

赵宁博, 杨佳佳, 赵英俊, 等. 航空高光谱支持下黑土地理化性质反演对地形因子的响应[J]. 江苏农业学报, 2020, 36(6): 1444-1451.
doi:10.3969/j.issn.1000-4440.2020.06.013

航空高光谱支持下黑土地理化性质反演对地形因子的响应

赵宁博¹, 杨佳佳², 赵英俊¹, 秦 凯¹, 杨越超¹, 崔 鑫¹

(1.核工业北京地质研究院, 遥感信息与图像分析技术国家级重点实验室, 北京 100029; 2.中国地质调查局沈阳地质调查中心, 辽宁 沈阳 110000)

摘要: 为了探讨地形因子对黑土地理化性质航空高光谱反演的影响, 本研究以黑龙江省海伦市典型黑土地为例, 开展 12 种地形因子与土壤有机质含量、全氮含量、全磷含量、全钾含量、pH 值、阳离子交换量、全盐量的相关性研究, 并对比不同建模方法在加入地形因子前后的预测精度。航空高光谱反演结果表明, 海拔、坡度、山谷指数、地形粗糙指数、起伏度与多种土壤理化指标显著相关, 在加入地形因子后, 支持向量机法和随机森林法的模型预测精度稳定提升, 而偏最小二乘法的模型预测精度却大幅下降。随机森林法的模型预测效果整体最好, 与纯光谱反射率的反演模型相比, 加入地形因子后全氮含量模型预测 R^2 的提升幅度最大(提升了 0.062), 7 种理化指标预测模型 R^2 的平均提升幅度为 0.036。总之, 地形因子对提升黑土地理化性质的航空高光谱反演精度有积极作用, 且适合用于机器学习方法进行反演。

关键词: 黑土地; 理化性质; 航空高光谱数据; 地形因子; 反演

中图分类号: TP79 **文献标识码:** A **文章编号:** 1000-4440(2020)06-1444-08

Response of physical and chemical properties inversion of black soil to terrain factors supported by airborne hyperspectral data

ZHAO Ning-bo¹, YANG Jia-jia², ZHAO Ying-jun¹, QIN Kai¹, YANG Yue-chao¹, CUI Xin¹

(1. National Key Lab of Remote Sensing Information and Image Analysis Technology, Beijing Research Institute of Uranium Geology, Beijing 100029, China; 2. Shenyang Center of Geological Survey, CGS, Shenyang 110000, China)

Abstract: In order to explore the effects of terrain factors on the airborne hyperspectral inversion of physical and chemical properties of black soil, the typical black soil area of Hailun in Heilongjiang province was selected as the research area. The correlation research of 12 terrain factors with soil organic matter content, total nitrogen content, total phosphorus content, total potassium content, pH, cation exchange capacity and total salt content was carried out, and the prediction accuracy of different modeling methods before and after adding terrain factors was compared. The results showed that there

was a significant correlation between various physical and chemical properties and the terrain factors such as altitude, slope, valley index, terrain roughness index, fluctuation degree. After adding terrain factors, the model prediction accuracy of support vector machine method and the random forest method was improved steadily, but the prediction accuracy of partial least squares method was greatly decreased. In this study, the modeling effect of

收稿日期: 2020-05-09

基金项目: 中国地质调查局东北黑土地 1: 250 000 土地质量地球化学调查项目 (DD20160316); 中国地质调查局兴凯湖平原及松辽平原西部土地质量地球化学调查项目 (DD20190520)

作者简介: 赵宁博 (1985-), 男, 河南郑州人, 硕士, 高级工程师, 主要从事遥感与地球化学综合研究。(E-mail) zhaoningbo1985@126.com

random forest method was the best. Compared with the pure spectral inversion model, the model prediction R^2 of seven indices was generally improved after adding terrain factors, with an average increase of 0.036, while the increase of total nitrogen content model was the largest (0.062). On the whole, the terrain factor plays an active role in improving the accuracy of airborne hyperspectral inversion of the physical and chemical properties of black soil, and it is suitable to use machine learning method for inversion.

Key words: black soil; physical and chemical properties; airborne hyperspectral data; terrain factors; inversion

土壤理化性质的空间分布特征受到成土母质、气候条件、植被类型、地形条件、人为因素等多方面因素的制约,其中地形条件不但支配着水热资源的重新分配,而且影响着土壤生态系统的物质循环过程和强度,对土壤理化性质有着重要影响^[1]。目前,地形因子与土壤理化性质的相关性研究在黄土高原^[2-5]、南方丘陵^[6-8]、华北及长江中下游平原^[9-12]、荒漠绿洲^[13-15]、东北黑土区^[16-19]等地均有开展,研究表明土壤理化参数的空间分布受到高程、坡度、坡向等地形因子不同程度的影响。

