

DOI: 10.11686/cyxb2016509

http://cyxb.lzu.edu.cn

崔霞, 宋清洁, 张瑶瑶, 胥刚, 孟宝平, 高金龙. 基于高光谱数据的高寒草地土壤有机碳预测模型研究. 草业学报, 2017, 26(10): 20-29.
CUI Xia, SONG Qing-Jie, ZHANG Yao-Yao, XU Gang, MENG Bao-Ping, GAO Jin-Long. Estimation of soil organic carbon content in alpine grassland using hyperspectral data. Acta Prataculturae Sinica, 2017, 26(10): 20-29.

基于高光谱数据的高寒草地土壤 有机碳预测模型研究

崔霞^{1*}, 宋清洁¹, 张瑶瑶¹, 胥刚², 孟宝平², 高金龙²

(1. 兰州大学资源环境学院西部环境教育部重点实验室, 甘肃 兰州 730000; 2. 草地农业生态系统
国家重点实验室, 兰州大学草地农业科技学院, 甘肃 兰州 730020)

摘要: 土壤退化是草地退化的更深层次指示, 运用遥感手段大面积测定土壤有机碳进而评估草地土壤状况有助于对草地退化状态的正确认识。以甘南州高寒草地土壤为研究对象, 使用 ASD 地物光谱仪, 在室内条件下对土壤样品进行可见光/近红外光谱测量, 分析 8 种光谱变换形式与土壤有机碳含量的相关性并选取特征波段, 利用 3 种多元回归方法(逐步多元线性回归、主成分回归、偏最小二乘回归), 通过验证样本的决定系数(R_0^2)、均方根误差(RMSE)和剩余估计偏差(RPD)来评价模型, 进而确定高寒草地土壤有机碳的最佳估测模型。结果表明, 微分变换方法可以显著提高光谱特征与土壤有机碳含量的相关性, 在所有变换形式中以光谱反射率的一阶微分与土壤有机碳含量相关性最好, 最大相关系数绝对值为 0.865; 基于光谱反射率一阶微分变换形式的 3 种多元回归方法对土壤有机碳均有极好的预测能力, 表明对于土壤有机碳的稳定监测来说光谱反射率的一阶微分是非常有效的变换形式; 综合考虑基于所有光谱变换形式的 3 种多元回归方法的预测结果, 偏最小二乘回归法具有高的 R_0^2 和 RPD, 同时具有低的 RMSE 值, 是研究区土壤有机碳估测的最优回归方法; 基于光谱反射率对数的一阶微分变换形式所建立的偏最小二乘回归模型具有相对较高的预测集决定系数($R_0^2=0.878$)、最大剩余估计偏差(RPD=2.946)和最小均方根误差(RMSE=7.520), 因此该模型为甘南高寒草地土壤有机碳的最优估测模型, 最优模型的 RPD 大于 2.5 说明该模型有足够的稳定性可以应用于其他地区土壤有机碳的估测。

关键词: 高光谱; 光谱预处理; 多元逐步线性回归; 主成分回归; 偏最小二乘回归

Estimation of soil organic carbon content in alpine grassland using hyperspectral data

CUI Xia^{1*}, SONG Qing-Jie¹, ZHANG Yao-Yao¹, XU Gang², MENG Bao-Ping², GAO Jin-Long²

1. Key Laboratory of Western China's Environmental Systems (Ministry of Education), College of Earth and Environmental Sciences, Lanzhou University, Lanzhou 730000, China; 2. State Key Laboratory of Grassland Agro-ecosystems, College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou 730020, China

Abstract: Soil degradation is often reflects grassland degradation. Monitoring soil organic carbon (SOC) content over large areas using remote sensing technology can help assess soil condition allowing better understanding of grassland degradation. Alpine grassland in the Gannan Prefecture was selected for this research. NIR-Visible spectral reflectance of grassland soil samples was measured using ASD (analytical spectral devices) spectroradiometer under laboratory conditions. Correlation analyses between eight transformations of soil spec-

收稿日期: 2016-12-30; 改回日期: 2017-03-31

基金项目: 国家自然科学基金(41401472)和兰州大学中央高校基本科研业务费专项资金(lzujbky-2015-140)资助。

作者简介: 崔霞(1984-), 女, 甘肃民勤人, 讲师, 博士。

* 通信作者 Corresponding author. E-mail: xiacui2006@163.com

tral reflectance and SOC content were undertaken and sensitive wavebands selected. Three multivariate regression techniques (stepwise multiple linear regression, SMLR, principal components regression, PCR, partial least squares regression, PLSR) were compared with the aim of identifying the best inversion model to estimate alpine grassland SOC. The determination coefficient of validation dataset (R_v^2), the root mean square error (RMSE), and the residual prediction deviation (RPD) were used to evaluate the models. The result indicated that differential transformation could improve the correlation between spectral characteristics and SOC content. The first derivative of reflectance had the best correlation with SOC content during transformation, the maximum correlation coefficient value was 0.865; Three multivariate regression models based on the first derivative of reflectance had excellent SOC prediction capability and recommended as a good spectral transformation for reliably estimating SOC. Comparing the multivariate regression techniques based on all transformations, PLSR performed best (high R_v^2 and RPD, low RMSE), which suggests that PLSR is the most suitable method for estimating SOC content in this study. The optimal SOC estimation model of Gannan alpine grassland was the combination of PLSR and the first derivative of log reflectance spectra $[(\lg R)']$, providing a relatively high coefficient of determination for the validation set ($R_v^2=0.878$), the highest residual prediction deviation (RPD=2.946) and the lowest root mean square error (RMSE=7.520). The RPD of the optimal model was higher than 2.5, which suggested that the model was robust and stable enough to be applied for estimation of SOC in other areas.

