

基于宽波段与窄波段综合光谱指数的 土壤有机质遥感反演

郑曼迪¹, 熊黑钢², 乔娟峰¹, 刘靖朝¹

¹新疆大学资源与环境科学学院绿洲生态重点实验室, 新疆 乌鲁木齐 830046;

²北京联合大学应用文理学院城市科学系, 北京 100083

摘要 对比基于宽波段、窄波段建立的土壤有机质(SOM)含量的预测模型以及空间格局分布的差异性,采用地面高光谱测量和土质分析验证利用卫星遥感数据监测土壤基本生态参数的可行性。以天山北麓的土壤为研究对象,运用宽波段、窄波段两种方式计算实测光谱反射率的综合光谱指数,与无人干扰区、人为干扰区的有机质进行相关性分析以及主成分分析,以相关系数和特征向量值都较优的综合光谱指数作为自变量,使用多元线性回归模型(MLR)以及偏最小二乘回归模型(PLSR)分别建立了无人干扰区及人为干扰区宽、窄波段的 SOM 高光谱预测模型,并进行模型验证、对比与优选,最后基于最佳模型对研究区进行 SOM 含量的空间格局反演和分析。结果显示,通过有机质与盐分指数、植被指数的相关性分析和主成分分析,挑选出了无人干扰区窄波段的盐分指数 2(SI₂)、盐分指数 3(SI₃)和比值植被指数(RVI)、归一化植被指数(NDVI)以及宽波段的盐分指数 1(SI₁)、SI₂ 和 RVI、NDVI,人为干扰区窄波段的 SI₁、SI₃ 和 RVI、NDVI 以及宽波段的 SI₁、SI₂、重归一化植被指数(RDVI),以此为自变量建立有机质含量的 MLR 以及 PLSR 模型。通过对比所建模型的精度可知,无论是无人干扰区还是人为干扰区,有机质预测模型精度最高的均是窄波段的 PLSR 模型,可决定系数、相对分析误差分别为 0.753、2.01 和 0.819、2.14。基于以上的最佳模型对研究区的 SOM 含量进行空间反演与分析可知:无人干扰区的有机质质量分数集中在小于 10×10^{-3} 范围内,呈现出中间低、四周高的趋势;而人为干扰区有机质的质量分数集中在 $10 \times 10^{-3} \sim 15 \times 10^{-3}$ 范围内,呈现出西南、东北低,中北部高的趋势。

关键词 遥感; 反演; 宽波段; 窄波段; 综合光谱指数

中图分类号 TP79 文献标识码 A

doi: 10.3788/LOP55.072801

Remote Sensing Inversion of Soil Organic Matter Based on Broad Band and Narrow Band Comprehensive Spectral Index

Zheng Mandi¹, Xiong Heigang², Qiao Juanfeng¹, Liu Jingchao¹

¹Key Laboratory of Oasis Ecology, College of Resource and Environment Sciences, Xinjiang University, Urumqi, Xinjiang 830046, China;

²College of Applied Arts and Science, Beijing Union University, Beijing 100083, China

Abstract Comparing the soil organic matter (SOM) prediction model based on the broad and narrow bands and the difference in spatial pattern distribution, we validate the feasibility of using satellite remote sensing data to monitor soil basic ecological parameters by ground hyperspectral measurement and analysis of soil. Taking the soil of Tianshan as the research object, we calculate comprehensive spectral index of the broad and narrow band, respectively, using correlation analysis and principal component analysis in the organic matter of unmanned interference area, human interference area, and choosing the comprehensive spectral index with better correlation coefficient and characteristic vector value as the independent variables, using multivariate linear regression model

收稿日期: 2017-12-12; 收到修改稿日期: 2018-01-12

基金项目: 国家自然科学基金(41671198)

作者简介: 郑曼迪(1993—),女,硕士研究生,主要从事干旱区资源与环境遥感方面的研究。E-mail: 762820677@qq.com

导师简介: 熊黑钢(1956—),男,博士,教授,博士生导师,主要从事资源环境方面的研究。

E-mail: heigang@buu.edu.cn(通信联系人)

(MLR) and partial least squares regression model (PLSR) to establish respectively the hyperspectral prediction model of SOM in broad and narrow band of unmanned interference area and human interference area. The validation, the comparison and selection the model are carried out. Finally, we analyze and inverse the spatial pattern of SOM content based on the best model of research area. Results show that, through the correlation analysis and principal component analysis of the organic matter and salinity index, vegetation index to establish the MLR and PLSR of organic matter component, we pick out the salinity index 2 (SI_2), salinity index 3 (SI_3) and ratio vegetation index (RVI), normalized difference vegetation index (NDVI) of narrow band, and the SI_1 , SI_2 , RVI, NDVI of broad band in unmanned interference area; we pick out SI_1 , SI_3 , RVI and NDVI of narrow band, and SI_1 , SI_2 , RVI and renormalized difference vegetation index (RDVI) of broad band. Taking these parameters as independent variables, we build MLR and PLSR models of soil organic matter. By comparing the precision of the models, we find that the PLSR model with narrow band has high precision in human or unmanned interference areas, and determinable coefficient and relative percent deviation are 0.753, 2.01 and 0.819 and 2.14, respectively. Spatial inversion and analysis of SOM in research area are carried out based on the best model above. The mass fraction of organic matter in unmanned interference area is concentrated in less than 10×10^{-3} , and presents the trend of low in middle and high in around. The mass fraction of organic matter in human interference is $10 \times 10^{-3} - 15 \times 10^{-3}$, presents the trend of low in southwest and northeast, and high around middle north region.

