

基于高光谱遥感的玉米叶片 SPAD 值估算模型研究

董哲 杨武德 张美俊 朱洪芬 王超

(山西农业大学旱作农业工程研究所, 030801, 山西太谷)

摘要 灌浆期玉米叶片叶绿素含量对玉米光合作用及产量形成具有重要作用。为通过高光谱特征准确、高效估测玉米叶片叶绿素含量,以 SPAD 值表征叶绿素相对含量,构建了基于光谱特征参数的传统回归模型、基于全谱和光谱特征参数的 PLSR 模型和 BP 神经网络模型,并进行了比较分析。结果表明:基于全谱构建的 PLSR 模型 SPAD 值拟合效果最好 ($R^2=0.910$, $RMSE=2.071$),而基于光谱特征参数所建立的 PLSR 模型拟合效果可达到与全谱 PLSR 模型相近的水平。但后者的实测值与预测值拟合效果 ($R^2=0.867$, $RMSE=2.581$, $RPD=2.628$) 优于全谱 PLSR 模型,且建模时间短,模型复杂程度降低。BP 神经网络模型相较于两种 PLSR 模型预测效果略差,但优于传统回归模型。综合来看,基于光谱特征参数建立的 PLSR 模型估测效果最好。

关键词 高光谱;玉米叶片;SPAD;偏最小二乘;BP 神经网络

叶绿素含量是衡量作物光合作用和生长发育的重要指标,具有获取和传输能量的重要作用,另外,由于植物叶片中的氮素大部分集中在叶绿素中,故叶绿素含量能够间接评价植物营养状态^[1]。因此,监测叶绿素含量的变化可以反映作物的长势及营养状况,为作物的田间管理提供依据。采用传统的化学分析方法测定叶绿素含量,不仅会破坏植物叶片组织,而且会耗费时间及试剂,且难以做到实时估测。

在实际研究和应用中,常用便携式叶绿素仪 (SPAD-502) 测量植物叶片的 SPAD 值来表征叶片的叶绿素含量,通过测定叶片在红光 (650nm) 和近红外光 (940nm) 两处波长范围内的透射率来反映叶片叶绿素含量,其 SPAD 值读数可以较准确地表征叶绿素含量,测定方法简便,在不同作物的叶绿素含量测定中广泛应用^[2-5]。

但利用叶绿素仪获取 SPAD 值时,需要人工将叶片插入,反复测量,难以适用于大范围农田的叶绿素含量监测。研究表明,叶绿素浓度与可见光和近红外光波段反射率有较显著的关系^[6-8]。因此,可以利用叶绿素含量和高光谱反射率的关系,建立基于叶片光谱的 SPAD 值估算模型。

作者简介:董哲,硕士研究生,主要从事农业信息技术研究

杨武德为通信作者,教授,研究方向为 3S 技术与作物生产、旱作栽培与作物生态研究

基金项目:国家自然科学基金 (31871571);山西省科技攻关项目 (20150311002-4)

收稿日期:2018-09-21;修回日期:2019-04-04

高光谱监测技术相对于便携式叶绿素仪具有简便、高效的优势,并且能够实现对作物的遥感监测,将高光谱监测技术应用于无人机遥感监测平台、田间车载监测平台等可以实现对大范围农田的叶绿素含量监测,具有非常大的发展潜力^[9]。叶片尺度的高光谱测量数据对于由叶绿素含量变化导致的光谱细微变化非常敏感,从叶片光谱特征中准确估算叶片叶绿素含量是从遥感影像获得冠层叶绿素含量的重要一步^[10]。

