

郭 松, 常庆瑞, 郑智康, 等. 基于无人机高光谱影像的玉米叶绿素含量估测[J]. 江苏农业学报, 2022, 38(4): 976-984.

doi: 10.3969/j.issn.1000-4440.2022.04.014

## 基于无人机高光谱影像的玉米叶绿素含量估测

郭 松, 常庆瑞, 郑智康, 蒋丹焱, 高一帆, 宋子怡, 姜时雨

(西北农林科技大学资源环境学院, 陕西 杨凌 712100)

**摘要:** 为实现玉米叶绿素含量的快速估测, 使用低空无人机搭载 S185 高光谱相机获取关中地区抽雄期玉米冠层高光谱影像, 并在地面同步测定采样点叶绿素含量(Chl 值), 以原始光谱和一阶导数光谱为基础构建模型, 分别通过任意 2 波段组合以及连续投影算法提取单因素建模参数和多因素建模参数, 对比分析各类模型(单因素回归模型、多元线性回归模型和狮群算法优化的核极限学习机模型)对抽雄期玉米 Chl 值的预测能力。结果表明, 原始光谱上 Chl 值的敏感波段主要集中于绿光波段和近红外波段, 一阶导数光谱上 Chl 值的敏感波段主要集中于近红外波段。原始光谱和一阶导数光谱中的最优单因素建模参数均为差值光谱指数(DSI), 相关系数分别为 0.71、0.68, 连续投影算法筛选的多因素建模参数分别为 14 个、8 个。原始光谱和一阶导数光谱下均为多因素模型估测效果优于单因素模型, 机器学习算法优于传统回归算法, 其中基于原始光谱的狮群算法优化的核极限学习机(ISO-KELM)回归模型是此次研究中的最优模型, 具有较好的填图精度, 其建模决定系数( $R^2$ )和验证  $R^2$  分别为 0.73、0.70, 平均相对误差(MRE)分别为 3.56%、3.53%。说明结合无人机高光谱影像与 ISO-KELM 可较好地估测田间抽雄期玉米冠层的叶绿素含量。

**关键词:** 玉米; 叶绿素含量; 无人机; 连续投影算法; 核极限学习机回归模型

**中图分类号:** S127; S513 **文献标识码:** A **文章编号:** 1000-4440(2022)04-0976-09

## Estimation of maize chlorophyll content based on unmanned aerial vehicle (UAV) hyperspectral images

GUO Song, CHANG Qing-rui, ZHENG Zhi-kang, JIANG Dan-yao, GAO Yi-fan, SONG Zi-yi, JIANG Shi-yu

(College of Natural Resources and Environment, Northwest A&F University, Yangling 712100, China)

**Abstract:** In order to achieve rapid estimation of maize chlorophyll content, a low-altitude unmanned aerial vehicle (UAV) equipped with S185 hyperspectral camera was used to obtain the hyperspectral images of maize canopy during the tasseling period in Guanzhong area, and the chlorophyll content (Chl value) of sampling point was measured synchronously on the ground. The models were built based on the primary spectra and first derivative spectra. Single-factor modeling parameters and multi-factor modeling parameters were extracted by any two bands combination and successive projections algorithm, respectively, and the prediction ability of various models (single-factor regression model, multiple linear regression model and kernel extreme learning machine model combined with lion swarm algorithm) on Chl value of maize at tasseling stage was compared and analyzed. The results showed that the sensitive bands of Chl value in the primary spectra were mainly concentrated in the green band and the near infrared band, and the sensitive bands of Chl value in the first derivative spectra were

mainly concentrated in the near infrared band. The optimal single-factor modeling parameters in the primary spectra and the first derivative spectra were difference spectral index (DSI), and the correlation coefficients were 0.71 and 0.68, respectively. The number of multi-factor modeling parameters selected by the successive projections algorithm was

收稿日期: 2021-10-25

基金项目: 国家高技术研究发展计划(“863”计划)项目(2013AA102401)