Key words remote sensing; inversion; broad range; narrow band; comprehensive spectral index

OCIS codes 280.0280; 280.4788; 280.4750

1 引言

土壤有机质(SOM)是土壤肥力的重要指标^[1],它包括各种动植物残体、微生物及维持生命活动的各种有机产物,是土壤的重要组成部分。遥感数据具有覆盖面积广、寻访周期快等优点,为遥感信息的定量研究开辟了崭新的途径^[2]。其中,光谱指数波段组合能部分消除环境背景(如非植被目标土壤、水体等)的光谱影响,在此基础上建立的多波段光谱植被指数所体现的光谱信息比单波段光谱具有更好的灵敏性,统计分析结果更加精确^[3-4]。近年来,部分国内外学者通过运用归一化指数以及差值指数对受人为干扰的土壤盐分、有机质、全氮、速效氮、含水量进行预测,预测效果都优于以原始光谱反射率以及以光谱变换为自变量的预测模型^[5]。也有学者得出了通过遥感影像获取的扩展增强型植被指数能够更好地预测土壤盐分的结论,并以此来讨论阜北地区的土壤盐渍化程度^[6]。

采用光谱指数技术预测土壤理化性质的研究已经较为成熟,但是绝大部分的研究都是以人为干扰的土壤作为研究对象,以实测光谱反射率或遥感影像反射率计算得到的各类指数为自变量建立预测模型,很少有研究针对同一背景下,在不同程度人为干扰的区域内,对比以宽、窄波段综合光谱指数为自变量建立的有机质含量模型的精度。本文运用宽、窄波段的土壤盐分指数及植被指数,建立有、无人干扰的 SOM 预测模型,并对比寻求其最佳的模型,为今后治理区域的土壤退化判断提供依据。

2 材料与方法

2.1 研究区概况

研究区位于天山北麓与准噶尔盆地南缘($87^{\circ}40' \sim 87^{\circ}70' E$, $44^{\circ}12' \sim 45^{\circ}00' N$),属于典型的温带大陆性气候,夏季炎热,冬季寒冷,降水稀少,年降水量为 164 mm,年蒸发潜力为 2000 mm 左右。土壤类型有灰漠土、龟裂土以及沙土^[7]。研究区分为 A、B 两区:A 区为无人干扰区,基本保持原貌,表层有梭梭、红柳、杂草等植物,表面光滑且有盐碱斑分布;B 区为人为干扰区,位于新疆生产建设兵团 102 团附近,多为人工种植的梭梭林地、榆树林地以及育苗地。由于是普通林地,所以只是翻耕并没有施加肥料。

2.2 样本采集

在 A 区由南向北布设 5 条间距为 600~800 m 的东西向采样线,B 区布设 6 条间距为 800~1000 m 的采样线。在每条采样线上选择 5 个具有代表性的点,其间距为 300~500 m,并对 55 个采样点进行 GPS 定位,如图 1 所示。采集时间为 2014

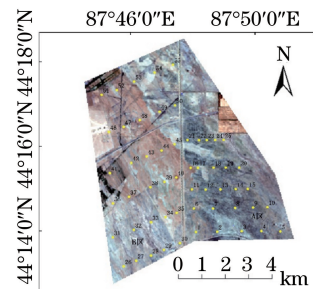


图 1 研究区采样点分布图

Fig. 1 Sampling point distribution map of the research area

年6月。采集深度为0~20 cm的土样,土样经过研磨、过筛等预处理之后,运用重铬酸钾容量法对有机质含量进行测定。

2.3 影像数据

选取2014年6月的Landsat 8陆地成像仪(OLI)影像,共有9个波段,包括ETM+传感器的所有波段,增加了蓝色波段(B1)和短红外波段(B9)两个波段,其中B8是分辨率为15 m的全色波段。云量为0,去除影像的B8、B9波段,保留分辨率为30 m的前7个波段进行处理。使用ENVI 5.1软件对所需区域进行裁剪、辐射定标及几何校正。由于获取影像信息时会受到大气中水汽、气溶胶等因素的影响而使得波段的噪声增加,信息模糊,所以使用FLAASH大气校正消除这些噪声,降低邻近像元之间的辐射干扰,同时也可以调整因人为抑制而导致的波谱平滑^[8],最后获取近似地表真实反射率的影像。

2.4 光谱采集与预处理

2.4.1 窄波段反射率的测定与处理

窄波段光谱测量采用美国ASD公司生产的FieldSpec[®] 3Hi-Res光谱仪,光谱的有效范围为350~2500 nm,分析软件采用ASD ViewSpecPro。光谱测量采用野外实测方法,为了避免天气对光谱造成的不利影响,实验时间选在11:00—15:00(北京时间),在晴朗少云且无风的天气下进行。每次采集光谱前对光谱仪进行白板校正,以去除暗电流的影响。采用25°视场角探头,在距土壤表面15 cm处以垂直角度进行光谱采集。在每个采样点周围2 m范围内采用梅花桩采样法选取5个位置采集光谱,每个位置重复测量10次。由于研究区的土壤背景相近,下垫面比较均一,所以得到的50条光谱曲线的平均值即为该采样点的实测光谱值。A、B两

区的土壤表面有部分植被覆盖,采集所得的光谱是土壤与少量植被的混合光谱。

在数据分析之前,通过Savitzky-Golay平滑法^[9]对光谱进行平滑处理,首先剔除处于紫外波段350~400 nm范围内光谱反射率及短波红外波段2400~2500 nm范围内的光谱反射率,因为这部分波段的信噪比较低^[10-11],信号受噪声的影响较大,数据的可信度降低。其次剔除位于水分吸收带的波段(1400~1500 nm和1900~2000 nm),因为此波段对光谱反演有机质含量的精准性有较大影响,不适合作为SOM含量高光谱估测模型的特征波段^[12]。