目前,高光谱技术在监测植物尤其是农作物的叶绿素含量上取得了很大进展,但应用高光谱数据反演叶绿素含量多采用某一种方法或多种方法之间的比较,反演方法之间缺乏集成,反演精度也有待提高。同时使用多种方法反演玉米叶片叶绿素含量并进行比较的研究尚不多见。为此,以灌浆期 (灌浆期是玉米生殖生长阶段的重要生育时期,持续时间长,便于监测叶绿素含量) 玉米叶片为研究对象,采集玉米叶片光谱信息,分别通过全谱、植被指数及红边参数构建玉米叶片 SPAD 值的估算模型,并对各估算模型的预测效果进行比较与评价,旨在探索估测效果良好、建模方便和实用性强的玉米叶片叶绿素含量估算模型,对后续遥感监测设备开发提供参考。

1 材料与方法

1.1 试验设计

试验于 2018 年 7-8 月在山西省晋中市太谷县

胡村镇现代农业示范区 (37° 12' N, 112° 28' E) 进行。该示范区主要种植夏玉米和冬小麦两种粮食作物, 且种植面积较大。夏玉米品种以晋单 86、强盛 388 和登海 679 等为主, 多采用一次性施足底肥, 自然方式生长, 生育期为 5 月初至 9 月中旬, 行距 50cm, 株距 35cm。采取大田随机采样方式, 选取能代表该区域生产特性、间隔 500~1 000m 的 14 个样点, 每个样点采集若干叶片, 尽量取玉米植株的不同叶位。

1.2 测量指标与方法

采样后立即将样本运回实验室, 同步进行室内叶片光谱与叶绿素 SPAD 值的测定。选用美国 ASD Field-Spec3.0 便携式高光谱仪测量叶片光谱, 波段范围 350~2 500nm。利用自带光源型手持叶片光谱探测器直接测量叶片光谱, 每片叶子均匀

采集 5 次以上光谱, 将平均值作为该样本光谱反射率数据, 每次测量前均用白板进行校正。与光谱测量同步, 使用叶绿素仪 (SPAD-502) 测定叶片的 SPAD 值, 每片叶子均匀测量 10 次, 取平均值作为样本的 SPAD 值, 实现与叶片光谱数据的一一对应。

试验共得到 120 个样本数据, 将该样本按 2:1 随机分为两部分, 80 份数据作为建模集, 40 份数据作为验证集。

1.3 数据分析

选取几种已被前人证明与叶片叶绿素含量有显著相关关系的植被指数与红边参数作为自变量, 包括归一化植被指数 NDVI、绿光归一化植被指数 GNDVI、比值植被指数 RVI、红边位置 λ_r 、红边峰度 k_{λ_r} 和红边偏度 s_{λ_r} (表 1)。

表 1 本试验所用光谱特征参数计算公式

Table 1 Spectral characteristic parameters formulas used in the study

光谱特征参数 Spectral characteristic parameter	定义 Definition	来源 Source
NDVI	$NDVI=(R_{750}-R_{705})/(R_{750}+R_{705})$	Gitelson 等 ^[11]
GNDVI	$GNDVI=(R_{750}-R_{550})/(R_{750}+R_{550})$	Gitelson 等 ^[11]
RVI	$RVI=R_{760}/R_{710}$	Penuelas 等 ^[12]
λ_r	红光范围内一阶导数光谱最大值对应的波长	Horler 等 ^[13]
k_{λ_r}	红光范围内一阶导数光谱的峰度	姚付启等 ^[14]
s_{λ_r}	红光范围内一阶导数光谱的偏度	姚付启等 ^[14]

选取决定系数 R^2 、均方根误差 RMSE、残留预测偏差 RPD 对模型进行综合评定。 R^2 越接近于 1, 模型的精度越高; RMSE 越小, 模型的稳定性越好。RPD 值是样本实测值的标准偏差与样本预测值的均方根误差 RMSE 的比值。RPD<1.000 时表明模型预测能力极差; 1.000<RPD<1.400 时表明模型预测能力较差; 1.400<RPD<1.800 时表明模型预测模型可用; 1.800<RPD<2.000 时表明模型预测能力较好; 2.000<RPD<2.500 时表明模型预测能力很好; RPD>2.500 时表明模型预测能力极好^[15]。