遥感技术凭借数据覆盖面广、时效性强等优势,在土壤理化指标提取和空间分布研究方面受到学者的广泛关注^[20-24]。基于地形因子与理化参数空间分布的相关性特征,研究人员尝试利用地形因子与遥感数据相结合的方式开展土壤养分的空间分布预测。例如,张素梅等^[25]利用多种地形因子和美国陆地卫星影像(ETM)提取的植被归一化指数(NDVI)开展有机质、全氮含量的预测研究,结果表明高程、坡度、地面粗糙度、河流动能指数和NDVI对养分多元回归模型贡献较大。刘焕军等^[26]利用Landsat8影像和地形数据开展土壤有机质含量的反演研究,结果表明引入地形因子(高程)后利用回归克里格法预测的有机质含量精度得到提高。综合目前的研究成果,在土壤养分遥感反演中加入地形因子后,预测模型的精度普遍得到了一定程度的提升,但是采用的遥感数据主要是多光谱数据,空间和光谱分辨率受到不同程度的制约,无法充分发挥光谱数据在反演中的作用。此外,该类研究的研究对象主要是土壤有机质及全氮,针对阳离子交换量、pH值等其他土壤理化指标的研究相对较少。

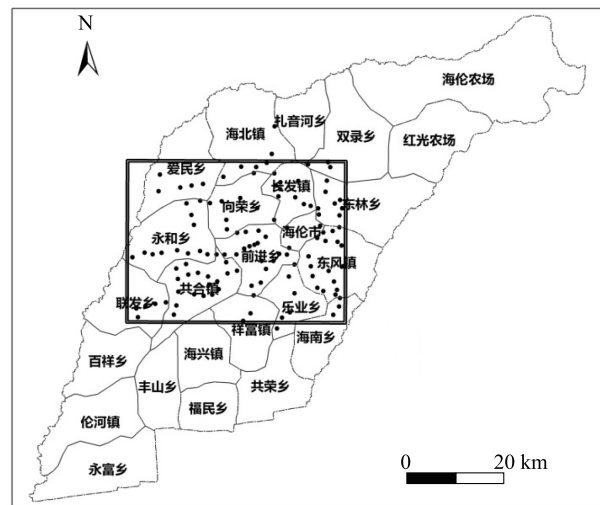
在已有研究的基础上,本研究拟对遥感数据源、研究指标及建模方法等进行改进,以航空高光谱数据作为遥感数据源,通过对比偏最小二乘、支持向量和随机森林等多种建模方法的研究结果,探讨地形因子对土壤有机质含量、全氮含量、全磷含量、全

钾含量、pH值、阳离子交换量、全盐量等理化指标反演效果的影响,评价地形因子在土壤理化性质航空高光谱反演中的作用,以期对土壤理化性质航空高光谱反演精度的提升提供理论依据和技术支持。

1 材料与方法

1.1 研究区概况

图1显示,研究区位于中国黑龙江省海伦市,海伦市地处松嫩平原东北端,小兴安岭西麓,位于东北黑土区的中心区域,地形为丘陵、漫岗,属于中温带大陆性气候,种植农作物主要为水稻、玉米和大豆。研究区面积约为1 500 km²,主要土壤类型为黑土、暗棕壤和草甸土等,平均海拔194 m,海拔落差达246 m,农田范围内海拔落差为136 m,地形特征较为明显。



□ 航空高光谱数据范围; □ 海伦行政区划界限; • 地面采样点

图1 研究区位置及数据分布

Fig.1 Schematic diagram of study area location and data distribution

1.2 航空高光谱数据获取及预处理

航空高光谱数据采用CASI/SASI成像光谱仪获取,选择Y-12飞机为搭载平台。CASI传感器谱

段范围为 380~1 050 nm, 空间分辨率 1.50 m, 波段数 72 个。SASI 传感器谱段范围为 950~2 450 nm, 空间分辨率为 3.75 m, 波段数 100 个。数据时相为 2018 年 5 月, 地表处于裸土期。

数据预处理主要包括辐射校正、几何校正、大气校正和光谱重建。辐射校正和几何校正采用航空成像光谱测量系统配备的软件进行, 其中辐射校正软件为 RCX (Radiometric calibration Xpress) 9.3.5.1, 几何校正软件为 Geocor 3.0。处理步骤包括: 第一, 辐射校正; 第二, 传感器姿态数据处理; 第三, GPS 定位数据处理; 第四, 姿态数据与定位数据时间同步与集成; 第五, 精细的几何校正。在此基础上, 利用 ENVI5.3 软件中的 FLAASH 大气辐射传输模型对航空高光谱遥感数据进行大气校正, 并利用地面黑、白 2 种定标布的光谱测量数据进一步消除大气成分造成的光谱辐射畸变, 最终获得地物光谱反射率, 为后续参数反演提供数据支撑。