作者简介: 郭 松(1996-), 男, 贵州贵阳人, 硕士研究生, 主要从事土地资源和空间信息技术研究。(E-mail) 1185716519@qq.com

通讯作者: 常庆瑞, (E-mail) changqr@nwsuaf.edu.cn

fourteen and eight, respectively. In both the primary spectrum and the first derivative spectrum, the estimation effect of the multi-factor model was better than that of the single-factor model, and the machine learning regression algorithm was better than the traditional regression algorithm. In this study, the kernel extreme learning machine combined with the lion swarm algorithm (LSO-KELM) regression model was the optimal model, which had good mapping accuracy. The modeling and validation determination coefficient ( $R^2$ ) were 0.73 and 0.70, respectively, and the mean relative errors (MRE) were 3.56% and 3.53%, respectively. The research results demonstrate that the combination of UAV hyperspectral images and LSO-KELM can better estimate the chlorophyll content of maize canopy during the tasseling period in the field.

**Key words:** maize; chlorophyll content; unmanned aerial vehicle (UAV); successive projections algorithm; kernel extreme learning machine regression model

叶绿素作为植被叶片的主要色素,是植被光合作用的核心,其含量能反映植被的长势,指导田间氮肥施用<sup>[1-2]</sup>。原子吸收光谱法、研磨法<sup>[3]</sup>、分光光度法<sup>[4]</sup>等传统的植被叶绿素含量检测方法不仅对样品本身有着不可逆的损害,还耗时耗力;SPAD仪<sup>[5]</sup>、植物多酚叶绿素测量计<sup>[6]</sup>等新型叶绿素含量检测仪器虽然能有效避免损害样品,增加检测效率,但是需要接触测量,不适合大规模样本和面尺度下的叶绿素含量检测<sup>[7]</sup>。近年来,高光谱遥感技术因其具有数据量大、精度高的特点,在作物生理生化参数反演研究方面逐渐被学者们所重视。然而,地面高光谱遥感存在灵活性较低,获取数据时人力、物力损耗较大的缺点;航天高光谱遥感对机场等基础设施和专业人员的技术水平要求比较高;航空高光谱遥感面临重访周期长,受大气影响较大,数据质量偏低的挑战<sup>[8-9]</sup>。上述缺陷限制了高光谱遥感在作物生理生化参数反演方面的应用,而低空无人机高光谱遥感的出现,使得上述问题得到一定解决,低空无人机高光谱影像的数据质量高,受大气影响小,较为便捷且学习成本低<sup>[10]</sup>。

目前,在基于低空无人机遥感平台反演植被生理生化参数方面,国内外学者做了大量研究。陈鹏飞等<sup>[11]</sup>发现对高空间分辨率无人机影像剔除土壤背景和计算纹理信息,均能提升棉花氮素估测模型的精度,建模决定系数从0.33上升至0.57。陈晓凯等<sup>[12]</sup>基于随机森林算法反演冬小麦叶面积指数(LAI),并对无人机搭载UHD185传感器获取的冬小麦影像进行填图,填图验证均方根误差为0.27。陶惠林等<sup>[13]</sup>通过无人机高光谱影像数据计算不同植被指数,运用多种回归方法建立冬小麦各生育期的长势监测指标(GMI)估测模型,结果表明开花期的多元回归模型精度最高,其归一化均方根误差为15.9%。Singhal等<sup>[14]</sup>结合核脊回归算法和无人机影像光谱信息,反演姜黄叶片的叶绿素

值,结果表明该机器学习算法具有良好的建模能力,最佳模型的决定系数达0.76。Colorado等<sup>[15]</sup>利用无人机影像提取光谱曲线,采取多元线性回归、支持向量回归以及神经网络回归估测水稻叶片氮素含量,发现机器学习算法(支持向量回归以及神经网络回归)的建模效果优于多元线性回归。综上所述,利用低空无人机影像数据反演植被生理生化参数的研究已经较为成熟,配备的传感器包括Rededge、UHD185等,使用的方法由传统回归算法向机器学习算法扩展,但是目前有关无人机搭载S185高光谱相机以及基于核极限学习机算法建模的报道较少。