2 结果与分析

2.1 估算玉米叶片 SPAD 值的传统回归模型

对选取的几种植被指数和红边参数分别与玉米叶片 SPAD 值进行相关性分析, 结果如表 2 所示。所有的光谱特征参数均与 SPAD 值呈极显著正相关关系, 相关系数在 0.811~0.935, 说明植被指数和红边参数均能够反演叶片 SPAD 值。其中, 红

边位置 λ_r 与 SPAD 值呈最大正相关, 相关系数为 0.935。

表 2 光谱特征参数与玉米叶片 SPAD 的相关系数 (n=80)

Table 2 Correlation coefficients between the spectral characteristic parameters and SPAD value of maize leaves (n=80)

参数 Parameter		相关系数 Correlation coefficient
植被指数 Vegetation index	NDVI	0.843**
	GNDVI	0.811**
	RVI	0.893**
红边参数 The trilateral parameter	λ_r	0.935**
	k_{λ_r}	0.872**
	s_{λ_r}	0.914**

注: “**” 表示在 0.01 水平上极显著相关

Note: “**” indicates extremely significant correlation at 0.01 level

基于各光谱特征参数分别构建玉米叶片 SPAD 值的拟合模型, 依据 R^2 高、RMSE 低的原则选出各参数的最优拟合模型 (表 3)。从建模效果来看,

基于红边位置 λ_r 建立的幂函数方程决定系数最高, R^2 为 0.880, 但其均方根误差 RMSE 较高, 模型稳定性较差。基于红边偏度 s_{λ_r} 建立的二次函数方程 R^2 为 0.847, 仅低于 λ_r , 且其 RMSE 最低, 模型稳定性较强, 表明通过 s_{λ_r} 建立的方程与 SPAD 值拟合效果最好。将基于各光谱特征参数建立的 SPAD 值估算模型分别代入验证集中, 对求得的

SPAD 预测值与实测值进行线性拟合 (表 3)。通过 λ_r 和 s_{λ_r} 求得的预测值和实测值拟合方程 R^2 均达到 0.800 以上, RPD 均大于 2.000, 表明估算模型的预测能力达到了很好的水平。其中, s_{λ_r} 的验证模型具有更低的 RMSE, 因此具有更高的 RPD, 综合来看为最优的光谱特征参数。基于 NDVI、RVI 和 k_{λ_r} 构建的估算模型预测能力也达到了可用

表 3 基于光谱特征参数的玉米叶片 SPAD 值估算模型
Table 3 SPAD value estimation model of maize leaves based on the spectral indexes

光谱特征参数 Spectral characteristic parameter	最优拟合模型 The optimal fitting equation	建模集 Modeling set		验证集 Validation set		
		R^2	RMSE	R^2	RMSE	RPD
NDVI	$y=12.696e^{2.6521x}$	0.728	3.686	0.636	4.072	1.500
GNDVI	$y=10.909e^{2.7102x}$	0.672	4.020	0.574	4.366	1.340
RVI	$y=20.758x^{0.9129}$	0.812	3.095	0.716	3.715	1.791
λ_r	$y=3E-104x^{36.805}$	0.880	3.774	0.854	3.327	2.005
k_{λ_r}	$y=640.3e^{1.7036x}$	0.784	3.432	0.714	3.657	1.739
s_{λ_r}	$y=87.296x^2+54.674x+46.509$	0.847	2.697	0.804	3.028	2.137

水平。

由于叶绿素的敏感波段主要集中于可见光和近红外光波段, 所以本研究中叶片光谱范围选择为 400~1 300nm。以玉米叶片原始光谱 400~1 300nm 波段内的反射率作为自变量, 以玉米叶片 SPAD 作为因变量, 采用交叉验证法构建基于全谱的 PLSR 模型, 通过比较各潜在因子数建模集与验证集的决定系数 R^2 、均方根误差 RMSE 和残留预测偏差 RPD, 确定最佳因子数为 4 (表 3)。