2.4.2 宽波段反射率的获取

采用波段平均法获取宽波段反射率,将各波段的实测光谱反射率值拟合为Landsat 8 OLI影像的B1~B7波段对应的宽波段反射率^[13],B1波段范围为433~453 nm,B2波段范围为450~515 nm,B3波段范围为525~600 nm,B4波段范围为630~680 nm,B5波段范围为845~885 nm,B6波段范围为1560~1660 nm,B7波段范围为2100~2300 nm^[14]。

2.5 综合光谱指数的计算

选取5种土壤盐分指数和6种植被指数。土壤盐分指数分别为:盐分指数1(SI₁)、盐分指数2(SI₂)、盐分指数3(SI₃)、归一化盐分指数(NDSI)、亮度指数(BI)^[15-16];植被指数分别为比值植被指数(RVI)、归一化植被指数(NDVI)、差值植被指数(DVI)、垂直植被指数(PVI)、重归一化植被指数(RDVI)、调整转换型土壤调节植被指数(ATSAVI)^[17-20]。综合光谱指数计算方法如表1所示。将实测光谱反射率 R_i (i 为对应的波长)代入表1得出由光谱计算得到的盐分指数与植被指数。其中: X 为调节因子,设为0.08; a 、 b 为土壤线系数,分别设置为1.0、0.5。

表1 综合光谱指数计算方法

Table 1 Calculation method of comprehensive spectral index

| Index type | Narrow band | Broad band |
|--|---|---|
| Salinity index | SI ₁ : $\sqrt{R_{456} R_{686}}$ | SI ₁ : B_2/B_4 |
| | SI ₂ : $\sqrt{R_{456}^2 + R_{686}^2 + R_{1373}^2}$ | SI ₂ : $(B_2 - B_4)/(B_2 + B_4)$ |
| | SI ₃ : $\sqrt{R_{456}^2 + R_{686}^2}$ | SI ₃ : $B_3 B_4/B_2$ |
| | NDSI: $(R_{686} - R_{1373})/(R_{686} + R_{1373})$ | NDSI: $(B_4 - B_5)/(B_4 + B_5)$ |
| | BI: $\sqrt{R_{533}^2 + R_{686}^2}$ | BI: $\sqrt{B_4^2 + B_5^2}$ |
| Vegetation index | RVI: R_{1137}/R_{710} | RVI: B_5/B_3 |
| | NDVI: $(R_{1137} - R_{710})/(R_{1137} + R_{710})$ | NDVI: $(B_5 - B_4)/(B_5 + B_4)$ |
| | DVI: $R_{1137} - R_{710}$ | DVI: $B_5 - B_4$ |
| | PVI: $(R_{1148} - aR_{807} - b)/\sqrt{1+a^2}$ | PVI: $(B_5 - aB_4 - b)/\sqrt{1+a^2}$ |
| | RDVI: $(R_{800} - R_{670})/(\sqrt{R_{800} + R_{670}})$ | RDVI: $(B_5 - B_4)/(\sqrt{B_5 + B_4})$ |
| ATSAVI: $a(R_{800} - aR_{670} - b)/[aR_{800} + R_{670} - ab + X(1+a^2)]$ | ATSAVI: $a(B_5 - aB_4 - b)/[aB_5 + B_4 - ab + X(1+a^2)]$ | |

2.6 模型检验

通过修正自由度的可决定系数 R^2 、均方根误差 R_{MSE} 、相对分析误差 R_{PD} 等指标衡量模型的精度^[21], 它们的表达式分别见(1)式、(2)式、(3)式。 R^2 反映预测值解释实测值变化的程度, R_{MSE} 用来检验模拟值与实测值的符合度。若 R^2 越大, R_{MSE} 越小, 说明模型的精度越高。当 $R_{PD} > 2.0$ 时, 表明模型具有良好的预测能力; 当 $1.4 < R_{PD} < 2.0$ 时, 说明模型可对样品进行粗略估测; 当 $R_{PD} < 1.4$ 时, 则表明模型无法对样品进行估测^[22]。

$$R^2 = \left[\frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2 (y_i - \bar{y})^2} \right]^2, \quad (1)$$

$$R_{MSE} = \sqrt{\frac{\sum_{i=1}^n (x_i - x'_i)^2}{n}}, \quad (2)$$

$$R_{PD} = \frac{\delta}{R_{MSE}}. \quad (3)$$

式中: x_i 为观测值; x'_i 为真实值; \bar{x} 为实测有机质含量的平均值; \bar{y} 为预测有机质含量的平均值; n 为观测次数; δ 为样本标准差。

3 结果与分析

3.1 SOM 与盐分的相关性分析

无人干扰区和人为干扰区有机质含量与盐分含量

表3 土壤盐分指数、植被指数与有机质含量的相关性

Table 3 Correlation between soil salinity index, vegetation index and organic matter content

| Index | Index type | Correlation coefficient | | | |
|------------------|-----------------|-------------------------|----------------|------------------|----------------|
| | | Narrow band | | Broad band | |
| | | Undisturbed area | Disturbed area | Undisturbed area | Disturbed area |
| Salinity index | SI ₁ | -0.650** | -0.699** | -0.608** | -0.607** |
| | SI ₂ | -0.657** | -0.607** | -0.568** | -0.565** |
| | SI ₃ | -0.716** | -0.584** | -0.635** | -0.511** |
| | NDSI | -0.494* | -0.439* | -0.428* | -0.376* |
| | BI | -0.631** | -0.539** | -0.574** | -0.461** |
| | RVI | 0.591** | 0.644** | 0.531** | 0.564** |
| Vegetation index | NDVI | 0.583** | 0.440* | 0.506** | 0.389* |
| | DVI | 0.497* | 0.451* | 0.426* | 0.394* |
| | PVI | 0.532** | 0.498** | 0.466* | 0.400** |
| | RDVI | 0.486* | 0.539** | 0.405* | 0.467** |
| | ATSAVI | 0.561** | 0.467** | 0.469* | 0.385* |

