

基于成像光谱技术的土壤剖面发生层划分初探

郑光辉^{1,2}, 焦彩霞², 上官晨曦², 吴文乾², 刘 祎³, 洪长桥^{2,4}

1. 南京信息工程大学气象灾害预测与评估协同创新中心, 江苏 南京 210044
2. 南京信息工程大学地理科学学院, 江苏 南京 210044
3. 南京信息工程大学遥感与测绘工程学院, 江苏 南京 210044
4. 南京大学地理与海洋科学学院, 江苏 南京 210023

摘 要 土壤剖面是土壤发生学研究的核心, 但过去几十年以来研究土壤剖面的技术并没有发生质的变化。成像光谱技术可以提供高空间、高光谱分辨率的土壤剖面数据, 能够弥补反射光谱技术采样深度间隔较大的不足, 用于定量研究土壤属性连续深度变化。以室内采集的土壤剖面成像光谱数据为研究对象, 采用支持向量机方法进行光谱数据主成分分类, 探讨成像光谱数据用于剖面发生层划分的可行性并分析影响因素。研究中首先定性分析各发生层平均光谱曲线形态特征, 然后通过分析剖面光谱数据主成分深度变化特征及其散点分布情况, 探讨其用于剖面发生层划分的可行性; 最后进行 1 000 次随机划分数据集并建模、预测以减小误差, 定量证明成像光谱数据用于土壤发生层划分的可行性, 并通过样本分类错误频率来分析影响分类精度的因素。研究结果表明, 受成土过程影响剖面内各发生层平均光谱曲线特征存在差异。成像光谱数据的主成分可以定量呈现土壤剖面深度方向上属性的连续变化及样本散点分布的集聚特征, 能较好反映发生层之间的差异性, 可以用于发生层划分。建模预测结果表明发生层的预测精度平均值达到 93.08%。同时发现, 光谱主成分分布相似区域的样本及位于发生层过渡区域的样本分类错误率较高。该研究为利用成像光谱技术进行土壤剖面发生层划分提供了理论依据, 为下一步进行剖面发生层制图奠定了技术基础。

关键词 土壤剖面; 发生层; 成像光谱; 支持向量机(SVM)

中图分类号: TP79; S151.9 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)03-0882-04

引 言

土壤剖面记录了土壤发生过程与历史, 是土壤时间序列及土壤发生学研究的核心。剖面内由成土作用形成的大致呈水平状态、具有独特理化及矿物特征的连续层次称为发生层。发生层的类型和数量是研究土壤发生过程及进行土壤分类的基础, 但在过去几十年中, 研究土壤剖面的技术和工具并没有发生质的变化^[1], 限制了土壤学研究与应用快速发展, 亟须发展快速测定、分析土壤剖面属性的新技术^[2-3]。近年来部分学者将反射光谱技术应用于土壤剖面研究, 多以发生层为研究对象^[4]或采用一定深度间隔采样方法^[5-6], 难以量化垂直方向土壤属性的连续变异。成像光谱技术是近几年发展起来的新技术^[7-8], 它为每个像元采集上百个连续窄波段反射率形成平滑的光谱曲线, 可以提供高空间、高光谱分

辨率的土壤剖面数据, 能够弥补反射光谱技术采样深度间隔较大的不足, 对于获取土壤属性连续深度变化特征、划分与识别发生层、进行土壤分类等都具有重要实践意义^[9-10]。本研究利用土壤剖面成像光谱的行平均光谱数据提取的主成分, 采用支持向量机(support vector machine, SVM)方法进行 1 000 次随机划分数据集、建模、预测计算过程, 探讨成像光谱数据用于剖面发生层划分的可行性并分析影响因素。

1 实验部分

1.1 土壤剖面样本采集与剖面发生层划分

土壤剖面采样地点为青海省海北藏族自治州门源回族自治县风匣口一无名滩地, 位于东经 101°18'1", 北纬 37°37'16", 海拔为 3 195 m, 属于高原亚寒带亚湿润气候, 年平均气温为 0.5 °C, 年平均降水量为 520 mm。采样地点的地形