2.2 基于 PLSR 的玉米叶片 SPAD 值估算模型

基于全谱构建的 PLSR 模型, 引入的自变量多, 导致计算量加大, 而且有些波段所含有用信息较少或者所含相同信息过多, 导致共线性问题及信息冗余, 影响预测模型的精度及稳定性。各光谱特征参数均与 SPAD 值呈极显著相关 (表 1), 将其作为自变量使用 PLSR 建模, 结果如表 4 所示。

表 4 基于 PLSR 的玉米叶片 SPAD 估算模型
Table 4 SPAD value estimation model of maize leaves based on PLSR model

模型 Model	建模集 Modeling set		验证集 Validation set		
	R^2	RMSE	R^2	RMSE	RPD
PLSR (全谱)	0.910	2.071	0.862	2.600	2.545
PLSR (光谱特征参数)	0.900	2.200	0.867	2.581	2.628

由表 4 可知, 基于 PLSR 构建的全谱模型和光谱特征参数模型, 建模集决定系数 R^2 均达到 0.900

以上, 其 SPAD 值的拟合效果均优于传统回归模型。从验证情况来看, 两种 PLSR 模型的预测值和实测值拟合方程的 R^2 均高于传统回归模型, 且 RMSE 更低, RPD 均在 2.500 以上, 表明模型的预测能力极好。基于光谱特征参数建立的 PLSR 模型虽然在建模效果上略逊于全谱 PLSR 模型, 但其验证效果要优于全谱 PLSR 模型, 能更有效地预测玉米叶片 SPAD 值, 且其计算量远远小于全谱 PLSR 模型, 建模更加快速。

2.3 基于 BP 神经网络的玉米叶片 SPAD 值估算

BP 神经网络具有预测性、容错性和实用性的特点, 更适合用于解释非线性问题。采用 BP 神经网络, 将上述光谱特征参数 6 个神经元作为输入层, 将玉米叶片 SPAD 值 1 个神经元作为输出层, 通过多次试错得出隐含层神经元数为 7 时训练结果相对较好。构建的 BP 神经网络估算模型最终结构为: 1 个输入层 (6 个神经元)、1 个隐含层 (7 个神经元) 及 1 个输出层 (1 个神经元)。设定最大学习迭代次数为 10 000, 学习精度为 0.01。使用 MATLAB 2010b 对 BP 神经网络进行训练, 建模并验证 (表 5), 其建模决定系数 R^2 同样在 0.900 以上, SPAD 值的拟合效果与 PLSR 模型相差不大, 同样优于传统回归模型。但其验证情况不如 PLSR 模型, 验证集 R^2 较小, RMSE 较大, RPD 小于 2.500, 预测能力没有达到极好水平。

表 5 基于 BP 神经网络的玉米叶片 SPAD 估算模型
Table 5 SPAD value estimation model of maize leaves based on BP neural network

建模集 Modeling set		验证集 Validation set		
R^2	RMSE	R^2	RMSE	RPD
0.904	2.165	0.822	2.861	2.305

