

激光与光电子学进展

光谱配置对最优波段组合算法预测土壤
有机质和电导率的影响

张焱, 丁建丽*, 张子鹏, 葛翔宇, 王瑾杰

新疆大学资源与环境科学学院, 新疆 乌鲁木齐 830046

摘要 可见光和近红外(Vis-NIR)光谱仪是一种实用的工具,可用于替代土壤物理和化学分析以评估土壤性质。最佳波段组合算法是一种通过考虑波段之间的相互作用信息来提取光谱变量的有效方法,但是对于实验室 Vis-NIR 光谱数据,该方法易受“维数灾难”的影响。提出一种适当地降低光谱配置(减少光谱带的数量和粗化光谱分辨率),以提高计算效率且不会显著影响预测精度的方法。首先,构建了 6 个光谱配置,即光谱波段的数量从 2001 个减小到 19 个,光谱分辨率从 3 nm 增加至 100 nm,并且光谱采样间隔等于光谱分辨率(均匀间隔采样)。然后,通过偏最小二乘回归结合最佳波段组合算法预测不同光谱配置下土壤有机质(SOM)和电导率(EC)的最佳光谱参数。结果表明:直到光谱分辨率为 60 nm 时(每个光谱 32 个波长),最佳波段组合算法与全波段光谱数据相比,可以提高预测精度。最佳波段组合算法在 1~20 nm 光谱分辨率下,预测精度没有显著差异(约为 2%);SOM[决定系数 R^2 等于 0.77,均方根误差(RMSE_p)等于 0.325%,性能四分位数范围的比率(RPIQ_v)等于 3]和 EC(R^2 等于 0.70, RMSE_p 等于 6.88 dS·m⁻¹, RPIQ_v 等于 2.21)分别在 20 nm 和 10 nm 的光谱分辨率下获得了最佳预测性能。

关键词 传感器; 可见光-近红外光谱; 光谱配置; 主成分分析; 最佳光谱参数; 偏最小二乘回归

中图分类号 S127

文献标志码 A

doi: 10.3788/LOP202158.2128001

Effect of Spectral Configuration on Soil Organic Matter and Electrical
Conductivity Predicted by Optimal Band Combination Algorithm

Zhang Yi, Ding Jianli*, Zhang Zipeng, Ge Xiangyu, Wang Jinjie

College of Resource and Environment Sciences, Xinjiang University, Urumqi, Xinjiang 830046, China

Abstract Visible light and near infrared (Vis-NIR) spectrometer is a practical tool that can be used for alternative soil physical and chemical analysis to assess soil properties. The optimal band combination algorithm is an effective method to extract spectral variables by considering the interaction information between bands. However, for laboratory Vis-NIR spectral data, this method is vulnerable to the “dimension disaster”. A method that appropriately reduces the spectral configuration (reducing the number of spectral bands and coarsening the spectral resolution) to improve the calculation efficiency without significantly affecting the prediction accuracy is proposed. First, six spectral configurations are constructed, which means that the number of spectral bands is reduced from 2001 to 19, the spectral resolution is increased from 3 nm to 100 nm, and the spectral sampling interval is equal to the spectral resolution (uniformly spaced sampling). Then, partial least squares regression combined with the optimal band combination algorithm is used to predict the optimal spectral parameters of soil organic matter (SOM) and electrical conductivity (EC) under different spectral configurations. The results show that until the spectral resolution is 60 nm (32 wavelengths per spectrum), the optimal band combination

收稿日期: 2021-03-02; 修回日期: 2021-03-02; 录用日期: 2021-03-08

基金项目: 国家自然科学基金(41771470, 41661046)、自治区重点实验室专项基金(2016D03001)

通信作者: *watarid@xju.edu.cn

algorithm can improve the prediction accuracy compared with the full band spectrum data. The best band combination algorithm has no significant difference in prediction accuracy under 1–20 nm spectral resolution (about 2%); SOM [determination coefficient(R_v^2) is 0.77, root mean square error(REMS_p) is 0.325%, ratio of performance to interquartiledistance(RPIQ_v) is 3] and EC (R_v^2 is 0.69, RMSE_p is 6.88 dS·m⁻¹, RPIQ_v is 2.21) obtain the best prediction performance at the spectral resolution of 20 nm and 10 nm.

Key words sensors; Vis-NIR spectrum; spectral configuration; principal component analysis; optimal spectral parameter; partial least squares regression

OCIS codes 280.4788; 300.6170; 300.6550

1 引言

土壤有机质(SOM)含量和电导率(EC)是衡量土壤肥力与盐度的关键指标。快速准确地监测 SOM 和 EC 对干旱地区土地资源的合理利用至关重要。近年来,可见光-近红外(Vis-NIR)光谱可以快速地定量分析某些特定土壤特性,且无需制备任何样品^[1-2]。因此,Vis-NIR 光谱被广泛认为是可以替代常规手段的土壤特性监测方法。

对土壤性质进行光谱建模时,输入数据大致可分为全波段光谱和筛选出的光谱变量两种。最佳波段组合算法可以从任一可能的波长组合进行提取,从而有效提高目标的灵敏度并降低无关波长的影响,目前已广泛应用于高光谱数据的变量选择^[3-4]。在实践中,对于 Vis-NIR 范围内的高光谱数据,该方法是通过两波段指标(即比率、差、归一化差、乘积和总和)实现的^[5]。许多研究表明三波段指数(TBI)的性能优于两波段指数,但该结果仅是基于特定范围的高光谱数据,例如 350~1000 nm, 300~1150 nm 和 400~1000 nm^[6-8]。目前,在 Vis-NIR 范围内很少使用三波段光谱指数,尤其在涉及高光谱数据时。

将高光谱数据集简化为简单的表示形式,有利于数据的有效存储、计算、分析和传输。Peng 等^[9]对实验室获得的 Vis-NIR 光谱进行了采样间隔为 1~10 nm 的重采样,结果表明,9 nm 是土壤有机碳估算的最佳光谱采样间隔。Yang 等^[10]证明了在土壤有机氮估算中,用 100 nm 带宽(每个光谱 21 个波长)建立的模型几乎与未经重采样光谱(每个光谱 2100 个波长)建立的模型表现相同。此外,Adeline 等^[11]证明了使用实验室光谱法在 3~60 nm 的光谱分辨率下可获得良好的粘土预测效果。

本文旨在评估光谱配置的可行性,即通过减少光谱数量和粗化光谱分辨率,来提高最佳波段组合算法的计算效率,且不影响预测精度。原始的 Vis-NIR 光谱具有 2001 个光谱带,重采样的光谱间隔为

1 nm,从初始数据库构建 6 个光谱配置,光谱分辨率等于光谱采样间隔(即间隔采样均匀)。在这些配置中,光谱带的数量从 2001 减小到 19,光谱分辨率从 3 nm 扩大到 100 nm。然后使用最佳波段组合算法,得出了不同光谱配置下土壤特性(SOM 和 EC)的最佳光谱参数。最后在不同的光谱配置中,通过偏最小二乘回归(PLSR)模型比较了最佳波段组合算法和全波段光谱数据预测的性能。

2 材料和方法

2.1 研究区和土壤数据收集

研究在中国新疆西北部的艾比湖地区(82°20′~83°55′ E, 44°23′~45°12′ N)进行,该地区是典型的大陆性干旱气候,年均降水稀少(平均年雨量 105 mm)且蒸发作用强烈(年蒸发量 1315 mm)。根据世界参考基准土壤分类可以得知,盐渍土是该地区的主要土壤类型。研究区的西南部是博尔塔拉河和精河的入湖口,这里土地类型主要是农田,主要农作物是棉花^[12-13]。

借鉴之前的实地采样经验^[3],约 2 km 间隔设置一个采样点,并根据野外调查时道路的通达性选择了一个相对平坦的区域进行采样,同时为避免过度的人为干扰,未采集处于农田、沟渠和垃圾场附近的土壤。在 2019 年 7 月 26 日—8 月 7 日,本实验组共采集了 255 个土壤样本。每个采样点,在 0~20 cm 深度范围内收集了 5 个子样本(质量约 2 kg)。