对土壤盐分指数与植被指数进行主成分分析, 结果如表4、5所示; 无人干扰区的窄波段盐分指数与植被指数的第一主成分的方差贡献率分别达到了77.531%、75.165%, 其中 SI₂、SI₃ 和 RVI、NDVI 的特征向量较高; 宽波段盐分指数和植被指数的第一主成

呈负相关, 相关系数分别达到 0.867 和 0.803, 且均呈现极显著的相关关系, 如表2所示(上标**表示在 0.01 水平上显著相关)。这是因为 SOM 大部分来源于动植物以及微生物残体等盐分过高的区域, 不利于生物生存, 即动植物、微生物残体质量过少, 有机质含量较少, 故而出现盐分越高, 有机质含量越少的现象。

表2 SOM 与盐分之间的相关性

Table 2 Correlation between SOM and salinity

| Area | Correlation coefficient |
|------------------|-------------------------|
| Undisturbed area | -0.867** |
| Disturbed area | -0.803** |

人为干扰区因为有频繁的人类活动, 例如翻耕、犁地等行为, 破坏了土壤表层结皮, 将下层土壤都翻至表面, 打破了原有土壤中盐分、有机质的分布状态, 使其分布更加复杂, 所以人为干扰区盐分与有机质的相关性小于无人干扰区。

3.2 土壤盐分指数、植被指数与有机质的相关性主成分分析

有、无人为干扰区的有机质含量与盐分指数、植被指数分别呈负相关、正相关, 如表3所示, 其中上标*表示在 0.05 水平上显著相关。有机质含量与盐分含量、盐分指数均呈负相关。无论植被的生长还是其残体的分解都有利于有机质含量的积累, 所以植被指数与有机质呈现正相关的关系。

分贡献率分别为 80.168%、78.631%, 其中 SI₁、SI₂ 和 RVI、NDVI 的特征向量值较高。人为干扰区窄波段与宽波段的盐分指数第一主成分的方差贡献率分别为 76.155%、81.361%, 其中 SI₁、SI₃ 和 SI₁、SI₂ 特征向量高于其他盐分指数的特征向量值; 窄、宽波段的植

表4 土壤盐分指数与有机质含量的主成分分析

Table 4 Principal components analysis of soil salinity index and organic matter content

| Area | Band | Principal components | Variance contribution rate / % | Index type | | | | |
|------------------|-------------|----------------------|--------------------------------|-----------------|-----------------|-----------------|--------|--------|
| | | | | SI ₁ | SI ₂ | SI ₃ | NDSI | BI |
| Undisturbed area | Narrow band | 1 | 77.531 | 0.386 | 0.416 | 0.576 | 0.250 | 0.463 |
| | | 2 | 12.334 | 0.298 | 0.149 | 0.476 | 0.041 | -0.227 |
| | | 3 | 6.315 | 0.168 | -0.211 | 0.367 | -0.274 | 0.349 |
| | | 4 | 3.641 | 0.477 | -0.443 | 0.074 | -0.116 | -0.441 |
| | | 5 | 0.179 | -0.516 | 0.224 | 0.449 | -0.336 | 0.013 |
| | Broad band | 1 | 80.168 | 0.641 | 0.553 | 0.496 | 0.117 | 0.330 |
| | | 2 | 13.547 | 0.316 | -0.367 | 0.366 | 0.229 | 0.216 |
| | | 3 | 5.118 | 0.333 | 0.379 | 0.196 | -0.064 | -0.014 |
| | | 4 | 1.009 | -0.017 | 0.109 | 0.416 | 0.079 | 0.296 |
| | | 5 | 0.158 | -0.255 | 0.057 | -0.331 | 0.146 | -0.351 |
| Disturbed area | Narrow band | 1 | 76.155 | 0.597 | 0.499 | 0.649 | 0.410 | 0.219 |
| | | 2 | 12.229 | 0.279 | 0.337 | -0.063 | 0.332 | -0.145 |
| | | 3 | 7.147 | -0.334 | 0.063 | 0.096 | 0.169 | 0.196 |
| | | 4 | 3.580 | 0.285 | -0.361 | 0.359 | 0.306 | 0.208 |
| | | 5 | 0.889 | 0.275 | 0.113 | -0.298 | 0.059 | -0.196 |
| | Broad band | 1 | 81.361 | 0.652 | 0.599 | 0.492 | 0.255 | 0.402 |
| | | 2 | 9.796 | 0.316 | 0.119 | -0.217 | 0.334 | -0.295 |
| | | 3 | 6.117 | -0.049 | 0.177 | 0.163 | 0.329 | -0.014 |
| | | 4 | 2.516 | 0.053 | -0.174 | 0.411 | 0.429 | 0.218 |
| | | 5 | 0.180 | -0.034 | 0.225 | 0.329 | -0.179 | 0.291 |