收稿日期: 2018-01-09, 修订日期: 2018-05-22

基金项目: 国家自然科学基金项目(41201215, 41877004), 江苏省大学生创新创业训练计划项目(201610300070Y), 江苏省研究生科研与实践创新计划项目(KYCX17_0892)资助

作者简介: 郑光辉, 1977 年生, 南京信息工程大学地理科学学院副教授 e-mail: zgh@nuist.edu.cn

为倾斜的冲积湖积扇平原,母质为黄土状沉积物,土壤类型为钙积草毡寒冻锥形土。采集 60 cm 深的土壤剖面,在室内自然风干后用于室内成像光谱测定。剖面发生层分为 4 层:0~5.2 cm 为 Oo 层,5.2~18.8 cm 为 A 层,18.8~50.4 cm 为 ABk 层,50.4~60 cm 为 Bw 层(图 1)。



图 1 样品剖面图

Fig. 1 Picture of sample profile

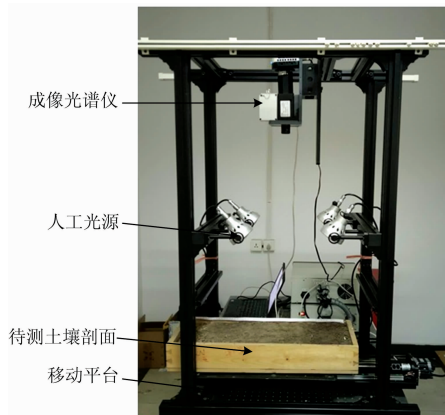


图 2 成像光谱仪器设置图

Fig. 2 Setup of imaging spectroscopy

1.2 土壤剖面成像光谱数据测定与预处理

自然风干后样品剖面放置于移动平台之上,在暗室内采集剖面成像光谱数据。仪器为加拿大 Lumenera 公司生产的 INFINITY V10E 成像光谱仪,光谱范围为 383.53~1 050.26 nm,波段数为 520 个,人工光源为 4 盏卤素灯。采集到的成像光谱数据经剔除四周木框及皮尺像元后为一幅 1 800 行×400 列的剖面图像。

在 ENVI 软件中使用监督分类方法剔除植物根系及阴影像元,将剩余像元取行平均值作为该深度的有效光谱数据,共得到 1 800 行光谱数据;每 5 行取平均值得到 360 个对应

深度平均光谱,即每条光谱曲线代表剖面深度方向约 0.17 cm 带状土样的光谱数据。进行 Savitzky-Golay 平滑处理(2 次多项式,11 点),去除两端噪声较大波段数据,保留 389~965 nm 共 452 个波段数据,然后提取光谱数据主成分用于建模和预测。

1.3 分类方法

支持向量机是由 Vapnik 于 1995 年首先提出的,它在解决小样本、非线性及高维模式识别中表现出许多特有的优势。台湾大学林智仁(Lin Chih-Jen)教授开发设计的 Libsvm 是一个简单、易于使用的 SVM 模式识别与回归的软件包,可以解决利用 SVM 方法进行分类的问题。北京师范大学的李洋在 Libsvm 基础上开发了 libsvm-mat 加强工具箱,增加了数据归一化、主成分(PCA)降维、参数寻优等功能。

1.4 样本集划分及建模预测

随机提取 50% 的样本(180 个)作为预测集,剩余样本(180 个)为建模集。在 Matlab(R2013a)平台下调用 Libsvm 工具箱,利用 Libsvm-mat 加强工具箱的格网参数寻优算法得到最佳参数 c 和 g ,以得到最优建模和预测结果。完成 1 000 次随机划分数数据集、建模及预测过程,以尽量减小随机划分数数据集可能产生的误差,统计分析 1 000 次建模及预测结果以探讨影响分类精度的因素。