2.2 理化分析

在室内,将所有土壤样品均化、风干(温度为 25~30 °C)、研磨(用玛瑙研钵),然后通过 100 目筛网(孔径为 0.149 mm)。SOM 含量通过重铬酸钾氧化外加热法进行测定,为了测量 EC 和 pH,在室温(25 °C)下,对比比例为 1:5 的土壤样品(质量为 20 g)和蒸馏水(体积为 100 mL)进行了 30 min 的预处理。然后从标本中提取渗滤液,分别使用数字多参数测量仪(Multi 3420 Set B, WTW GmbH)和数字 pH 仪(PES-3E, Shanghai

INESA Scientific Instrument Co., Ltd.)测量 EC 和 pH。

2.3 实验室 Vis-NIR 光谱

实验室 Vis-NIR 光谱数据是使用分光辐射计 (ASD FieldSpeco[®]3, Malvern Panalytical Ltd.) 在 350~2500 nm 的反射范围内在暗室中测得的。350~1000 nm 波段的光谱分辨率为 3 nm, 光谱采样间隔为 1.4 nm; 1000~2500 nm 波段的光谱分辨率为 10 nm, 光谱采样间隔为 2 nm。最后用 ViewSpecPro (Version 6.20, Malvern Panalytical Ltd.) 将反射在两个光谱段内的光谱重新采样为 1 nm, 从而产生了 2151 个光谱并输出。将处理后的土壤样品放在培养皿 (直径为 10 cm, 高度为 2 cm) 中, 然后用直尺将其表面均匀化, 把 50 W 功率卤素灯安装在距样品约 30 cm 的三脚架上, 天顶角为 15°, 将 10° 视场的探头置于培养皿中心上方 15 cm 处。每测量 10 次, 用白板 (Spectralon[®], Malvern Panalytical Ltd.) 对光谱仪进行校准, 每个土壤样品重复测量 10 次, 并将平均值作为代表光谱。由于光谱边缘存在大量噪声, 因此将 350~399 nm 和 2401~2500 nm 的光谱忽略, 再将剩余的光谱 (400~2400 nm) 命名为 ASD_1/1, 因为它在两个光谱范围内 (即 400~1000 nm 和 1000~2400 nm) 的光谱采样间隔均为 1 nm。

2.4 实验室 Vis-NIR 光谱配置

光谱配置中, 必须定义以下 3 个参数: 光谱分辨率 (SR)、频谱采样间隔 (SSI)、光谱带的数量 (N)。本实验组假设光谱配置的 SR 等于 SSI, 即均匀间隔采样。因此, 通过以下两个步骤构建了光谱配置。

1) 根据不同的 SR 计算高斯响应函数, 表达式为

$$f(\lambda, \sigma) = \exp\left[-\frac{(\lambda - \lambda_c)^2}{2\sigma^2}\right], \quad (1)$$

式中: λ 是不同 SSI 下的光谱步长; λ_c 是某些 SSI 的光谱响应范围内的中心波段; σ 是 f 的宽度, 表达式为

$$\sigma = \frac{R_{SR}}{2\sqrt{2\ln 2}}. \quad (2)$$

2) 根据 SSI 和 $f(\lambda, \sigma)$, 可以得出光谱配置 ASD_1/1 降低的反射率为

$$R = \frac{\sum_{\lambda_s}^{\lambda_e} r(\lambda) f(\lambda, \sigma)}{\sum_{\lambda_s}^{\lambda_e} f(\lambda, \sigma)}, \quad (3)$$

式中: s 和 k 分别是光谱范围内起始带和结束带的波段号; $r(\lambda)$ 代表反射率。图 1 为 SR 等于 SSI 时的高斯响应函数示意图。

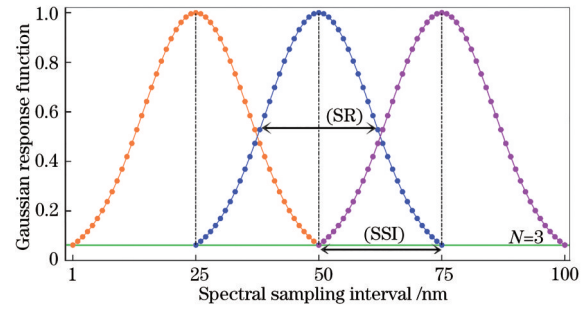


图 1 $N=3, R_{SR}=25 \text{ nm}, I_{SSI}=25 \text{ nm}$ 时的高斯响应函数
Fig. 1 Gaussian response function when $N=3,$
 $R_{SR}=25 \text{ nm}, I_{SSI}=25 \text{ nm}$

每个光谱配置都记录为 Con_X/Y, 其中 X 是 400~1000 nm 处的光谱分辨率, Y 是 1000~2400 nm 处的光谱分辨率, 如表 1 所示。6 个光谱配置的 N 值从 2001 降至 19, SSI 分别在 400~1000 nm 和 1000~2400 nm 的光谱范围内从 3~100 nm 和 10~100 nm 变化。

表 1 7 种光谱配置的描述

Table 1 Description of 7 spectral configurations

Configuration	N	400~1000 nm		1000~2400 nm	
		SSI	SR	SSI	SR
ASD_1/1	2001	1	3	1	10
Con_3/10	339	3	3	10	10
Con_10/10	199	10	10	10	10
Con_20/20	99	20	20	20	20
Con_40/40	49	40	40	40	40
Con_60/60	32	60	60	60	60
Con_100/100	19	100	100	100	100

2.5 主成分分析

主成分分析用于光谱数据, 可以揭示和可视化不同光谱配置的分佈模式, 并检测光谱离群值。首先, 在每种光谱配置中, 通过主成分分析全光谱数据, 用该方法选择主成分数量时是以方差贡献率大于 95% 为标准的; 其次, 将每个光谱配置中提取的主成分投影到不同的坐标空间中; 最后, 基于每个土壤样品在不同投影空间上的距离计算 Hotelling T2 椭圆, 以识别光谱离群值。位于 Hotelling T2 椭圆之外的数据点被定义为离群值, 并从数据集中删除。

2.6 最佳波段组合算法

使用 5 种组合类型探究光谱与土壤特性的关系, 这 5 种类型的 TBI 已成功用于叶片氮和地上生物量以及黄锈病标识^[7,14-15]。本实验组使用的 TBI 参考了 Zhang 等^[16]的研究。

$$I_{TBI_1}(R_1, R_2, R_3) = R_1 / (R_2 + R_3), \quad (4)$$

$$I_{TBI_2}(R_1, R_2, R_3) = (R_1 - R_2) / (R_1 + R_3), \quad (5)$$

$$I_{TBI_3}(R_1, R_2, R_3) = (R_1 - R_2) / (R_2 - R_3), \quad (6)$$

$$I_{TBI_4}(R_1, R_2, R_3) = R_1 / (R_2 \times R_3), \quad (7)$$

$$I_{TBI_5}(R_1, R_2, R_3) = (R_1 - R_2) / (R_1 - 2R_2 + R_3), \quad (8)$$

式中： R_1, R_2 和 R_3 是400~2400 nm范围内的光谱值，且 $R_1 \neq R_2 \neq R_3$ 。最佳波段组合算法是在MATLAB 2019 b软件中进行编程实现的。最佳光谱参数的推导过程包括3个步骤：1)在建模集中，遍历每个TBI与土壤特性(SOM&EC)之间的相关性(r)；2)对于每个TBI， $|r|$ 值按降序排列，保留前1% $|r|$ 对应的波段组合用于后续分析；3)将建模和验证集之间具有最小 r 的波段组合定义为最佳光谱参数。

2.7 土壤特性反演模型建立

为了对土壤特性反演建模并验证模型，本实验组在剔除异常值后，将每个光谱结构中的剩余样本分为

两组。根据土壤特性的实测值，按升序对正常样品进行分类，将具有偶数和奇数的样本分别分配给建模数据集和验证数据集再进行方差分析，以验证土壤性质的统计指标是否在建模和验证数据集之间存在显著差异。本实验组用PLSR建立土壤特性和光谱之间的关系，设定自变量为光谱变量，因变量为土壤属性值(即SOM和EC)。在建模过程中，用投影变量的重要性(VIP)来确定光谱变量的重要性，若特定变量的VIP值超过1，则认为该变量很重要。采用MATLAB 2019 b进行PLSR建模，预测准确性用决定系数(R^2)、均方根误差(RMSE)和性能四分位数范围的比率(RPIQ)3个参数进行评估。通常 R^2 和RPIQ越高，RMSE越低，表明模型具有越好的预测能力。