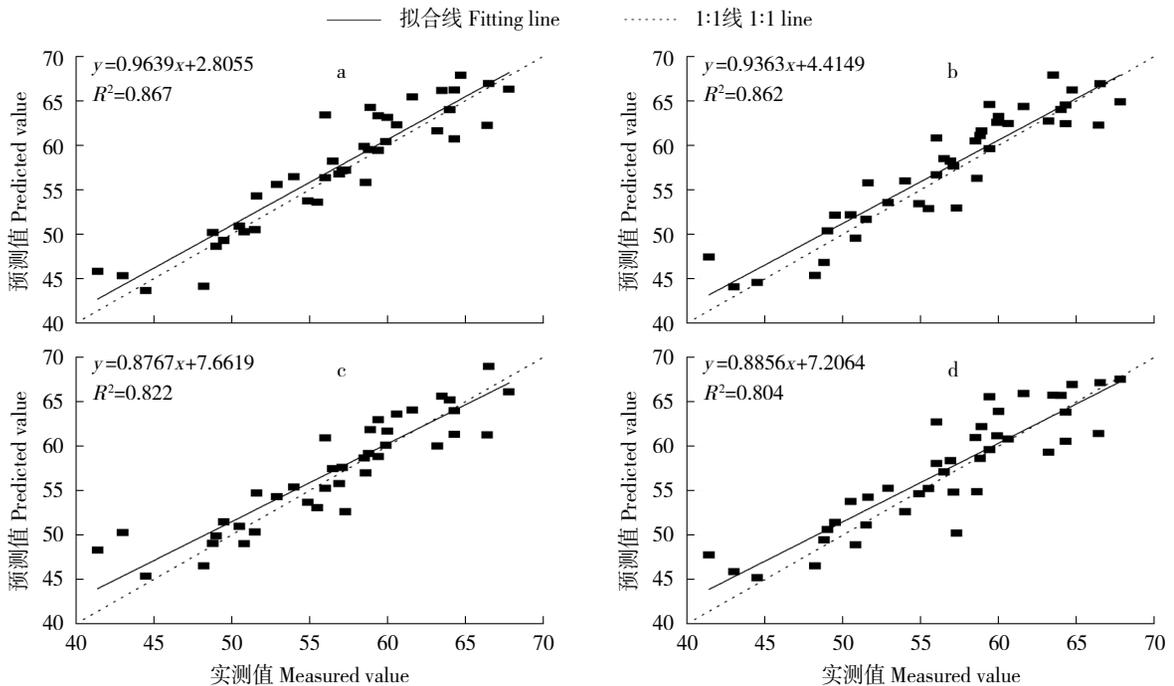
2.4 SPAD 值估算模型预测效果比较

表 6 列出了基于光谱特征参数构建的 PLSR 模型、基于全谱建立的 PLSR 模型、BP 神经网络模型、最优光谱特征参数估算模型的建模及验证情

况。结果表明，基于全谱的 PLSR 模型建模效果最好，其建模决定系数 R^2 最高，RMSE 最低，其次为 BP 神经网络模型。就验证情况来看，基于光谱特征参数构建的 PLSR 模型效果最好，其验证集决定系数 R^2 和 RPD 最高，RMSE 最低，该模型的构建计算量小，建模时间短。BP 神经网络模型验证情况较两种 PLSR 模型略差，但优于红边偏度 $s_{\lambda r}$ 构建的估算模型。综合建模与验证效果可知，基于光谱特征参数构建的 PLSR 模型预测效果最好。各模型预测结果如图 1 所示。

表 6 不同估算模型预测效果比较
Table 6 Comparison of prediction effects of different estimation models

模型 Model	建模集 Modeling set		验证集 Validation set		
	R^2	RMSE	R^2	RMSE	RPD
PLSR (光谱特征参数) PLSR (spectral characteristic parameter)	0.900	2.200	0.867	2.581	2.628
PLSR (全谱) PLSR (full spectrum)	0.910	2.071	0.862	2.600	2.545
BP 神经网络 BP neural network	0.904	2.165	0.822	2.861	2.305
多项式 Polynomial ($s_{\lambda r}$)	0.847	2.697	0.804	3.028	2.137



a. PLSR (光谱特征参数), b. PLSR (全谱), c. BP 神经网络, d. 多项式 ($s_{\lambda r}$)
a. PLSR (spectral characteristic parameter), b. PLSR (full spectrum), c. BP neural network, d. Polynomial ($s_{\lambda r}$)

图 1 基于不同估算模型的 SPAD 实测值与预测值 1:1 关系图

Fig.1 The relationships between measured value and predicted value of SPAD based on different estimation models

