

基于PLSR的土壤颜色预测方法及其与色系转换法的对比研究

李怡春^{1, 2} 潘 恺^{1, 2} 王昌昆¹ 刘 娅¹ 吴士文^{1, 2} 刘 杰^{1, 2}
徐爱爱^{1, 2} 潘贤章^{1†}

(1 土壤与农业可持续发展国家重点实验室(南京土壤研究所), 南京 210008)

(2 中国科学院大学, 北京 100049)

摘 要 传统的土壤颜色测定主要采用蒙塞尔比色卡比对, 精度高, 但费时费力。近年来尝试采用色系转换法预测土壤颜色, 方法较为简便。基于土壤高光谱反射率和偏最小二乘回归(PLSR)方法进行土壤颜色预测, 并与色系转换法进行对比研究。采集皖赣鄂交界地区76个不同颜色的土壤样品, 分别采用PLSR及色系转换法进行土壤颜色预测, 并与实测结果进行了对比。结果表明, PLSR交叉验证的 R^2_{cv} 分别达到0.62、0.61和0.75, 测定值标准偏差与标准预测误差的比值(RPD)分别达到1.94、1.67和2.15, 说明PLSR模型用于土壤颜色预测是可行的; 其均方根误差(RMSE)仅为1.32、0.55和0.97个单位, 较色系转换法的RMSE分别低0.94、1.24 和0.95个单位, 其HV/C整体预测误差 ΔE 的平均值为1.91, 较色系转换法的平均值低5.16, 说明PLSR方法预测土壤蒙塞尔颜色较色系转换法更优。该方法为土壤颜色的获取提供了一种新的途径。

关键词 土壤颜色; 色系转换; 光谱; 偏最小二乘回归

中图分类号 S151.9 **文献标识码** A

土壤颜色是土壤的一种重要属性, 能反映土壤组成和肥力状况, 甚至能指示土壤年龄, 并表征某些土壤过程。不同类型土壤常具有不同的颜色, 因而土壤颜色能辅助野外识别土壤类型^[1]。土壤颜色与其他土壤属性关系密切。当有机质含量较高时, 土壤通常呈现暗棕色或黑色^[2]。当赤铁矿的含量较高时土壤呈现红色, 而针铁矿含量较高时土壤呈现黄色^[3]。土壤颜色受水分影响很大, 通常湿态较风干态颜色更暗^[4], 这个特性经常用于定性描述土壤含水率。土壤颜色信息也可用于土壤质量评估, 研究认为, 土壤色调、明度和彩度等颜色信息是石灰性土壤分布区森林立地评估的最佳指标

之一^[5]。

人眼依靠视网膜上的视锥细胞和视杆细胞来感知不同的颜色, 根据细胞敏感光波的波长分为红锥、绿锥和蓝锥, 通过三种锥细胞的组合可实现颜色的分辨^[6]。工业界常用红(R)、绿(G)、蓝(B)3个颜色通道的变化以及它们相互之间的叠加来描述颜色, 由于RGB色彩可以通过电子设备直接获取, 比较方便。而描述土壤颜色最常用的色系是蒙塞尔色系^[7], 其与RGB颜色系统之间差异很大, 因此, 实际使用过程中往往需要进行色系之间的转换。色系转换法的基本思路是首先将RGB反射率值转换到CIE XYZ坐标系下, 分别得到XYZ三刺

* 科技部基础性专项课题(2015FY110700S5)、STS项目(KFJ-SW-STS-168)资助 Supported by the Basic Work of Ministry of Science and Technology of China (No. 2015FY110700S5), the STS Program of CAS (No. KFJ-SW-STS-168)

† 通讯作者 Corresponding author, E-mail: panxz@issas.ac.cn

作者简介: 李怡春(1992—), 女, 河南南阳人, 硕士研究生, 主要从事土壤遥感应用研究。E-mail: ycLi@issas.ac.cn

收稿日期: 2018-03-30; 收到修改稿日期: 2018-06-14; 优先数字出版日期(www.cnki.net): 2018-08-17

激值,然后再转换为蒙塞尔色系HV/C值^[8]。该方法的优点是,可以将蒙塞尔比色卡的颜色引入计算机自动配色测色,实现不同色度系统之间的快速转换和色域匹配^[9],但缺点在于土壤颜色受多方面因素的影响,用色系转换法来预测土壤颜色时,忽略了土壤属性和土壤表面状态对土壤颜色测量的影响。此外,由于需经过多次转换获得蒙塞尔颜色值,各个转换环节均会带来一些误差。因此,运用色系转换法预测土壤颜色存在一些不可避免的问题,有必要探索一些新的预测方法。

近年来广泛用于土壤属性光谱预测建模的偏最小二乘回归(Partial Least Squares Regression, PLSR)方法取得了很好的预测效果。刘娅等^[10]研究表明,PLSR方法可以比较准确地预测土壤电导率和滨海盐土地地区土壤盐分含量。在应用PLSR预测游离铁含量大于20 g kg⁻¹的土壤样品时,该模型可以较好地预测游离铁的含量^[11]。PLSR方法对野外实测光谱处理也有较好的效果,郭燕等^[12]利用野外实测光谱分别进行全波段、归一化光谱指数和特征波段的PLSR建模,均可以准确预测有机质含量。刘秀英等^[13]研究表明PLSR模型是估测黄绵土钾含量的最佳模型。然而将PLSR方法应用于土壤颜色预测是否可行仍需探索。由于土壤有机质的敏感波段大部分集中在波长为400~1 000

nm的光谱范围内,并且主要集中于640~790 nm^[14-15],氧化铁的光谱响应波段为550~650 nm和750~950 nm^[16]。而人眼识别红、绿、蓝三色分别对应于620~750 nm、495~570 nm、450~495 nm^[17],这些波段与土壤有机质、氧化铁等的光谱敏感波段重叠。由此可见,利用可见光-近红外的不同波段对土壤颜色进行预测具有较大的可行性。

