

基于高光谱的石楠叶片叶绿素含量估算模型

何桂芳, 吴 见, 彭 建, 谷双喜

(滁州学院 地理信息与旅游学院, 安徽 滁州 239000)

摘 要:准确估算叶绿素含量对于植物生长监测、产量预测、生境的适宜性评价具有重要作用。为寻求叶片叶绿素含量的高精度估算模型,以石楠为对象,实测叶片叶绿素含量和反射光谱反射率,对原始光谱进行变换并计算植被指数,通过相关性分析挑选特征波段,运用多元逐步线性回归和偏最小二乘回归建立叶绿素预测模型。结果表明:1)FDR 的逐步线性回归模型和偏最小二乘模型优于 R、1/R、LR、SDR;2)DNDVI(R_{645} , R_{1370})的指数函数模型为估算叶绿素含量的最佳单变量模型;3)DRI(R_{747} , R_{1464})与 RI(R_{733} , R_{944})的逐步线性回归模型精度最高,验证结果的决定系数 R^2 为 0.955,均方根误差 RMSE 为 3.145。因此,该模型可以实现叶片叶绿素含量的准确估算,从而为实现高光谱技术监测植被叶绿素含量变化提供依据。

关键词:叶绿素含量;高光谱;偏最小二乘;逐步线性回归;植被指数

中图分类号:S792.24

文献标志码:A

文章编号:1001-7461(2022)01-0025-08

Estimation Models of Chlorophyll Contents in *Photinia serrulata* Leaves Based on Hyperspectral Data

HE Gui-fang, WU Jian, PENG Jian, GU Shuang-xi

(School of Geographic Information and Tourism, Chuzhou University, Chuzhou 239000, Anhui, China)

Abstract: Accurate estimation of chlorophyll contents plays important roles in plant growth monitoring, yield prediction and habitat suitability evaluation. In order to find a high-precision estimation model, *Photinia phoebe* was taken as the research object to measure the chlorophyll content and reflectance of the leaves. Firstly, the original spectra were transformed and the vegetation indexes were calculated. Secondly, the characteristic bands were selected through correlation analyses. Then the chlorophyll prediction models was established by stepwise linear regression and partial least square regression (PLSR). The results showed that 1) the stepwise linear regression model and partial least square model of FDR were better than R, 1/R, LR and SDR. 2) The exponential model of DNDVI (R_{645} , R_{1370}) was the best single variable model for estimating chlorophyll contents. 3) The stepwise linear regression model of DRI (R_{747} , R_{1464}) and RI (R_{733} , R_{944}) was the most accurate model, and the validation results showed that R^2 was 0.955, and RMSE was 3.145. Therefore, we concluded that the gradual linear regression model of vegetation index could accurately estimate the leaf chlorophyll content. The results provide a basis for monitoring the change of chlorophyll contents in vegetation by hyperspectral technology.

Key words: chlorophyll content; hyper-spectral; partial least square; stepwise linear regression; vegetation index

叶绿素作为植物进行光合作用的主要色素,其 含量与光合作用能力、植物胁迫以及健康状况关系

收稿日期:2021-02-21 修回日期:2021-03-12

基金项目:安徽省教育厅高校自然科学研究重点项目(KJ2019A0634);安徽省教育厅高校自然科学研究重点项目(KJ2018A0434);安徽省自然科学基金(1808085QC72)。

第一作者:何桂芳,讲师,硕士。研究方向:3S应用。E-mail:heguifang@chzu.edu.cn

密切^[1]。快速、准确地掌握植株叶绿素含量,对于植物生长状态监测、产量预测、生境的适宜性评价等都具有非常重要的意义^[2]。检测叶片叶绿素的常用方法有分光光度计法、原子吸收法、高效液相色谱法等,这些方法不仅具有破坏性,而且繁琐,还只能测有限样本^[3]。高光谱分析技术具有快速、无损、成本低的优点,能直接对植被进行微弱光谱差异的定量分析^[4]。因此,国内外有很多学者借助高光谱对叶绿素含量进行估算。目前,高光谱估算叶绿素含量的研究主要有2类:第1类使用原始光谱及各种变换光谱对叶片叶绿素含量进行估算。冯海宽等^[5]、黄慧等^[6]、李媛媛等^[7]、胥喆等^[8]对原始光谱、倒数之对数、一阶微分和连续统去除等变量,尼加提·卡斯木等^[9]利用反射率、吸收深度等光谱参数,邓小蕾等^[10]利用原始光谱及一阶微分的小波包去噪光谱等构建叶绿素反演模型。这类研究都是对单个敏感波段的光谱进行建模,精度可能会在一定程度上受到制约。第2类使用任意2波段构建的植被指数对叶绿素含量进行估算。何彩莲等^[11]、姜海玲等^[12]直接使用已经发表的固定波段的植被指数,Tumbo *et al*^[13]利用比值植被指数,程志庆等^[14]利用归一化植被指数,孙红等^[15]利用差值植被指数,武旭梅等^[16]、罗丹等^[17]利用比值、差值、归一化植被指数及土壤调节指数,陈秀青等^[18]、王鑫梅等^[19]用原始光谱、一阶导数的归一化差分和比值植被指数,马文勇等^[20]、刘文雅等^[4]、李春等^[21]利用红边位置植被指数,构建 SPAD 估算模型。这类研究使用的反演方法主要有随机森林、偏最小二乘、线性回归、BP 神经网络、主成分分析、支持向量机等,根据自变量和因变量的个数选用合适的反演方法,都能够得到较高的精度模型。因此,利用高光谱分析技术对植被叶片叶绿素含量估算是可行可靠的,只是针对不同的植物和不同反演方法所选用的最佳波段和最佳估算变量不尽相同^[22]。

以往研究大多集中在水稻、玉米、小麦等农田经济作物和苹果等果蔬上,对于园林绿化常用的绿色植物研究较少,本研究选取具有观赏和药用价值的常绿阔叶灌木石楠(*Photinia serrulata*)为对象,通过实测光谱反射率和叶绿素含量,选择叶片尺度的原始光谱及4种变换光谱和差值型、比值型、归一化值型的6种常用植被指数,采用逐步线性回归和偏最小二乘方法建立叶绿素含量的高光谱反演模型,并对比单波段构建的模型与植被指数回归模型,寻求叶绿素最佳反演模型,为实现高光谱技术快速、大面积监测绿色植物的长势和营养监测提供理论依据和技术支撑,为园林绿化的养护和管理提供参考。