表5 植被指数与有机质含量的主成分分析

Table 5 Main component analysis of vegetation index and organic matter content

| Area | Band | Principal components | Variance contribution rate / % | Index type | | | | |
|------------------|-------------|----------------------|--------------------------------|-----------------|-----------------|-----------------|--------|--------|
| | | | | SI ₁ | SI ₂ | SI ₃ | NDSI | BI |
| Undisturbed area | Narrow band | 1 | 75.165 | 0.682 | 0.597 | 0.254 | 0.416 | 0.279 |
| | | 2 | 13.476 | 0.357 | 0.296 | 0.021 | 0.339 | 0.211 |
| | | 3 | 7.210 | 0.196 | 0.441 | 0.419 | -0.316 | -0.016 |
| | | 4 | 3.019 | -0.397 | 0.111 | -0.052 | 0.174 | 0.252 |
| | | 5 | 1.104 | 0.226 | 0.049 | 0.334 | 0.296 | -0.385 |
| | | 6 | 0.026 | -0.014 | 0.275 | -0.077 | 0.274 | 0.452 |
| | Broad band | 1 | 78.631 | 0.634 | 0.519 | 0.336 | 0.224 | 0.394 |
| | | 2 | 10.485 | 0.431 | 0.318 | 0.152 | 0.362 | -0.132 |
| | | 3 | 6.441 | 0.271 | 0.063 | 0.416 | 0.118 | -0.240 |
| | | 4 | 2.719 | 0.088 | -0.349 | 0.114 | -0.055 | 0.017 |
| | | 5 | 1.294 | 0.345 | -0.334 | 0.264 | 0.117 | 0.361 |
| | | 6 | 0.430 | 0.029 | 0.169 | 0.441 | -0.337 | 0.249 |
| Disturbed area | Narrow band | 1 | 79.330 | 0.716 | 0.395 | 0.451 | 0.688 | 0.201 |
| | | 2 | 10.511 | 0.315 | 0.051 | 0.366 | -0.255 | 0.411 |
| | | 3 | 5.170 | 0.084 | 0.339 | -0.410 | 0.294 | -0.344 |
| | | 4 | 3.229 | -0.387 | 0.019 | 0.315 | 0.201 | 0.451 |
| | | 5 | 1.219 | 0.198 | -0.119 | 0.395 | -0.312 | 0.294 |
| | | 6 | 0.541 | 0.011 | -0.378 | -0.308 | 0.331 | 0.449 |
| | Broad band | 1 | 77.318 | 0.681 | 0.254 | 0.298 | 0.597 | 0.462 |
| | | 2 | 13.006 | 0.254 | 0.441 | 0.408 | 0.419 | 0.319 |
| | | 3 | 5.519 | -0.294 | 0.327 | -0.106 | 0.415 | 0.287 |
| | | 4 | 2.441 | -0.350 | 0.219 | 0.214 | 0.344 | -0.019 |
| | | 5 | 1.096 | 0.326 | 0.371 | 0.416 | -0.281 | -0.144 |
| | | 6 | 0.620 | -0.014 | 0.222 | -0.341 | 0.430 | 0.317 |

被指数第一主成分的方差贡献率都大于 77%，且特征向量较高的分别是 RVI、NDVI 和 RVI、RDVI。第一主成分中的特征向量值较高说明其对有机质含量的贡献最大，包含的信息也较为丰富。一般情况下，在主成分反演中，将选取主成分分析后累计方差贡献率大于 90% 的主成分分量作为反演的参量^[23]。

在相关性分析与主成分分析中，盐分指数、植被指数与有机质含量相关性较好的各两种指数与特征向量值较高的各两种指数是一致的，入选的盐分指

表 6 全样本土壤有机质统计分析结果

Table 6 Statistical analysis results of full-sample soil organic matter

| Area | Composition | Number of samples | Range / (g·kg ⁻¹) | Mean / (g·kg ⁻¹) | Variance / (g·kg ⁻¹) |
|------------------|--------------|-------------------|-------------------------------|------------------------------|----------------------------------|
| Undisturbed area | Modeling set | 15 | 9.588-23.034 | 13.744 | 12.311 |
| | Test set | 10 | 7.645-17.928 | 13.302 | 12.290 |
| | Full sample | 25 | 7.645-23.034 | 13.567 | 11.839 |
| Disturbed area | Modeling set | 18 | 6.376-21.768 | 11.432 | 11.424 |
| | Test set | 12 | 6.455-16.439 | 11.674 | 11.808 |
| | Full sample | 30 | 6.376-21.768 | 11.537 | 11.204 |

3.4 基于窄、宽波段综合光谱指数的有机质预测模型

在盐分指数、植被指数与有机质相关性分析和主成分分析中，各选择与有机质含量相关性较好且特征向量值较高的两个指数，用以建立有机质含量

表 7 基于实测光谱反射率的 SOM 含量预测模型

Table 7 SOM content prediction model based on the measured spectral reflectivity

| Independent variables | Undisturbed area | | | | | Disturbed area | | | | |
|-----------------------|------------------|-----------|----------|-----------|----------|----------------|-----------|----------|-----------|----------|
| | Modeling set | | Test set | | R_{PD} | Modeling set | | Test set | | R_{PD} |
| | R^2 | R_{MSE} | R^2 | R_{MSE} | | R^2 | R_{MSE} | R^2 | R_{MSE} | |
| M-MLR | 0.742 | 2.6123 | 0.699 | 2.8416 | 1.83 | 0.692 | 2.7516 | 0.674 | 2.8861 | 1.76 |
| M-PLSR | 0.819 | 1.6524 | 0.792 | 2.0611 | 2.14 | 0.753 | 1.8874 | 0.765 | 2.0674 | 2.01 |
| N-MLR | 0.661 | 3.2019 | 0.648 | 3.5179 | 1.76 | 0.627 | 3.5149 | 0.631 | 3.6716 | 1.68 |
| N-PLSR | 0.713 | 2.6418 | 0.704 | 2.7729 | 1.95 | 0.694 | 2.8879 | 0.688 | 3.0076 | 1.84 |

整体上，无人干扰区的建模效果优于人为干扰区，原因是人类活动严重影响了土壤表面的结皮，土壤中的水分损失严重，而水分的丢失直接影响到了盐分的迁移，打破了土壤中盐分、有机质的原始分布状况，加之耕翻、犁地等耕作方式将深层的土壤外翻至表层，使其分布更为复杂，难以预测，从而使人为干扰区的模型精度小于无人干扰区。