Key words: hyperspectral; spectral pre-processing; stepwise multiple linear regression; principal components regression; partial least squares regression

甘南藏族自治州位于青藏高原东北部,地处黄河、长江源头,是甘肃省重要的畜牧业生产基地,也是全国主要少数民族集聚的草原牧区之一。20 世纪 80 年代以来,甘南州生态环境不断恶化,90%的天然草地存在不同程度的退化,重度退化草地及中度退化草地面积分别约占全州草地面积的 30%和 50%^[1],亟待科学监理方法的运用与提升。通过遥感手段实时动态监测甘南草地退化的时空特征,对及时掌握其现状和开展针对性管理具有重要作用。草地退化按层次可分为植被的退化和土壤的退化,尽管土壤退化滞后于植被退化,却是比植被退化更深层次和更严重的退化^[2]。而在目前评价草地退化程度及采取改良措施时,对土壤的地位及作用还没有给予应有的重视^[3]。因此在综合评价草地退化状况时,应考虑整合土壤退化指标。

有机碳含量是反映土壤状况的最重要指标之一。它既能指示土壤养分状况,同时还是形成土壤结构的重要因素,直接影响土壤肥力、持水能力、土壤抗侵蚀能力和土壤容重等土壤特性^[4],其变化状况可以指示土壤退化与否^[5]。遥感手段测定土壤有机碳(soil organic carbon, SOC)克服了农化分析方法耗时、耗资大、费力、周期长的缺点^[6-7],一直倍受学术界研究者的关注。随着遥感技术的发展,高光谱遥感在土壤养分监测中发挥越来越大的作用。自 20 世纪 80 年代开始土壤光谱研究从最初的对光谱曲线的定性分析发展到对土壤组分含量的定量反演,至今学者们利用高光谱技术估测土壤有机碳含量,对土壤光谱进行了详细分析,并建立了相应的预测模型,取得了较为满意的结果。如 Al-Abbas 等^[8]发现土壤光谱反射率与有机质含量呈显著负相关;Galvão 等^[9]研究表明土壤反射率在 0.55~0.70 μm 处的吸收峰主要是由有机质引起的;Cambou 等^[10]采用 PLSR(partial least squares regression)法结合野外自制取土工具所测量的 ASD(analytical spectral devices)野外高光谱数据对法国土壤有机碳进行了预测,发现预测模型准确,预测值与实测值的 R^2 (determination coefficient)达到 0.75,剩余估计偏差(relative prediction deviation)为 2.0。以上的研究表明高光谱遥感由于光谱分辨率高,对土壤养分信息进行定量获取是不可或缺的数据源。

利用高光谱技术来获取土壤有机碳含量的研究对象主要为农田^[11-13],对草地尤其是有机碳含量较高的高寒草地研究较少。现有的高光谱遥感在草地研究中的应用主要集中在对草地植被的研究方面如生物量估算、草地种类识别、草地化学成分估测^[14-15],对土壤的研究较少,在仅有的对于草地土壤高光谱研究中,也是只对草地不同

类型土壤反射光谱特性进行分析^[16],没有对草地养分进行定量分析。因此,本研究利用美国 ASD FieldSpecPro (FR)地物光谱仪在室内条件下对经过处理的高寒草地土壤样品进行光谱测量,通过对土壤样品光谱反射率不同变换形式与有机碳含量进行相关性分析,进而选取特征波段;通过比较主成分回归(principal components regression, PCR)、逐步多元线性回归(stepwise multiple linear regression, SMLR)和偏最小二乘回归(partial least squares regression, PLSR)3种方法对土壤有机碳的模拟效果,建立高寒草地土壤光谱变量与有机碳含量的多元回归模型,探索高寒草地土壤有机碳快速定量遥感估算方法。

1 材料与方法

1.1 研究区概况

甘南藏族自治州位于甘肃省西南部,地处青藏高原与黄土高原过渡带,其行政区划范围 $33^{\circ}06' - 35^{\circ}44' N$, $100^{\circ}46' - 104^{\circ}44' E$ 。海拔 1100~4900 m,大部分在 3000 m 以上;属于典型的高原大陆性气候,年降水量在 500 mm 以上;年均温仅 $1 \sim 3^{\circ}C$, $\geq 10^{\circ}C$ 的气温持续期仅两个多月,全年平均日照时数 2200~2400 h^[17];甘南州畜牧业资源丰富,有牧草地 272 万 hm^2 ,占该州国土总面积的 61.9%,该区水草茂盛,被誉为亚洲最好的牧场之一。

1.2 土样样本采集

2015 年 7 月在甘南州不同退化程度的高寒草地选取典型样地(图 1),样地选取在地势平坦,草地植被分布比较均一能代表较大范围植被覆盖状况的区域。共采集土壤表层土 52 份,采用 GPS 定位坐标,同时记录样地的基本特征信息。在 $30 m \times 30 m$ 的样地中采用五点混合取样法获得一个样本,装入密封袋内,带回实验室自然风干,将风干的样品取出少部分过筛后采用重铬酸钾氧化-外加加热法测定土壤有机碳含量^[18];另外一部分过 2 mm 孔筛,用于高光谱数据采集。