对预处理后的影像数据进行地物分类, 提取农田空间分布范围内的影像数据, 分类方法采用 ENVI 软件扩展工具中的随机森林分类法。

表 1 部分地形因子含义描述

Table 1 Description of some terrain factors

| 地形因子 | 含义描述 |
|---------------|---|
| 起伏度 (Relief) | $Relief = elev_{max} - elev_{min}$ 式中, $elev_{max}$ 为邻域中最大高程; $elev_{min}$ 为邻域中最小高程。 |
| 地形粗糙指数 (TRI) | $TRI = \sqrt{\frac{\sum (Z_i - Z_c)^2}{8}}$ 式中, Z_i 为邻域栅格 i 上的高程值; Z_c 为中心栅格的高程值。 |
| 高程百分比指数 (EPI) | $EPI = \frac{100}{N_c} \times \text{count}(Z_i < Z_c)$ 式中, Z_i 为邻域栅格 i 上的高程值; Z_c 为中心栅格的高程值; N_c 为邻域中的栅格数目; count 函数用于统计 $Z_i < Z_c$ 的数量。 |
| 高程起伏比 (ER) | $ER = \frac{elev_{mean} - elev_{min}}{elev_{max} - elev_{min}}$ 式中, $elev_{max}$ 为邻域中最大高程; $elev_{min}$ 为邻域中最小高程; $elev_{mean}$ 为邻域中高程均值。 |
| 表面曲率指数 (Cs) | $Cs = \frac{\sum [(Z_i - \bar{Z}) / dist]}{n}$ 式中, Z_i 为邻域栅格 i 上的高程值; \bar{Z} 为邻域中的高程均值; $dist$ 为 Z_i 到中心点的距离; n 为参加统计的栅格数。 |
| 山谷指数 (VI) | 山谷指数是借鉴数学形态学中开/闭运算的思想对数字高程模型进行变换 (TOPHAT 变换), 对变换后的结果结合规则判断每一栅格相对处于沟谷的情况。 |
| 地形特征指数 (TCI) | $TCI = Cs \times \lg(SCA)$ 式中, Cs 为表面曲率指数; SCA 为单位等高线上汇流面积。TCI 可用于土壤景观单元的划定以及对景观中水/物质流潜力的定量估计。 |

从研究区的海拔和坡度分布情况 (图 2) 可以看出, 研究区海拔分布整体上呈东高西低的趋势, 在坡度空间分布上, 研究区中部和西部分布着明显的坡

1.3 地面采样与分析

在获取航空数据时同步进行地面土壤采样, 综合土壤类型、地形地貌及农作物类型等进行均匀布点, 采集地表 0~20 cm 的土壤样品共 124 个。采用的分析指标包括有机质含量、全氮含量、全磷含量、全钾含量、阳离子交换量、pH 值和全盐量。分析方法如下: 有机质含量采用硫酸-重铬酸钾消解法滴定; 全氮含量采用石墨消解仪消解, 凯氏定氮仪测量; 全磷含量和全钾含量采用熔融制样-X 射线荧光光谱法测量; 阳离子交换量采用乙酸铵交换法测定; pH 值采用 PHS-25 酸度计测定; 全盐量采用电导法测定。

1.4 地形因子提取

基于研究区 30 m 分辨率的数字高程模型 (DEM) 数据, 共提取了 12 种地形因子, 其中海拔、坡度、坡向、剖面曲率、平面曲率在 ArcGIS10.2 软件中计算, 高程起伏比、表面曲率指数、山谷指数、高程百分比指数、地形粗糙指数、起伏度、地形特征指数在 SimDTA 软件^[27]中计算, 后 7 种地形因子的含义见表 1。

度带状高值, 这主要是由农田开垦造成的。

地形因子与土壤理化性质含量间的数理统计分析主要利用 SPSS 22 软件进行。

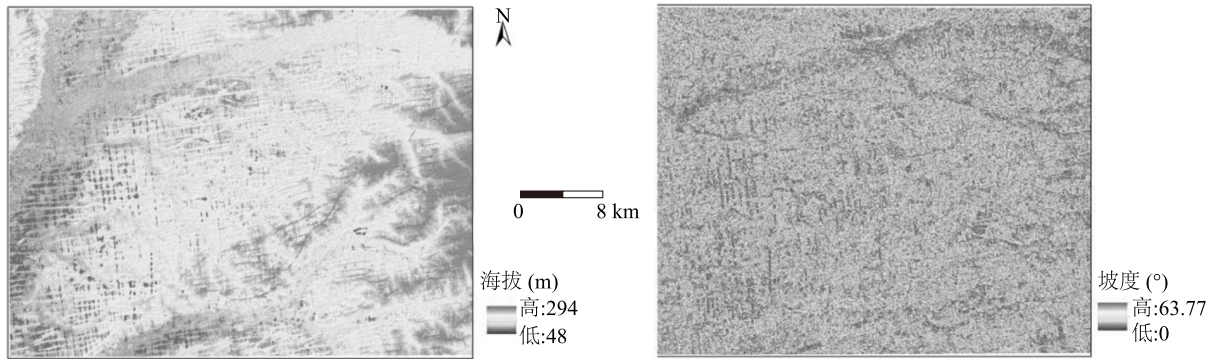


图2 研究区海拔及坡度图

Fig.2 Elevation and slope map of the study area

1.5 土壤理化性质航空高光谱反演

将航空数据与地面采样点进行空间叠加,在航空高光谱影像中以采样点位置为中心,周围 3×3 像元范围的光谱反射率进行平均后作为该点的光谱数据。去除影像边缘的采样点数据后,共整理 121 组样本数据进行反演,按空间分布均匀挑选 80 组样本为建模样本,剩余 41 组样本为验证样本。建模方法采用了线性方法中应用效果较好的偏最小二乘法^[28]以及机器学习中的支持向量机法和随机森林法,模型评价指标包括决定系数(R^2)和均方根误差(RMSE)。