因此,本研究主要探索基于PLSR方法预测土壤蒙塞尔颜色的建模预测效果,并与色系转换法进行比较,以期土壤蒙塞尔颜色的获取提供一种更加快速和准确的方法。

1 材料与方法

1.1 研究区概况

研究区位于皖南山区和豫章平原地区,主要位于安徽省的宿松县、望江县、东至县和祁门县,以及江西省的鄱阳县和景德镇郊区(图1),经纬度范围在116° 8' ~117° 15' E, 29° 17' ~30° 12' N之间。该地区年平均气温16℃~17℃,年降水量1 350~1 400 mm,属于典型的亚热带季风性湿润气候。采样区地形多样,包括山地、丘陵和平原,山地土地利用以林地为主,丘陵以林灌草利用为主,平原多为耕地。

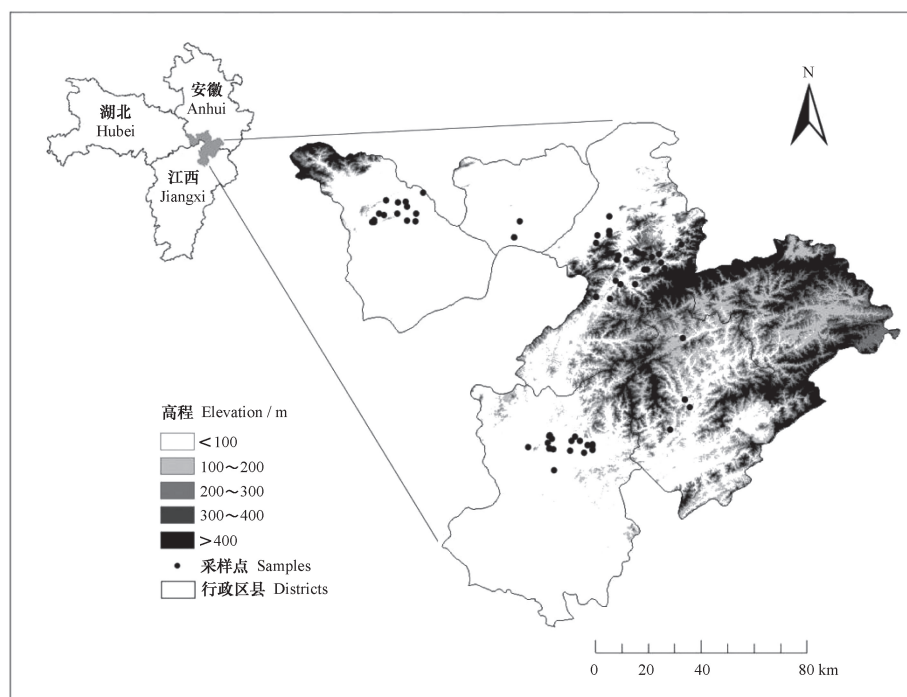


图1 采样点分布图

Fig. 1 Distribution map of soil sampling points

1.2 样品采集与分析

为了尽可能增加土壤颜色的多样性, 采样设计既考虑了主要的土壤类型, 又考虑了不同的利用方式。采样点土壤类型包含红壤、水稻土、黄棕壤、潮土、紫色土和黄壤(对应的中国土壤系统分类分别是黏化湿润富铁土、铁聚水耕人为土、铁质湿润淋溶土、浅色潮湿锥形土、酸性紫色湿润锥形土和铁质常湿淋溶土)^[18]。土地利用类型包括裸地、稀疏幼林地、灌丛草地、水田和旱地等。土壤样品采用五点法采集, 取样深度为0~15 cm, 共采集样品76个。样品风干后去除植物残体后, 分别过10目和100目筛。本实验中, pH、有机质、全氮、全磷和全钾测定方法参照《土壤农业化学分析方法》^[19]。土壤颗粒组成使用激光粒度分析仪测定。

1.3 土壤光谱测定与Munsell颜色获取

将过2 mm筛的土壤样品放入直径为7 cm、厚度为1 cm的涂黑的圆形铝盒中, 表面刮平后进行高光谱测试。高光谱采集采用美国ASD FieldSpec4 Hi-Res 地物光谱仪, 此光谱仪的波段范围为350~2 500 nm。土壤样品测定过程中使用两个50W的卤素灯, 分别放置于与垂直方向成45°角的斜上方, ASD探头视场角为8°, 探头放置于样品上方40 cm处。每次进行光谱测量之前, 要对实验仪器进行暗电流校正和白板校正, 白板校正使用的是25 cm×25 cm的漫反射标准参考板。为保证测量结果不受土壤表面状态的干扰, 每次测量旋转样品台90°, 每个方向测量五条光谱曲线, 共20条光谱曲线, 以这20条光谱曲线的平均值作为样品的测试值^[20]。ASD测量首先获得的土壤表面每个波段的亮度值, 通过拼接校正后获得土壤表面在不同波段的光谱反射率值。在进行波段范围选择时人眼对颜色的感知主要与波长1 000 nm之前的波段有关, 在色系转换算法中亦只用到红、绿、蓝三种波段, 因此, 本研究选取的波段范围为400~1 000 nm。为减少噪声误差, 对该波段范围内的光谱进行平滑处理, 平滑处理的方式为Savitzky-Golay多项式法(二次多项式, 9点平滑)。

采用中国科学院南京土壤研究所1989年编制的中国标准土壤色卡在自然光下(光线强度4 300~4 500 Lux时)对土壤样品进行比色, 土壤颜色的记录方式为HV/C。光强测试计选用SMART SENSOR数字光照度计。为了方便计算, 利用色

系转换与PLSR法预测蒙塞尔颜色值时, 通常将蒙塞尔H值用数值表达, 在蒙塞尔色度系统中, H代表色调值, 依次参照文献[21]中的方法将H值转换为数字形式, 2.5YR=12.50、5YR=15.00、7.5YR=17.50、10YR=20.00、2.5Y=22.50、5Y=25.00。V和C的值可以直接用于计算和预测, 不需要做转换。