本研究拟以关中地区抽雄期玉米为研究对象,使用无人机搭载S185高光谱相机获取研究区的遥感影像,结合地面采样点实测的叶绿素含量,构建传统回归模型和核极限学习机回归模型,从中选取鲁棒性、泛化性较好的玉米叶绿素含量估测模型并填图,以期为关中地区的绿色农业发展提供技术、决策支持。

## 1 材料与方法

### 1.1 研究区概况和试验设计

研究区位于陕西省杨陵区西北农林科技大学试验农场(108.092°E, 36.303°N),地处关中平原中部,属于温带季风气候,四季分明,年均气温12.85℃,平均降水量630.35 mm,平均海拔480 m,试验地土壤为壤土,水肥充足,适宜耕作。此次试验开展于2021年6月20日,试验共设置20个小区,小区面积30 m<sup>2</sup>(5 m×6 m),每个小区选取3个采样点并记录其GPS坐标(图1),施用氮、磷、钾3种肥料,其中氮肥和钾肥均设5个水平,磷肥仅设1个水平,共10个处理,每个处理重复2次,氮肥处理N<sub>1</sub>~N<sub>5</sub>,纯氮施用量分别为0 kg/hm<sup>2</sup>、50 kg/hm<sup>2</sup>、100 kg/hm<sup>2</sup>、150 kg/hm<sup>2</sup>、200 kg/hm<sup>2</sup>,钾肥处理K<sub>1</sub>~K<sub>5</sub>的K<sub>2</sub>O施用量分别为0 kg/hm<sup>2</sup>、144 kg/hm<sup>2</sup>、288

kg/hm<sup>2</sup>、432 kg/hm<sup>2</sup>、576 kg/hm<sup>2</sup>, 磷肥处理 P<sub>1</sub> 的 P<sub>2</sub>O<sub>5</sub> 施用量为 1 125 kg/hm<sup>2</sup>, 所有肥料播种前全部施入, 后期不追肥, 田间管理与当地一般大田一致。10 个处理分别为 N<sub>1</sub>P<sub>1</sub>K<sub>2</sub>、N<sub>2</sub>P<sub>1</sub>K<sub>2</sub>、N<sub>3</sub>P<sub>1</sub>K<sub>2</sub>、N<sub>4</sub>P<sub>1</sub>K<sub>2</sub>、N<sub>5</sub>P<sub>1</sub>K<sub>2</sub>、N<sub>2</sub>K<sub>1</sub>、N<sub>2</sub>K<sub>2</sub>、N<sub>2</sub>K<sub>3</sub>、N<sub>2</sub>K<sub>4</sub>、N<sub>2</sub>K<sub>5</sub>。