2 结果与讨论

2.1 土壤剖面发生层平均光谱曲线特征分析

计算与各发生层深度对应的平均光谱曲线并绘图(图 3)。Oo 层与 A 层光谱曲线在 389~700 nm 波段范围内几乎重合,在 700 nm 后的反射率开始出现分异;这两条曲线在 500~800 nm 波段范围内均呈现下凹的形状特征,说明这两个发生层内土壤有机碳含量较高^[11];ABk 层光谱曲线在 500~850 nm 波段范围内反射率整体升高且光谱曲线较为平直,说明此层内有机碳含量较前两层有所降低;由于土壤碳酸钙含量与光谱反射率呈正相关^[12],该层内碳酸钙含量较高也是其反射率升高的可能原因。Bw 层光谱反射率最高且曲线整体呈上凸的形状,说明该发生层内有机碳含量最低^[12]。

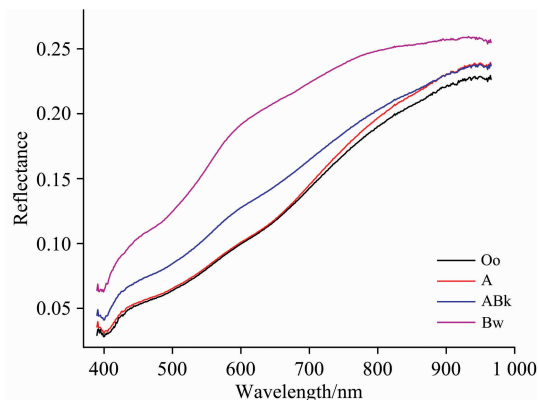


图 3 土壤剖面内各发生层平均光谱曲线图

Fig. 3 Average spectral curves of four horizons in sample profile

2.2 土壤剖面光谱数据主成分特征分析

将平滑及 5 行平均处理后的光谱数据进行主成分分析,提取前 2 个主成分(PC1 解释了 95%, PC2 解释了 5%),分别绘制主成分深度变化图和主成分散点分布图(图 4 和图 5)。可以看出,由于前 2 个主成分包含了原光谱数据的几乎全部信息,较好反映了剖面深度方向上各发生层的属性变化,说明光谱数据主成分可以用于土壤发生层划分。总体来讲,PC1 包含了原光谱绝大部分信息,在剖面深度方向上变异较显著,PC2 在深度方向变异较小。各发生层的主成分分布特征有所不同: Oo 层与 A 层大部分样本的 PC1 为负值而 PC2 为正值, Oo 层 PC1 随深度增加而减小, A 层 PC1 随深度增加而增大,但这两层的 PC1 变化幅度相似,可能会导致分类错误率偏大; ABk 层光谱数据的 PC1 和 PC2 在 0 值附近呈近似对称分布; Bw 层光谱数据的 PC1 均大于 0 且变异较大、PC2 变异较小。

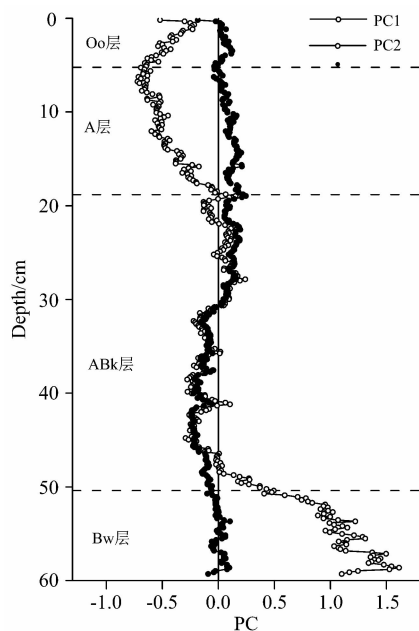


图 4 土壤剖面光谱主成分深度变化图
Fig. 4 Changes of principal components of profile spectroscopy with depth