1.4 色系转换法

(1) RGB值的计算 RGB值采用红、绿和蓝三种颜色对应波段范围(620~750 nm、495~570 nm、450~495 nm)内的反射率均值(式(1)~式(3)), 用于估算HV/C值。

$$R = \frac{1}{131} \sum_{i=620}^{750} \lambda_i \quad (1)$$

$$G = \frac{1}{76} \sum_{i=495}^{570} \lambda_i \quad (2)$$

$$B = \frac{1}{46} \sum_{i=450}^{495} \lambda_i \quad (3)$$

式中, i 表示对应的波段, λ_i 表示 i 波段对应的反射率值, R 、 G 、 B 分别表示红、绿、蓝三波段的反射率值。

(2) 色系转换 将计算得到的 R 、 G 、 B 值采用式(4), 首先转化为国际标准色系坐标CIE XYZ下的XYZ值^[22]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119194 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4)$$

然后, 根据Miyahara和Yoshida^[23]提出的转换公式完成从CIE XYZ到蒙塞尔颜色的转换(式(5)~式(14))

$$f(X_c) = 11.559X_c^{\frac{1}{3}} - 1.695 \quad (5)$$

$$f(Y) = 11.396Y^{\frac{1}{3}} - 1.610 \quad (6)$$

$$f(Z) = 11.510Z_c^{\frac{1}{3}} - 1.691 \quad (7)$$

式中, $X_c=1.020X$ 、 $Z_c=0.487Z$ 、

$$H_1 = f(X_c) - f(Y) \quad (8)$$

$$H_2 = 0.4(f(Z_2) - f(Y)) \quad (9)$$

令 $\theta = \cot(H_2 / H_1)$,

$$S_1 = (8.398 + 0.832 \times \cos(\theta))H_1 \quad (10)$$

$$S_2 = (-6.102 - 1.323 \times \cos(\theta))H_2 \quad (11)$$

H、V和C的表达式为:

$$H = \left| \tan^{-1} \left(\frac{S_2}{S_1} \right) \times \frac{100}{2\pi} \right| \quad (12)$$

$$V = f(Y) \quad (13)$$

$$C = \sqrt{S_1^2 + S_2^2} \quad (14)$$

色系转换法中评价模型精度时使用的参数为均方根误差 (Root mean square error, RMSE)^[24]、平均值和 ΔE , ΔE 参照文献 [21] 中提出 HV/C 颜色精度评价参数, 该参数由式 (15) 针对每个样品计算得到, 再计算所有样品的均值。

$$\Delta E = \sqrt{2C_1C_2 \times \left[1 - \cos \left(\frac{2\pi}{100} \Delta H \right) \right] + (\Delta C)^2 + (4\Delta V)^2} \quad (15)$$

式中, C_1 和 C_2 表示预测和实测的彩度值, ΔH 表示预测和实测的色调值之差, ΔV 表示预测和实测的明度值之差, ΔC 表示预测和实测的彩度值之差。RMSE 和 ΔE 均值越小, 表示模型精度越高。

1.5 偏最小二乘法

PLSR 是基于因子提取的多元回归数学优化模型。首先对光谱矩阵与颜色属性进行双线性分解, 之后在变量系统中提取一系列对因变量有最佳解释能力的新综合变量 (即因子提取), 因子确定时要求光谱矩阵分解得到的因子与含量矩阵分解得到的因子相关性最大, 最后利用提取的因子进行回归分析^[25]。建模过程采用交叉验证的验证方法, 即给定一个含有 i 个样品的矩阵 X , 将其中某个样品暂时剔除在 PLSR 建模计算范围之外, 用保留下来的 $i-1$ 个样品计算模型参数向量, 最后用这 $i-1$ 个样品计算模型参数向量并预测被剔除的样品的颜色坐标值。重复该运算流程直至所有样品均被剔除和预测过一次^[26]。在本研究中使用该方法分别建立土壤实测 H、V 和 C 值与波长为 400 ~ 1 000 nm 的反射率值之间的 PLSR 模型。

PLSR 中评价建模精度时使用建模集交叉验证决定系数 (Determination coefficient of cross validation, R_{cv}^2)、交叉验证均方根误差 (Root mean square error, $RMSE_{cv}$)、测定值标准偏差与标准预测误差的比值 (Ratio of standard deviation to standard prediction error, RPD)^[27], 式 (16) 和 ΔE 。

$$RPD = \frac{SD}{RMSE} \quad (16)$$

式中, SD 表示数据集实测值标准差。

R_{cv}^2 与 RPD 越高, 且 $RMSE_{cv}$ 越低, 则模型越好。已有研究结果表明, 当 $RPD \geq 2.50$ 时, 该模型可以精确预测; 当 $2.50 > RPD \geq 2.00$ 时, 模型可以很好地预测; 当 $2.00 > RPD \geq 1.80$ 时, 模型可以较好地预测; 当 $1.80 > RPD \geq 1.40$ 时, 该模型可以近似地预测; 当 $1.40 > RPD \geq 1.00$ 时, 该模型可以区分变量值的大小; 当 $RPD < 1.00$ 时, 该模型不适合预测^[28]。

计算中涉及的色系转换在 MATLAB R2012a 中完成, PLSR 的建模及验证在 The Unscrambler 10.1 中完成, 图表制作由 OriginPro 9.0、ArcGIS 10.2 完成。

2 结 果

2.1 土壤属性及光谱特征

土壤样品基本属性的统计特征如表 1 所示。土壤 pH 介于 3.96 ~ 7.60, 平均值为 4.84, 表明大部分土壤为酸性土壤。有机质平均值为 $20.57 \text{ g} \cdot \text{kg}^{-1}$, 土壤肥力总体较高。在土壤颗粒组成中, 粉粒含量较高, 土壤质地主要是壤质砂土、粉 (砂) 壤土、壤土等。表 2 为不同土壤类型全铁和有机质含量的统计值, 有机质含量由高到低依次为黄棕壤、水稻土、黄壤、紫色土、潮土和红壤。全铁含量由高到低依次为黄壤、红壤、潮土、黄棕壤、水稻土和紫色土。