1 材料与方法

1.1 数据采集与处理

研究区为滁州市琅琊山景区,采集健康成熟的单株石楠叶片,在每棵树上采集3片叶片,并将叶片编号后同时进行光谱与叶绿素含量的测量。叶片光谱值采用美国 ASD Field Spec 光谱仪实测获取,波长 350~2 500 nm,采样间隔为 1 nm,共有 2 150 个波段。每个样点光谱数据记录 8 次,取每个样点的 8 条光谱数据平均值,作为该样点的光谱数据。叶绿素测定采用 SPAD-502 叶绿素仪,该仪器采用光电无损检测方法测得的 SPAD(soil and plant analyzer development)值与叶绿素含量具有很高的相关性,常用于表征叶绿素含量^[23]。选择晴朗的无风、无云天气 9:00—11:00,先按照说明对仪器校准后,再对每片树叶的叶尖、叶中和叶基 3 个部位测量叶片光谱,同时在该样点上进行 SPAD 值测量。

分别对每一片树叶 3 个样点的 SPAD 值与光谱数据取平均值,作为该叶片的 SPAD 值与对应的光谱反射率。研究共测得 59 份叶片样本,随机选取 40 份作为建模样本,剩余 19 份作为验证样本。各类型样本的 SPAD 值统计特征见表 1。其中,建模样本的 SPAD 值为 3.70~50.00,区间分布合理,变异程度较大,能保证所建模型的适应范围;验证样本和建模样本的统计特征相差不大,能够验证模型的可靠性。

表 1 SPAD 值的统计特征

Table 1 Statistical description of SPAD

样本类型	样本数	最小值	最大值	平均值	标准差 /(mg·g ⁻¹)
总样本	59	3.70	58.90	23.36	12.27
建模样本	40	3.70	50.00	23.48	11.65
验证样本	19	5.30	58.90	23.11	13.82

1.2 研究方法

1.2.1 光谱变换 在原始光谱反射率(reflectance, R)的基础上,计算其倒数 $1/R$,倒数之对数(logarithm of reciprocal, LR,公式中用 L_R 表示)、一阶微分(first order differential of reflectance, FDR,公式中用 F_{DR} 表示)、二阶微分(second-order differential of reflectance, SDR,公式中用 S_{DR} 表示)4种指标,计算公式见式(1)~式(3)。原始光谱经过 LR 变换后能够减少因为光照条件变化引起的乘性因素的影响,经过 FDR 变换后可以消除背景噪声的干扰,分解出混合的重叠峰,从而提高光谱的分辨率和灵敏度,找到相关性更高的波段。

$$L_R = \ln(1/R) \quad (1)$$

$$F_{DR} = (R_{i+1} - R_{i-1}) / 2\Delta\lambda \quad (2)$$

$$S_{DR} = (R'_{i+1} - R'_{i-1}) / 2\Delta\lambda \quad (3)$$

式中: R_{i-1} 表示上一个波段的反射率, R_{i+1} 表示下一个波段的反射率, R' 表示一阶微分, $\Delta\lambda$ 为采样间隔。

1.2.2 植被指数 光谱的植被指数是 2 个或者多个波段光谱反射率的组合。通过对估算叶绿素模型的植被指数进行分析,选取 6 个常用植被指数^[16]: 比值指数(ratio vegetation index, RI, 公式中用 R_1 表示)、差值指数(difference vegetation index, VI, 公式中用 V_1 表示)、归一化指数(normalized difference vegetation Index, NDVI, 公式中用 N_{DVI} 表示)、光谱反射率一阶微分比值指数(ratio index of derivative, DRI, 公式中用 D_{RI} 表示)、光谱反射率一阶微分差值指数(difference vegetation index of derivative, DVI, 公式中用 D_{VI} 表示)、光谱反射率一阶微分归一化指数(normalized difference vegetation index of derivative, DNDVI, 公式中用 D_{DNDVI} 表示),植被指数公式见表 2^[24]。

1.2.3 逐步线性回归和偏最小二乘回归 逐步线性回归在构建方程时,先考虑自变量对因变量的作

用显著程度大小,按照从大到小的原则逐个引入方程。每引入一个变量都将检验其显著性 F , 对不符合的变量进行筛选、剔除,确保每次引入新的变量之前方程中只含有显著性的自变量。逐步线性回归可以避免多元线性回归模型中的个别自变量贡献率较小的问题,筛选出显著性因子强的自变量^[25]。偏最小二乘回归(partial least squares regression, PLSR)主要研究多因变量或单因变量对多自变量的回归建模,但是单因变量的偏最小二乘回归模型在日常分析中最为常见^[12]。对于该方法的具体介绍可参考 P. Geladi *et al*^[26] 和 S. Wold *et al*^[27]。偏最小二乘回归适合变量个数很多且存在多重共线性同时样本数较少。在光谱建模过程中,PLSR 能有效辨识光谱信息与噪声,降低光谱维数,减少数据冗余。PLSR 的重点在于主成分个数的确定,主成分的个数会直接影响模型的稳健性,主成分个数过少会出现欠拟合情况,主成分个数过大会出现过拟合情况^[28]。PLSR 模型在 SPSS 中的实现过程参考王国华等^[29]。

表 2 植被指数运算公式

Table 2 Spectral index in this study

植被指数	计算公式	说明
RI	$R_1 = R_i / R_j$	式中: i 与 j 代表光谱任意波段; R_i 、 R_j 为 i 、 j 波段的光谱反射率; R'_i 、 R'_j 为 i 、 j 波段的光谱反射率一阶微分
VI	$V_1 = R_i - R_j$	
NDVI	$N_{DVI} = (R_i - R_j) / (R_i + R_j)$	
DRI	$D_{RI} = R'_i / R'_j$	
DVI	$D_{VI} = R'_i - R'_j$	
DNDVI	$D_{DNDVI} = (R'_i - R'_j) / (R'_i + R'_j)$	

1.2.4 精度验证 模型构建与验证分析在 SPSS23.0 中完成,选用的模型精度验证的 2 个指标决定系数(coefficient of determination, R^2),均方根误差(root mean square error, RMSE, 公式中用 R_{MSE} 表示),其计算公式:

$$R^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / \sum_{i=1}^n (y_i - \bar{y})^2 \quad (4)$$

$$R_{MSE} = \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / n} \quad (5)$$