3 结 果

3.1 不同光谱配置的主成分分析

图2(a)~(g)为在不同的坐标系中不同光谱配

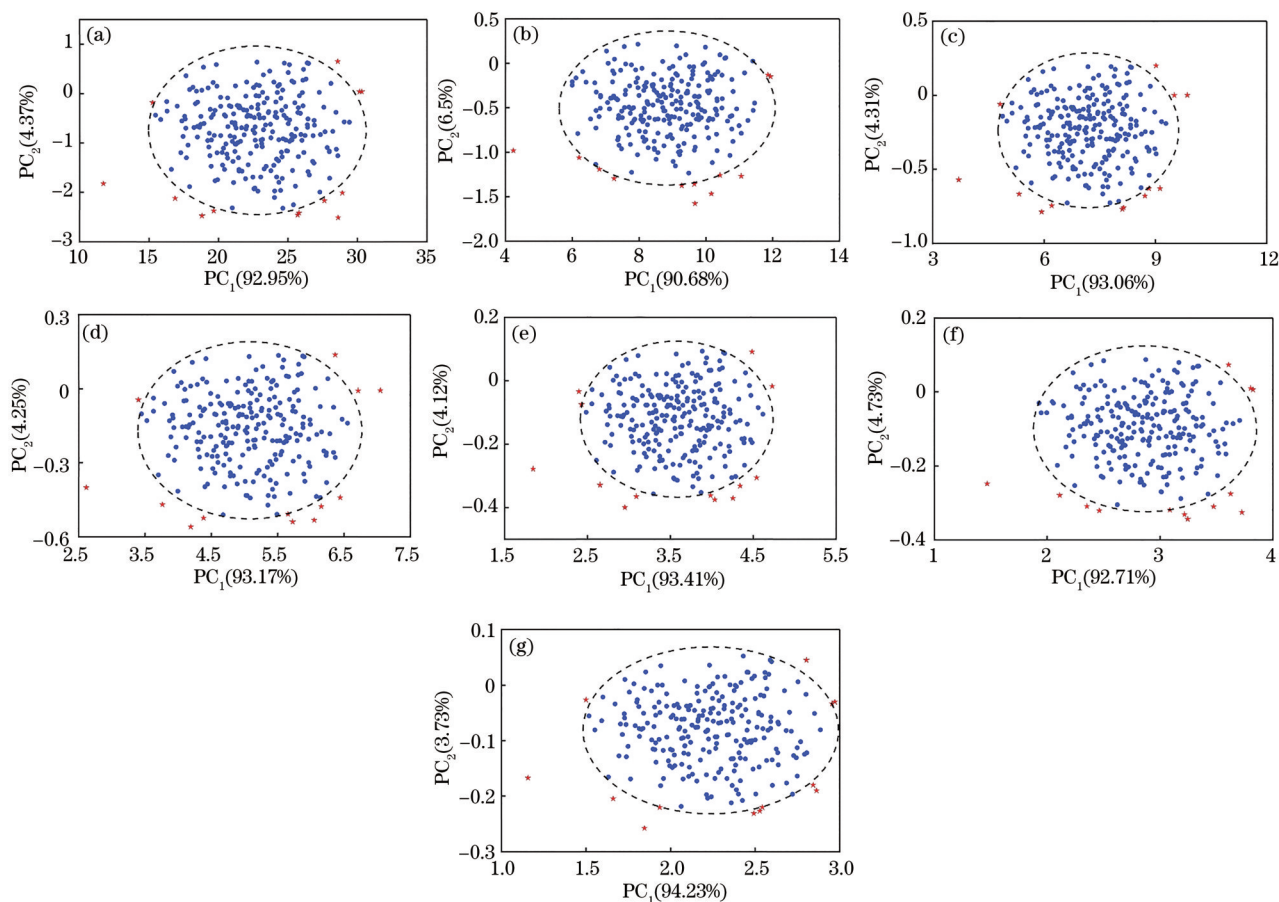


图2 不同光谱配置下255个样本的主成分投影。(a) ASD_1/1; (b) Con_3/10; (c) Con_10/10; (d) Con_20/20; (e) Con_40/40; (f) Con_60/60; (g) Con_100/100

Fig. 2 Principal component projections of 255 samples for different spectral configurations. (a) ASD_1/1; (b) Con_3/10; (c) Con_10/10; (d) Con_20/20; (e) Con_40/40; (f) Con_60/60; (g) Con_100/100

置下 255 个样本的主成分投影。总体而言,前两种主成分的光谱变化均大于 95%,而与光谱配置无关。Shi 等^[17]将第一主成分(PC₁)和第二主成分(PC₂)分别称为光谱反射强度和光谱形状。光谱信号的方差随光谱配置降低而减小,如图 2(c)~(d)所示。同时 Hotelling T²椭圆将每个子图划分为两个区域,有效区域位于黑色虚线内,并且将该区域中的样本点定义为正常样本(圆点);黑色虚线之外为无效区域,并且这些点被标识为光谱离群值(五角星),在每个子图中确定 13 个异常值,并将其从后续分析中删除。

3.2 土壤特性的描述统计

研究区土壤 pH 值范围为 6.91~10.33(即从中性到非常碱性),平均值为 8.29,表明研究区土壤总体呈碱性。图 3 是去除异常值后每种光谱配置中土壤特性的建模集、验证数据集的箱型图和相应的描述统计信息。总体而言,相同土壤特性之间的密度差异很小,如图 3(a)和图 3(d)所示。在每种光谱配置中,EC 的范围(0.21~72.07 dS·m⁻¹)比 SOM 的范围(0.03%~4.10%)宽。Wilding^[18]将变异系数(CV)分为 3 类:低方差(小于 0.15),中方差(0.15~0.35)和高方差(大于 0.35)。SOM(大于 0.6)和 EC(大于或等于 0.9)的 CV 在每种光谱配置中均较高,如图 3(a)和图 3(d)所示,可能是土壤母质的差异导致的。图 3(a)~(c)的 SOM 和图 3(d)~(e)的 EC 都具有正偏度(即 Ske 大于 0),表明分布包括许多低值和相对较少的高值。建模集和验证数据集的描述性统计[最小值(Min),均值(Mean),最大值(Max),标准差(Std),变异系数(CV)和偏度(Ske)]之间的单因素方差分析结果表明,集合的分布无显著差异,即建模集和验证集的数据之间无显著差异。

3.3 最佳波段组合算法

图 4 分别展示了不同含量 SOM 和 EC 下的土壤反射率光谱,分为 6 个不同等级,取 n 个光谱反射率的平均值, n 为实验次数。所有光谱曲线形状相似,先急剧上升后缓慢变化,同时光谱反射率随着 SOM 含量升高有明显下降趋势。其中,SOM 土壤反射率光谱从上到下的含量依次为 <0.6%、0.6%~1%、1%~2%、2%~3%、3%~4%、>4%;EC 土壤反射率光谱从上到下的含量依次为 0~2 dS·m⁻¹、2~4 dS·m⁻¹、4~8 dS·m⁻¹、8~16 dS·m⁻¹、>16 dS·m⁻¹。

最佳波段组合算法计算了 5 个 TBI 与不同光谱配置的土壤特性之间的相关性,如表 2 和表 3 所示,为了更好地比较这些值,最佳光谱参数的相关性以 R^2 的形式表示,发现土壤特性并非总是在 ASD_1/1 和 Con_100/100 下获得最佳和最差的 R^2 值。总体而言,除了 Con_40/40 至 Con_100/100 下的 TBI₂ 外,SOM 的三波段指数准确性总体上高于 EC,随着光谱配置降低,SOM 的 R^2 在 Con_40/40 之前的变化不明显;而 EC 三波段指数的 R^2 变化并无相似现象。对于 SOM 的 Con_60/60,与之前的其他 TBI 相比,TBI₂ 在建模集中准确性相对较低(R^2 值之间的差距小于 0.1),而 EC 的 Con_60/60 所对应的 TBI₂ 的准确性要优于其他 TBI(R^2 值之间的差距介于 0.11~0.45)。表 2 和表 3 总结了每个最佳光谱参数的波段组合,表中 SOM 的 600,620,700,760,1800,2100,1400,1480,1490,1880,2160,2170,2180 nm,EC 的 2120,2180,2280,2320,2360 nm 出现了多次,这对光谱建模很重要。