无论是宽、窄波段，还是有、无人干扰区，PLSR 的模型精度都优于 MLR，原因是多元线性回归很难消除变量间的多重相关性，会造成模型的正规方程系数矩阵出现严重的病态性，进而使模型参数的最小二乘失真，回归系数估计值的稳定性也将变得很差。但是偏最小二乘回归在处理样本容量小、解释

数 SI_1 、 SI_2 和 SI_3 较其他盐分指数对盐度信息更为敏感。入选的植被指数是土壤和大气调整类植被指数，这是因为加入土壤和大气调节因子之后的植被指数可以有效抵消土壤和大气的影 响，使建模效果更好^[24]。

3.3 不同人为干扰程度的有机质预测模型

选用 KS(Kennard-Stone)算法^[25]计算出各个样品有机质含量之间的欧氏距离，按照 3:2 的比率将样品划分为建模集和检验集，如表 6 所示。

多元线性回归模型(MLR)和偏最小二乘回归模型(PLSR)，如表 7 所示，其中 M 代表窄波段，N 代表宽波段。为了比较不同干扰程度下 SOM 含量的预测模型精度，计算了各个模型的决定系数 R^2 、 R_{MSE} 以及 R_{PD} 3 项指标。

变量多、变量间存在严重多重相关性方面具有优势，而且可以同时实现回归建模数据结构简化以及两组变量间的相关性分析^[26]。

宽波段与窄波段建立模型的 R_{PD} 均大于 1.4，都可以对区域内有机质的含量进行粗略估计，其中有、无人干扰区预测效果最好的都是窄波段的 PLSR， R^2 分别达到 0.753、0.819， R_{PD} 分别为 2.01、2.14，如表 7 所示。窄波段建立的模型精度均大于宽波段建立的模型精度，这可能是因为宽波段范围内包含了许多波段的反射率信息，其中包括对有机质响应敏感和不敏感的波段，平均之后对有机质响应敏感的波段信息会被淡化，所以模型精度小于窄波段。

3.5 基于窄、宽波段综合光谱指数的有机质空间反演

使用 ENVI 5.1 软件的 Band Math 模块,将有、无人干扰区采样点对应的反射率代入到基于综合光谱指数的最佳预测模型中,反演得到研究区土壤有机质的空间分布格局。将研究区土壤有机质的质量分数划分为 5 个等级:有机质质量分数在 $0 \sim 5 \times 10^{-3}$ 范围内的等级为“低”,在 $5 \times 10^{-3} \sim 10 \times 10^{-3}$

范围内的等级为“较低”,在 $10 \times 10^{-3} \sim 15 \times 10^{-3}$ 范围内的等级为“中等”,在 $15 \times 10^{-3} \sim 20 \times 10^{-3}$ 范围内的等级为“较高”,在 $20 \times 10^{-3} \sim 25 \times 10^{-3}$ 范围内的等级为“高”。

使用密度分割方法按照 5 个等级的划分范围,得到不同级别有机质含量的空间格局分布图,如图 2 所示。

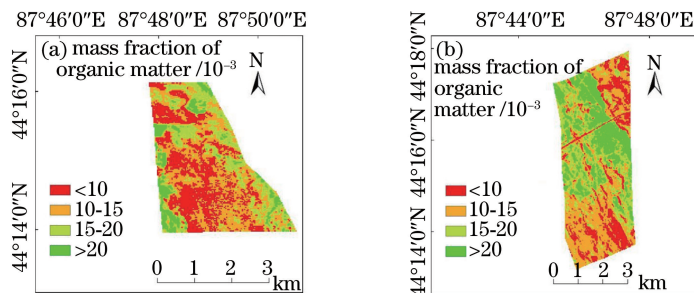


图 2 有、无人干扰区有机质含量空间格局分布图。(a)人为干扰区;(b)无人干扰区

Fig. 2 Spatial pattern distributions of organic matter content in areas with or without human interference.

(a) Area with human interference; (b) area without human interference

整体上,无人干扰区的有机质含量较低。原因是该研究区常年干旱少雨、盐渍化严重,土壤常年处于疏干状态,普遍缺少有机胶结体,有机质不易形成团粒结构;盐分过多会派生出许多不良土壤性状而使土壤肥力不能发挥,危害植物生长发育^[27]。加之该区域没有人为干扰,土壤中微生物的活性较低,土壤中的动植物残体没有促进有机质的累积过程,使得该区域内的有机质含量较少。

无人干扰区有机质含量分布呈现中间低、四周高的格局。对影像进行重分类,计算各级别有机质分布区域的面积,有机质质量分数小于 10×10^{-3} 的面积较大,占无人干扰区的 45.8%。人为干扰区的有机质含量较无人干扰区有较大增加,原因是人为干扰区的植物覆盖量、植物种类要多于无人干扰区,有利于土壤中有有机质的累积。该区域的有机质分布呈现西北、东南含量较高,东北、西南较低的格局,有机质质量分数为 $10 \times 10^{-3} \sim 15 \times 10^{-3}$ 的面积最大,占人为干扰区总面积的 44.6%。人为干扰区中北部的地表全部被翻耕,并被整平且修成“田”字格的育苗地,打破了原有有机质含量的分布;人类活动增加了土壤中微生物的活性,加速了动植物残体的分解速度,使得有机质含量大大增加。该区域的南部是林龄小于 20 a 的人工梭梭林地,由于从 20 a 林龄开始,梭梭生长趋于缓慢,尽管小于 20 a 林龄的梭梭生长较旺盛,但是土壤中的全氮、有机质含量较低^[28]。