3 讨论

本研究利用高光谱数据，研究了基于单一光谱特征参数和多元统计分析方法估算玉米叶片 SPAD 值的效果，所选研究区域范围大，水肥条件

不均，玉米品种多样，长势有差异，数据获取的背景条件不尽相同，避免了单一条件因素的影响，因此增强了估算模型的普适性。

本研究结果表明，不同反演方法的 SPAD 值估算效果不同。基于植被指数的估算模型以 RVI

为最优,可能由于特征波段相除消除了光谱曲线漂移的影响,减小了误差^[5]。红边参数中以红边偏度 $s_{\lambda r}$ 最优,本研究中各样本的红边偏度大多为正,说明红边峰值形状呈右偏态分布,直观表现为右边的尾部相对于左边长。根据姚付启等^[14]的研究,随着叶片叶绿素的增减,红边峰值形状会相应地上下左右移动,说明红边峰值形状可以用来估算叶绿素含量,且红边偏度较传统的红边位置预测效果更好。以植被指数建立的估算模型整体上不如基于红边参数建立的模型估算效果好,可能由于从一阶微分光谱中提取的红边参数消除了噪声影响,更能体现反射率极值的波长位置,较原始光谱特征波段组合与植物生理参数的相关性更好^[16-17]。虽然传统回归模型预测精度不如PLSR模型与BP神经网络,但其模型简单,求解方便,适用于一些对预测精度要求不高的场合。

多元统计分析方法PLSR模型和BP神经网络利用多个与SPAD值相关性高的变量,保证了输入变量对SPAD值的贡献,从而提高了模型的预测精度。基于光谱特征参数构建的PLSR模型相较于全谱PLSR模型,剔除了大量与叶绿素含量不相关的波段,在很大程度上降低了模型的复杂程度,提高了模型的预测精度和稳定度,达到了简化模型的目的。基于BP神经网络建立的模型预测效果与PLSR模型相比略差,与汤旭光等^[18]的研究结果一致,可能与构建神经网络模型时设置的训练参数有关,而且BP神经网络无法准确确定最佳隐含层神经元数,需经过多次训练试错来确定最好的训练结果,故其实用性还有待进一步探究。

本研究针对灌浆期玉米叶片SPAD值建立了估算模型,其能否应用于玉米整个生育期叶片SPAD值的监测还有待进一步验证。此后还将会继续探讨冠层尺度的叶绿素含量反演,以便为卫星遥感应用提供技术支持。

4 结论

本研究采集了灌浆期玉米叶片高光谱数据,构建了基于光谱特征参数的玉米叶片SPAD值传统回归模型,以及PLSR和BP神经网络两种多元统计模型,并对各模型的预测效果进行了比较分析。结果表明,传统回归模型中以基于 $s_{\lambda r}$ 建立的多项式回归方程为最优,利用多元统计分析方法建立

的模型均优于传统回归模型。其中以基于光谱特征参数建立的PLSR模型验证效果最好($R^2=0.867$, $RMSE=2.581$, $RPD=2.628$),预测能力达到极好水平,且其计算量小、建模快,为最优的玉米叶片SPAD估算模型。