式中: n 为样本数; \hat{y}_i 为预测值; \bar{y} 为预测值的平均值; y_i 为实测值。决定系数 R^2 用来评价回归模型的拟合优度,取值 0~1,越靠近 1,说明模型的拟合程度越高,稳定性越好。均方根误差 RMSE 用来衡量预测值和观测值之间的偏差,用来检验模型的预测能力, RMSE 越小,说明模型估算能力越高^[30]。

2 结果与分析

2.1 不同叶片叶绿素含量光谱曲线特征分析

对 59 个样本按照叶绿素含量进行升序排列,采

用等间隔方法选取 6 个样本,其 SPAD 值分别为 3.7、14.17、23.4、37.4、46.6、58.9,得到对应的 6 条原始光谱曲线(图 1)。该曲线存在以下特征:1)波长 350~500 nm,光谱反射率区别较小;2)可见光波段 500~760 nm,光谱反射率波动趋势变化较大,不同叶绿素含量的曲线峰值波段有所差别;3)在 760~1 360 nm,光谱反射率达到最大值,曲线比较平稳,6 条曲线趋于平行;4)650~700 nm 存在明显的叶绿素吸收谷,1 360~1 470 nm 存在明显的水分吸收谷,但是部分曲线的吸收深度和吸收面积存在差异。

2.2 相关分析

2.2.1 叶绿素含量和变换光谱的相关性分析 分别计算叶绿素含量与 R 、 $1/R$ 、 LR 、 FDR 、 SDR 之间的相关系数,其相关性分析见图 2。

在 0.01 显著性水平上,原始光谱 R 在 560~720 nm 显著负相关,最大相关系数波长 699 nm ($r = -0.8009$),峰值带宽较小;光谱倒数 $1/R$ 在 580~720 nm 显著正相关,最大相关系数波长 700 nm ($r =$

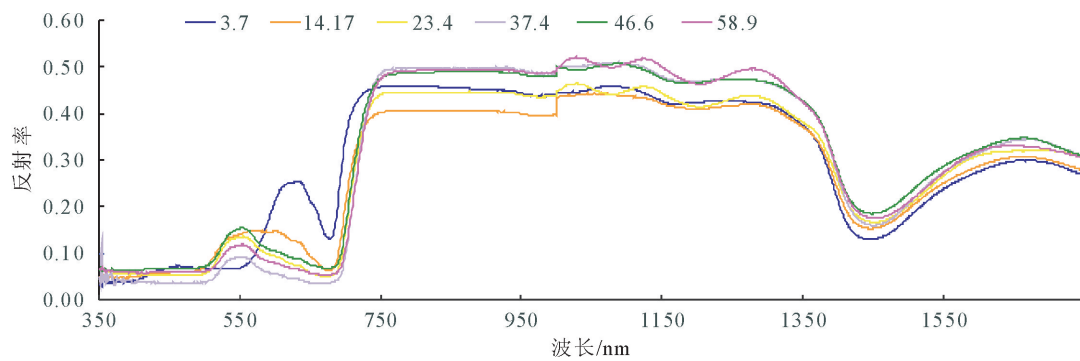


图1 不同 SPAD 值对应的原始光谱曲线

Fig. 1 Original spectrum curves of different SPAD values of tree leaves

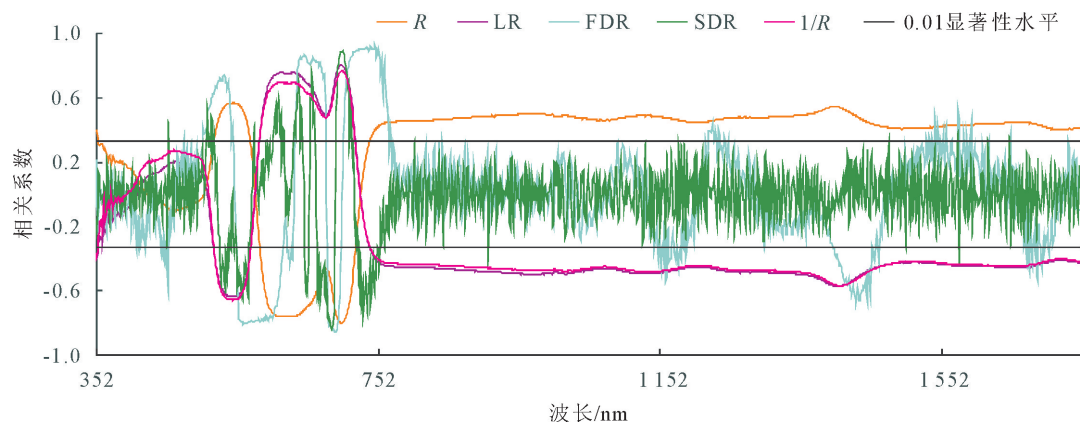


图2 SPAD 值与原始光谱、倒数、一阶微分、二阶微分光谱的相关系数

Fig. 2 Relationship between the SPAD values and spectral reflectance and reciprocal spectral reflectance and the first-derivative spectral reflectance and the second-derivative spectral reflectance

0.766 2);光谱倒数之对数 LR 与 $1/R$ 曲线几乎重合,但相关性比 $1/R$ 稍高,最大相关系数波长 698 nm($r=0.804\ 4$);光谱一阶微分 FDR 消除了相邻波段的影响,相关性比 R 有所增强,相关性变化剧烈,最大相关系数波长 747 nm($r=0.928\ 3$),其峰值带宽较大;SDR 最大相关系数波长 701 nm($r=0.894\ 1$),其峰值带宽较小。各种光谱反射率与叶绿素含量之间相关系数最大波长均在可见光波段。