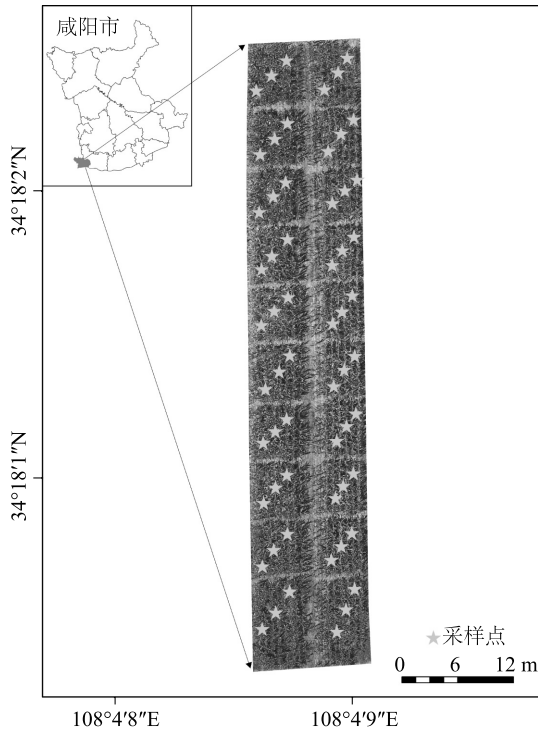


图 1 研究区玉米试验田位置

Fig.1 Location of maize test field in the study area

## 1.2 测定项目和方法

**1.2.1 无人机高光谱影像和 RGB 影像的获取** 无人机高光谱影像数据通过 M600 Pro 无人机(大疆创新科技有限公司产品)搭载 S185 高光谱相机(Cubert 公司产品)获取。S185 高光谱相机采用画幅式成像技术, 结合了高光谱数据的精确性和快照成像的高速性, 可在 0.001 s 内采集整个视场内的高光谱影像数据, 影像光谱范围为 450~998 nm, 包含 138 个波段, 光谱采样间隔 4 nm。影像采集于 2021 年 8 月 20 日, 此时玉米处于抽雄期, 是其生长发育的关键时期<sup>[16]</sup>, 无人机飞行时间为 14:00–16:00, 天气晴朗无风, 飞行前进行黑白板校正, 设定飞行高度为 50 m, 飞行速度为 5 m/s, 航向与旁向重叠度均为 80%, 镜头垂直向下, 影像地面分辨率为 1.4 cm。由于获取的无人机高光谱影像缺失地理参考, 故使用自带 GPS 仪的精灵 4 RTK 无人机(大疆创新科技有

限公司产品)同步获取研究区 RGB 影像, 设定的飞行参数与 M600 Pro 无人机一致, 用于高光谱影像的地理配准和几何校正。

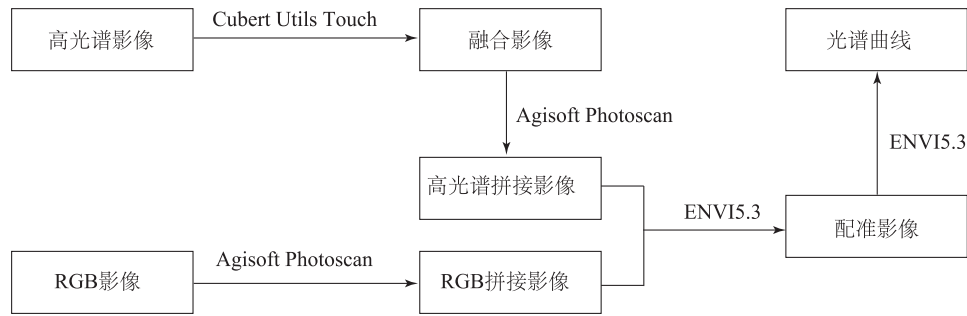
**1.2.2 叶绿素含量测定** 玉米叶片叶绿素含量使用植物多酚-叶绿素测量计 Dualex scientific+™(上海泽泉科技股份有限公司产品)测定, 测定项目中的 Chl 值代表叶绿素含量<sup>[6]</sup>, 单位为 μg/cm<sup>2</sup>, 测定时间为无人机影像采集后的 1 h 内, 每个采样点选择 3 片健康、无病虫害的冠层叶片, 于叶尖、中部、基部 3 个位置分别测量 3 个 Chl 值, 每个采样点获取 27 个 Chl 值, 取平均值作为该采样点的 Chl 值。

## 1.3 数据处理

在 S185 高光谱相机处理软件 Cubert Utils Touch 中进行高光谱影像的波段融合, 并输出为单张 Multiff 格式的高光谱影像, 使用 Agisoft Photoscan 软件分别将单张 Multiff 影像与精灵 4 RTK 无人机获取的 RGB 影像拼接为研究区的完整影像。在软件 ENVI 5.3 中以精灵 4 RTK 无人机获取的 RGB 影像作为基准影像, 以无人机高光谱影像作为待校准影像, 实施地理配准和几何校正, 使高光谱影像具有正确的地理坐标, 并消除几何变形, 然后对高光谱影像进行平滑滤波以减小影像数据采集过程中的随机噪声, 依据采样点的 GPS 坐标, 在高光谱影像上生成半径为 30 cm 的圆形感兴趣区(ROI), 通过 ROI 提取出各采样点的反射光谱曲线。在前人研究结果<sup>[17]</sup>的基础上, 本次研究提取的光谱曲线范围在 450~850 nm, 影像处理流程如图 2 显示。依据从小到大的原则对采集的各采样点 Chl 值进行排序, 通过分层抽样法以 3:1 的比例获取建模样本 45 个, 验证样本 15 个。