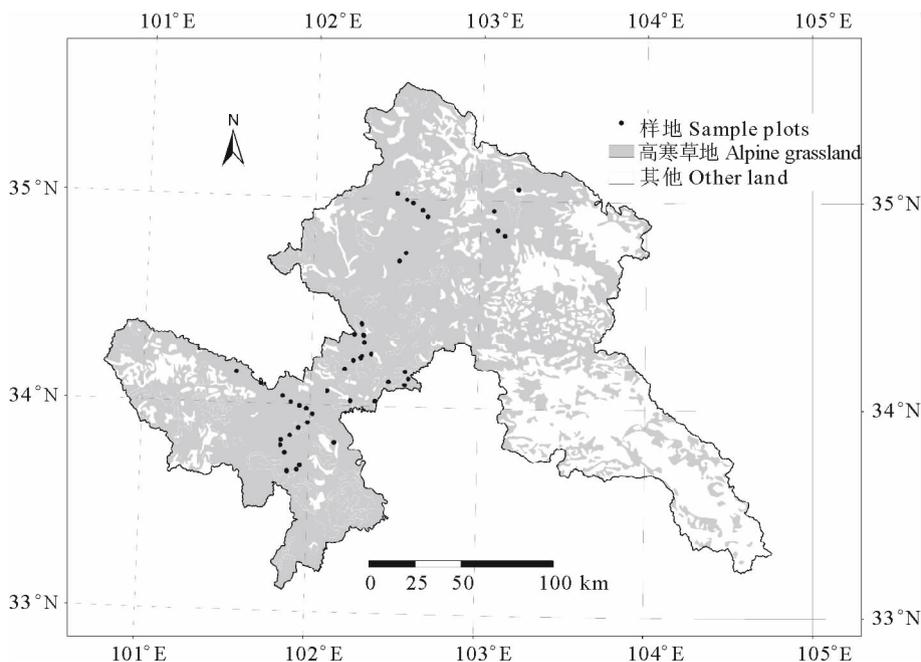


图 1 甘南样地分布图

Fig. 1 Location of the plots in the Gannan region

1.3 光谱测定

土壤光谱反射率的测定采用美国 ASD 公司生产的 FieldSpecPro FR 地物光谱仪。波谱范围为 350~2500 nm,采样间隔为 1.4 nm(350~1000 nm)和 2.0 nm(1000~2500 nm)。将过筛后的土样装满直径为 10 cm 的容器中,土样厚度约为 1.5 cm,光谱测量在暗室内进行。光源为能够提供平行光的 50 W 的卤素灯,距土壤表面 30

cm,光源的天顶角为 15° ,采用 5° 视场角光纤探头,探头位于土壤表面垂直上方 15 cm 处。在测定前及每测定 6 个土壤样本后进行一次白板校正,每个土壤样本采集 12 条光谱曲线^[19]。

1.4 建模集和验证集的划分

参考 Shi 等^[19]对土壤全 N 含量高光谱预测中建模集与验证集划分的方法,将所有土壤样本有机碳含量按从小到大的顺序排列,根据总样本大小选定一个数字 n (本文中 n 的取值范围为 1~17),将第 n 和第 $2n$ 个样本作为建模样本,第 $3n$ 个样本作为验证样本(表 1),这种划分方法可以保证所有区间的有机碳含量均能出现在建模集和验证集中。

表 1 土壤有机碳含量统计特征

Table 1 Statistical characteristics of soil carbon content

| 样本类型 Type of dataset | 样本数 Observations number | 最小值 Min (g/kg) | 最大值 Max (g/kg) | 均值 Mean (g/kg) | 标准差 STD (g/kg) |
|--------------------------|----------------------------|-------------------|-------------------|-------------------|-------------------|
| 总体样本 Whole dataset | 52 | 14.327 | 109.629 | 48.618 | 22.602 |
| 建模样本 Calibration dataset | 35 | 14.327 | 109.629 | 48.500 | 23.136 |
| 验证样本 Validation dataset | 17 | 17.169 | 108.469 | 48.860 | 22.154 |

STD: Standard deviation.

1.5 光谱数据预处理

光谱数据的预处理主要包括去噪和数据变换处理。首先使用 ViewSpec 软件中的 Splice Correction 修正功能对土壤反射光谱数据进行修正,并将每个土壤样本的 12 条光谱曲线算术平均后得到该样本的实际反射光谱数据(R)。为减少实验室光学环境差异和磨样过筛的影响,对所有的光谱数据采用 Savitzky-Golay 平滑处理^[20],为提高光谱灵敏度,对所有波段光谱数据进行一阶微分(R')、倒数($1/R$)、倒数的一阶微分 $[(1/R)']$ 、对数($\lg R$)、对数的一阶微分 $[(\lg R)']$ 、对数的倒数($1/\lg R$)、对数的倒数的一阶微分 $[(1/\lg R)']$ 的变换,同时结合 R 共 8 种光谱变量用来建立土壤有机碳含量预测模型。

1.6 预测模型建立与验证

利用单相关分析,得到 8 种光谱变量与土壤有机碳含量的相关系数,选择相关系数大于 0.6 的特征波段,运用 SPSS 17.0 数据处理软件对特征波段与土壤有机碳含量进行 SMLR 分析及 PCR 分析,同时利用 Matlab R2015a 软件进行 PLSR 分析,建立土壤有机碳含量高光谱预测模型,并对预测模型进行验证。

预测模型的验证主要选用验证样本的决定系数(determination coefficient of validation dataset, R_v^2)、均方根误差(root mean squares error, RMSE)以及剩余估计偏差(relative prediction deviation, RPD)为模型评价参数, R_v^2 、RMSE 及 RPD^[21]的计算公式如下:

$$R_v^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (2)$$

$$RPD = STD/RMSE \quad (3)$$

其中: \hat{y}_i 为验证集中第 i 个样本的估测值; y_i 为验证集中第 i 个样本的实测值; \bar{y} 为验证集中实测值的平均值;STD 为验证集中实测值的标准差。 R_v^2 越高, RMSE 越小,表明模型估算的准确性越高,反之则模型估算的准确性越差。当 $RPD < 1.0$ 时,表明模型预测能力很差,模型不可靠;当 $1.0 < RPD < 1.4$ 时,表明模型预测能力较差,仅能对样品高含量和低含量进行粗略估测;当 $1.4 < RPD < 1.8$ 时,表明模型预测能力较好,可以对样本进行估测;当 $1.8 < RPD < 2.0$ 时,表明模型预测能力好,可对样本进行定量估测,当 $2.0 < RPD < 2.5$ 时,表明模型具有很好的定量预测能力;当 $RPD > 2.5$ 时,模型具有极好的预测能力^[20]。