本研究中的反演包括 3 个过程,首次反演的模型自变量只包含光谱反射率,通过对地形和理化性质进行关联分析后,在模型中加入若干指示效果较好的地形因子进行第二次反演,然后在模型中加入全部地形因子进行第三次反演,并对 3 次反演效果

进行对比。

2 结果与分析

2.1 土壤理化性质及反演统计结果

对地面采样点的土壤理化性质进行统计(表 2),根据土壤养分的评价标准^[29],研究区土壤有机质、全氮、全磷、全钾含量整体处于较为丰富的状态;阳离子交换量(CEC)较高,说明研究区土壤有较强的保肥能力;根据全盐量的评价标准^[30],研究区土壤整体上属于非盐渍化状态。

变异程度方面,全盐量的变异系数(CV)最高,全钾含量的 CV 最低。根据变异系数分级标准^[31]:弱变异, $CV < 0.1$; 中等变异, $CV = 0.1 \sim 1.0$; 强变异, $CV > 1.0$ 。全钾含量和 pH 值属于弱变异,其余指标属于中等变异。

表2 地面采样点土壤理化指标描述统计

Table 2 Description and statistics of soil physical and chemical properties at ground sampling points

| 土壤理化指标 | 最小值 | 最大值 | 均值 | 标准差 | 变异系数 |
|-----------------|--------|--------|--------|-------|-------|
| 有机质含量(%) | 2.138 | 7.930 | 4.236 | 0.608 | 0.144 |
| 全氮含量(%) | 0.107 | 0.357 | 0.172 | 0.036 | 0.209 |
| 全磷含量(g/kg) | 0.510 | 1.355 | 0.926 | 0.116 | 0.125 |
| 全钾含量(%) | 2.260 | 2.690 | 2.502 | 0.076 | 0.030 |
| pH 值 | 5.150 | 8.450 | 6.366 | 0.580 | 0.091 |
| 阳离子交换量(cmol/kg) | 21.200 | 35.400 | 26.641 | 2.714 | 0.102 |
| 全盐量(mS/cm) | 0.056 | 0.497 | 0.161 | 0.067 | 0.416 |

样本数为 121 个。

利用地面采样数据和偏最小二乘法建立预测模型,获得各参数的航空高光谱反演数据,表 3 显示,各参数的预测集 R^2 与建模集 R^2 相比均有小幅度的

降低,整体结果较为稳定;各参数间比较,有机质含量模型 R^2 最高,全氮含量、全钾含量、pH 值、CEC 模型的 R^2 较为接近,全盐量模型的 R^2 最低。

表 3 航空高光谱反演模型参数

Table 3 Parameters of airborne hyperspectral inversion model

| 数据集类型 | 模型指标 | 有机质含量 | 全氮含量 | 全磷含量 | 全钾含量 | pH 值 | 阳离子交换量 | 全盐量 |
|-------|-------|---------|---------|---------|---------|---------|---------|---------|
| 建模集 | RMSE | 0.402 0 | 0.029 5 | 0.063 0 | 0.041 0 | 0.339 0 | 1.632 0 | 0.044 0 |
| | R^2 | 0.799 0 | 0.746 0 | 0.671 0 | 0.735 0 | 0.705 0 | 0.736 0 | 0.613 0 |
| 预测集 | RMSE | 0.439 0 | 0.032 1 | 0.071 0 | 0.043 0 | 0.354 0 | 1.868 0 | 0.049 0 |
| | R^2 | 0.707 0 | 0.687 0 | 0.637 0 | 0.683 0 | 0.665 0 | 0.687 0 | 0.546 0 |

R^2 : 决定系数; RMSE: 均方根误差。

进一步统计反演结果的基本参数,为了兼顾空间精度和后续计算的便捷性,在农田空间范围内按照 100 m×100 m 的网度布设统计样点 9 334 个,将各

土壤理化指标的反演结果赋值给样点,结果见表 4。对比表 2 和表 4 的结果可以看到,各土壤理化指标的数据参数较为接近。

表 4 土壤理化性质的反演结果

Table 4 Inversion results of soil physical and chemical properties

| 土壤理化指标 | 最小值 | 最大值 | 均值 | 标准差 | 变异系数 |
|-----------------|--------|--------|--------|-------|-------|
| 有机质含量(%) | 1.834 | 8.635 | 4.178 | 0.644 | 0.177 |
| 全氮含量(%) | 0.103 | 0.373 | 0.174 | 0.035 | 0.144 |
| 全磷含量(g/kg) | 0.463 | 1.534 | 0.890 | 0.121 | 0.130 |
| 全钾含量(%) | 2.234 | 2.713 | 2.513 | 0.073 | 0.029 |
| pH 值 | 5.030 | 8.300 | 5.936 | 0.603 | 0.102 |
| 阳离子交换量(cmol/kg) | 16.573 | 43.580 | 26.824 | 1.781 | 0.066 |
| 全盐量(mS/cm) | 0.049 | 0.493 | 0.163 | 0.064 | 0.393 |