3.4 模型的建立和验证

用所提评价因子对 PLSR 建模之后的结果和相对应的全波段光谱数据结果的预测性能进行了比较,如表 4 和表 5 所示, R^2 的下标 c 和 v 分别代表建模集和验证集,而 RMSE 中下标 cv 和 p 也分别代表建模集和验证集。总体而言,两种土壤特性(SOM 和 EC)除了 Con_60/60 和 Con_100/100 中的 SOM 之外,所提取的最佳光谱参数的模型精度都优于全波段光谱数据,具体表现在 R^2 和 RPIQ_v 值较高且 RMSE_p 较小。对于提取的最佳光谱参数(表 4 和表 5),从 ASD_1/1 到 Con_20/20,SOM 和 EC 的预测性能对光谱配置较不敏感,而从 Con_40/40 到 Con_100/100,SOM 的精度显著降低(R^2 从 0.72 降至 0.35),但 EC 的精度变化较小(R^2 等于 0.65 或 0.66)。SOM 和 EC 的最佳预测结果分别出现在 Con_20/20 (RMSE_p 等于 0.325%,RPIQ_v 等于 3)和 Con_10/10 (RMSE_p 等于 6.88,RPIQ_v 等于 2.21)中,而不是 ASD_1/1。无论 SOM 和 EC 的光谱配置如何,潜在变量的数量(LVs)均保持为 5,而全波段光谱数据中,SOM 的 LVs 在 12 到 15 之间,EC 的 LVs 保持为 4。

从图 5 可以看出,不同的土壤属性均在 Con_20/20 之后运算效率明显提高,且不损失算法精度,同时最佳波段明显优于全波段组合。运算时间大大缩减也证明了在 Con_20/20 之前存在着一定程度的计算冗余。

图 6 给出了不同 PLSR 模型估测的土壤特性

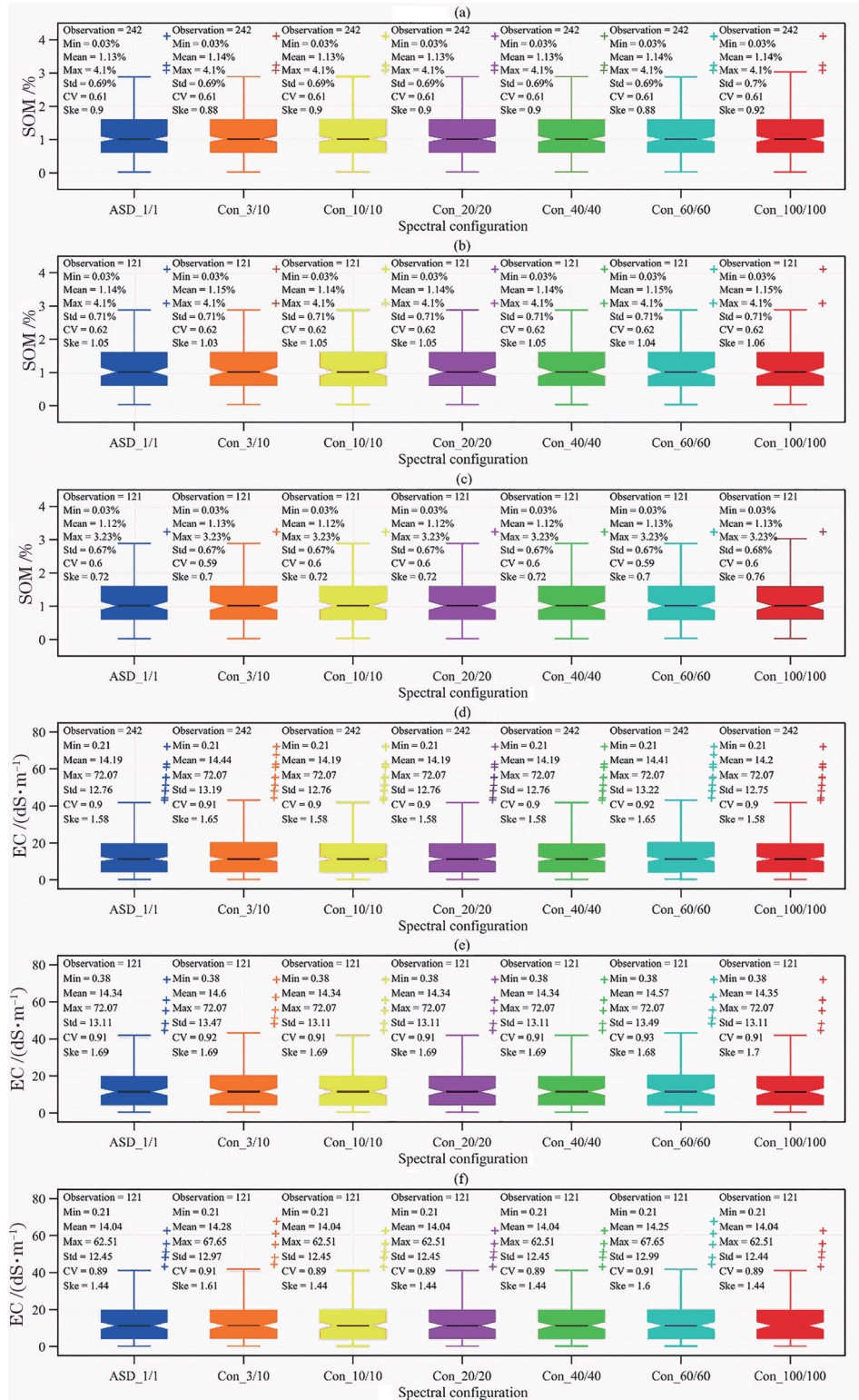


图 3 剔除异常值后的每个光谱配置的土壤特性箱型图以及相应描述统计。(a)~(c) 整个数据集、建模数据集和验证数据集的 SOM 最小值、均值、最大值、标准差、变异系数、偏度；(d)~(f) 整个数据集、建模数据集和验证数据集的 EC 最小值、均值、最大值、标准差、变异系数、偏度

Fig. 3 Soil characteristic box diagram of each spectral configuration after excluding outliers and the corresponding descriptive statistics. (a)–(c) SOM Min, Mean, Max, Std, CV, Ske of entire, calibration, and validation datasets; (d)–(f) EC Min, Mean, Max, Std, CV, Ske of entire, calibration, and validation datasets

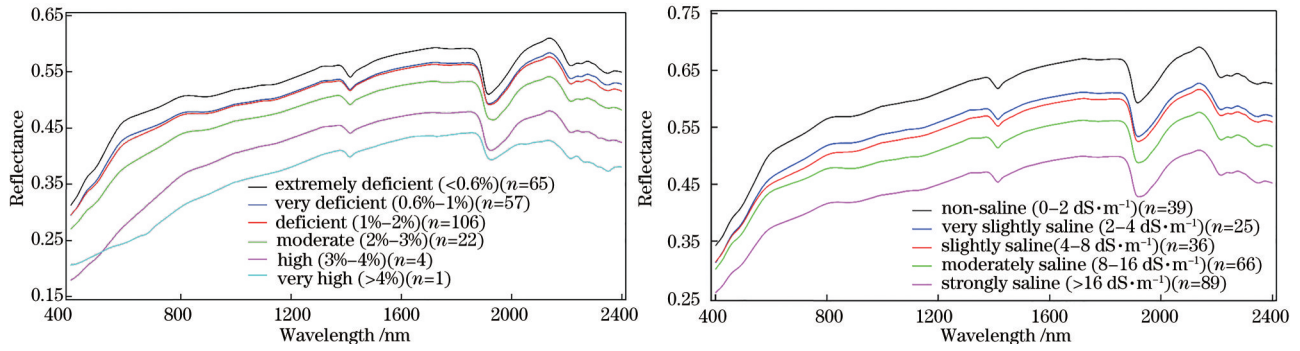


图 4 土壤反射率光谱。(a)不同含量SOM;(b)不同含量EC

Fig. 4 Soil reflectance spectra. (a) SOM with different content; (b) EC with different content

表 2 从最佳波段组合算法得出的各个频谱配置中SOM的最佳频谱参数

Table 2 Best frequency-spectrum parameters of SOM in each spectral configuration derived from optimal band combination algorithm