研究区内的 76 个风干样本的实测土壤颜色 (Munsell HV/C) 如下表 3 所示, 色调主要集中在 12.50 到 20.00 之间的黄红色调中, 明度范围在 4.00 至 8.00 之间, 彩度值的变化范围在 1.00 至 8.00 之间。由于比色时土壤呈风干态, 所以样品偏亮、明度值偏高。土壤彩度值变化范围较大, 样品的色彩浓度差异较大。

表1 采样点土壤基本属性

Table 1 Basic properties of the sampling soils					
属性 Properties	范围 Range	平均值 Mean	标准差 SD	峰度 Kurtosis	偏度 Skewness
pH	3.96 ~ 7.60	4.84	0.7	6.65	2.4
有机质 Soil organic matter/ (g·kg ⁻¹)	4.48 ~ 45.45	20.57	10.83	0.16	0.67
黏粒 Clay/%	8.71 ~ 39.10	22.05	6.12	-0.1	0.06
粉粒 Silt/%	13.51 ~ 77.52	64.05	12.4	3.62	-1.85
砂粒 Sand/%	0.10 ~ 59.01	13.67	12.26	3.02	1.97

表2 颜色相关的不同类型土壤属性信息

Table 2 Color-dependent properties of soils different in type				
类型 Soil type	属性 Properties	范围 Range	平均值 Mean	标准差 SD
红壤 Red soil	有机质 Soil organic matter/ (g·kg ⁻¹)	4.48 ~ 34.37	13.79	7.83
	全铁 Total iron/ (g·kg ⁻¹)	23.10 ~ 106.20	57.72	18.20
水稻土 Paddy soil	有机质 Soil organic matter/ (g·kg ⁻¹)	19.31 ~ 41.58	31.56	7.45
	全铁 Total iron/ (g·kg ⁻¹)	26.90 ~ 62.22	40.05	11.23
黄棕壤 Yellow-brown soil	有机质 Soil organic matter/ (g·kg ⁻¹)	21.61 ~ 44.70	33.16	16.33
	全铁 Total iron/ (g·kg ⁻¹)	35.21 ~ 49.40	43.3	5.98
潮土 Fluvo-aquic soil	有机质 Soil organic matter/ (g·kg ⁻¹)	21.22 ~ 21.37	21.3	0.11
	全铁 Total iron/ (g·kg ⁻¹)	52.11 ~ 54.34	53.32	9.10
紫色土 Purple soil	有机质 Soil organic matter/ (g·kg ⁻¹)	7.28 ~ 32.91	23.13	8.64
	全铁 Total iron/ (g·kg ⁻¹)	20.80 ~ 59.33	37.49	12.97
黄壤 Yellow soil	有机质 Soil organic matter/ (g·kg ⁻¹)	8.82 ~ 45.45	29.74	14.14
	全铁 Total iron/ (g·kg ⁻¹)	52.69 ~ 78.90	64.02	10.04

表3 实测HV/C值基本信息

Table 3 Statistics of the measured HV/C values				
	样品数 Sample size	范围 Range	平均值 Mean	标准差 SD
蒙塞尔 H Munsell H	76	20.00 ~ 12.50	18.16	2.14
蒙塞尔 V Munsell V	76	4.00 ~ 8.00	6.68	0.88
蒙塞尔 C Munsell C	76	1.00 ~ 8.00	4.45	1.89

土壤样本在400 ~ 1 000 nm波长内原始光谱反射率总体呈现上升的趋势，本研究区土壤中均含有一定量的氧化铁，在850 ~ 950 nm处存在不同程度的氧化铁吸收峰，黄壤和红壤氧化铁含量较高，在光谱曲线中存在较为明显的吸

收峰。
2.2 色系转换计算土壤颜色结果
本研究中，以色系转换方法计算的H、V、C值为预测值的平均值与范围，以人眼观察值为实测值计算RMSE，结果如表4所示：

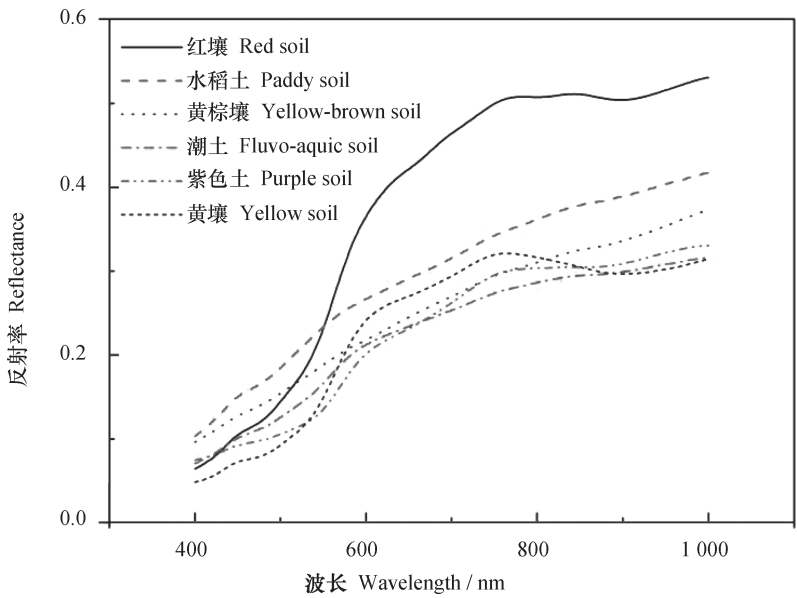


图2 不同土壤样品原始光谱

Fig. 2 Original spectra of soil samples

表4 色系转换预测土壤蒙塞尔颜色结果

Table 4 Statistics of soil Munsell colors predicted with the color space conversion method

	样本量	范围	平均值	RMSE
	Sample size	Range	Mean	
H	76	13.95 ~ 23.57	19.25	2.26
V	76	3.69 ~ 6.00	4.99	1.79
C	76	4.24 ~ 8.30	5.79	1.92