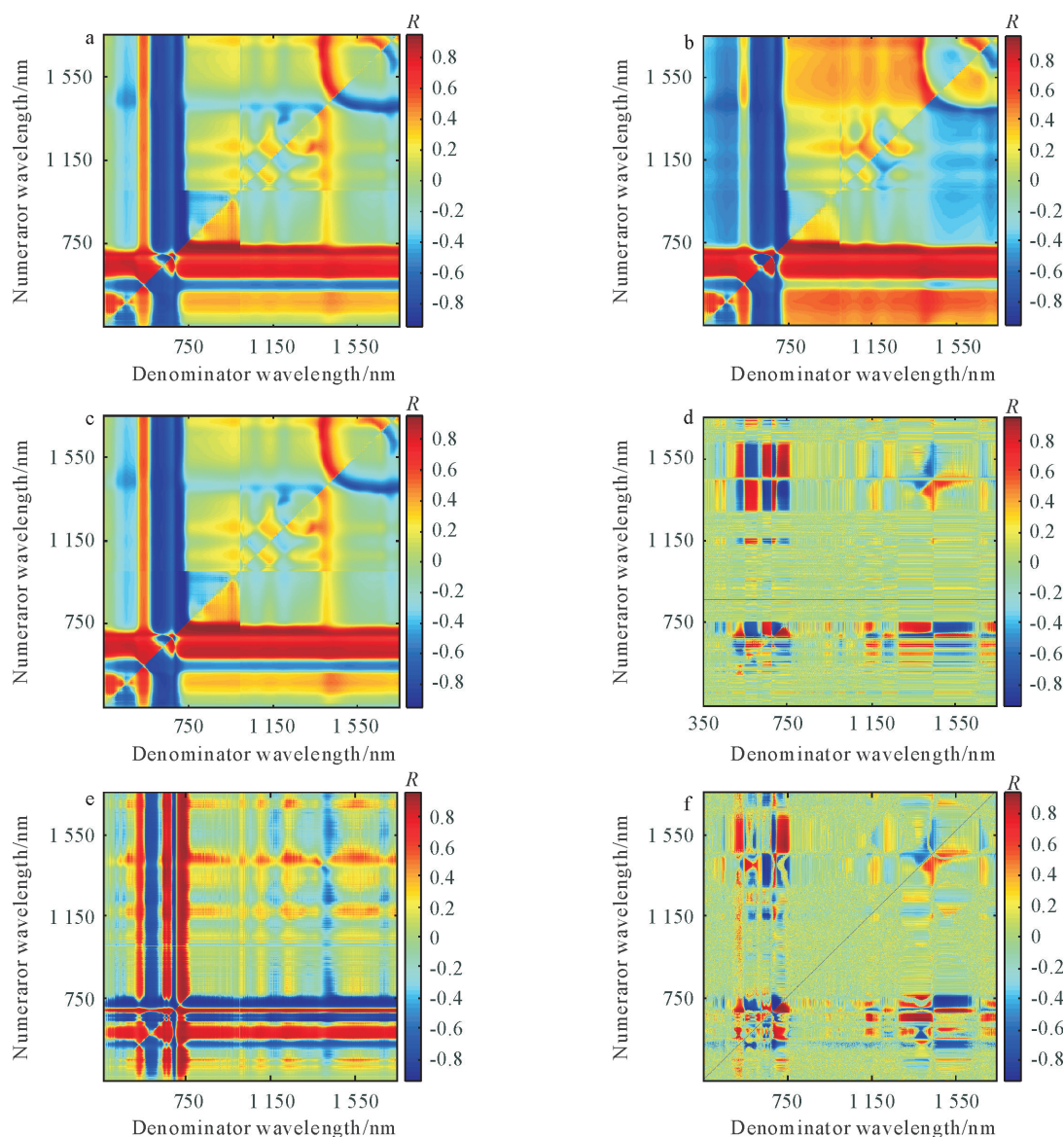
在可见光波段 350~760 nm,叶绿素含量与各种光谱反射率之间相关性有正有负,变化明显。其中,LR 与 $1/R$ 2 条曲线差异最小,部分重合。FDR、SDR 的相关系数变化较大,在正负之间波动,与 R 的相关系数相比有所增加,说明 FDR、SDR 可以将 R 在可见光波段的叶绿素光谱吸收特征反映出来。在 760~1 700 nm,R、LR 与 $1/R$ 3 条曲线趋于平缓,在 1 400 nm 左右的水分吸收带有少量增强,FDR、SDR 2 条曲线变化仍然较大,相关性时正时负,在水分吸收带时变化较为明显。

对叶绿素含量与 R、 $1/R$ 、LR、FDR、SDR 的相关系数进行显著性检验。在 0.01 显著性水平上的波段有:R 的 519~569、741~1 718 nm; $1/R$ 的 583~719 nm;LR 的 518~719、740~1 518 nm;

FDR 的 503~546、634~677、702~769、1 221~1 236 nm;SDR 的 527~567、575、604~760 nm。

2.2.2 叶绿素含量与植被指数相关性分析 研究了 350~2 000 nm 范围内任意 2 波段组合的 6 种植被指数与石楠(PS)叶片的叶绿素含量值的相关性。为突出显示最佳波段组合,利用 MATLAB 软件分析并制作相关性矩阵二维图(图 3),红色到蓝色表示高正相关到高负相关。

波长之间的相关分析表明,可见光波段(350~760 nm)的 2 个波段之间相关性相对较强,特别是红光波段相关系数最高。近红外短波阶段(780~1 100 nm)光谱数据冗余较多,近红外长波阶段相关性有所增强。对比不同植被指数,RI、VI、NDVI 的敏感波段组合相似,DRI、DVI、DNDVI 的敏感波段组合相似。对相关系数进行统计分析,挑选相关系数的最大绝对值所在波段作为特征波段。结果表明,RI 的最佳波段组合为 RI(R_{733} , R_{944}),相关系数 0.952 6,VI 的最佳波段组合为 VI(R_{732} , R_{980}),相关系数 0.956 2,NDVI 的最佳波段组合为 NDVI(R_{732} , R_{931}),相关系数 0.952 4,DRI 的最佳波段组合为 DRI(R_{747} , $R_{1\ 464}$),相关系数 0.950 2,DVI 的最佳波段组合为 DVI(R_{747} , $R_{1\ 464}$),相关系数 0.944 8,



注: a. SPAD and RI; b. SPAD and VI; c. SPAD and NDVI; d. SPAD and DRI; e. SPAD and DVI; f. SPAD and DNDVI。

图 3 不同植被指数与 SPAD 的相关系数矩阵

Fig. 3 Correlation matrix of relationship between the SPAD values and spectral index calculated by different spectra reflectance

DNDVI 的最佳波段组合为 $DNDVI(R_{645}, R_{1370})$, 相关系数 0.936 9。

2.3 叶片叶绿素高光谱估算模型构建

2.3.1 基于逐步线性回归和 PLSR 的叶片叶绿素估算模型 在通过显著性检验的光谱曲线中, 挑选相关系数为波峰或波谷的波段作为特征波段。在 R 曲线中, 特征波段为 615、630、664、699 nm; $1/R$ 曲线中, 特征波段为 546、560、593、631、659、700 nm; LR 曲线中, 特征波段为 546、558、621、631、664、698 nm; FDR 曲线中, 特征波段为 532、536、541、553、560、611、636、647、655、660、680、691、705、735 nm; SDR 曲线中, 特征波段为 567、639、658、685、693、701、708、736 nm。