## 1.4 模型建立

通过各采样点的原始光谱以及一阶导数光谱获取建模参数。使用任意 2 波段组合的方式构建单因素建模参数, 包括归一化光谱指数(NDSI)、差值光谱指数(DSI)和比值光谱指数(RSI)<sup>[18]</sup>; 基于连续投影算法(SPA)构建多因素建模参数。单因素模型的构建方式有线性回归、指数回归、幂回归、对数回归以及多项式回归, 多因素模型的构建方式包括多元线性回归(MLR)以及核极限学习机回归(KELMR), 对比各类模型的精度, 从中筛选出反演效果最佳的模型。所有模型的建立以及高光谱遥感影像的填图均在软件 Matlab 的环境下实现。



Cubert Utils Touch、Agisoft Photoscan、ENVI5.3 为处理数据的软件名称。

图2 影像处理流程

Fig.2 Processing flows of images

模型精度使用决定系数( $R^2$ ) (公式 1) 和平均相对误差( $MRE$ ) (公式 2) 进行检验,其中  $n$  为样本数量, $\hat{y}_i, y_i$  分别表示第  $i$  个采样点 Chl 值的预测值和实测值, $\bar{y}_i$  代表建模集或验证集 Chl 值实测值的平均值, $R^2$  越接近 1,  $MRE$  越小,表示模型精度越好。

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

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

表1 玉米叶片叶绿素含量(Chl 值)统计特征

Table 1 Statistical characteristics of chlorophyll content (Chl value) in maize leaves

样本类型	样本数量	Chl 值 ( $\mu\text{g}/\text{cm}^2$ )			变异系数 (%)
		最大值	最小值	平均值	
总数据集	60	54.3	42.0	47.86	6.61
建模集	45	53.9	42.0	47.75	6.64
验证集	15	54.3	42.9	48.20	6.88

Chl 值表示叶绿素含量 ( $\mu\text{g}/\text{cm}^2$ )。

2.1.2 Chl 值与光谱反射率的相关性 图3显示,原始光谱反射率与 Chl 值相关性曲线变化较为平缓,且所有波段反射率与 Chl 值均呈现负相关,通过 0.01 相关性检验的波段主要位于 500~624 nm 以及 710~850 nm,在 766 nm 波长处相关性最佳,相关系数为-0.61。一阶导数光谱反射率与 Chl 值相关性曲线的走势波动性较大,通过 0.01 相关性检验的波段在 504~654 nm 零星分布,在 686~766 nm 集中分布,相关性最佳的波长为 714 nm,相关系数为 0.64。整体上看,在通过 0.01 相关性检验的波长数量上,原始光谱要大于一阶导数光谱,但是在相关性大小方面,一阶导数的最大相关系数绝对值要高于原始光谱。

## 2 结果与分析

### 2.1 玉米叶片 Chl 值光谱特征分析

2.1.1 玉米叶片 Chl 值基本统计特征 供试玉米叶片 Chl 值统计结果(表 1)表明,建模集和验证集的各项统计特征均与总数据集较为接近,3 类样本的 Chl 值平均值和变异系数分别为 47.75~48.20、6.61%~6.88%,说明使用分层抽样法划分的建模集和验证集代表性较好<sup>[19]</sup>,各集群所反映的采样点 Chl 值较一致。