其中色调H的范围介于13.95~23.57，而实测H值范围介于12.50~20.00，二者之间有一定差距，从均值来看，预测平均值为19.25，其实测值为18.16，本研究中的土壤主要集中在黄红色调中，因此，预测结果大约偏黄一个色调。明度V的预测范围介于3.69~6.00，而实测值范围介于4.00~8.00，预测平均值为4.99，实测值为6.68，可见，预测值较实测值偏暗。彩度值C的预测范围介于4.24~8.30，实测值的范围介于1.00~8.00，色系转换法预测彩度值C存在一定的偏差；C预测均值为5.79，而实测均值为4.45，C均值偏低约1.34个单元。因此，预测值不足以表达土壤颜色的丰富程度。

从转换结果（图3）来看，色调H的预测值与实测值的回归系数 R^2 仅为0.30，预测结果较差。明度V回归系数 R^2 达到0.67，虽然预测值与实测值存在显著正相关的关系，但预测V值明显低于1:1线，说明预测出的明度偏暗。彩度C预测值与实测值的回归系数 R^2 达到0.61，但C值的预测范围较

窄。从预测结果中可以得出，色系转换法预测的颜色与真实值有较大差异。

基于色系转换的方法预测土壤颜色时，土壤色调普遍偏黄，明度偏暗，彩度值的预测范围较窄，该结论与Rossel等^[29]2009年的研究结果相同。因此，色系转换法虽然可以预测土壤颜色的实测值，但整体效果并不理想。

2.3 PLSR建模预测土壤颜色结果

从模型预测结果来看（表5），预测范围与实测范围非常接近，H预测值的范围介于14.42~21.96，预测范围偏高，但在中国标准土壤色卡的划分中该误差不足一个色调。V的预测值范围介于5.12~8.21，范围下限较实测值高约一个单元，上限与实测值上限大致相同；彩度值C的预测范围介于0.40~8.70，与实测值的彩度范围十分接近。PLSR方法预测的H、V、C的平均值与实测平均值一致。可见，PLSR预测值的平均值与实测值的平均值基本相同。

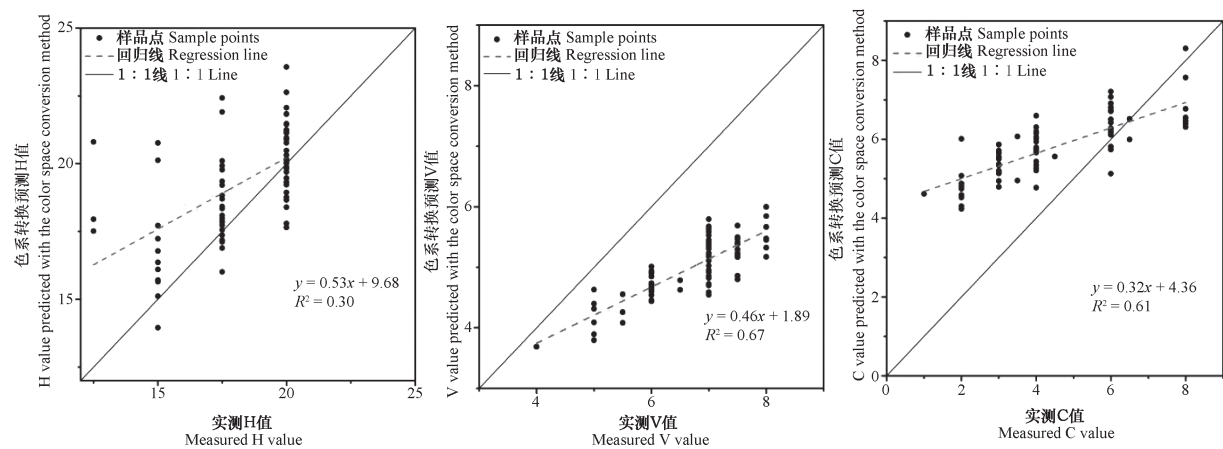


图3 色系转换与实测值对比图

Fig. 3 Measured HV/C versus predicted HV/C using color space conversion method

表5 PLSR预测土壤蒙塞尔颜色结果

Table 5 Statistics of soil Munsell colors predicted with PLSR							
	样本量	范围	平均值	因子数	R^2_{cv}	RPD	RMSE _{cv}
	Sample size	Range	Mean	Number			
H	76	14.42 ~ 21.96	18.16	6	0.62	1.94	1.32
V	76	5.12 ~ 8.21	6.68	2	0.61	1.67	0.55
C	76	0.40 ~ 8.70	4.45	3	0.75	2.15	0.97

利用PLSR法分别建立土壤色调、明度、彩度值与反射率回归模型，PLSR模型在预测土壤蒙塞尔颜色的H、V、C值中，交叉验证决定系数 R^2_{cv} 分别达到0.62、0.61和0.75，RPD分别达到1.94、1.67、2.15，说明PLSR方法可以很好地预测土壤的彩度值，可以较好地预测土壤色调值，可以近似地预测土壤明度值。

表6 不同预测方法ΔE统计结果

Table 6 Statistics of ΔEs predicted with different methods			
	范围	平均值	标准差
	Range	Mean	SD
色系转换	1.73 ~ 11.43	7.07	2.14
Color space conversion			
PLSR	0.31 ~ 5.63	1.91	1.12

色系转换法预测的蒙塞尔颜色的RMSE达到2.26、1.79和1.92个误差单元，而PLSR法预测蒙塞尔颜色H、V和C值的RMSE_{cv}仅为1.32、0.55和0.97个误差单元，PLSR法的误差与色系转换法相比明显偏低。表6中所示，利用ΔE对H、V、C三个因数进行综合分析可得，平均值和预测范围显著下降，PLSR预测误差明显小于色系转换法。因此，PLSR方法更适合预测土壤的蒙塞尔颜色。

从回归分析结果（图4）可知，PLSR预测H与实测H的决定系数 R^2 为0.69，相比于色系转换法变异明显减小，说明色调H的预测精度有较大幅度的提升。明度值V与色系转换法对明度值的预测结果相似，但回归线与1：1线更接近，明度值的预测范围较色系转换法更大。彩度值C的预测值与实测值 R^2 达到0.78，回归方程与1：1线非常接近，说明预测值与真实值很接近，基于PLSR方法预测C值的