Configuration	Spectral variable					Maximum determination coefficient				
						(R^2) calibration/validation				
ASD_1/1	TBI ₁ (R ₂₁₇₀ ,	TBI ₂ (R ₄₀₄ ,	TBI ₃ (R ₁₈₇₆ ,	TBI ₄ (R ₁₄₁₉ ,	TBI ₅ (R ₁₈₈₀ ,	0.62/	0.59/	0.68/	0.68/	0.65/
	R ₁₈₈₃ ,R ₂₃₂₂)	R ₆₁₂ ,R ₅₅₂)	R ₂₁₇₀ ,R ₈₄₇)	R ₂₃₇₉ ,R ₉₁₆)	R ₆₅₃ ,R ₂₁₇₀)	0.58	0.34	0.65	0.69	0.64
Con_3/10	TBI ₁ (R ₁₇₈₀ ,	TBI ₂ (R ₄₀₄ ,	TBI ₃ (R ₁₄₈₀ ,	TBI ₄ (R ₁₀₁₀ ,	TBI ₅ (R ₁₄₈₀ ,	0.63/	0.61/	0.68/	0.66/	0.65/
	R ₁₄₈₀ ,R ₁₇₆₀)	R ₆₄₁ ,R ₅₈₇)	R ₁₇₇₀ ,R ₆₀₈)	R ₁₈₆₀ ,R ₉₅₉)	R ₁₇₇₀ ,R ₇₂₂)	0.58	0.36	0.67	0.64	0.65
Con_10/10	TBI ₁ (R ₁₈₈₀ ,	TBI ₂ (R ₄₁₀ ,	TBI ₃ (R ₁₈₈₀ ,	TBI ₄ (R ₁₄₂₀ ,	TBI ₅ (R ₁₈₈₀ ,	0.62/	0.59/	0.67/	0.67/	0.64/
	R ₂₁₇₀ ,R ₁₉₁₀)	R ₆₂₀ ,R ₅₇₀)	R ₂₁₇₀ ,R ₇₅₀)	R ₂₃₈₀ ,R ₉₃₀)	R ₂₁₇₀ ,R ₆₂₀)	0.58	0.34	0.68	0.56	0.65
Con_20/20	TBI ₁ (R ₁₈₈₀ ,	TBI ₂ (R ₄₂₀ ,	TBI ₃ (R ₁₈₈₀ ,	TBI ₄ (R ₁₄₂₀ ,	TBI ₅ (R ₁₈₈₀ ,	0.62/	0.59/	0.66/	0.66/	0.64/
	R ₂₁₈₀ ,R ₂₃₂₀)	R ₆₀₀ ,R ₅₈₀)	R ₂₁₈₀ ,R ₆₂₀)	R ₂₃₈₀ ,R ₉₄₀)	R ₆₈₀ ,R ₂₁₈₀)	0.63	0.42	0.69	0.64	0.59
Con_40/40	TBI ₁ (R ₁₈₈₀ ,	TBI ₂ (R ₄₄₀ ,	TBI ₃ (R ₁₄₀₀ ,	TBI ₄ (R ₁₄₀₀ ,	TBI ₅ (R ₁₄₀₀ ,	0.51/	0.47/	0.56/	0.58/	0.53/
	R ₂₁₆₀ ,R ₁₈₄₀)	R ₆₀₀ ,R ₆₆₀)	R ₂₁₆₀ ,R ₉₆₀)	R ₂₂₄₀ ,R ₁₀₈₀)	R ₉₆₀ ,R ₂₁₆₀)	0.56	0.34	0.54	0.54	0.54
Con_60/60	TBI ₁ (R ₂₂₇₀ ,	TBI ₂ (R ₄₇₀ ,	TBI ₃ (R ₉₅₀ ,	TBI ₄ (R ₁₄₉₀ ,	TBI ₅ (R ₈₉₀ ,	0.34/	0.41/	0.49/	0.48/	0.47/
	R ₁₉₁₀ ,R ₆₅₀)	R ₆₅₀ ,R ₅₉₀)	R ₂₂₁₀ ,R ₁₄₉₀)	R ₂₃₃₀ ,R ₁₁₃₀)	R ₂₂₁₀ ,R ₁₄₉₀)	0.3	0.15	0.42	0.43	0.41
Con_100/100	TBI ₁ (R ₁₄₀₀ ,	TBI ₂ (R ₁₄₀₀ ,	TBI ₃ (R ₁₈₀₀ ,	TBI ₄ (R ₁₄₀₀ ,	TBI ₅ (R ₁₅₀₀ ,	0.32/	0.35/	0.41/	0.53/	0.38/
	R ₇₀₀ ,R ₂₃₀₀)	R ₇₀₀ ,R ₆₀₀)	R ₁₉₀₀ ,R ₂₂₀₀)	R ₂₃₀₀ ,R ₈₀₀)	R ₆₀₀ ,R ₅₀₀)	0.1	0.13	0.2	0.27	0.06

表 3 从最佳波段组合算法得出的各个频谱配置中EC的最佳频谱参数

Table 3 Best frequency-spectrum parameters of EC in each spectral configuration derived from optimal band combination algorithm

Configuration	Spectral variable					Maximum determination coefficient				
						(R^2) calibration/validation				
ASD_1/1	TBI ₁ (R ₂₂₉₂ ,	TBI ₂ (R ₈₂₆ ,	TBI ₃ (R ₈₀₃ ,	TBI ₄ (R ₂₃₆₄ ,	TBI ₅ (R ₂₁₇₉ ,	0.33/	0.58/	0.42/	0.37/	0.37/
	R ₂₃₅₅ ,R ₆₃₄)	R ₂₀₆₉ ,R ₆₉₈)	R ₁₀₉₉ ,R ₂₁₆₉)	R ₂₃₁₆ ,R ₂₂₄₁)	R ₂₁₃₄ ,R ₂₁₀₂)	0.27	0.5	0.54	0.26	0.34
Con_3/10	TBI ₁ (R ₁₉₆₀ ,	TBI ₂ (R ₇₅₀ ,	TBI ₃ (R ₁₉₂₀ ,	TBI ₄ (R ₁₉₆₀ ,	TBI ₅ (R ₇₂₈ ,	0.29/	0.6/	0.47/	0.3/	0.32/
	R ₁₈₉₀ ,R ₁₉₇₀)	R ₁₄₈₀ ,R ₁₃₈₀)	R ₉₅₀ ,R ₈₉₃)	R ₁₉₁₀ ,R ₁₈₅₀)	R ₁₇₃₀ ,R ₄₁₃)	0.37	0.61	0.51	0.39	0.16
Con_10/10	TBI ₁ (R ₂₂₉₀ ,	TBI ₂ (R ₁₁₁₀ ,	TBI ₃ (R ₂₁₈₀ ,	TBI ₄ (R ₂₃₆₀ ,	TBI ₅ (R ₂₁₈₀ ,	0.33/	0.58/	0.44/	0.36/	0.36/
	R ₂₃₆₀ ,R ₆₉₀)	R ₂₃₁₀ ,R ₆₂₀)	R ₁₁₅₀ ,R ₈₉₀)	R ₂₃₂₀ ,R ₂₂₂₀)	R ₂₁₃₀ ,R ₂₁₀₀)	0.26	0.5	0.55	0.25	0.37
Con_20/20	TBI ₁ (R ₂₂₈₀ ,	TBI ₂ (R ₉₄₀ ,	TBI ₃ (R ₂₁₂₀ ,	TBI ₄ (R ₂₃₆₀ ,	TBI ₅ (R ₂₁₈₀ ,	0.31/	0.57/	0.42/	0.34/	0.35/
	R ₂₃₆₀ ,R ₆₆₀)	R ₂₁₂₀ ,R ₅₈₀)	R ₁₀₈₀ ,R ₇₆₀)	R ₂₃₂₀ ,R ₂₂₂₀)	R ₂₁₄₀ ,R ₂₁₀₀)	0.26	0.5	0.51	0.26	0.3
Con_40/40	TBI ₁ (R ₂₂₈₀ ,	TBI ₂ (R ₇₆₀ ,	TBI ₃ (R ₂₁₂₀ ,	TBI ₄ (R ₂₃₆₀ ,	TBI ₅ (R ₂₂₈₀ ,	0.3/	0.57/	0.4/	0.29/	0.26/
	R ₂₃₆₀ ,R ₆₄₀)	R ₂₂₈₀ ,R ₆₄₀)	R ₁₀₈₀ ,R ₇₆₀)	R ₂₃₂₀ ,R ₂₂₈₀)	R ₂₁₆₀ ,R ₁₈₀₀)	0.31	0.51	0.45	0.28	0.22
Con_60/60	TBI ₁ (R ₂₂₇₀ ,	TBI ₂ (R ₁₄₃₀ ,	TBI ₃ (R ₁₃₇₀ ,	TBI ₄ (R ₂₂₇₀ ,	TBI ₅ (R ₁₀₁₀ ,	0.11/	0.56/	0.38/	0.11/	0.22/
	R ₂₂₁₀ ,R ₂₀₃₀)	R ₂₂₁₀ ,R ₂₂₁₀)	R ₁₄₉₀ ,R ₁₀₇₀)	R ₂₁₅₀ ,R ₁₉₁₀)	R ₂₂₁₀ ,R ₄₇₀)	0.15	0.62	0.41	0.28	0.13
Con_100/100	TBI ₁ (R ₁₈₀₀ ,	TBI ₂ (R ₁₂₀₀ ,	TBI ₃ (R ₁₃₀₀ ,	TBI ₄ (R ₁₃₀₀ ,	TBI ₅ (R ₁₉₀₀ ,	0.21/	0.54/	0.38/	0.2/	0.25/
	R ₂₂₀₀ ,R ₆₀₀)	R ₂₃₀₀ ,R ₁₉₀₀)	R ₂₂₀₀ ,R ₁₄₀₀)	R ₂₃₀₀ ,R ₉₀₀)	R ₂₁₀₀ ,R ₁₈₀₀)	0.12	0.56	0.21	0.04	0.11