2 结果与分析

2.1 土壤有机碳含量与光谱变量的相关性分析

8种光谱变量与土壤有机碳含量在部分波段达到很好的相关性,相关系数绝对值 $|r|$ 最大值均在0.6以上(图2)。R与土壤有机碳含量在1580~1890 nm和2200~2224 nm波段呈微弱的正相关,在其他波段呈负相关,且存在高相关性的宽波段,lgR和土壤有机碳含量的相关性与R相似,1/R、1/lgR与土壤有机碳含量的相关性,与R和土壤有机碳相关性相反。R经1/R、1/lgR变换后相关系数的绝对值并没有得到提高,但对应的微分变换形式使得部分波段与土壤有机碳的相关性得到了改善,尤其是在近红外区域部分波段经微分变换后由原来的不显著水平而增强为极显著水平($P=0.01$),且相关系数数值变化剧烈,呈正负波动起伏,说明微分变换可将近红外波段范围内一些含有机碳光谱吸收特征的信息被释放出来。微分变换形式与土壤有机碳有很强的相关性,相关系数绝对值最大值依次为0.865(R')、0.854 $[(\lg R)']$ 、0.799 $[(1/\lg R)']$ 和0.787 $[(1/R)']$,且相关系数绝对值最大值均出现在1913 nm处。

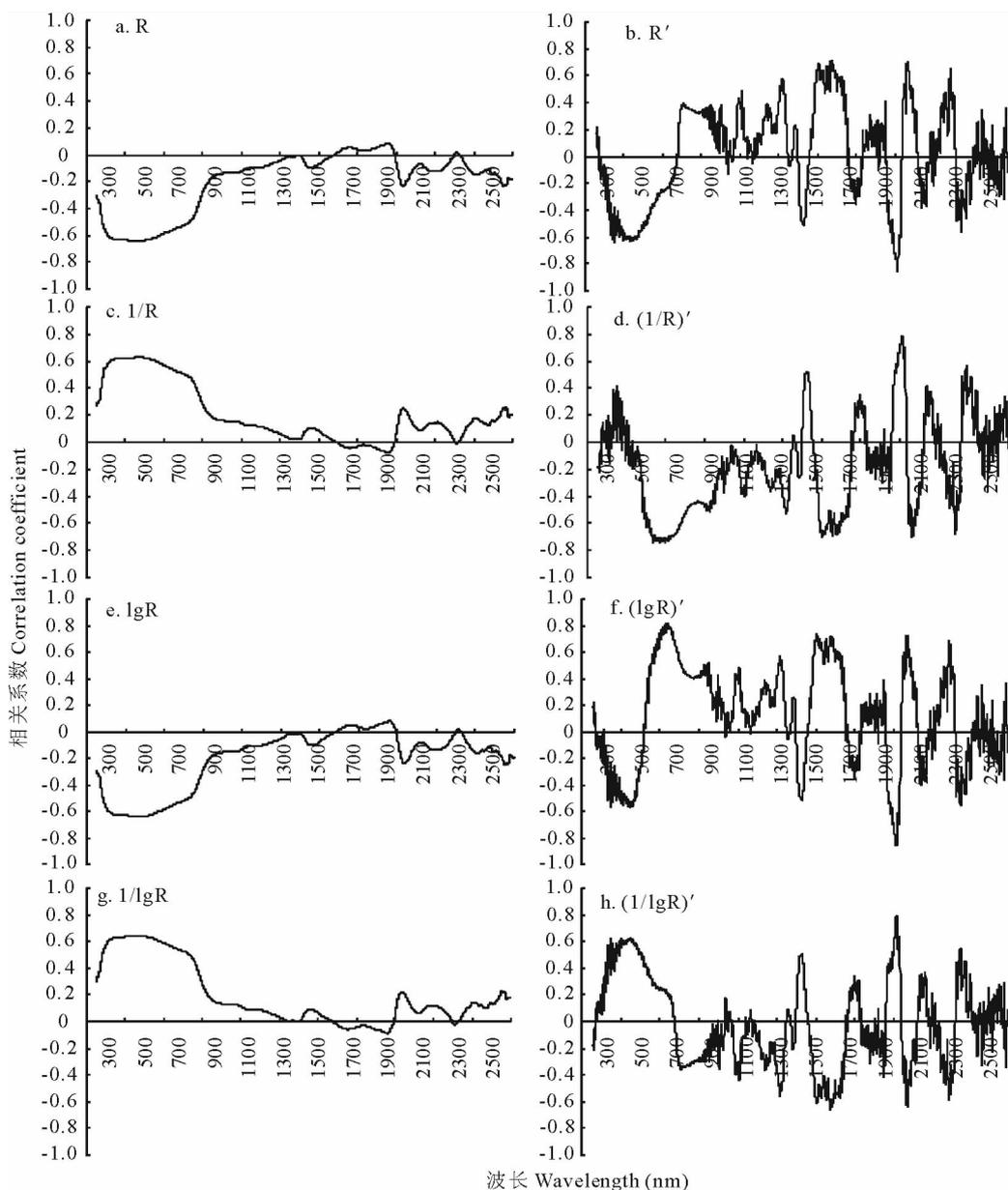


图2 光谱反射率及不同变换形式与土壤有机碳含量的相关性分析($n=52$)

Fig. 2 Correlation analysis on SOC contents and raw reflectance and pre-processing transformations ($n=52$)

2.2 高寒草地土壤有机碳含量 SMLR、PCR、PLSR 模型建立

表 2~表 4 显示了基于 8 种光谱变量的土壤有机碳含量 SMLR、PCR、PLSR 模型的精度评价结果。在 SMLR 模型中 8 种光谱变量建立的模型 R_v^2 都很高(均在 0.7 以上),但 RMSE 和 RPD 的值却参差不齐。其中基于 R' 的模型 R_v^2 最高,为 0.885, RMSE 值最低(10.101 g/kg), RPD 最高,为 2.193,是 8 种模型中唯一一个大于 2 的,说明该模型具有很好的定量预测能力。因此,该模型为 SMLR 模型中的最佳模型。

$$SOC = 0.227 - 76190.805 \times b_{1913} - 217348.163 \times b_{529} + 185766.881 \times b_{601} \quad (4)$$