样本数为 9 334 个。

2.2 土壤理化指标与地形因子的相互关系

2.2.1 相关系数 计算土壤样品理化指标与地形因子间的相关系数(表 5),主要特征如下:pH 值、全氮含量与海拔均呈极显著负相关,有机质含量、CEC 与海拔呈显著负相关,全磷含量与海拔呈显著正相关,其中 pH 值与海拔的相关性最高;全磷含量与坡

度呈极显著负相关,全氮含量与坡度呈显著负相关;全磷含量与地形粗糙指数呈极显著负相关,全钾含量与地形粗糙指数呈显著正相关;全磷含量与起伏度呈极显著负相关,有机质含量与起伏度呈显著负相关;其他地形因子与土壤理化指标间没有呈现出显著的相关性。

表 5 土壤样品理化指标与地形因子间的相关性

Table 5 Correlation between physical and chemical properties of ground samples and terrain factors

| 土壤理化指标 | 相关系数 | | | | | | | | | | | |
|--------|-----------|-----------|--------|--------|--------|--------|--------|--------|---------|-----------|-----------|--------|
| | 海拔 | 坡度 | 坡向 | 剖面曲率 | 平面曲率 | 高程起伏比 | 表面曲率指数 | 山谷指数 | 高程百分比指数 | 地形粗糙指数 | 起伏度 | 地形特征指数 |
| 有机质含量 | -0.213 * | -0.092 | 0.052 | 0.100 | 0.029 | 0.063 | 0.030 | -0.033 | 0.050 | -0.157 | -0.228 * | 0.018 |
| 全氮含量 | -0.254 ** | -0.127 * | -0.029 | -0.097 | 0.066 | 0.013 | 0.011 | -0.012 | 0.030 | -0.075 | -0.068 | 0.014 |
| 全磷含量 | 0.100 * | -0.175 ** | 0.007 | -0.068 | 0.048 | -0.015 | -0.009 | 0.034 | -0.053 | -0.141 ** | -0.168 ** | -0.003 |
| 全钾含量 | 0.025 | 0.084 | -0.018 | 0.045 | -0.071 | -0.019 | -0.052 | 0.012 | 0.013 | 0.117 * | 0.098 | -0.071 |
| pH 值 | -0.384 ** | 0.079 | -0.073 | 0.042 | -0.015 | -0.003 | -0.046 | -0.043 | 0.096 | 0.088 | 0.085 | -0.051 |
| 阳离子交换量 | -0.230 * | -0.009 | 0.049 | 0.064 | -0.066 | 0.188 | 0.155 | -0.086 | 0.066 | 0.034 | 0.071 | 0.147 |
| 全盐量 | 0.034 | -0.025 | -0.040 | -0.041 | -0.101 | 0.095 | 0 | 0.043 | -0.116 | -0.081 | -0.080 | 0.015 |

* 表示在 0.05 水平上显著相关; ** 表示在 0.01 水平上显著相关。

利用得到的航空高光谱反演数据计算土壤理化指标与地形因子的相关系数(表6),主要表现如下:除全盐量外,其余理化指标与海拔均呈极显著相关,其中只有全钾含量与海拔呈极显著正相关;全氮含量、全钾含量、pH 值、CEC 与坡度均呈极显著相关;有机质含量和全盐量与地形粗糙指数和起伏度呈显

著相关,其他理化指标与地形粗糙指数和起伏度均呈极显著相关;关于高程起伏比、山谷指数、高程百分比指数,土壤样品理化指标与其均没有呈现出显著的相关性(表5),而利用航空高光谱反演数据计算的这些理化指标中有部分理化指标与其存在显著或极显著相关性。

表 6 航空高光谱反演的土壤理化指标与地形因子的相关性

Table 6 Correlation between soil physical and chemical properties retrieved from airborne hyperspectral data and terrain factors

| 土壤理化指标 | 相关系数 | | | | | | | | | | | |
|--------|-----------|-----------|--------|--------|--------|----------|--------|----------|----------|-----------|-----------|--------|
| | 海拔 | 坡度 | 坡向 | 剖面曲率 | 平面曲率 | 高程起伏比 | 表面曲率指数 | 山谷指数 | 高程百分比指数 | 地形粗糙指数 | 起伏度 | 地形特征指数 |
| 有机质含量 | -0.223 ** | -0.116 * | 0.004 | 0.014 | 0.016 | 0.009 | 0.007 | 0.036 ** | 0.017 | -0.120 * | -0.112 * | 0.013 |
| 全氮含量 | -0.198 ** | -0.137 ** | 0.005 | 0.009 | 0.020 | -0.008 | 0.005 | 0.044 ** | 0.017 | -0.149 ** | -0.145 ** | 0.011 |
| 全磷含量 | -0.118 ** | 0.124 * | -0.016 | -0.003 | -0.010 | 0.020 | -0.008 | 0.042 ** | -0.006 | 0.142 ** | 0.152 ** | -0.008 |
| 全钾含量 | 0.095 ** | 0.150 ** | 0.001 | -0.008 | -0.006 | 0.011 | -0.011 | -0.014 | -0.009 | 0.182 ** | 0.186 ** | -0.018 |
| pH 值 | -0.317 ** | 0.142 ** | 0.019 | 0.013 | 0.011 | 0.035 ** | 0.014 | -0.001 | 0.018 | 0.163 ** | 0.175 ** | 0.011 |
| 阳离子交换量 | -0.392 ** | 0.142 ** | 0.016 | 0.013 | 0.019 | 0.027 * | 0.008 | 0.036 ** | 0.028 ** | 0.163 ** | 0.180 ** | 0.007 |
| 全盐量 | 0.046 | -0.092 | 0.052 | 0.100 | 0.029 | 0.063 | 0.030 | -0.033 | 0.050 | -0.157 * | -0.228 * | 0.018 |