分别以各种形式光谱特征波段的反射率作为自变量, 以叶绿素含量为因变量, 在 SPSS 中构建逐步

线性回归模型和偏最小二乘模型。构建逐步回归方程时, 先进行共线性诊断, 确保入选的自变量之间不存在共线性。设置入选条件为显著性 F 检验的概率值小于 0.05。构建偏最小二乘模型时, 根据潜在因子的方差解释比例, 确定最佳的潜在因子数。为检验模型的可靠性和实用性, 采用决定系数 R^2 和均方根误差 RMSE 对模型进行评定, 结果见表 3。

在叶绿素的逐步线性回归模型中, 对 R 进行 LR、FDR、SDR 变换后, 建模集和验证集的 R^2 都变大了, RMSE 都减小了, 其中 FDR 变换的光谱对叶绿素含量预测效果最好。

在叶绿素的最小偏二乘模型中, 基于变换光谱建立的模型较于原始光谱有效提高了反演精度, 验证集的 R^2 变大了, RMSE 均减小, 但是建模集的 $1/R$ 、SDR 变换后, 2 个精度指标变化不太一致, 模

型精度最高的仍然是 FDR。从建模集的精度来看, R、SDR 的逐步线性回归模型优于偏最小二乘模型。从验证集的精度来看, LR 的偏最小二乘模型优于逐步线性回归模型。综合来看, 最优模型为 FDR 的逐步线性回归模型。

因此, 对 R 进行 LR、FDR、SDR 变换, 能够提高叶绿素含量预测模型的精度和稳定性。

2.3.2 基于植被指数的单变量叶绿素估算模型

选择线性函数、二次函数、对数函数、逆函数、指数函数分别与 6 种植被指数拟合, 构建叶绿素含量的单变量估算模型。通过比较决定系数 R^2 和均方根误差 RMSE, 探索叶绿素含量的最佳单变量估算模型, 结果见表 4。

以上各种模型中, 基于同一植被指数所建立的 5 种模型中, 线性函数模型普遍具有较高的决定系数和较低的均方根误差, 模型精度较高。二次函数

表 3 基于原始光谱和变换光谱的叶绿素估算模型及精度

Table 3 SPAD estimation models and accuracies based on original spectra and conversion spectra

光谱指标	逐步线性回归					偏最小二乘				
	入选波段 /nm	建模集		验证集		因子数	建模集		验证集	
		R^2	RMSE	R^2	RMSE		R^2	RMSE	R^2	RMSE
R	699/664/630	0.758	5.970	0.841	7.706	3	0.736	5.681	0.838	7.755
1/R	700/593	0.689	6.669	0.932	5.151	4	0.7	5.97	0.919	5.597
LR	700/546/659/631	0.802	5.471	0.882	6.528	5	0.787	4.957	0.899	6.08
FDR	735/680	0.819	5.084	0.958	4.057	4	0.859	4.096	0.931	5.201
SDR	701/693	0.894	5.358	0.908	5.957	2	0.786	7.452	0.886	6.592

表 4 SPAD 的单变量拟合模型及精度

Table 4 SPAD estimation models and accuracies based on single spectral index

植被指数	函数类型	模型	建模集		验证集	
			R^2	RMSE	R^2	RMSE
RI	线性函数	$y = 275.684 - 269.923 \times x$	0.885	3.994	0.930	4.030
	二次函数	$y = 73.318 \times x - 183.821 \times x^2 + 115.749$	0.885	4.035	0.920	4.256
	对数函数	$y = 6.180 - 251.368 \times \ln(X)$	0.884	4.024	0.937	3.873
	逆函数	$y = -226.91 + 233.52/x$	0.882	4.073	0.942	3.742
	指数函数	$y = 7\,659\,601.763 \times e^{((-13.748)x)}$	0.867	0.221	0.887	5.671
VI	线性函数	$y = 9.871 + 500.580 \times x$	0.889	3.933	0.941	3.805
	二次函数	$y = 585.021 \times x - 1\,491.581 \times x^2 + 9.379$	0.893	3.920	0.918	4.333
	对数函数	—	—	—	—	—
	逆函数	$y = 23.984 - 0.004/x$	0.042	11.553	0.070	5.616
	指数函数	$y = 10.155 \times e^{(25.305 \times x)}$	0.857	0.230	0.888	14.02
NDVI	线性函数	$y = 4.222 + 473.587 \times x$	0.884	4.039	0.943	3.691
	二次函数	$y = 582.603 \times x - 1\,303.782 \times x^2 + 2.624$	0.887	4.024	0.919	4.235
	对数函数	—	—	—	—	—
	逆函数	$y = 23.412 + 0.009/x$	0.021	11.679	0.270	13.871
	指数函数	$y = 7.592 \times e^{(24.070 \times x)}$	0.861	0.228	0.883	6.106
DRI	线性函数	$y = 3.580 + 10.956 \times x$	0.899	3.750	0.921	4.218
	二次函数	$y = 11.281 \times x - 0.087 \times x^2 + 3.363$	0.899	3.799	0.920	4.2
	对数函数	$y = 18.018 + 14.207 \times \ln(X)$	0.812	5.132	0.807	5.92
	逆函数	$y = 31.287 - 8.334/x$	0.494	8.392	0.531	9.317
	指数函数	$y = 7.480 \times e^{(0.547 \times x)}$	0.846	0.239	0.833	12.675
DVI	线性函数	$y = 14.924 + 20\,925.874 \times x$	0.854	4.513	0.952	2.945
	二次函数	$y = 24\,053.357 \times x - 3\,610\,614.862 \times x^2 + 15.181$	0.859	4.482	0.911	4.03
	对数函数	—	—	—	—	—
	逆函数	$y = 23.401$	0.015	11.714	—	—
	指数函数	$y = 13.082 + 1\,063.065^x$	0.832	0.249	0.899	7.77
DNDVI	线性函数	$y = 43.827 \times x + 17.308$	0.850	4.567	0.940	4.022
	二次函数	$y = 49.425 \times x - 17.962 \times x^2 + 17.928$	0.856	4.538	0.904	4.586
	对数函数	—	—	—	—	—
	逆函数	$y = 23.458 + 0.002/x$	0.000	11.803	0.007	13.484
	指数函数	$y = 14.832 \times e^{(2.195 \times x)}$	0.805	0.268	0.960	2.832