式中: SOC 是有机碳含量的预测值, b_{1913} , b_{529} , b_{601} 是 3 个波段反射率的一阶微分变换值。

PCR 模型中基于微分变换形式建立的模型具有相对高的 R_v^2 和 RPD 值及相对较低的 RMSE, 这种模型明显优于未进行微分变换的光谱变量所建立的模型。且在所有预测模型中, 基于 $(\lg R)'$ 模型的 R_v^2 值最高, 为 0.853, 均方根误差 RMSE 值最低(8.229 g/kg), 剩余估计偏差最高为 2.692, 说明该模型具有极好的预测能力, 基于 $(\lg R)'$ 的模型为 PCR 模型中的最佳模型。

表 2 土壤有机碳含量 SMLR 预测模型精度评价

Table 2 Summary statistics of accuracy for SOC estimating models produced by SMLR

| R 变换形式 transformations of R | 验证集 Validation dataset | | |
|--------------------------------|------------------------|-------------|-------|
| | R_v^2 | RMSE (g/kg) | RPD |
| R | 0.710 | 14.298 | 1.549 |
| R' | 0.885 | 10.101 | 2.193 |
| 1/R | 0.754 | 17.295 | 1.281 |
| $(1/R)'$ | 0.803 | 17.625 | 1.257 |
| lgR | 0.754 | 14.630 | 1.514 |
| $(\lg R)'$ | 0.818 | 19.582 | 1.131 |
| 1/lgR | 0.717 | 14.806 | 1.496 |
| $(1/\lg R)'$ | 0.785 | 17.516 | 1.265 |

RMSE: 均方根误差 The root mean square error.

在 PLSR 模型中, 1/lgR 经过微分变换后的模型没有提高反而降低了模型的精度, 而 R、1/R 和 lgR 经过微分变换后模型的精度有了很大的提升。所有模型中基于 $(\lg R)'$ 模型的 R_v^2 值最高, 为 0.878, 均方根误差 RMSE 值最低(7.520 g/kg), 剩余估计偏差最高为 2.946, 因此基于 $(\lg R)'$ 建立的模型为 PLSR 模型中的最佳模型。

3 种模型中, SMLR 的 R_v^2 较高, 均大于 0.7, 但 RPD 相对较低且 RMSE 普遍较高, 因此 SMLR 模型由于其准确度低不适合对本研究区土壤有机碳进行精确预测; 而在 PCR 模型中微分变换形式对应的 R_v^2 较高(大于 0.65), RPD 均大于 1.4, RMSE 相对较低, 大部分微分变换形式所对应的 PCR 模型比相应的 SMLR 模型有了明显的改善, 其中利用 $(\lg R)'$ 建立的 PCR 模型, RMSE 比对应的 SMLR 模型降低了 11.353

g/kg, 同时 RPD 提高了 1.561。对于 PCA 模型来说光谱微分变换是对光谱处理的有效手段, 可以增强土壤有机

表 3 土壤有机碳含量 PCR 预测模型精度评价

Table 3 Summary statistics of accuracy for SOC estimating models produced by PCR

| R 变换形式 transformations of R | 验证集 Validation dataset | | |
|--------------------------------|------------------------|-------------|-------|
| | R_v^2 | RMSE (g/kg) | RPD |
| R | 0.457 | 15.835 | 1.399 |
| R' | 0.773 | 10.920 | 2.030 |
| 1/R | 0.464 | 15.734 | 1.408 |
| $(1/R)'$ | 0.660 | 12.730 | 1.740 |
| lgR | 0.557 | 14.349 | 1.544 |
| $(\lg R)'$ | 0.853 | 8.229 | 2.692 |
| 1/lgR | 0.461 | 15.772 | 1.405 |
| $(1/\lg R)'$ | 0.720 | 11.686 | 1.896 |

表 4 土壤有机碳含量 PLSR 预测模型精度评价

Table 4 Summary statistics of accuracy for SOC estimating models produced by PLSR

| R 变换形式 transformations of R | 验证集 Validation dataset | | |
|--------------------------------|------------------------|-------------|-------|
| | R_v^2 | RMSE (g/kg) | RPD |
| R | 0.332 | 19.405 | 1.142 |
| R' | 0.858 | 9.332 | 2.374 |
| 1/R | 0.701 | 11.742 | 1.887 |
| $(1/R)'$ | 0.760 | 11.630 | 1.905 |
| lgR | 0.741 | 10.984 | 2.017 |
| $(\lg R)'$ | 0.878 | 7.520 | 2.946 |
| 1/lgR | 0.667 | 12.272 | 1.805 |
| $(1/\lg R)'$ | 0.494 | 16.771 | 1.321 |

碳含量的预测精度;PLSR 模型除了利用 R 和 $(1/\lg R)'$ 建模时效果不佳外,其他变换形式建立的 PLSR 模型在 3 种模型中精度最高, R_v^2 普遍大于 0.65, RMSE 最低,RPD 最高,且均大于 1.8。相比较其他两种模型,PLSR 模型的 RMSE 降低幅度和 RPD 增加的幅度都很大,尤其是利用 $(\lg R)'$ 建立的预测模型比 SMLR 模型的 RMSE 降低 12.062 g/kg,RPD 提高了 1.815。

2.3 高寒草地土壤有机碳含量最佳预测模型选择

对于 R 的不同变换形式中,利用 R' 建立的 SMLR、PCR 和 PLSR 三种不同回归模型差异不大,RPD 均大于 2(2.193、2.03、2.374),说明利用 R' 能很好地预测土壤有机碳含量,3 种模型验证集的 R_v^2 都很高,分别为 0.858、0.773、0.885,同时 RMSE 也较低(10.101、10.92、9.332 g/kg)说明利用 R' 是土壤有机碳含量预测中的稳定参数。

基于 R' 的 SMLR 模型和基于 $(\lg R)'$ 的 PCR 及 PLSR 模型分别是 3 种回归模型中的最优土壤有机碳含量预测模型,这 3 个回归模型光谱预测值与实测值之间均具有很好的相关性(表 2~表 4、图 3),数据点都较为均匀地分布在 1:1 直线的两侧,预测效果较好。总体来说由于基于 $(\lg R)'$ 的 PLSR 模型验证样本的 R_v^2 较高,均方根误差 RMSE 最小,剩余估计偏差 RPD 最高,因此该模型是研究区高寒草地土壤有机碳含量的最佳预测模型。