4 结 论

(1) 无人干扰区和人为干扰区中的有机质含量与盐分含量均呈极显著相关,相关系数分别达到 -0.867 和 -0.803 ,人类活动加速了土壤中的水分散失,加之降水较少,土壤中的盐分都聚集在表面,盐分过高不利于动植物的生存与生长,所以动植物、微生物残体质量过少,有机质含量较少。加之人类的耕作行为导致土壤中的盐分、有机质分布变得复杂,所以人为干扰区盐分与有机质的相关性小于无人干扰区。

(2) 在有机质含量与盐分指数、植被指数的相关性分析和主成分分析中,无人干扰区窄波段的 SI_2 、 SI_3 和 RVI 、 $NDVI$ 以及宽波段的 SI_1 、 SI_2 和 RVI 、 $NDVI$,人为干扰区窄、宽波段的 SI_1 、 SI_3 和 SI_1 、 SI_2 以及 RVI 、 $NDVI$ 和 RVI 、 $RDVI$ 与有机质含量的相关系数、特征向量值较高,用来建立有机质含量的 MLR 以及 PLSR,其中有、无人干扰区窄波段 PLSR 的 R^2 分别达到 0.753、0.819, R_{PD} 分别为 2.01、2.14。

(3) 基于最佳模型对研究区的土壤有机质含量进行空间反演与分析,反演所得的土壤有机质含量与样本的统计性结果基本一致,反演结果较为准确,人为干扰区的有机质含量要高于无人干扰区。无人干扰区的有机质分布呈现中间低、四周高的趋势,有机质的质量分数主要集中在小于 10×10^{-3} 的

范围内;人为干扰区的分布格局是西南、东北低,中北部高,有机质的质量分数主要集中在 $10 \times 10^{-3} \sim 15 \times 10^{-3}$ 的范围内。

(4) 无论是 MLR、还是 PLSR,窄波段的模型精度均略高于宽波段,但是宽波段模型的 R_{PD} 均大于 1.65,且无人为干扰区、人为干扰区 PLSR 模型的 R_{PD} 分别达到了 1.95、1.84,可以较为准确地预测土壤有机质的含量,说明宽波段模型对大面积、高精度地预测有机质含量具有较大的潜力。在下一步工作中,需要充分考虑其他因素的综合影响,建立引入其他理化参数的有机质光谱预测模型,对从高分辨率的遥感影像中提取的各类参数进行改进,大幅增加采样点的个数,增大研究区域面积或选取更多的影响因子进行分析,以提升预测的普适性和精准性。

参 考 文 献

- [1] Fang S W, Yang M H, Zhao X M, *et al.* Spectral characteristics and quantitative estimation of SOM in red soil typical of Ji'an Country, Jiangxi Province [J]. *Acta Pedologica Sinica*, 2014, 51(5): 1003-1010.
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——以江西省吉安为例[J]. *土壤学报*, 2014, 51(5): 1003-1010.
- [2] Wu Y Q, Zhou Y, Long Y L. Small target detection in hyperspectral remote sensing image based on adaptive parameter SVM [J]. *Acta Optica Sinica*, 2015, 35(9): 0928001.
吴一全, 周杨, 龙云淋. 基于自适应参数支持向量机的高光谱遥感图像小目标检测[J]. *光学学报*, 2015, 35(9): 0928001.
- [3] Pisek J, Chen J M. Comparison and validation of MODIS and VEGETATION global LAI products over four BigFoot sites in North America [J]. *Remote Sensing of Environment*, 2007, 109(1): 81-94.
- [4] Sims D A, Gamon J A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages [J]. *Remote Sensing of Environment*, 2002, 81(2): 337-354.
- [5] Zhang J J, Tian Y C, Yao X, *et al.* The spectral index for estimating soil OM, TN and AN content simultaneously using near-infrared spectroscopy [J]. *Acta Pedologica Sinica*, 2012, 49(1): 50-59.
张娟娟, 田永超, 姚霞, 等. 同时估测土壤全氮、有机质和速效氮含量的光谱指数研究 [J]. *土壤学报*, 2012, 49(1): 50-59.
- [6] Wang F, Ding J L, Wei Y, *et al.* Sensitivity analysis of soil salinity and vegetation indices to detect soil salinity variation by using Landsat series images: applications in different oases in Xinjiang, China [J]. *Acta Ecologica Sinica*, 2017, 37(15): 5007-5022.
王飞, 丁建丽, 魏阳, 等. 基于 Landsat 系列数据的盐分指数和植被指数对土壤盐度变异性的响应分析——以新疆天山南北典型绿洲为例 [J]. *生态学报*, 2017, 37(15): 5007-5022.
- [7] Zhang L J, Yue M, Zhang Y D, *et al.* Characteristics of plant community species diversity of oasis desert ecotone in Fukang, Xinjiang [J]. *Scientia Geographica Sinica*, 2003, 23(3): 329-334.
张林静, 岳明, 张远东, 等. 新疆阜康绿洲荒漠过渡带植物群落物种多样性特征 [J]. *地理科学*, 2003, 23(3): 329-334.
- [8] Yao F, Guli · Jiapaer, Bao A M, *et al.* Damage assessment of the vegetable types based on remote sensing in the open coalmine of arid desert area [J]. *China Environmental Science*, 2013, 33(4): 707-713.
姚峰, 古丽·加帕尔, 包安明, 等. 基于遥感技术的干旱荒漠区露天煤矿植被群落受损评估 [J]. *中国环境科学*, 2013, 33(4): 707-713.
- [9] Zhao F, Peng Y K. Measurement of iodine value of pork's subcutaneous adipose tissue by interval partial least square and Raman spectroscopy [J]. *Chinese Journal of Lasers*, 2017, 44(11): 1111001.
赵芳, 彭彦昆. 区间偏最小二乘法结合拉曼光谱测定猪肉皮下脂肪的碘值 [J]. *中国激光*, 2017, 44(11): 1111001.
- [10] Fabio C, Angelo P, Federico S, *et al.* Evaluation of the potential of the current and forthcoming multispectral and hyperspectral imagers to estimate soil texture and organic carbon [J]. *Remote Sensing of Environment*, 2016, 179: 54-65.
- [11] Xie X L, Sun B, Hao H T. Relationship between visible-near infrared reflectance spectroscopy and heavy metal of soil concentration [J]. *Acta Pedologica Sinica*, 2007, 44(6): 982-993.
解宪丽, 孙波, 郝红涛. 土壤可见光-近红外反射光谱与重金属含量之间的相关性 [J]. *土壤学报*, 2007, 44(6): 982-993.
- [12] Chen Z. Relationship model among water content, bulk density and reflectivity of different soil [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2012, 28(4): 76-81.
陈祯. 不同土壤含水量、体积质量及光谱反射率的关