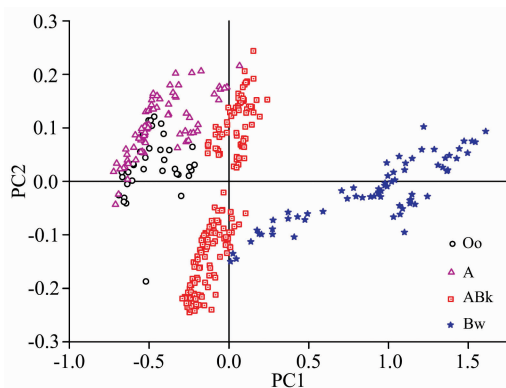


图 5 土壤剖面光谱主成分散点分布图
Fig. 5 Scatter plot of principal components of profile spectroscopy

2.3 土壤剖面光谱数据分类结果与分析

由于每次建模过程中均利用交叉验证方法进行网格参数寻优,在最佳参数 c 和 g 情况下建立模型及实现类别预测,分类结果的精度较高。在 1 000 次运算结果中,建模精度平均值为 95.21%、预测精度平均值为 93.08%(表 1),即 180 个预测样本中平均 167.54 个样本类别预测正确、仅 12.46 个样本类别预测错误。

表 1 建模及预测分类结果统计表
Table 1 Classification results of calibration and prediction

	Min ⁺	Max ⁺	Mean ⁺	Std.	CV
建模精度/%	91.67	98.89	95.21	1.20	1.34
预测精度/%	85.56	97.78	93.08	1.86	1.99

⁺依次为最小值、最大值、平均值; Std. 为标准差; CV 为变异系数,单位: %

为进一步分析哪些样本最容易被错误分类,我们统计了 1 000 次预测结果中各样本被错误分类的频率并绘制了分类错误样本频率深度分布图(图 6)。由图 6 可知,分类错误情况出现最多的是 Oo 层(0~5.2 cm)范围内的样本,主要原因可能是其主成分与 A 层(5.2~18.8 cm)主成分数值分布范围相似,容易出现错误分类的情况。另外,在发生层过渡的位置如深度为 18.8 和 51 cm 附近样本分类错误率也较高,说明在发生层过渡区域由于土壤属性渐变而使其光谱特征相似,易导致分类错误。

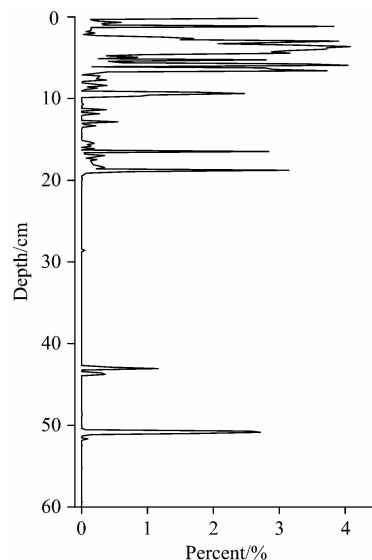


图 6 预测结果中分类错误样本频率深度分布图
Fig. 6 Percentage of wrong classification with depth in prediction results

3 结 论

由土壤剖面成像光谱数据获取固定深度间隔光谱曲线,采用支持向量机方法对光谱数据所提取的 2 个主成分进行建

模及分类预测,探讨成像光谱数据用于剖面发生层划分的可行性及影响因素。结果表明,所提取的光谱数据主成分能够表征土壤属性在垂直方向上的连续变化,较好反映出发生层之间的差异性,可以用于发生层划分。通过网格参数寻优方法使1 000次分类运算结果精度较高,预测精度平均值为

93.08%。分类错误率较高的样本多位于Oo层、A层及发生层过渡区域,说明在该区域范围内土壤属性相近或渐变使其光谱特征相似,易出现分类错误。该研究为利用成像光谱数据进行土壤剖面发生层划分与识别提供了理论依据,为下一步进行剖面发生层制图奠定了技术基础。