表 4 通过组合不同光谱配置中的最佳光谱参数/全光谱参数获得的 SOM 预测结果

Table 4 Predicted results of SOM obtained by combining best spectral parameters/all-optical spectral parameters in different spectral configurations

Configuration	LVs	R_c^2	RMSE _{cv} / %	RPIQ _c	R_v^2	RMSE _p / %	RPIQ _v	Time / s
ASD_1/1	5/12	0.76/0.66	0.349/0.416	2.78/2.33	0.76/0.69	0.350/0.382	2.78/2.55	>24 h/174.806
Con_3/10	5/12	0.77/0.66	0.339/0.415	2.87/2.35	0.76/0.69	0.332/0.381	2.94/2.56	553.184/138.728
Con_10/10	5/14	0.76/0.66	0.348/0.417	2.79/2.33	0.77/0.69	0.359/0.380	2.71/2.56	117.189/109.263
Con_20/20	5/14	0.74/0.67	0.357/0.410	2.72/2.37	0.77/0.68	0.325/0.392	3.00/2.48	15.941/55.348
Con_40/40	5/15	0.70/0.65	0.385/0.427	2.52/2.27	0.72/0.68	0.693/0.390	1.41/2.50	2.973/5.535
Con_60/60	5/12	0.63/0.64	0.431/0.427	2.26/2.28	0.6/0.64	0.464/0.417	2.11/2.34	1.353/2.222
Con_100/100	5/14	0.60/0.59	0.452/0.458	2.14/2.12	0.35/0.54	0.567/0.463	1.72/2.11	0.674/0.736

表 5 通过组合不同光谱配置中的最佳光谱参数/全光谱参数获得的 EC 预测结果

Table 5 Prediction results of EC obtained by combining best spectral parameters/all-optical spectral parameters in different spectral configurations

Configuration	LVs	R_c^2	RMSE _{cv} / (dS·m ⁻¹)	RPIQ _c	R_v^2	RMSE _p / (dS·m ⁻¹)	RPIQ _v	Time / s
ASD_1/1	5/4	0.70/0.52	7.12/9.05	2.12/1.67	0.68/0.51	7.06/8.79	2.15/1.73	>24 h/172.844
Con_3/10	5/4	0.71/0.51	7.27/9.42	2.14/1.65	0.70/0.51	7.04/9.17	2.15/1.65	554.677/128.822
Con_10/10	5/4	0.70/0.52	7.17/9.08	2.11/1.67	0.69/0.49	6.88/8.96	2.21/1.70	120.039/105.332
Con_20/20	5/4	0.72/0.52	6.94/9.09	2.18/1.66	0.69/0.49	6.92/8.95	2.19/1.70	15.864/57.023
Con_40/40	5/4	0.67/0.52	7.46/9.05	2.03/1.67	0.65/0.48	7.33/9.10	2.07/1.67	2.431/5.236
Con_60/60	5/4	0.67/0.52	7.74/9.32	2.02/1.68	0.66/0.49	7.87/8.96	1.93/1.70	1.062/2.169
Con_100/100	5/4	0.66/0.50	7.61/9.27	1.99/1.63	0.66/0.49	7.21/8.98	2.11/1.69	0.566/0.739

(SOM&EC)相对于实测土壤特性的散点图。与全波段光谱建模相比,垂直线较长说明最佳波段组合算法的预测差异受到了光谱本质的影响,如

图 6(a)、(b)所示。但是从全波段光谱数据中建立的 PLSR 模型分别低估了 2 个百分点和 25 dS·m⁻¹ 以上的 SOM 和 EC 值,如图 6(c)、(d)所示。因此最

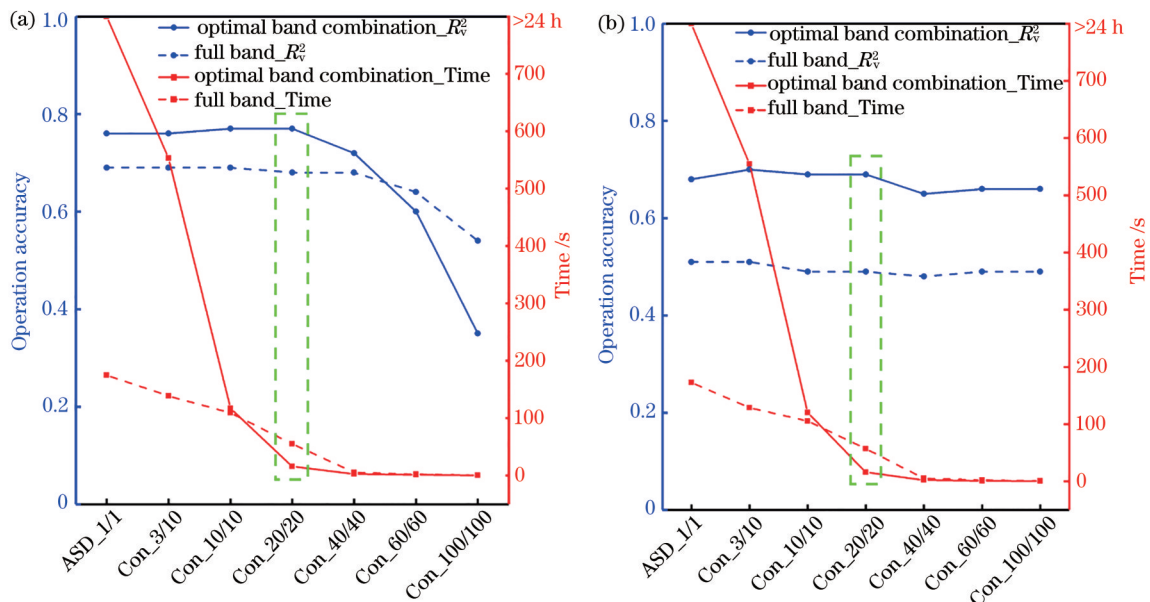


图 5 不同波段组合下 SOM 和 EC 的运算精度以及运算时间对比。(a) SOM; (b) EC

Fig. 5 Comparison of operation accuracy and operation time between SOM and EC under different band combinations.