注: y 表示 SPAD 值, x 代表对应的植被指数。

模型的 R^2 和 RMSE 都很接近。除了 RI 的逆函数 $R^2 > 0.6$ 外,其他几个逆函数的 R^2 均小于 0.6,表明逆函数不能进行叶绿素估算。对数函数与指数函数的 R^2 和 RMSE 有的最好,有的较差。从建模集来看,精度最优的模型是 DRI 的线性模型,从验证

集来看,最优的是 DNDVI 的指数模型。

3.3.3 基于植被指数的多变量叶绿素估算模型

以叶绿素含量为因变量,6 种植被指数为自变量,在 SPSS 中构建逐步线性回归模型,并利用验证样本对建立的模型进行精度评价(表 5)。

表 5 基于植被指数的叶绿素估算模型及精度

Table 5 SPAD estimation models and accuracies based on spectral indexes

方法	模型	建模集		验证集	
		R^2	RMSE	R^2	RMSE
逐步线性回归	$132.348 + 6.198 \times \text{DRI} - 128.567 \times \text{RI}$	0.930	3.155	0.955	3.145
偏最小二乘	$77.370 + 7.735 \times \text{DRI} - 76.749 \times \text{RI} - 5.208 \times \text{DNDVI} + 102.522 \times \text{NDVI} + 81.325 \times \text{VI} - 4.586.830 \times \text{DVI}$	0.928	2.957	0.948	3.449

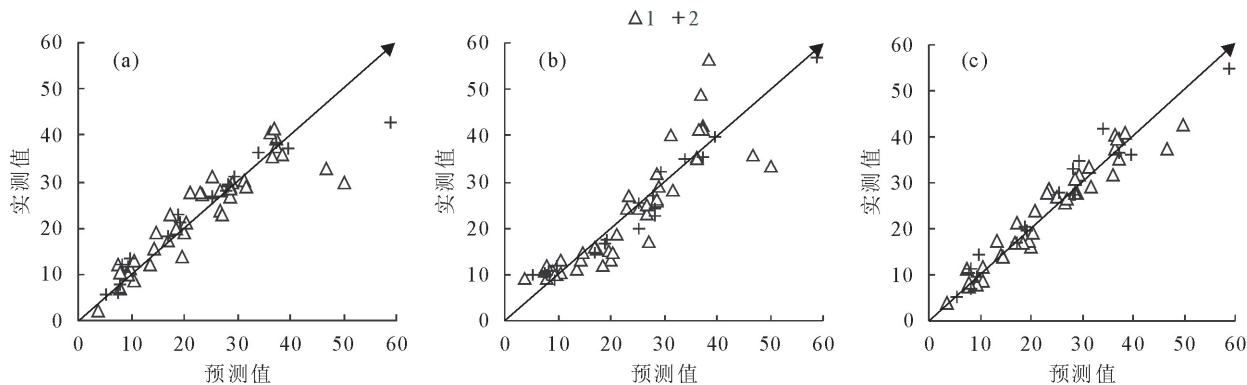
逐步线性回归模型最终仅入选了 DRI 和 RI 2 个变量,其建模集的 R^2 达到了 0.93,说明模型预测效果非常高, RMSE 为 3.145,说明预测精度也很理想,同时检验精度也达到了 0.955。

2.4 模型精度比较

依据上述分析,5 种光谱指标的逐步线性回归模型和偏最小二乘模型的 R^2 取值在 0.700 ~ 0.958,预测精度最高为 FDR 的逐步线性回归模型。6 种植被指数的 5 类函数模型中,除逆函数外,其余

4 类函数 R^2 取值在 0.8 ~ 0.96,预测精度最高为 DNDVI 的指数模型。6 种植被指数构建的逐步线性回归模型最佳为 DRI+RI, R^2 取值 0.955。

为了直观展示叶绿素估算模型的拟合度和可靠性,绘制了预测值与实测值之间的 1 : 1 关系图(图 4),可以看出基于植被指数 DRI+RI 的线性回归模型的样本点集中在 1 : 1 直线的两侧,比单光谱指标和单植被指数模型的精度都要高。



注:(a)-FDR 模型预测值与实测值比较;(b)-DNDVI 指数模型预测值与实测值比较;(c)-DRI+RI 线性回归模型预测值与实测值比较;1-建模样本;2-验证样本。

图 4 叶绿素含量最优模型的实测值与预测值 1 : 1 关系

Fig. 4 1 : 1 relationship between measured and predicted values of the optimal SPAD estimation models

3 结论与讨论

通过对入选波段的统计表明,石楠叶绿素含量相关性最高的波段在红波段(622 ~ 770 nm);利用逐步线性回归和 PLSR 构建的各种变换光谱与叶绿素含量的预测模型结果具有一致性,光谱的一阶微分变换后能有效增强有价值波段信息,提高模型的精度;使用 5 种函数拟合的 6 种植被指数的预测模型中,仅 RI 的 5 种函数精度都较高,其他 5 种植被指数的对数函数和逆函数不能预测叶绿素含量;石楠叶片叶绿素估算最佳模型是 RI 和 DRI 的线性模型,说明比值植被指数比差值、归一化值植被指数有

更好的预测能力。因此,本研究成果能为高光谱遥感技术在园林绿化中的应用提供依据。

对原始光谱与 4 种变换光谱构建叶绿素含量的逐步线性回归和偏最小二乘模型,发现基于 FDR 的逐步线性回归和偏最小二乘模型均最优。这可能与 FDR 消除了相邻波段的干扰有关,使得光谱的吸收特征信息被释放,相关系数最大值提高了 0.13,从而改善了模型精度。由于 R、1/R、LR、SDR 入选的特征波段有限,使得其偏最小二乘模型精度低于逐步线性回归,而 FDR 构建的偏最小二乘模型精度与逐步线性模型较为接近,这可能与其入选特征波段较多有关。从试验结果分析可知,偏最小二乘模型