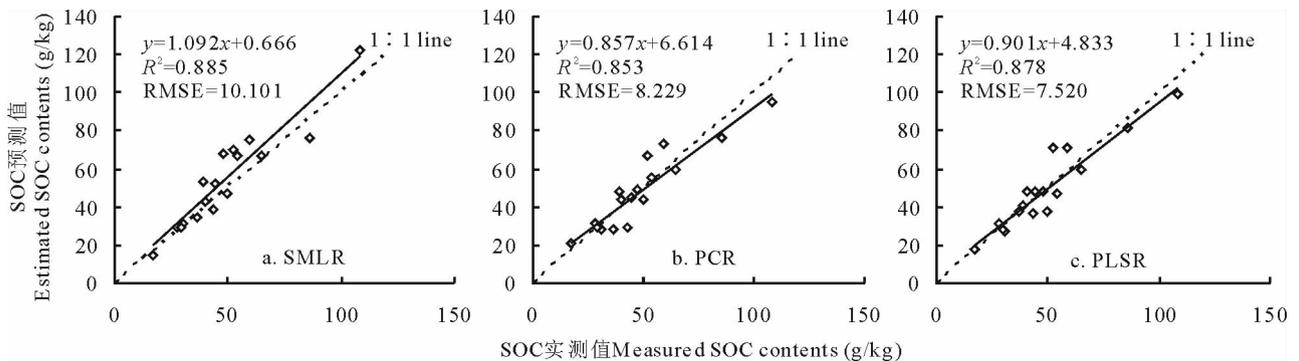


图 3 土壤有机碳 SMLR、PCR、PLSR 回归最优模型实测值与预测值比较

Fig. 3 Comparison of measured SOC contents and estimated SOC contents by SMLR、PCR、PLSR regression model

3 讨论

3.1 光谱微分变换对光谱与土壤有机碳含量相关性的改进

对光谱进行微分变换后可明显增强其与土壤有机碳含量的相关性,很多研究结果也表明微分光谱能较好地消除土壤质地、成土母质等潜在因素对光谱的影响,可以增强光谱对有机碳的敏感性,使得一些原本被遮蔽的土壤有机碳光谱反射特征显现出来^[12,22-23]。在本研究中近红外波段对光谱微分变换形式更加敏感,光谱微分变换前近红外波段与土壤有机碳含量的 $|r|$ 低(小于 0.4),且均低于可见光波段与土壤有机碳含量对应的 $|r|$ 值,但经过微分变换后相关系数有明显的增强, $|r|$ 超过可见光对应的值。这进一步验证了一些研究的结论:近红外波段对于估算土壤有机碳含量具有相当大的潜力^[24]。

光谱的其他变换形式对于相关系数影响不大,因而倒数变换、对数变换对于高寒草地土壤有机碳高光谱预测模型的改进来说意义不大。在所有的微分变换中, $|r|$ 的最大值均出现在 1913 nm 处,说明此波段是高寒草地土壤有机碳反演的敏感波段。

3.2 高寒草地土壤有机碳含量预测光谱变量及最优监测模型的选择

本研究涉及 8 种光谱变量,经精度评价发现利用 R' 建立的 SMLR、PCR 和 PLSR 三种不同回归模型均能很好地预测土壤有机碳含量,说明 R' 是预测土壤有机碳含量比较稳定的参数,且在现有的对于土壤参数反演研究中 R' 是使用最广泛的参数之一^[20,25],因此 R' 可以有效提高土壤参数预测模型的可靠性与普适性。综合考虑 8 种光谱变量的 3 种回归模型的精度验证结果,本研究中 $(\lg R)'$ 的 PLSR 模型是所有模型中的最优估算模型,验证

集的 R_v^2 达到 0.878, 说明利用土壤的高光谱特性对高寒草地土壤有机碳含量进行估测具有可行性。对于不同区域不同土壤类型有机碳含量估测研究中 $(\lg R)'$ 使用也很广泛^[12-13, 26], 说明 $(\lg R)'$ 对于土壤有机碳含量监测来说是有效的光谱变换形式。本研究以甘南高寒草地土壤为对象, 采用可见光/近红外高光谱技术, 在室内进行土样的光谱采集与分析, 虽然野外采集的土壤样本是从不同退化程度的高寒草地获取的, 包含不同的土壤类型, 土壤有机碳含量范围大(表 1), 但建立的土壤有机碳含量反演最优模型的 RPD 大于 2.5, 说明该模型有足够的稳定性可以应用于其他类似的土壤有机碳含量监测中。高寒草地土壤有机碳含量监测模型的建立说明可见光/近红外高光谱技术在土壤养分监测中有良好的应用前景, 可以快速、准确、动态和经济地获取土壤养分, 同时随着高光谱遥感的发展, 结合土壤养分监测模型, 为土壤养分快速定量反演打下良好的基础, 利用机载和星载高光谱遥感数据有望实现大区域高寒草地土壤养分制图。

3.3 不同建模方法比较

3 种模型中, PLSR 模型验证精度总体来说较高, 对于高寒草地土壤有机碳含量的预测是比较理想的建模方法。PLSR 在处理样本容量小、自变量多、变量间存在严重多重相关性方面具有独特的优势, 集中了普通多元线性回归、主成分分析、典型相关分析 3 种分析方法的优点, 能更有效地提取和利用光谱中的重要信息^[6, 27]。PCR 模型对于微分变换的光谱变量来说预测效果较好, PCR 模型结合光谱微分变换形式可以得到更高的模型估测精度; SMLR 模型在 3 种模型中 RMSE 值相对较高、RPD 相对较低, 模型预测效果较差, 可能由于土壤有机质成分复杂, 光谱特性很难用几个波段解释^[27]。