References

- [1] Hartemink A E, Minasny B. *Geoderma*, 2014, 230-231: 305.
- [2] Viscarra Rossel R A, Brus D J, Lobsey C, et al. *Geoderma*, 2016, 265: 152.
- [3] Stockmann U, Minasny B, McBratney A B. *Geoderma*, 2014, 216: 48.
- [4] Dematté J A M, Terra F D S. *Geoderma*, 2014, 217-218: 190.
- [5] Fajardo M, McBratney A, Whelan B. *Geoderma*, 2016, 263: 244.
- [6] Weindorf D C, Zhu Y, McDaniel P, et al. *Geoderma*, 2012, 189-190: 268.
- [7] Gomez C, Oltra-Carrió R, Bacha S, et al. *Remote Sens. Environ.*, 2015, 164: 1.
- [8] Franceschini M H D, Dematté J A M, Terra F D S, et al. *Int. J. Appl. Earth Obs. and Geoinformation*, 2015, 38: 358.
- [9] Steffens M, Buddenbaum H. *Geoderma*, 2013, 195-196: 122.
- [10] LI Shuo, WANG Shan-qin, SHI Zhou(李 硕, 汪善勤, 史 舟). *Acta Pedol. Sin.* (土壤学报), 2015, 52(5): 1014.
- [11] Zheng Guanghui, RYU Dongryeol, Jiao Caixia, et al. *Pedosphere*, 2016, 26(1): 130.
- [12] HONG Chang-qiao, ZHENG Guang-hui, CHEN Chang-chun(洪长桥, 郑光辉, 陈昌春). *Acta Pedol. Sin.* (土壤学报), 2016, 53(5): 1120.

Horizon Classification in Soil Profile Using Imaging Spectroscopy

ZHENG Guang-hui^{1,2}, JIAO Cai-xia², SHANGGUAN Chen-xi², WU Wen-qian², LIU Yi³, HONG Chang-qiao^{2,4}

1. Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science & Technology, Nanjing 210044, China
2. School of Geographic Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China
3. School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China
4. School of Geographic and Oceanographic Sciences, Nanjing University, Nanjing 210023, China

Abstract The soil profile is one of the core topics in pedogenesis research, but traditional pedological observations of soil profiles rely on the use of visible light and a toolbox that has not changed in the past decades. The imaging spectroscopy can provide high-resolution spatial and spectral soil profile information, which gives continuous depth functions of soil properties and compensates for the large gap between the sampling depths of reflectance spectroscopy. The objective of this study is to analyze the classification of soil horizon in a profile by investigating the spectral data of imaging spectroscopy collected in the laboratory. The support vector machine (SVM) method was used to classify the spectral data, and the feasibility and influence factors of the imaging spectroscopy for classification were studied. Firstly, the morphological characteristics of the average spectral curves of each horizon in sample profile were analyzed qualitatively. Secondly, depth dynamic and scatter plot of principal components were qualitatively analyzed to explain the feasibility of horizon classification using profile imaging spectroscopy. Finally, One thousand times computations were carried out to reduce the classification error by partitioning random dataset and building prediction model. The prediction results can quantitatively testify the feasibility and the influence factors were discussed by the percentage of wrong classification in prediction results. The results indicated that the four average spectral curves in sample profile differed and reflected the variation in depth derived from pedogenic processes. The principal components of the imaging spectral data showed the continuous change in the depth direction of the soil profile and the grouping feature in scatter plot, which proved that imaging spectroscopy reflected the difference between the genetic horizons and can be used for the horizon classification. The average accuracy of classification prediction reached 93.08%. Moreover, it was found that the sample with similar scattering distribution and locating transition region were classified in wrong classes easily. This study provides a theoretical basis for horizon classification, and proves that imaging spectroscopy is a potential technology for mapping soil profile.

Keywords Soil profile; Horizon; Imaging spectroscopy; Support vector machine (SVM)

(Received Jan. 9, 2018; accepted May 22, 2018)