(a) SOM; (b) EC

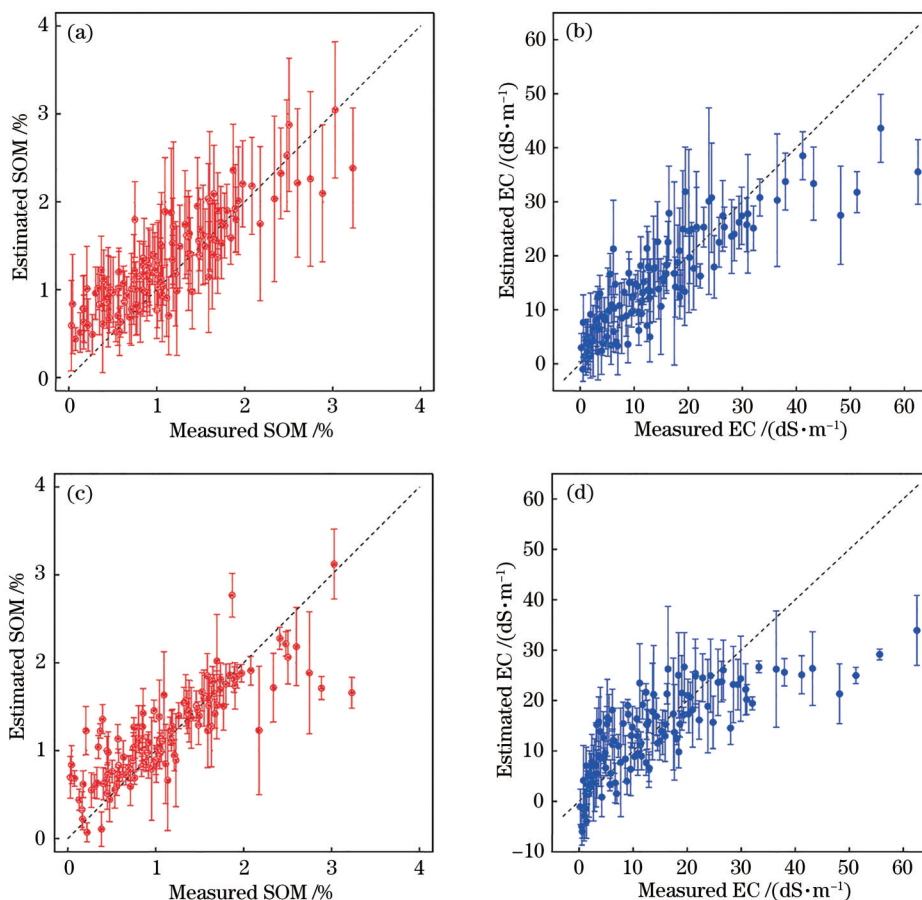


图 6 不同 PLSR 模型模拟的土壤特性与实测值的散点图。(a)(b)最佳光谱;(c)(d)全波段光谱
 Fig. 6 Scatter plots of soil characteristics and measured values simulated by different PLSR models. (a)(b) Optimal spectra;
 (c)(d) full-band spectra

佳波段组合算法改善了这种情况,表明模型的准确性得到了提高。

图 7 为从 PLSR 模型获得的 5 个 TBI 的 VIP 值,

该图表示了不同光谱配置 TBI 的方差及在 SOM 或 EC 光谱估计中预测因子的重要性。从图 7(a)可以看出,SOM 中所有 TBI 的 VIP 值均大于 0.5 且小于

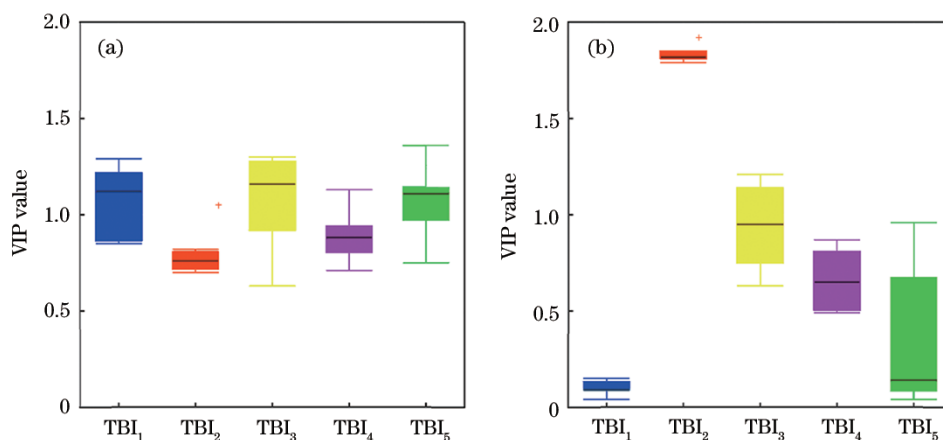


图 7 从不同的 PLSR 模型获得的 VIP 值,以最佳光谱参数作为预测因子,用于估算土壤性质。(a) SOM; (b) EC
 Fig. 7 VIP values obtained from different PLSR models, best spectral parameters used as predictors to estimate soil properties.
 (a) SOM; (b) EC

1.5, 其中 TBI_3 表现最佳, 其次是 TBI_1 , 表现最差的是 TBI_2 (VIP 值小于 1)。从图 7(b) 可以看出, EC 中每个 TBI 的 VIP 值的范围和平均值均发生了变化, 表明与 EC 相关的预测变量在不同光谱配置之间不稳定, 同时 TBI_2 (VIP 值大于 1.5) 是 EC 最重要的预测指标, 其次是 TBI_3 。

4 讨 论

高光谱数据包含大量连续的光谱信息, 可以估算一些土壤特性, 例如 SOM^[1] 和 EC^[19], 但是高光谱数据估测的土壤性质在很大程度上具有空间异质性^[5]。最佳波段组合算法是一种强大的数学工具, 可以有效考虑波段之间的相互作用信息, 将无关波长的影响最小化, 并增强特性含量与光谱特征之间的关系^[20-21]。大多数成果主要研究了两个波段^[22-23], 因此最佳波段组合算法在 Vis-NIR 光谱学中的潜在应用受到的关注很少。Shi^[17] 指出, 350~950 nm 的 Vis-NIR 光谱数据是土壤砷含量的敏感波段, TBI 的预测准确性高于两波段指数。在本研究中, TBI 的准确性(即表 2 中的 TBI_3 和表 3 中的 TBI_1) 大于或等于在 ASD_1/1 下的全波段光谱数据(表 6 和表 7) 的预测性能。此外, 一些学者^[24] 已经提出了四波段指数, 但是对于高光谱数据, 这样的指数会在几何上增加数据量并大幅降低计算效率。因此, 在使用最佳波段组合算法时, 尤其是对于 Vis-NIR 光谱数据, 必须综合考虑效率和准确性。

Zhu 等^[25-26] 发现, 仅包含两个或三个波段的光谱指数可能无法提供令人满意的高光谱信息, 而多元高光谱指数的组合可能是更好的选择。Hong 等^[20] 还指出对于机载高光谱数据, 结合 5 个光谱指数的最佳光谱参数显然比使用单个光谱指数对土壤有机碳(SOC)的检测更敏感。另外, 与全波段光谱数据相比, 该技术提高了预测精度。这些结果大体上与所提方法的结论一致。也就是说, 对于所研究的土壤特性, 无论光谱结构如何, 综合使用 5 个 TBI (表 4 和表 5) 的预测精度均明显优于使用单个指标(表 2 和表 3) 的预测精度。此外, 与 Con_60/60 之前的全光谱数据(表 6 和表 7) 相比, 所提方法提取了有效的预测变量并提高了预测准确性。用最佳光谱组合算法得出的 SOM 和 EC 的最佳估计结果分别出现在 Con_20/20 和 Con_10/10 下(表 4 和表 5)。通常, 从最佳光谱参数中选择的波长大多数与已知吸收特征的位置大致相等。对于 SOM, 选择的波长

分别对应针铁矿(420, 600, 620 nm)、赤铁矿(680 nm)、有机物(580, 2180 nm)、水(940, 1420, 1880 nm)和碳水化合物(2320, 2380 nm), 这些波长对 SOM 具有一定的解释性。对于 EC, 选择的波长分别对应针铁矿(620 nm)、赤铁矿(690, 890, 1050 nm)、水(1140 nm)、有机物(2100, 2180 nm) (相关偏差为 ± 10 nm)、粘土矿物[高岭土(2150 nm)、蒙脱石(2220, 2290 nm)、伊利石(2340, 2360 nm)]^[27-28]。根据美国农业部(USDA) 的土壤盐分分类标准, 本实验组超过 70% 的样品处于低盐碱度($0\sim 2$ dS \cdot m⁻¹) 至中度盐碱度($8\sim 16$ dS \cdot m⁻¹) 的范围内。因此, EC 较低这一发现可能与粘土或矿物有关; 而不同最佳光谱参数的波段组合不一致是光谱配置、指数公式和土壤性质的差异造成的(表 2 和表 3)。