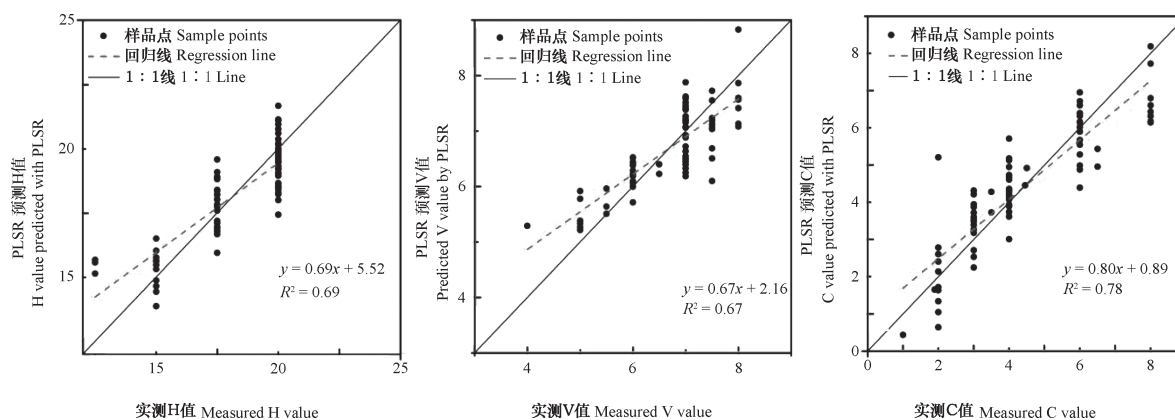


图4 PLSR预测值与实测值对比图

Fig. 4 Measured HV/C versus predicted HV/C using PLSR

效果明显优于色系转换法。

3 讨论

本研究用色系转换与PLSR法预测土壤颜色,结果表明,PLSR法预测土壤的H、V、C值的效果较好,该方法估测的土壤蒙塞尔颜色值较为可靠,整体预测精度高于色系转换方法,预测结果与真实值差异小于色系转换方法。由于PLSR法与色系转换法存在较大的不同,所以预测效果存在一些差异,从土壤光谱、影响因素、计算过程和测量误差四个方面分析差异产生的原因。

从土壤光谱来看,由于色系转换法仅采用了肉眼能够感知的RGB三个波段范围光谱信息,数据处理中由于无法获取肉眼的光谱响应函数,故仅对各个波段反射率做了简单平均,使其损失了较多有效信息。同时由于色系转换采用的波段仅有红、绿、蓝3个波段,相对于原始光谱来说,存在信息过度压缩的问题。而PLSR方法运用了400~1 000 nm的全部波段反射率信息,对于光谱信息的利用更为全面。因此,PLSR方法可以更好地进行土壤颜色预测。

从土壤颜色的影响因素来看,一般认为土壤有机质含量、土壤针铁矿、赤铁矿含量,以及颗粒组成和研磨程度对颜色的判定和光谱反射率有较大的影响^[30]。从有机质含量来看,于士凯等^[31]认为647 nm、750 nm、876 nm均是土壤有机质的敏感波段。从铁氧化物来看,依据Schwertmann的研究^[32],土壤中的某些矿物,例如针铁矿、赤铁矿也与土壤颜色息息相关,Brown等^[33]提出由于针铁矿与赤

铁矿的原子构型不同,导致在430 nm和550 nm处的敏感波段的反射率曲线形态存在一定差异,赤铁矿是一段“平缓曲线”,而针铁矿则出现“反射峰”。PLSR法将这些细微差异完全保留,并运用于颜色预测的建模过程中,而色系转换法则忽略了土壤颜色产生的机理,仅从色系坐标和色彩表达的角度进行推算,虽然这对色彩学研究有较大帮助,但由于土壤颜色的表达是一个复杂的过程,简单的方法并不一定完全适用于土壤颜色研究。

从计算过程来看,色系转换方法只选取红、绿、蓝三波段范围内的反射率信息,转化步骤繁杂,每一步转换均会存在一些不可避免的误差;而PLSR法预测土壤颜色较为直接,该方法以误差最小和决定系数最高为目标进行建模,从而可以获得能够充分揭示数据内涵的模型。此外,蒙塞尔比色卡比色时对土壤颗粒研磨的粗细程度和样品表面状态没有过多的要求,而不同土壤样品表面粗糙度和土壤质地存在差异,反射率受到一定的影响,这些因素也会导致土壤预测时的误差。因此,相比之下色系转换计算土壤蒙塞尔颜色的误差来源更多。本研究中提出的PLSR预测土壤颜色有一定的优势和较为广阔的应用前景,对快速获取土壤颜色提供了一种新的方法。

本研究仅限于安徽南部、江西北部 and 湖北东部地区的土壤,样本的颜色主要集中在黄红色调,未来将扩大土壤类型,提高本研究方法的适应性。从研究方法上而言,研究中仅使用了PLSR方法和色系转换方法,关于其他色调土壤的研究和其他预测方法的探索将在后续展开。此外,在自然状态下土壤颜色的判定还受到土壤水分含量等客观因素的干

扰。因此, 后续研究将考虑土壤水分等因素对土壤颜色预测的影响。

4 结 论

本文探讨了色系转换法与PLSR方法对土壤蒙塞尔颜色的色调H、明度V和彩度C的预测效果, 并进行了对比分析, PLSR的均方根误差分别较色系转换法低0.94、1.24 和0.95, 其 ΔE 的平均值为1.91, 色系转换法的平均值为7.07, 色系转换法预测结果偏离实测值更远, 表明PLSR方法对土壤HV/C的预测效果优于色系转换法。该方法可以应用于不同类型、质地的土壤颜色的预测, 为土壤颜色的获取提供了一条新的途径。

致 谢 感谢国家科技基础条件平台-国家地球系统科学数据共享平台-土壤科学数据中心(<http://soil.geodata.cn>)及中国土壤数据库(<http://vdb3.soil.csd.cn/>)提供数据支撑。