* 表示在 0.05 水平上显著相关; ** 表示在 0.01 水平上显著相关。

2.2.2 不同坡度、坡向类别土壤的理化性质差异
由于坡度和坡向具有比较明确的分类及类别意义,因此对土壤理化性质在不同坡度、坡向类别间的变化规律进行探讨。

坡度分为 5 个等级,分别为:1 级($\leq 2.0^{\circ}$)、2 级($2.1^{\circ} \sim 6.0^{\circ}$)、3 级($6.1^{\circ} \sim 15.0^{\circ}$)、4 级($15.1^{\circ} \sim 25.0^{\circ}$)、5 级($>25.0^{\circ}$)。计算不同坡度级别中土壤理化性质的航空高光谱反演均值,发现各指标在不同坡度级别中绝对值变化幅度均较小。而在相对趋势上,全氮含量和有机质含量与坡度整体呈负相关,全磷含量、全钾含量、pH 值、CEC 与坡度呈正相关,这与表 6 的结果也较为吻合。

坡向分为 4 类,分别是:阴坡($0^{\circ} \sim 45.0^{\circ}$, $315.1^{\circ} \sim 360.0^{\circ}$)、半阴坡($45.1^{\circ} \sim 90.0^{\circ}$, $270.1^{\circ} \sim 315.0^{\circ}$)、半阳坡($90.1^{\circ} \sim 135.0^{\circ}$, $225.1^{\circ} \sim 270.0^{\circ}$)、阳坡($135.1^{\circ} \sim 225.0^{\circ}$)。计算不同坡向类别中土壤理化指标的航空高光谱反演均值,发现各指标对应的坡向组间差异明显小于坡度组间差异,坡向组间变化趋势也没有呈现明显的规律性。

2.2.3 地形因子筛选 挑选与各理化指标显著或极显著相关的地形因子参与反演建模,各指标筛选的地形因子见表 7。整体上,海拔、坡度、山谷指数、地形粗糙指数、起伏度出现的频率较高。

表 7 筛选出的与各类土壤理化指标相关显著的地形因子

Table 7 The selected terrain factors related to various physical and chemical properties

| 土壤理化指标 | 相关显著的地形因子 |
|--------|-------------------------------------|
| 有机质含量 | 海拔、坡度、山谷指数、地形粗糙指数、起伏度 |
| 全氮含量 | 海拔、坡度、山谷指数、地形粗糙指数、起伏度 |
| 全磷含量 | 海拔、坡度、山谷指数、地形粗糙指数、起伏度 |
| 全钾含量 | 海拔、坡度、地形粗糙指数、起伏度 |
| pH 值 | 海拔、坡度、高程起伏比、地形粗糙指数、起伏度 |
| 阳离子交换量 | 海拔、坡度、高程起伏比、山谷指数、高程百分比指数、地形粗糙指数、起伏度 |
| 全盐量 | 地形粗糙指数、起伏度 |

2.3 地形因子对建模的影响

为了更全面地讨论地形因子对反演模型的影响,将模型的自变量分为光谱反射率、光谱反射率+筛选的地形因子、光谱反射率+全部地形因子 3 类。由于模型的预测精度比建模精度更具有实际应用价值,因此以下着重以预测精度为例进行讨论。

利用前期试验应用效果较好的偏最小二乘法进行建模,图 3 显示,加入筛选的地形因子后 7 种理化

指标的模型预测 R^2 均大幅下降,其中 pH 值和 CEC 模型预测 R^2 的降幅最大,pH 值的模型预测 R^2 由 0.665 降为 0.149,CEC 的模型预测 R^2 由 0.687 降为 0.107,加入所有地形因子后,除 CEC 外其余土壤理化指标的模型预测 R^2 进一步下降。通过检查偏最小二乘法计算过程的数据发现,单纯利用光谱反射率建模时模型精度随着主成分数量增加而逐渐提升,达到最高点后随主成分数量增加而有所降低;加入地形因子后模型精度随主成分数量增加而一直降低,认为地形因子与光谱反射率之间的相关程度较低,二者对主成分分析计算造成了干扰,从而影响了偏最小二乘法模型预测的精度。

利用支持向量机法和随机森林法进行对比研究,支持向量机法的核函数采用径向基函数(RBF),随机森林法参数中的决策树数目为5 000,创建分枝所需变量数目为3。图3显示,单纯利用光谱反射