4 结论

1) 对光谱进行微分变换的方法, 可以明显提高光谱特征与土壤有机碳含量的相关性, 在所有变换形式中, 以 R' 与土壤有机碳相关性最好, $|r|$ 最大值为 0.865, 其次是 $(\lg R)'$ 、 $(1/\lg R)'$ 、 $(1/R)'$, $|r|$ 最大值分别为 0.854、0.799 和 0.787, 这 4 种变换形式 $|r|$ 最大值均出现在 1913 nm 处, 说明 1913 nm 是土壤有机碳含量预测的敏感波段。

2) 基于 R' 的 SMLR、PCR 和 PLSR 回归模型 R_v^2 较高, 依次为 0.885、0.773、0.858; 均方根误差均较低, 依次为 10.101、10.920、9.332 g/kg; 而剩余估计偏差均达到 2 以上, 说明用 R' 对土壤有机碳含量有很好的定量估测能力, 用 R' 来估测高寒草地土壤有机碳含量是一个相对稳定的参数。在本研究中基于 $(\lg R)'$ 建立的 PLSR 模型具有高的 R_v^2 和 RPD 值以及低的 RMSE 值, 是甘南高寒草地土壤有机碳含量的最优预测模型。

3) 3 种多元回归方法中, PLSR 模型具有较高的 R_v^2 和 RPD 及较小的 RMSE 是理想的土壤有机碳含量的预测建模方法。

参考文献 References:

- [1] Wang W H. The problem and countermeasure of grassland in Gannan Prefecture. *Yangtze River*, 2009, 40(7): 36-37.
王文浩. 甘南草原面临的问题及对策. *人民长江*, 2009, 40(7): 36-37.
- [2] Li S L, Chen Y J, Guan S Y, *et al.* Relationships between soil degradation and rangeland degradation. *Journal of Arid Land Resources and Environment*, 2002, 16(1): 92-95.
李绍良, 陈有君, 关世英, 等. 土壤退化与草地退化关系的研究. *干旱区资源与环境*, 2002, 16(1): 92-95.
- [3] Gao Y Z, Han X G, Wang S P. The effects of grazing on grassland soils. *Acta Ecologica Sinica*, 2004, 24(4): 790-797.
高英志, 韩兴国, 汪诗平. 放牧对草原土壤的影响. *生态学报*, 2004, 24(4): 790-797.
- [4] Lu R K. Principles of Soil-Plant Nutrition and Fertilization[M]. Beijing: Chemical Industry Press, 1998.
鲁如坤. 土壤-植物营养学[M]. 北京: 化学工业出版社, 1998.
- [5] Zhou H K, Zhao X Q, Wen J, *et al.* The characteristics of soil and vegetation of degenerated alpine steppe in the Yellow River Source Region. *Acta Prataculturae Sinica*, 2012, 21(5): 1-11.
周华坤, 赵新全, 温军, 等. 黄河源区高寒草原的植被退化与土壤退化特征. *草业学报*, 2012, 21(5): 1-11.
- [6] Yu L, Hong Y S, Geng L, *et al.* Hyperspectral estimation of soil organic matter content based on partial least squares regres-

- sion. Transactions of the Chinese Society of Agricultural Engineering, 2015, 31(14): 103-109.
- 于雷, 洪永胜, 耿雷, 等. 基于偏最小二乘回归的土壤有机质含量高光谱估算. 农业工程学报, 2015, 31(14): 103-109.
- [7] Meng J H, Wu B F, Du X, *et al.* A review and outlook of applying remote sensing to precision agriculture. Remote Sensing for Land & Resources, 2011, (3): 1-7.
- 蒙继华, 吴炳方, 杜鑫, 等. 遥感在精准农业中的应用进展及展望. 国土资源遥感, 2011, (3): 1-7.
- [8] Al-Abbas A H, Swain P H, Baumgardner M F. Relating organic matter and clay content to the multispectral radiance of soils. Soil Science, 1972, 114(6): 447-485.
- [9] Galvão L S, Pizarro M A, Epiphany J C N. Variations in reflectance of tropical soils: spectral-chemical composition relationships from AVIRIS data. Remote Sensing of Environment, 2001, 75(2): 245-255.
- [10] Cambou A, Cardinael R, Kouakoua E, *et al.* Prediction of soil organic carbon stock using visible and near infrared reflectance spectroscopy (VNIRS) in the field. Geoderma, 2016, 261: 151-159.
- [11] Ji W J, Li X, Li C X, *et al.* Using different data mining algorithms to predict soil organic matter based on visible-near infrared spectroscopy. Spectroscopy and Spectral Analysis, 2012, 32(9): 2393-2398.
- 纪文君, 李曦, 李成学, 等. 基于全谱数据挖掘技术的土壤有机质高光谱预测建模研究. 光谱学与光谱分析, 2012, 32(9): 2393-2398.
- [12] Wang X P, Zheng X P, Liu F J, *et al.* Analysis and extraction of soil organic matter information based on hyperspectral remote sensing. Geospatial Information, 2012, 10(5): 75-78.
- 王小攀, 郑晓坡, 刘福江, 等. 高光谱遥感土壤有机质含量信息提取与分析. 地理空间信息, 2012, 10(5): 75-78.
- [13] Chen H Y, Zhao G X, Zhang X H, *et al.* Improving estimation precision of soil organic matter content by removing effect of soil moisture from hyperspectra. Transactions of the Chinese Society of Agricultural Engineering, 2014, 30(8): 91-100.
- 陈红艳, 赵庚星, 张晓辉, 等. 去除水分影响提高土壤有机质含量高光谱估测精度. 农业工程学报, 2014, 30(8): 91-100.
- [14] Zhou L, Xin X P, Li G, *et al.* Application progress on hyperspectral remote sensing in grassland monitoring. Pratacultural Science, 2009, 26(4): 20-27.
- 周磊, 辛晓平, 李刚, 等. 高光谱遥感在草原监测中的应用. 草业科学, 2009, 26(4): 20-27.
- [15] Qian Y R, Yu J, Jia Z H, *et al.* Extraction and analysis of hyper-spectral data from typical desert grassland in Xinjiang. Acta Prataculturæ Sinica, 2013, 22(1): 157-166.
- 钱育蓉, 于炯, 贾振红, 等. 新疆典型荒漠草地的高光谱特征提取和分析研究. 草业学报, 2013, 22(1): 157-166.
- [16] Xu J, Wu H Q, Jiang P A, *et al.* Soil spectral property of Xinjiang Grassland. Acta Agrestia Sinica, 2014, 22(5): 980-985.
- 胥静, 武红旗, 蒋平安, 等. 新疆典型草地土壤光谱特征研究. 草地学报, 2014, 22(5): 980-985.
- [17] Cui X, Liang T G, Liu Y. Modeling of aboveground biomass of grassland using remotely sensed MOD09GA data. Journal of Lanzhou University: Natural Sciences, 2009, 45(5): 79-87.
- 崔霞, 梁天刚, 刘勇. 基于 MOD09GA 产品的草地生物量遥感估算模型. 兰州大学学报: 自然科学版, 2009, 45(5): 79-87.
- [18] Bao S D. Soil and Agriculture Chemistry Analysis[M]. Beijing: China Agriculture Press, 1999: 30-34.
- 鲍士旦. 土壤农化分析[M]. 北京: 中国农业出版社, 1999: 30-34.
- [19] Shi T, Cui L, Wang J, *et al.* Comparison of multivariate methods for estimating soil total nitrogen with visible/near-infrared spectroscopy. Plant and Soil, 2013, 366: 363-375.
- [20] 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. Science China: Earth Sciences, 2014, 44(5): 978-988.
- 史舟, 王乾龙, 彭杰, 等. 中国主要土壤高光谱反射特性分类与有机质光谱预测模型. 中国科学: 地球科学, 2014, 44(5): 978-988.
- [21] 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. Soil Biology & Biochemistry, 2014, 68: 337-347.
- [22] Sun J Y, Li M Z, Zheng L H, *et al.* Real time analysis of soil moisture, soil organic matter, and soil total nitrogen with NIR spectra. Spectroscopy and Spectral Analysis, 2006, 26(3): 426-429.
- 孙建英, 李民赞, 郑立华, 等. 基于近红外光谱的北方潮土土壤参数实时分析. 光谱学与光谱分析, 2006, 26(3): 426-429.
- [23] Hou Y J, Tashpolat T, Mamat S, *et al.* Estimation model of desert soil organic matter content using hyperspectral data. Transactions of the Chinese Society of Agricultural Engineering, 2014, 30(16): 113-120.
- 侯艳军, 买买提, 沙吾提, 等. 荒漠土壤有机质含量高光谱估算模型. 农业工程学报, 2014, 30(16): 113-120.
- [24] Morra M J, Hall M H, Freeborn L L. Carbon and nitrogen analysis of soil fractions using near-infrared reflectance spectroscopy. Soil Science Society of America Journal, 1991, 55(1): 288-291.