参 考 文 献

- [1] Han P, Dong D, Zhao X, et al. A smart phone-based soil color sensor: For soil type classification. *Computers and Electronics in Agriculture*, 2016, 123: 232—241
- [2] Liu J, Sui Y, Yu Z, et al. High throughput sequencing analysis of biogeographical distribution of bacterial communities in the black soils of northeast China. *Soil Biology and Biochemistry*, 2014, 70: 113—122
- [3] Hu X F, Du Y, Guan C L, et al. Color variations of the Quaternary Red Clay in southern China and its paleoclimatic implications. *Sedimentary Geology*, 2014, 303: 15—25
- [4] Thompson J A, Bell J C. Color index for identifying hydric conditions for seasonally saturated mollisols in Minnesota. *Soil Science Society of America Journal*, 1996, 60 (6) : 1979—1988
- [5] Sánchez-Marañón M, Delgado R, Párraga J, et al. Multivariate analysis in the quantitative evaluation of soils for reforestation in the Sierra Nevada (southern Spain) . *Geoderma*, 1996, 69 (3/4) : 233—248
- [6] 张云. 基于生物视觉认知机理的特征提取及其应用研究. 武汉: 华中科技大学, 2014
Zhang Y. Feature extraction based on biological visual cognition and its applications on computer vision (In Chinese) . Wuhan: Huazhong University of Science and Technology, 2014
- [7] Munsell A H. *Munsell book of color*. The United States: Munsell Color Co, 1950
- [8] Gómez-Robledo L, López-Ruiz N, Melgosa M, et al. Using the mobile phone as Munsell soil-colour sensor: An experiment under controlled illumination conditions. *Computers and Electronics in Agriculture*, 2013, 99: 200—208
- [9] 徐芙姗, 徐海松, 王勇. 孟塞尔色序系统与CIE1931标准色度系统转换新算法. *光子学报*, 2007, 36 (4) : 650—654
Xu F S, Xu H S, Wang Y. A new conversion algorithm between munsell color order system and CIE 1931 standard colorimetric system (In Chinese) . *Acta Photonica Sinica*, 2007, 36 (4) : 650—654
- [10] 刘娅, 潘贤章, 王昌昆, 等. 基于可见-近红外光谱的滨海盐土土壤盐分预测方法. *土壤学报*, 2012, 49 (4) : 824—829
Liu Y, Pan X Z, Wang C K, et al. Prediction of coastal saline soil salinity based on VIS-NIR reflectance spectroscopy (In Chinese) , *Acta Pedologica Sinica*, 2012, 49 (4) : 824—829
- [11] 魏昌龙, 赵玉国, 郭登巍, 等. 基于光谱分析的土壤游离铁预测研究. *土壤*, 2014, 46 (4) : 678—683
Wei C L, Zhao Y G, Wu D W, et al. Prediction of soil free iron oxide content based on spectral analysis (In Chinese) . *Soils*, 2014, 46 (4) : 678—683
- [12] 郭燕, 纪文君, 吴宏晦, 等. 基于野外Vis-NIR光谱的土壤有机质预测与制图. *光谱学与光谱分析*, 2013, 33 (4) : 1135—1140
Guo Y, Ji W J, Wu H H, et al. Estimation and mapping of soil organic matter based on Vis-NIR reflectance spectroscopy (In Chinese) . *Spectroscopy and Spectral Analysis*, 2013, 33 (4) : 1135—1140
- [13] 刘秀英, 石兆勇, 常庆瑞, 等. 黄绵土钾含量高光谱估算模型研究. *土壤学报*, 2018, 55 (2) : 325—337
Liu X Y, Shi Z Y, Chang Q R, et al. Hyperspectral model for estimation of soil potassium content in Loessal soil (In Chinese) . *Acta Pedologica Sinica*, 2018, 55 (2) : 325—337
- [14] Gunsaulis F R, Kocher M F, Griffis C L. Surface structure effects on close-range reflectance as a function of soil organic matter content. *Transactions of the ASAE*, 1991, 34 (2) : 641—649
- [15] 侯艳军, 塔西甫拉提·特依拜, 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
Hou Y J, Tashpolat T, Mamat S, et al. Estimation model of desert soil organic matter content using