率建模时,支持向量机法对各土壤理化指标的模型预测 R^2 普遍低于偏最小二乘法,加入筛选的地形因子后模型预测 R^2 均稳定提升,平均提升幅度为 0.031,有机质含量的提升幅度最大(提升了 0.060),继续加入全部地形因子后模型预测 R^2 仍有较小幅度提升,但最终预测精度小于单纯利用光谱反射率的偏最小二乘法。

图3显示,仅利用光谱反射率建模时,随机森林法的预测结果中全氮含量和全磷含量的模型预测 R^2 稍低于偏最小二乘法。加入筛选的地形因子后,各理化指标的模型预测 R^2 均有明显提升,加入全部地形因子后各理化指标模型预测 R^2 仍进一步微弱提升。与纯光谱反射率模型相比,加入全部地形因子后各理化指标模型预测 R^2 的平均提升幅度为 0.036,其中全氮含量的模型预测 R^2 提升幅度最大(提升了 0.062)。

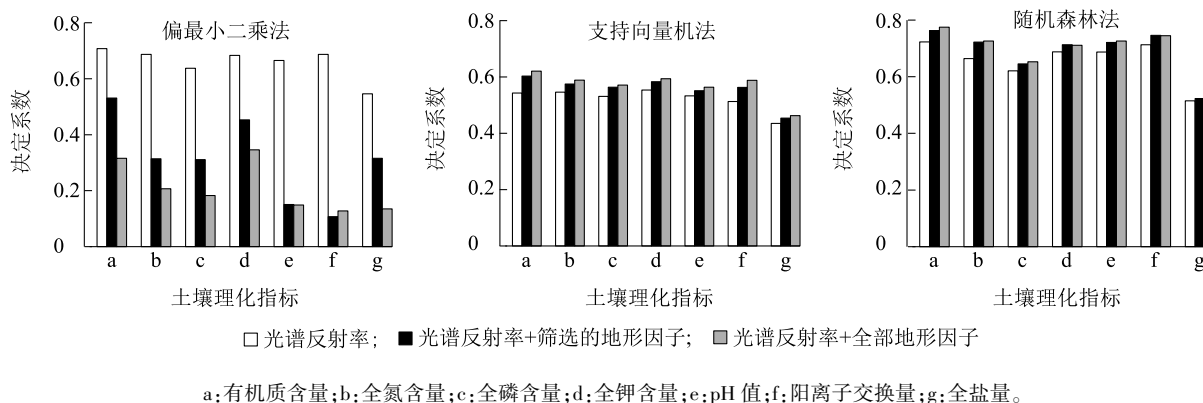


图3 地形因子对反演模型预测精度的影响

Fig.3 Effects of terrain factors on the prediction accuracy of inversion model

3 讨论

研究区12种地形因子中,海拔、坡度、山谷指数、地形粗糙指数、起伏度与土壤理化指标有较高的相关性,这为探讨地形因子对土壤理化性质反演的影响提供了依据。

基于地形因子与土壤理化指标的相关性分析,筛选与各个指标相关的地形因子加入反演计算,并利用偏最小二乘法、支持向量机法和随机森林法探讨地形因子对土壤理化性质航空高光谱反演的影响。结果表明,模型自变量中加入筛选的地形因子后,支持向量机法和随机森林法的模型预测 R^2 均有稳定提升,继续加入全部12种地形因子后, R^2 仍有

微弱提升,说明地形因子对提升模型精度有贡献,而且筛选的地形因子对模型的贡献更大。

偏最小二乘法与支持向量机法、随机森林法不同,加入筛选的地形因子后偏最小二乘法的模型预测 R^2 大幅度下降,加入全部地形因子后 R^2 进一步下降,分析原因可能是地形因子与光谱反射率之间的相关程度较低,二者对偏最小二乘法计算过程中的主成分分析造成了干扰,从而影响了模型精度。支持向量机法和随机森林法属于机器学习方法,模型的学习能力强,更有利于从高维数据中提取有效信息。横向对比3种方法的模型精度,随机森林法的建模效果整体上最好,因为其采用的是多个决策树的投票机制,适合处理高维数据,而支持向量机法受

核函数类型的影响较大,在本研究中的预测精度稍低。

本研究结果对后续建立更为完善的土壤理化指标航空高光谱综合反演模型有启发意义。航空高光谱技术凭借着光谱分辨率高、数据全覆盖等优势,为土壤理化指标的大范围反演制图提供了数据支撑,而为了提高该技术的实际应用能力,必须尽可能提升模型的反演精度和可靠性,在后续研究中需要引入更多相关变量开展综合反演模型研究,继续提升模型的反演效果和实际应用能力。

参考文献:

- [1] 魏孝荣,邵明安. 黄土高原沟壑区小流域不同地形下土壤性质分布特征[J]. 自然资源学报, 2007, 22(6):946-953.
- [2] 魏孝荣,邵明安. 黄土沟壑区小流域不同地形下土壤微量元素分布特征[J]. 环境科学, 2009, 30(9):2741-2746.
- [3] 孙文义,郭胜利. 黄土丘陵沟壑区小流域土壤有机碳空间分布及其影响因素[J]. 生态学报, 2010, 31(6):1604-1616.
- [4] 赵明月,赵文武,钟莉娜. 土地利用和环境因子对表层土壤有机碳影响的尺度效应——以陕北黄土丘陵沟壑区为例[J]. 生态学报, 2014, 34(5):1105-1113.
- [5] 薛志婧,侯晓瑞,程 曼,等. 黄土丘陵区小流域尺度上土壤有机碳空间异质性[J]. 水土保持学报, 2011,25(3):160-163.
- [6] 赵 越,罗志军,赵 杰,等. 峡江县土壤养分空间特征及与地形因子的相关性[J]. 水土保持研究, 2018, 25(5):128-133.
- [7] 陈桂香,高灯州,曾从盛,等.福州市农田土壤养分空间变异特征[J].地球信息科学学报,2017,19(2):216-224.
- [8] 邓欧平,周 稀,黄萍萍,等.川中紫色丘区土壤养分空间分异与地形因子相关性研究[J].资源科学,2013,35(12):2434-2443.
- [9] 王淑英,路 苹,王建立,等.不同研究尺度下土壤有机质和全氮的空间变异特征——以北京市平谷区为例[J].生态学报, 2008,28(10):4957-4964.
- [10] 贾振宇,张俊华,丁圣彦,等.基于 GIS 和地统计学的黄泛区土壤磷空间变异——以周口为例[J].应用生态学报,2016, 27(4):1211-1220.
- [11] 赵明松,张甘霖,李德成,等.苏中平原南部土壤有机质空间变异特征研究[J].地理科学,2013,33(1):83-89.
- [12] 宋 轩,李立东,寇长林,等. 黄河小流域土壤养分分布及其与地形的关系[J]. 应用生态学报, 2011, 22(12):3163-3168.
- [13] 徐 莉,闫俊杰,陈 晨,等. 新疆察布查尔县农耕地土壤养分空间特征及其与地形因子的关系[J]. 水土保持通报, 2018, 38(4):44-51.
- [14] 王雪梅,柴仲平,武红旗.典型干旱荒漠绿洲区耕层土壤养分空间变异[J].水土保持通报,2016,36(1):51-56.
- [15] 唐梦迎,丁建丽,夏 楠,等. 干旱区典型绿洲土壤有机质含量分布特征及其影响因素[J]. 土壤学报, 2017, 54(3): 759-766.
- [16] 高凤杰,鞠铁男,吴 啸,等. 黑土耕作层土壤 pH 空间变异及自相关分析[J]. 土壤, 2018,50(3):566-573.
- [17] 杜 可,王 乐,张淑香,等. 黑土区县域土壤养分空间分布特征及其影响因子[J]. 植物营养与肥料学报, 2018, 24(6): 1465-1474.
- [18] 高凤杰,马泉来,韩文文,等. 黑土丘陵区小流域土壤有机质空间变异及分布格局[J]. 环境科学, 2016, 37(5):325-332.
- [19] 张孝存,郑粉莉,安 娟,等. 典型黑土区坡耕地土壤侵蚀对土壤有机质和氮的影响[J]. 干旱地区农业研究, 2013, 31(4): 182-186.
- [20] 邢东兴,王明军,陈玲侠,等. 桃树遥感辨识的最佳时相与方法[J].江苏农业学报,2019,35(4):919-926.
- [21] 王利军,郭 燕,王来刚,等. 农情遥感信息服务与野外监测云平台构建[J].南方农业学报,2018,49(6):1242-1248.
- [22] 张自超,韩 宇,陈 建,等. 基于无人机遥感数据的生态渠系信息提取[J]. 排灌机械工程学,2018,36(10):1006-1011.
- [23] 王术波,韩 宇,陈 建,等.基于深度学习的无人机遥感生态灌区杂草分类[J]. 排灌机械工程学,2018,36(11):1137-1141.
- [24] 单 捷,孙 玲,王志明,等. GF-1 影像遥感监测指标与冬小麦长势参数的关系[J]. 江苏农业学报, 2019, 35(6):1323-1333.
- [25] 张素梅,王宗明,张 柏,等. 利用地形和遥感数据预测土壤养分空间分布[J]. 农业工程学报, 2010, 26(5):188-194.
- [26] 刘焕军,谢雅慧,潘 越,等. 黑土区田块土壤有机质空间分异及分布研究[J]. 土壤通报, 2018, 49(6): 1445-1451.
- [27] 秦承志,卢岩君,包黎莉,等. 简化数字地形分析软件(SimDTA)及其应用——以嫩江流域鹤山农场的坡位模糊分类应用为例[J]. 地球信息科学学报, 2009, 11(6): 737-743.
- [28] 沈从旺,徐丽华.土壤 pH 值和全钾含量高光谱反演方法比较[J]. 江苏农业学报, 2020, 36(1): 92-98.
- [29] 中华人民共和国自然资源部.土地质量地球化学评价规范: DZ/T 0295-2016 [S].北京:地质出版社,2016:42.
- [30] 吕贻忠,李保国. 土壤学实验[M]. 北京:中国农业出版社, 2010.
- [31] 杨劲松,姚荣江. 黄河三角洲地区土壤水盐空间变异特征研究[J]. 地理科学, 2007, 27(3): 348-353.

(责任编辑:王 妮)