LI Guan-wen, GAO Xiao-hong, XIAO Neng-wen, et al(李冠稳, 高小红, 肖能文, 等). Acta Optica Sinica(光学学报), 2019, 39(9): 361.
- [13] Asim Biswas. Catena, 2018, 160: 24.
- [14] Shen L, Gao M, Yan J, et al. Remote Sensing, 2020, 12(7): 1206.
- [15] Qiao X X, Wang C, Feng M C, et al. Spectroscopy Letters, 2017, 50(3): 156.
- [16] Li R, Gao M, Zexin X U, et al. Hyperspectral Estimation of Soil Organic Matter in Apple Orchard Based on CWT. IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2021, 734(1): 012030.

Hyperspectral Estimation of Soil Organic Matter Content Based on Continuous Wavelet Transformation

Yumiti Maiming¹, WANG Xue-mei^{1, 2*}

1. College of Geographic Science and Tourism, Xinjiang Normal University, Urumqi 830054, China

2. Xinjiang Uygur Autonomous Region Key Laboratory "Xinjiang Arid Lake Environment and Resources Laboratory", Urumqi 830054, China

Abstract The hyperspectral estimation of soil organic matter content can quickly and accurately monitor soil fertility and provide a scientific basis for proper fertilization of modern agricultural production. Taking the cultivated soil in the delta oasis of Weigan-Kuqa river, Xinjiang as the research object, the original spectral reflectance (R) of the collected 98 soil samples was subjected to the traditional reciprocal logarithm $\lg(1/R)$, the first-order differential (R') and the reciprocal logarithm first-order differential $[\lg(1/R)]'$ mathematical transformations, and continuous wavelet transformation (Continuous Wavelet Transformation, CWT) processing based on Bior1.3 as the wavelet mother function through different scale decomposition. Correlation analysis was conducted between the treatment results and the measured soil organic matter content to screen out the characteristic bands and wavelet coefficients closely related to soil organic matter content under various transformations ($p < 0.01$). With the original spectral reflectance, the characteristic band reflectance and the sensitive wavelet coefficient under different transformation treatments as independent variables and the soil organic matter content as dependent variables, partial least squares regression and support vector machine regression was used to estimate models of soil organic matter content. The results showed that: (1) Various spectral transformation methods can effectively improve the sensitivity between the spectrum and the content of soil organic matter. The correlation between the soil spectral reflectance and the organic matter content after continuous wavelet transformation has been significantly improved, and the correlation coefficient has been increased from 0.39 to 0.54 ($p < 0.01$). (2) The support vector machine regression model built by the traditional $[\lg(1/R)]'$ transformation has a higher coefficient of determination (R^2) than the model built by $\lg(1/R)$ and R' transformation, showing the reciprocal logarithm first-order differential transformation can help improve the accuracy of the estimation model, and the accuracy and stability of the support vector machine regression model were higher than that of the partial least squares regression model. (3) After continuous wavelet transformation decomposition, the estimation accuracy and stability of the models were obviously improved by using the sensitive wavelet coefficients of the original spectral reflectance at different scales as independent variables. The decision coefficient (R^2), root mean square error (RMSE) and relative analysis error (RPD) of the CWT-23-SVMR model were 0.84, 1.48 and 2.11 respectively. The model has the highest accuracy and excellent predictive ability. After multiple transformation processing, hyperspectral data can effectively remove white noise. In contrast, continuous wavelet transformation processing is more suitable for mining effective soil information than the traditional mathematical transformation method, to realize the effective separation of approximate features and detailed features of spectral signals, and the established inversion model can more accurately estimate soil organic matter content.

Keywords Continuous wavelet transformation; Hyperspectral estimation; Decomposition scale; Soil organic matter; The delta oasis of Weigan-Kuqa rivers

* Corresponding author

(Received Nov. 26, 2020; accepted Sep. 15, 2021)