参考文献

- [1]Curran P J, Windham W R, Gholz H L. Exploring the relationship between reflectance red edge and chlorophyll concentration in slash pine leaves. *Tree Physiology*, 1995, 15(3):203-206.
- [2]李哲,张飞,陈丽华,等. 光谱指数的植物叶片叶绿素含量估算模型. *光谱学与光谱分析*, 2018, 38(5): 1533-1539.
- [3]王晓星. 夏玉米冠层光谱特征及其生理生态参量的高光谱估算模型. 杨凌:西北农林科技大学, 2015.
- [4]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. *农业工程学报*, 2016, 32(16): 135-142.
- [5]丁希斌,刘飞,张初,等. 基于高光谱成像技术的油菜叶片SPAD值检测. *光谱学与光谱分析*, 2015, 35(2): 486-491.
- [6]Thenkabail P S, Enclona E A, Ashton M S, et al. Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. *Remote Sensing of Environment*, 2004, 91(3): 354-376.
- [7]Thenkabail P S, Smith R B, Pauw E D. Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization. *Photogrammetric Engineering & Remote Sensing*, 2002, 68(6): 607-621.
- [8]Chan C W, Paelinckx D. Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sensing of Environment*, 2008, 112(6): 2999-3011.
- [9]Gitelson A A, Zur Y, Chivkunova O B, et al. Assessing carotenoid content in plant leaves with reflectance spectroscopy. *Photochemistry & Photobiology*, 2010, 75(3): 272-281.
- [10]刘海启,李召良. 高光谱植被遥感. 北京:中国农业科学技术出版社, 2015: 85-86.
- [11]Gitelson A, Kaufman Y J, Stark R, et al. Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*, 2002, 80(1): 76-87.
- [12]Penuelas J, Baret F, Filella I. Semiempirical indexes to assess carotenoids chlorophyll-a ratio from leaf spectral reflectance. *Photosynthetica*, 1995, 31(2): 221-230.
- [13]Horler D N H, Barber J, Darch J P, et al. Approaches to detection of geochemical stress in vegetation. *Advances in Space Research*, 1983, 3(2): 175-179.
- [14]姚付启,张振华,杨润亚,等. 基于红边参数的植被叶绿素含量高光谱估算模型. *农业工程学报*, 2009, 25(S2): 123-129.
- [15]马赵扬,杜昌文,周健民. 土壤碳酸钙中红外光声光谱特征及其应用. *光谱学与光谱分析*, 2012, 32(5): 1255-1258.
- [16]陈志强,王磊,白由路,等. 整个生育期玉米叶片SPAD高光谱预测模型研究. *光谱学与光谱分析*, 2013, 33(10): 2838-2842.
- [17]宋开山,张柏,李方,等. 高光谱反射率与大豆叶面积及地上鲜生物量的相关分析. *农业工程学报*, 2005, 21(1): 36-40.
- [18]汤旭光,宋开山,刘殿伟,等. 基于可见/近红外反射光谱的大豆叶绿素含量估算方法比较. *光谱学与光谱分析*, 2011, 31(2): 371-374.

Estimation Models of Maize Leaf SPAD Value Based on Hyperspectral Remote Sensing

Dong Zhe, Yang Wude, Zhang Meijun, Zhu Hongfen, Wang Chao

(Institute of Dry Farming Engineering, Shanxi Agricultural University, Taigu 030801, Shanxi, China)

Abstract The leaf chlorophyll content at the filling stage plays an important role in photosynthesis and yield formation of maize. To estimate the maize leaf chlorophyll content accurately and efficiently, hyperspectral remote sensing was used by taking SPAD value as the relative chlorophyll content. Conventional regression models based on spectral characteristic parameters, PLSR models based on full spectrum and spectral characteristic parameters, and BP neural network model were constructed and compared. The results showed that SPAD value of PLSR model based on full spectrum had the best fitting effect ($R^2=0.910$, RMSE=2.071). The fitting effect of PLSR model based on spectral characteristic parameters was close to that of PLSR model based on full spectrum. But the measured and predicted values of the latter fitting effect ($R^2=0.867$, RMSE=2.581, RPD=2.628) was better than that of full spectrum PLSR model, and the PLSR model based on spectral characteristic parameters had short modeling time and low complexity. The prediction effect of the BP neural network model was worse than the two PLSR models, but better than the conventional regression models based on spectral characteristic parameters. In brief, the estimation effect of PLSR model based on spectral parameters was the best.

Key words Hyperspectral; Maize leaf; SPAD; Partial least square regression; BP neural network