- [25] Guo Y, Cheng Y Z, Wang L G, *et al.* Predication and mapping of soil organic matter content using hyperspectra and GF-1 multi-spectral. *Chinese Journal of Soil Science*, 2016, 47(3): 537-542.
郭燕, 程永政, 王来刚, 等. 利用高光谱和 GF-1 模拟多光谱进行土壤有机质预测和制图研究. *土壤通报*, 2016, 47(3): 537-542.
- [26] Zhou P, Wang R S, Yan B K, *et al.* Extraction of soil organic matter information by hyperspectral remote sensing. *Progress in Geography*, 2010, 27(5): 27-34.
周萍, 王润生, 阎柏琨, 等. 高光谱遥感土壤有机质信息提取研究. *地理科学进展*, 2010, 27(5): 27-34.
- [27] Liu L, Shen R P, Ding G X. Studies on the estimation of soil organic matter content based on hyper-spectrum. *Spectroscopy and Spectral Analysis*, 2011, 31(3): 762-766.
刘磊, 沈润平, 丁国香. 基于高光谱的土壤有机质含量估算研究. *光谱学与光谱分析*, 2011, 31(3): 762-766.

《草业学报》 欢迎投稿 欢迎订阅

《草业学报》由中国科协技术协会主管,中国草学会、兰州大学主办,国内外公开发行的高级学术期刊。2015年《草业学报》的影响因子为 3.743,在全国被统计的 1989 种期刊中排名第 2 位,现入编全国中文核心期刊,荣获“百强报刊”,“百种中国杰出学术期刊”,“中国精品科技期刊”,“中国科技论文在线优秀期刊一等奖”。

《草业学报》主要报道国内外草业科学及其相关领域,如畜牧学、农学、林学、经济学等领域的高水平理论研究和技术创新成果,发表国内外草业领域创新性的研究论文,刊载学术价值较高的草业科学专论、综述、评论等,探讨草业发展的新理论与新构思,是草业新秀成长的园地,推动草业科学发展的论坛。其读者对象主要是从事农林牧渔、园林绿化、生态环境、国土资源等领域的科研管理及教学等专业人员。

本刊为:中国科学引文数据库(CSCD)核心期刊,中国科技论文统计源期刊,英国 CABI 文摘数据库来源期刊,《中国生物学文摘》中国生物学文摘数据库收录期刊,中国核心期刊(遴选)数据库收录期刊,《中国学术期刊(光盘版)》全文收录期刊,《万方数据—数字化期刊群》入网期刊。

《草业学报》为月刊,全铜版印刷,逢每月 20 日出版,大 16 开本,220 页,定价 25 元,全年 300 元。国内邮发代号:54-84,全国各地邮局均可订阅,若错过订期,可在本编辑部直接办理订阅。

邮发代号:54-84;国内刊号:CN 62-1105/S;国际统一刊号:ISSN 1004-5759;

地址:兰州市嘉峪关西路 768 号《草业学报》编辑部;邮政编码:730020;

网址:<http://cyxb.lzu.edu.cn>;电话、传真:0931-8913494;E-mail:cyxb@lzu.edu.cn