从 ASD_1/1 到 Con_20/20 提取的最佳光谱参数的准确性未见明显变化(表 4 和表 5)。同样, Tian 等^[29] 的发现揭示了三个新提出的光谱指数在 1~25 nm 的带宽内保持几乎恒定的值。但是, 对于 ASD_1/1, 每个获取的最佳光谱参数都需要遍历 2001 个波段的立方组合, 并且计算复杂度是 Con_10/10 的 1000 倍以上和 Con_20/20 的 8000 倍以上。因此, 适当地降低光谱配置而不显著影响最佳频带参数的灵敏度, 可以增加 Vis-NIR 高光谱数据方法的有效性和便利性。并不总是在初始(1 nm) 和最粗糙(100 nm) 光谱配置下, 提取到的最佳光谱参数为最佳和最差精度, 光谱反射率随基于高斯响应函数的光谱分辨率的变化而变化, 如(1) 式所示; 因此高光谱分辨率的精度不一定比粗分辨率的精度高。

5 结 论

通过便携式光谱辐射计测量的光谱和用于评估 SOM 和 EC 的最佳波段组合算法 PLSR 模型评价了 7 种光谱配置的预测能力。通常, 在 60 nm 的光谱分辨率之前, 最佳波段组合算法得出的预测精度要优于全光谱数据得出的预测精度。最佳波段组合算法的预测性能在 3~20 nm 的分辨率下稳定且准确, 并且光谱带的数量从 2001 年减小至 19。SOM 和 EC 分别在 20 nm 和 10 nm 的光谱分辨率下获得最佳预测性能(SOM, R^2 等于 0.77, RMSE_p 等于 0.325%, RPIQ_v 等于 3; EC, R^2 等于 0.70, RMSE_p 等于 6.88 dS \cdot m⁻¹, RPIQ_v 等于 2.21)。这些发现可

能有助于减小 Vis-NIR 光谱数据的数量并开发更准确有效的预测模型,尤其是在样本数量较大的情况下,例如土壤光谱库。此外,所提方法可以应用于其他高光谱成像传感器,以提高数字土壤制图效率。

参 考 文 献

- [1] Nocita M, Stevens A, Toth G, et al. Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach[J]. *Soil Biology and Biochemistry*, 2014, 68: 337-347.
- [2] Viscarra Rossel R A, Behrens T, Ben-Dor E, et al. A global spectral library to characterize the world's soil[J]. *Earth-Science Reviews*, 2016, 155: 198-230.
- [3] Wang J, Ding J, Abulimiti A, et al. Quantitative estimation of soil salinity by means of different modeling methods and visible-near infrared (VIS-NIR) spectroscopy, Ebinur Lake Wetland, Northwest China [J]. *PeerJ*, 2018, 6: e4703.
- [4] Hong Y S, Shen R L, Cheng H, et al. Cadmium concentration estimation in peri-urban agricultural soils: using reflectance spectroscopy, soil auxiliary information, or a combination of both?[J]. *Geoderma*, 2019, 354: 113875.
- [5] 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.
- [6] Wang W, Yao X, Yao X F, et al. Estimating leaf nitrogen concentration with three-band vegetation indices in rice and wheat[J]. *Field Crops Research*, 2012, 129: 90-98.
- [7] Li F, Mistele B, Hu Y C, et al. Optimising three-band spectral indices to assess aerial N concentration, N uptake and aboveground biomass of winter wheat remotely in China and Germany[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 92: 112-123.
- [8] Kong W P, Huang W J, Casa R, et al. Off-nadir hyperspectral sensing for estimation of vertical profile of leaf chlorophyll content within wheat canopies[J]. *Sensors*, 2017, 17(12): E2711.
- [9] Peng X T, Shi T Z, Song A H, et al. Estimating soil organic carbon using VIS/NIR spectroscopy with SVMR and SPA methods[J]. *Remote Sensing*, 2014, 6(4): 2699-2717.
- [10] Yang H, Kuang B, Mouazen A M. Quantitative analysis of soil nitrogen and carbon at a farm scale using visible and near infrared spectroscopy coupled with wavelength reduction[J]. *European Journal of Soil Science*, 2012, 63(3): 410-420.
- [11] Adeline K R M, Gomez C, Gorretta N, et al. Predictive ability of soil properties to spectral degradation from laboratory Vis-NIR spectroscopy data[J]. *Geoderma*, 2017, 288: 143-153.
- [12] Zhang Z P, Ding J L, Wang J Z, et al. Quantitative estimation of soil organic matter content using three-dimensional spectral index: a case study of the Ebinur Lake basin in Xinjiang[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(5): 1514-1522.
张子鹏, 丁建丽, 王敬哲, 等. 利用三维光谱指数定量估算土壤有机质含量: 以新疆艾比湖流域为例[J]. *光谱学与光谱分析*, 2020, 40(5): 1514-1522.
- [13] Wang J Z, Ding J L, Yu D L, et al. Machine learning-based detection of soil salinity in an arid desert region, Northwest China: a comparison between Landsat-8 OLI and Sentinel-2 MSI[J]. *Science of the Total Environment*, 2020, 707: 136092.
- [14] Tian Y C, Yao X, Yang J, et al. Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground- and space-based hyperspectral reflectance[J]. *Field Crops Research*, 2011, 120(2): 299-310.
- [15] Zheng Q, Huang W J, Cui X M, et al. Identification of wheat yellow rust using optimal three-band spectral indices in different growth stages[J]. *Sensors*, 2018, 19(1): E35.
- [16] 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.
- [17] Shi Z, Wang Q L, Peng J, et al. Development of a national VNIR soil-spectral library for soil classification and prediction of organic matter concentrations[J]. *Science China Earth Sciences*, 2014, 57(7): 1671-1680.
- [18] Wilding L P. Spatial variability: its documentation, accomodation and implication to soil surveys[J]. *Spatial Variations*, 1985, 1(1): 166-187.
- [19] Nawar S, Buddenbaum H, Hill J, et al. Modeling

- and mapping of soil salinity with reflectance spectroscopy and landsat data using two quantitative methods (PLSR and MARS)[J]. *Remote Sensing*, 2014, 6(11): 10813-10834.
- [20] 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.
- [21] Jin X L, Song K S, Du J, et al. Comparison of different satellite bands and vegetation indices for estimation of soil organic matter based on simulated spectral configuration[J]. *Agricultural and Forest Meteorology*, 2017, 244/245: 57-71.
- [22] Abulaiti Y, Sawut M, Maimaitiaili B, et al. A possible fractional order derivative and optimized spectral indices for assessing total nitrogen content in cotton[J]. *Computers and Electronics in Agriculture*, 2020, 171: 105275.
- [23] Das B, Manohara K K, Mahajan G R, et al. Spectroscopy based novel spectral indices, PCA- and PLSR-coupled machine learning models for salinity stress phenotyping of rice[J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2020, 229: 117983.
- [24] Thenkabail P, Smith R B, de Pauw E. Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization[J]. *Photogrammetric Engineering and Remote Sensing*, 2002, 68(6): 607-621.
- [25] Zhu C M, Zhang Z P, Wang H W, et al. Assessing soil organic matter content in a coal mining area through spectral variables of different numbers of dimensions[J]. *Sensors*, 2020, 20(6): 1795.
- [26] Shi T Z, Liu H Z, Chen Y Y, et al. Estimation of arsenic in agricultural soils using hyperspectral vegetation indices of rice[J]. *Journal of Hazardous Materials*, 2016, 308: 243-252.
- [27] Dematté J A M, Horák-Terra I, Beirigo R M, et al. Genesis and properties of wetland soils by VIS-NIR-SWIR as a technique for environmental monitoring [J]. *Journal of Environmental Management*, 2017, 197: 50-62.
- [28] Knadel M, Viscarra Rossel R A, Deng F, et al. Visible-near infrared spectra as a proxy for topsoil texture and glacial boundaries[J]. *Soil Science Society of America Journal*, 2013, 77(2): 568-579.
- [29] Tian Y C, Yao X, Yang J, et al. Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground- and space-based hyperspectral reflectance[J]. *Field Crops Research*, 2011, 120(2): 299-310.