- 系模型[J]. 农业工程学报, 2012, 28(4): 76-81.
- [13] Li P. Hyperspectral estimation and remote sensing retrieval of soil water regime in the Yellow River Delta[D]. Taian: Shandong Agricultural University, 2016: 26.
李萍. 黄河三角洲土壤含水量状况的高光谱估测与遥感反演[D]. 泰安: 山东农业大学, 2016: 26.
- [14] Xu H Q, Tang F. Analysis of new characteristics of the first Landsat 8 image and their eco-environmental significance [J]. Acta Ecologica Sinica, 2013, 33(11): 3249-3257.
徐涵秋, 唐菲. 新一代 Landsat 系列卫星: Landsat 8 遥感影像新增特征及其生态环境意义[J]. 生态学报, 2013, 33(11): 3249-3257.
- [15] Khan N M, Rastokuev V, Sato Y, *et al.* Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators [J]. Agricultural Water Management, 2005, 77: 96-109.
- [16] Inakwu O A Odeh, Onus A. Spatial analysis of soil salinity and soil structural stability in a semiarid region of New South Wales, Australia [J]. Environmental Management, 2008, 42(2): 65-78.
- [17] Chen H Y, Zhao G X, Chen J C, *et al.* Remote sensing inversion of saline soil salinity based on modified vegetation index in estuary area of Yellow River [J]. Transactions of the Chinese Society of Agricultural Engineering, 2015, 31(5): 107-114.
陈红艳, 赵庚星, 陈敬春, 等. 基于改进植被指数的黄河口区盐渍土盐分遥感反演[J]. 农业工程学报, 2015, 31(5): 107-114.
- [18] Baret F, Guyot G. Potentials and limits of vegetation indices for LAI and APAR assessment [J]. Remote Sensing of Environment, 1991, 35(2/3): 161-173.
- [19] Schlerf M, Atzberger C, Hill J. Remote sensing of forest biophysical variables using HyMap imaging spectrometer data [J]. Remote Sensing of Environment, 2005, 95(2): 177-194.
- [20] Roujean J L, Breon F M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements [J]. Remote Sensing of Environment, 1995, 51(3): 375-384.
- [21] Cai L H, Ding J L. Inversion of soil moisture content based on hyperspectral multi-scale decomposition [J]. Laser & Optoelectronics Progress, 2018, 55(1): 013001.
蔡亮红, 丁建丽. 基于高光谱多尺度分解的土壤含水量反演 [J]. 激光与光电子学进展, 2018, 55(1): 013001.
- [22] Viscarra R V, McGlynn R N, McBratney A B. Determining the composition of mineral-organic mixes using UV-vis-NIR diffuse reflectance spectroscopy [J]. Geoderma, 2006, 137(1/2): 70-82.
- [23] Jiang X Q, Ye Q, Lin Y, *et al.* Inverting study on soil water content based on harmonic analysis and hyperspectral remote sensing [J]. Acta Optica Sinica, 2017, 37(10): 1028001.
姜雪芹, 叶勤, 林怡, 等. 基于谐波分析和高光谱遥感的土壤含水量反演研究 [J]. 光学学报, 2017, 37(10): 1028001.
- [24] Li X, Ding J L. Soil moisture monitoring based on measured hyperspectral index and HSI image index [J]. Transactions of the Chinese Society of Agricultural Engineering, 2015, 31(19): 68-75.
李相, 丁建丽. 基于实测高光谱指数与 HSI 影像指数的土壤含水量监测 [J]. 农业工程学报, 2015, 31(19): 68-75.
- [25] Liang X, Ji H Y, Wang P X, *et al.* Study on building quantitative analysis model for chlorophyll in winter wheat with reflective spectrum using MSC-ANN algorithm [J]. Spectroscopy and Spectral Analysis, 2010, 30(1): 188-191.
梁雪, 吉海彦, 王鹏新, 等. 用 MSC-ANN 方法建立冬小麦叶片叶绿素与反射光谱的定量分析模型研究 [J]. 光谱学与光谱分析, 2010, 30(1): 188-191.
- [26] Tang Q Y. DPS data processing system [M]. Beijing: Science Press, 2010: 2.
唐启义. DPS 数据处理系统 [M]. 北京: 科学出版社, 2010: 2.
- [27] Masoud A A, Koike K. Arid land salinization detected by remotely sensed land cover changes: a case study in the Siwa region, NW Egypt [J]. Journal of Arid Environments, 2006, 66(1): 151-167.
- [28] Dong Z L. The study of artificial Haloxylon forest distribution of soil carbon and nitrogen storage and its influence factors in arid desert area [D]. Beijing: Chinese Academy of Forestry Sciences, 2014: 36.
董志玲. 干旱荒漠区人工梭梭林土壤碳氮储量分布规律及影响因子研究 [D]. 北京: 中国林业科学研究院, 2014: 36.