更适合变量多且变量间存在较大共线性的情况,今后可尝试用所有显著性波段构建该模型。

6种植被指数与叶绿素含量间具有极显著相关性,最大相关系数都大于0.93,均高于原始光谱及变换光谱。研究表明,逆函数和对数函数只适用于RI,线性、二次、指数函数适用于6种植被指数,且6种植被指数构建的逐步线性回归模型中,RI(R_{733} , R_{944})和DRI(R_{747} , R_{1464})2个变量建立的模型精度优于单植被指数变量模型。说明多个植被指数结合也可以提高叶绿素含量的估算精度。

参考文献:

- [1] 郭超凡,郭道宇.基于可见光波段包络线去除的湿地植物叶片叶绿素估算[J].生态学报,2016,36(20):6538-6546.
GUO C F, GUO X Y. Estimation of wetland plant leaf chlorophyll content based on continuum removal in the visible domain[J]. Acta Ecologica Sinica, 2016, 36(20): 6538-6546. (in Chinese)
- [2] 宫兆宁,赵雅莉,赵文吉,等.基于光谱指数的植物叶片叶绿素含量的估算模型[J].生态学报,2014,34(20):5736-5745.
GONG Z N, ZHAO Y L, ZHAO W J, et al. Estimation model for plant leaf chlorophyll content based on the spectral index content[J]. Acta Ecologica Sinica, 2014, 34(20): 5736-5745. (in Chinese)
- [3] 王丽爱,马昌,周旭东,等.基于随机森林回归算法的小麦叶片SPAD值遥感估算[J].农业机械学报,2015,46(1):259-265.
- [4] 刘文雅,潘洁.基于神经网络的马尾松叶绿素含量高光谱估算模型[J].应用生态学报,2017,28(4):1128-1136.
- [5] 冯海宽,杨福芹,杨贵军,等.基于特征光谱参数的苹果叶片叶绿素含量估算[J].农业工程学报,2018,34(6):182-188.
FENG H K, YANG F Q, YANG G J, et al. Estimation of chlorophyll content in apple leaves base on spectral feature parameters[J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(6): 182-188. (in Chinese)
- [6] 黄慧,王伟,彭彦昆,等.利用高光谱扫描技术检测小麦叶片叶绿素含量[J].光谱学与光谱分析,2010,30(7):1811-1814.
- [7] 李媛媛,常庆瑞,刘秀英,等.基于高光谱和BP神经网络的玉米叶片SPAD值遥感估算[J].农业工程学报,2016,32(16):135-142.
LI Y Y, CHANG Q R, LIU X Y, et al. Estimation of maize leaf SPAD value based on hyper spectrum and BP neural network[J]. Transactions of the Chinese Society of Agricultural Engineering, 2016, 32(16): 135-142. (in Chinese)
- [8] 胥喆,舒清杰,杨凯博,吴娇娇,张亚红.基于非成像高光谱的高山松叶绿素估算模型研究[J].西北林学院学报,2017,32(2):73-78,179.
XU Z, SHU Q T, YANG K B, et al. Estimation models of chlorophyll in *Pinus densata* based on non-imaging hyperspectrum[J]. Journal of Northwest Forestry University, 2017, 32(2): 73-78, 179.
- [9] 尼加提·卡斯木,师庆东,王敬哲,等.基于高光谱特征和偏最小二乘法的春小麦叶绿素含量估算[J].农业工程学报,2017,33(22):208-216.
- [10] 邓小蕾,李民赞,郑立华,等.基于反射光谱预处理的苹果叶片叶绿素含量预测[J].农业工程学报,2014,30(14):140-147.
- [11] 何彩莲,郑顺林,周少猛,等.基于高光谱植被指数的马铃薯叶片叶绿素含量估测模型[J].华南农业大学学报,2016,37(5):45-49.
- [12] 姜海玲,杨杭,陈小平,等.利用光谱指数反演植被叶绿素含量的精度及稳定性研究[J].光谱学与光谱分析,2015,35(4):975-981.
JIANG H L, YANG H, CHEN X P, et al. Research on accuracy and stability if inversing vegetation chlorophyll content by spectral index method[J]. Spectroscopy and Spectral Analysis, 2015, 35(4): 975-981. (in Chinese)
- [13] TUMBO S D, WAGNER D G, HEINEMANN P H. On-the-go sensing of chlorophyll status in corn[J]. Transactions of the ASAE, 2002, 45(4): 1207-1215.
- [14] 程志庆,张劲松,孟平,等.杨树叶片叶绿素含量高光谱估算模型研究[J].农业机械学报,2015,46(8):264-271.
- [15] 孙红,李民赞,张彦娥,等.玉米生长期叶片叶绿素含量检测研究[J].光谱学与光谱分析,2010,30(9):2488-2492.
SUN H, LI M Z, ZHANG Y E, et al. Detection of corn chlorophyll content using canopy spectral reflectance[J]. Spectroscopy and Spectral Analysis, 2010, 30(9): 2488-2492. (in Chinese)
- [16] 武旭梅,常庆瑞,落莉莉,等.水稻冠层叶绿素含量高光谱估算模型[J].干旱地区农业研究,2019,37(3):238-243.
- [17] 罗丹,常庆瑞,齐雁冰,等.基于光谱指数的冬小麦冠层叶绿素含量估算模型研究[J].麦类作物学报,2016,36(9):1225-1233.
LUO D, CHANG Q R, QI Y B, et al. Estimation model for chlorophyll content in winter wheat canopy based on spectral indices[J]. Journal of Triticeae Crops, 2016, 36(9): 1225-1233. (in Chinese)
- [18] 陈秀青,杨琦,韩景晔,等.基于叶冠尺度高光谱的冬小麦叶片含水量估算[J].光谱学与光谱分析,2020,40(3):891-897.
- [19] 王鑫梅,牟洪香,李春友,等.107杨叶片叶绿素含量高光谱反演的研究[J].西北林学院学报,2015,30(3):35-38,44.
WANG X M, MOU H X, LI C Y, et al. Inversion of *Populus × euramericana* cv. "74/76" chlorophyll concentration based on the hyper-spectrum[J]. Journal of Northwest Forestry University, 2015, 30(3): 35-38, 44.
- [20] 马文勇,王训明.基于高光谱分析的草地叶绿素含量估算研究进展[J].地理科学进展,2016,35(1):25-34.
MA W Y, WANG X M. Progress on grassland chlorophyll content estimation by hyperspectral analysis[J]. Progress in Geography, 2016, 35(1): 25-34. (in Chinese)
- [21] 李春,刘卫国,丁旭,等.基于红边参数的芦苇叶绿素的反演研究[J].生态科学,2017,36(3):66-73.
LI C, LIU W G, DING X, et al. Inversion model for chlorophyll content of *Phragmites* under different sample areas based on red edge parameter[J]. Ecological Science, 2017, 36(3): 66-73. (in Chinese)