- hyperspectral data (In Chinese). Transactions of the Chinese Society of Agricultural Engineering, 2014, 30 (16): 113—120
- [16] Ben-Dor E, Banin A. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. Soil Science Society of America Journal, 1995, 59 (2): 364—372
- [17] Miyazaki T, Kondrashev S L, Tsuchiya T. Visual pigment genes and absorbance spectra in the Japanese sardine *Sardinops melanostictus* (Teleostei: Clupeiformes). Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, 2018, 218: 54—63
- [18] 陈志诚, 龚子同, 张甘霖, 等. 不同尺度的中国土壤系统分类参比. 土壤, 2004, 36 (6): 584—595
Chen Z C, Gong Z T, Zhang G L, et al. Correlation of soil taxa between Chinese Soil Genetic classification and Chinese Soil Taxonomy on various scales (In Chinese). Soils, 2004, 36 (6): 584—595
- [19] 鲁如坤. 土壤农业化学分析方法. 北京: 中国农业科学技术出版社, 2000
Lu R K. Analytical methods for soil and agro-chemistry (In Chinese). Beijing: China Agricultural Science and Technology Press, 2000
- [20] 刘娅, 潘贤章, 王昌昆, 等. 基于差异化光谱指数的盐渍土水分含量预测——以滨海盐土为例. 土壤, 2016, 48 (2): 381—388
Liu Y, Pan X Z, Wang C K, et al. Prediction of saline soil moisture content based on differential spectral index: A case study of coastal saline soil (In Chinese). Soils, 2016, 48 (2): 381—388
- [21] Rossel R A V, Minasny B, Roudier P, et al. Colour space models for soil science. Geoderma, 2006, 133 (3): 320—337
- [22] Wyman C, Sloan P P, Shirley P. Simple analytic approximations to the CIE XYZ color matching functions. Journal of Computer Graphics Techniques, 2013, 2 (2): 1—11
- [23] Miyahara M, Yoshida Y. Mathematical transform of (R, G, B) color data to Munsell (H, V, C) color data. visual communications and image processing' 88: Third in a series. International Society for Optics and Photonics, 1988: 650—657
- [24] 王昌昆, 潘贤章, 周睿, 等. 应用基于 PLSR 的土壤-环境模型预测土壤属性. 土壤学报, 2012, 49 (2): 237—245
Wang C K, Pan X Z, Zhou R, et al. Prediction of soil properties using PLSR-based soil-environment models (In Chinese). Acta Pedologica Sinica, 2012, 49 (2): 237—245
- [25] 王森, 解宪丽, 周睿, 等. 基于可见光-近红外漫反射光谱的红壤有机质预测及其最优波段选择. 土壤学报, 2011, 48 (5): 1083—1089
Wang M, Xie X L, Zhou R, et al. Determination of soil organic matter in red soils using vis-NIR diffuse reflectance spectroscopy and selection of optimal spectral bands (In Chinese). Acta Pedologica Sinica, 2011, 48 (5): 1083—1089
- [26] Liu Y, Pan X Z, Wang C K, et al. Predicting soil salinity with vis-NIR spectra after removing the effects of soil moisture using external parameter orthogonalization. PLoS One, 2015, 10 (10): e0140688
- [27] Liu Y, Pan X Z, Shi R J, et al. Predicting soil salt content over partially vegetated surfaces using non-negative matrix factorization. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8 (11): 5305—5316
- [28] Murphy R J, Monteiro S T. Mapping the distribution of ferric iron minerals on a vertical mine face using derivative analysis of hyperspectral imagery (430–970 nm). ISPRS Journal of Photogrammetry and Remote Sensing, 2013, 75: 29—39
- [29] Rossel R A V, Cattle S R, Ortega A, et al. In situ measurements of soil color, mineral composition and clay content by vis-NIR spectroscopy. Geoderma, 2009, 150 (3): 253—266
- [30] 赵其国, 史学正, 等. 土壤资源概论. 北京: 科学出版社, 2007: 93—117
Zhao Q G, Shi X Z, et al. Introduction to soil resource (In Chinese). Beijing: Science Press, 2007: 93—117
- [31] 于士凯, 姚艳敏, 王德营, 等. 基于高光谱的土壤有机质含量反演研究. 中国农学通报, 2013, 29 (23): 146—152
Yu S K, Yao Y M, Wang D Y, et al. Studies on the inversion of soil organic matter content based on hyper-spectrum (In Chinese). Chinese Agricultural Science Bulletin, 2013, 29 (23): 146—152
- [32] Schwertmann U. Relations between iron oxides, soil color, and soil formation. Soil Science Society of America Journal, 1993, 57 (4): 51—69
- [33] Brown A J, Sutter B, Dunagan S. The MARTE VNIR imaging spectrometer experiment: Design and analysis. Astrobiology, 2008, 8 (5): 1001—1011

PLSR-Based Prediction of Soil Color and Its Comparison with Color Space Conversion Method

LI Yichun^{1,2} PAN Kai^{1,2} WANG Changkun¹ LIU Ya¹ WU Shiwen^{1,2} LIU Jie^{1,2} XU Aiai^{1,2} PAN Xianzhang^{1†}

(1 Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China)

(2 University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract 【Objective】The color of a soil may, to a certain extent, reflect degree in development, type and fertility of the soil. Traditionally, soil color is measured with the Munsell colorimetry, which, though quite high in accuracy is time-consuming and low in efficiency. It is, therefore, essential to explore for a quick and accurate method to measure soil colors. Nowadays remote sensing and proximal sensing methods can be used to obtain soil information, and numerous attempts have been made to extract soil color information from soil spectra. For that end, color space conversion (CSC) method is a commonly used one. It uses mathematical formulas to match colors between different coordinates, so as to realize prediction of soil colors. The first step of this method is to extract average reflectance values of the RGB bands from spectral reflectance and then converts them into XYZ values in the CIE XYZ coordinate, and further into HV/C values in the munsell coordinate. In this paper, a novel method was introduced to predict soil colors using partial least squares regression (PLSR) of hyperspectral reflectance of soils, and then comparison was made between PLSR and CSC in prediction accuracy. 【Method】A total of 76 soil samples different in colors were collected in the bordering area of Anhui, Jiangxi and Hubei Province for the study, covering soil types e.g. red soil (Argi-Udic Ferrosols), paddy soil (Fe-accumuli-Stagnic Anthrosols), yellow-brown soil (Ferri-Udic Argosols), fluvo-aquic soil (Ochri-Aquic Cambosols), purple soil (Dystric Purpli-Udic Cambosols) and yellow soil (Ali-Perudic Argosols) in the study area. After being air-dried, the soil samples were determined in color through color matching with the Munsell color system, and their spectral reflectance was acquired simultaneously with the aid of the ASD spectrometer. Then PLSR and CSC was applied separately to predict colors of the soil samples. 【Result】Results show that the PLSR model can be well used to predict soil Hue (H), Value (V), and Chroma (C) with cross validation coefficient (R_{cv}^2) being 0.62, 0.61 and 0.75 respectively, and RPD being 1.94, 1.67 and 2.15 respectively, which suggests that it is feasible to use the PLSR method to predict soil colors and that the mean square root error (RMSE) of H, V and C predicted with PLSR was only 1.32, 0.55 and 0.97 units, respectively, and 0.94, 1.24 and 0.95 lower than their respective ones predicted with the CSC method. The former, being 1.91, was 5.16 lower than the latter in mean ΔE , the mean HV/C comprehensive index. Analysis of reasons for that reveals that PLSR uses the spectral reflectance information of all the bands, while CSC makes use of mean reflectance of Red, Green and Blue bands only. Furthermore, certain errors inevitably occur in every step of the conversion of CSC. 【Conclusion】Therefore, it could be concluded that the PLSR method is superior to the CSC method in predicting Munsell color of a soil. And compared the conventional soil color measuring methods, this one saves time and labor by a large margin. So this method opens up a new way for quick soil color acquisition via soil spectrum.

Key words Soil munsell color; Color space conversion method; Spectroscopy; Partial least squares regression

(责任编辑：檀满枝)