- grassland in Loess Hilly Region of Ningxia[J]. Northwest Agricultural Sciences, 2013, 22(1):200-206. (in Chinese)
- [16] 王红利,郭蓉,王红红,等. 铅锌矿山废弃地人工植被恢复研究[J]. 中国林副产品, 2017, 3(6):27-28.
- [17] 王红利,李军. 徽废弃铅锌浮选尾矿库表层植被培育[J]. 中国林副产品, 2017, 4(8):75-76.
- [18] 杨海裕,张宋智,刘小林,等. 秦岭西段天然落叶阔叶林乔木种间关联性[J]. 生态学报, 2012, 31(10):2513-2520.
YANG H Y, ZHANG S Z, LIU X L, *et al.* Interspecific associations among tree species in natural deciduous broadleaved forests in western Qinling, northwest China[J]. Chinese Journal of Ecology, 2012, 31(10):2513-2520. (in Chinese)
- [19] 安定国. 甘肃省小陇山高等植物志[M]. 兰州:甘肃民族出版社, 2002.
- [20] 张建贵,王理德,姚拓,等. 祁连山高寒草地不同退化程度植物群落结构与物种多样性研究[J]. 草业学报, 2019, 28(5):15-25.
ZANG J G, WANG L D, YAO T, *et al.* Plant community structure and species diversity differences in alpine grassland in the Qilian Mountains with different levels of degradation[J]. Acta Prataculturae Sinica, 2019, 28(5):15-25. (in Chinese)
- [21] 何芳兰,刘世增,李昌龙,等. 甘肃河西戈壁植物群落组成特征及其多样性研究[J]. 干旱区资源与环境, 2016, 30(4):74-78.
HE F L, LIU S Z, LI C L, *et al.* Study on the composition and diversity of Gobi Desert plant community in Hexi Gansu Province[J]. Arid Area Resources and Environment, 2016, 30(4):74-78. (in Chinese)
- [22] 党荣理,潘晓玲. 西北干旱荒漠区种子植物科的区系分析[J]. 西北植物学报, 2002, 22(1):24-32.
DANG R L, PAN X L. Floristic analysis of the spermatophyte family in the northwestern arid desert region[J]. Acta Botanica Boreali-Occidentalia Sinica, 2002, 22(1):24-32. (in Chinese)
- [23] 安俊珍. 风化型土质金矿尾矿库植被恢复研究[D]. 武汉:华中农业大学, 2010.
- [24] 崔静,黄佳健,陈云明,等. 黄土丘陵区人工柠条林下草本植物物种多样性研究[J]. 西北林学院学报, 2018, 33(3):14-20.
CUI J, HUANG J J, CHEN Y M, *et al.* Biodiversity of herbaceous species under caragana microphylla plantations in Loess Hilly Region[J]. Journal of Northwest Forestry University, 2018, 33(3):14-20. (in Chinese)
- [25] SU Y Z, ZHANG T H, LI Y L, *et al.* Changes in soil properties after changes in soil properties after establishment of *Artemisia halodendron* and *Caragana microphylla* on shifting sand dunes in semiarid Horqin sandy land, northern China[J]. Environmental Management, 2005, 36:272-281.
- [26] 王国梁,刘国彬,刘芳,等. 黄土沟壑区植被恢复过程中植物群落组成及结构变化[J]. 生态学报, 2003, 23(12):2550-2557.
WANG G I, LIU G B, LIU F, *et al.* Changes in composition and structure of plant communities during the course of restoration at loess gully region[J]. Acta Ecologica Sinica, 2003, 23(12):2550-2557. (in Chinese)

(上接第 32 页)

- [22] 艾金泉,陈文惠,陈丽娟,等. 冠层水平互花米草叶片光合色素含量的高光谱遥感估算模型[J]. 生态学报, 2015, 35(4):1175-1186.
AI J Q, CHEN W H, CHEN L J, *et al.* Hyperspectral remote sensing estimation models for foliar photosynthetic pigment contents at canopy level in an invasive species, *Spartina alterniflora*[J]. Acta Ecologica Sinica, 2015, 35(4):1175-1186. (in Chinese)
- [23] 徐晋,蒙继华. 农作物叶绿素含量遥感估算的研究进展与展望[J]. 遥感技术与应用, 2016, 31(1):74-85.
XU J, MENG J H. Overview on estimating crop chlorophyll content with remote sensing[J]. Remote Sensing Technology and Application, 2016, 31(1):74-85. (in Chinese)
- [24] 彭建,徐飞雄,邓凯,等. 琅琊山景区不同叶绿素条件下树种叶片光谱差异分析[J]. 光谱学与光谱分析, 2018, 38(6):1839-1849.
- [25] 罗庆辉,徐泽源,许仲林. 天山雪岭云杉林生物量估测及空间格局分析[J]. 生态学报, 2020, 40(15):5288-5297.
LUO Q H, XU Z Y, XU Z L. Estimation and spatial pattern analysis of biomass of piceaschrenkiana forests[J]. Acta Ecologica Sinica, 2020, 40(15):5288-5297. (in Chinese)
- [26] GELADI P, KOWALSKI B R. Partial least-squares regression; a tutorial[J]. Anal Chim Acta, 1986, 185:1-17.
- [27] WOLD S, SJÖSTRÖM M, ERIKSSON L. PLS-regression; a basic tool of chemometrics[J]. Chemometrics & Intelligent Laboratory Systems, 2001, 58:109-130.
- [28] 陶培峰,王建华,李志忠,等. 基于高光谱的土壤养分含量反演模型研究[J]. 地质与资源, 2020, 29(1):68-75, 84.
- [29] 王国华,张虎,魏岳嵩. 偏最小二乘回归在 SPSS 软件中的实现[J]. 统计与决策, 2017(7):67-71.
- [30] 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[J]. 农业工程学报, 2015, 31(14):103-109.
YU L, HONG Y S, GENG L, *et al.* Hyperspectral estimation of soil organic matter content based on partial least squares regression[J]. Transactions of the Chinese Society of Agricultural Engineering, 2015, 31(14):103-109. (in Chinese)