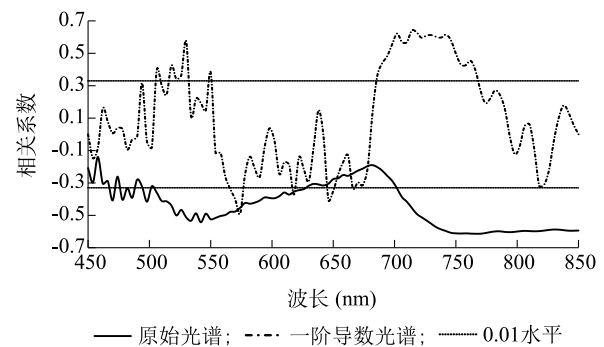


图3 不同类型光谱反射率与 Chl 值的相关性

Fig.3 Correlation between different types of spectral reflectance and Chl values

2.1.3 Chl 值与光谱指数的相关性 根据光谱指数计



算公式(表 2),分析统计 Chl 值与任意 2 波段组合光谱指数的相关性,绘制相关系数等势图(图 4),选择相关性最好的波长组合方式确定敏感光谱指数。从不同类型光谱指数来看,仅差值光谱指数(*DSI*)存在连片的相关性较高的区域,主要分布在 700~850 nm,

而归一化光谱指数(*NDSI*)、比值光谱指数(*RSI*)的相关系数偏低且分布较为零散。从相关性最佳来看,原始光谱和一阶导数光谱中均为 *DSI* 与 Chl 值的相关性大于 *NDSI* 和 *RSI*,相关系数分别为 0.71、0.68,表明 *DSI* 是此次研究中的最优光谱指数。

表 2 光谱指数计算方法

Table 2 Calculation method of spectral index

光谱类型	光谱指数	计算公式	入选波长	相关系数
原始光谱	比值光谱指数( <i>RSI</i> )	$R_i/R_j$	502、534	0.54 **
	差值光谱指数( <i>DSI</i> )	$R_i-R_j$	602、682	0.71 **
	归一化光谱指数( <i>NDSI</i> )	$(R_i-R_j)/(R_i+R_j)$	502、534	0.53 **
一阶导数光谱	比值光谱指数( <i>RSI</i> )	$R_i/R_j$	506、778	0.46 **
	差值光谱指数( <i>DSI</i> )	$R_i-R_j$	502、714	0.68 **
	归一化光谱指数( <i>NDSI</i> )	$(R_i-R_j)/(R_i+R_j)$	506、526	0.50 **

$R_i$ 表示波段  $i$  的反射率; $R_j$ 表示波段  $j$  的反射率。\*\*表示在 0.01 水平显著相关。

## 2.2 单因素模型构建及精度评价

以表 2 中的光谱指数作为自变量,玉米叶片 Chl 值为因变量,构建单因素估算模型,应用线性回归、2 次曲线、指数回归和对数回归 4 种方法,筛选出各光谱指数的最佳模型(表 3)。从不同类型光谱指数来看,*DSI* 的建模效果最好;从不同光谱类型来

看,原始光谱构建模型整体优于一阶导数光谱。由此可得,单因素模型中的最优模型是基于原始光谱的 *DSI* 模型,其建模、验证  $R^2$  均为 0.52, *MRE* 分别为 4.09%、4.36%,使用单因素估测玉米 Chl 值时,应优先考虑该模型。

表 3 不同光谱类型的单因素模型

Table 3 Single factor models based on different types of spectra

模型类型	模型表达式	建模精度		验证精度	
		$R^2$	<i>MRE</i> (%)	$R^2$	<i>MRE</i> (%)
PS- <i>RSI</i>	$Chl = 139.19 RSI^2 - 124.52 RSI + 73.08$	0.33 **	4.64	0.30 **	4.44
PS- <i>DSI</i>	$Chl = 56.90 e^{-0.15 DSI}$	0.52 **	4.09	0.52 **	4.36
PS- <i>NDSI</i>	$Chl = 59.99 e^{0.84 NDSI}$	0.30 **	4.71	0.29 **	4.51
FDS- <i>RSI</i>	$Chl = 48.70 e^{-0.11 RSI}$	0.22 **	6.50	0.25 *	5.31
FDS- <i>DSI</i>	$Chl = -8.27 \ln DSI + 48.72$	0.43 **	4.31	0.45 **	4.10
FDS- <i>NDSI</i>	$Chl = -2.96 NDSI + 45.45$	0.25 **	4.72	0.35 **	5.12

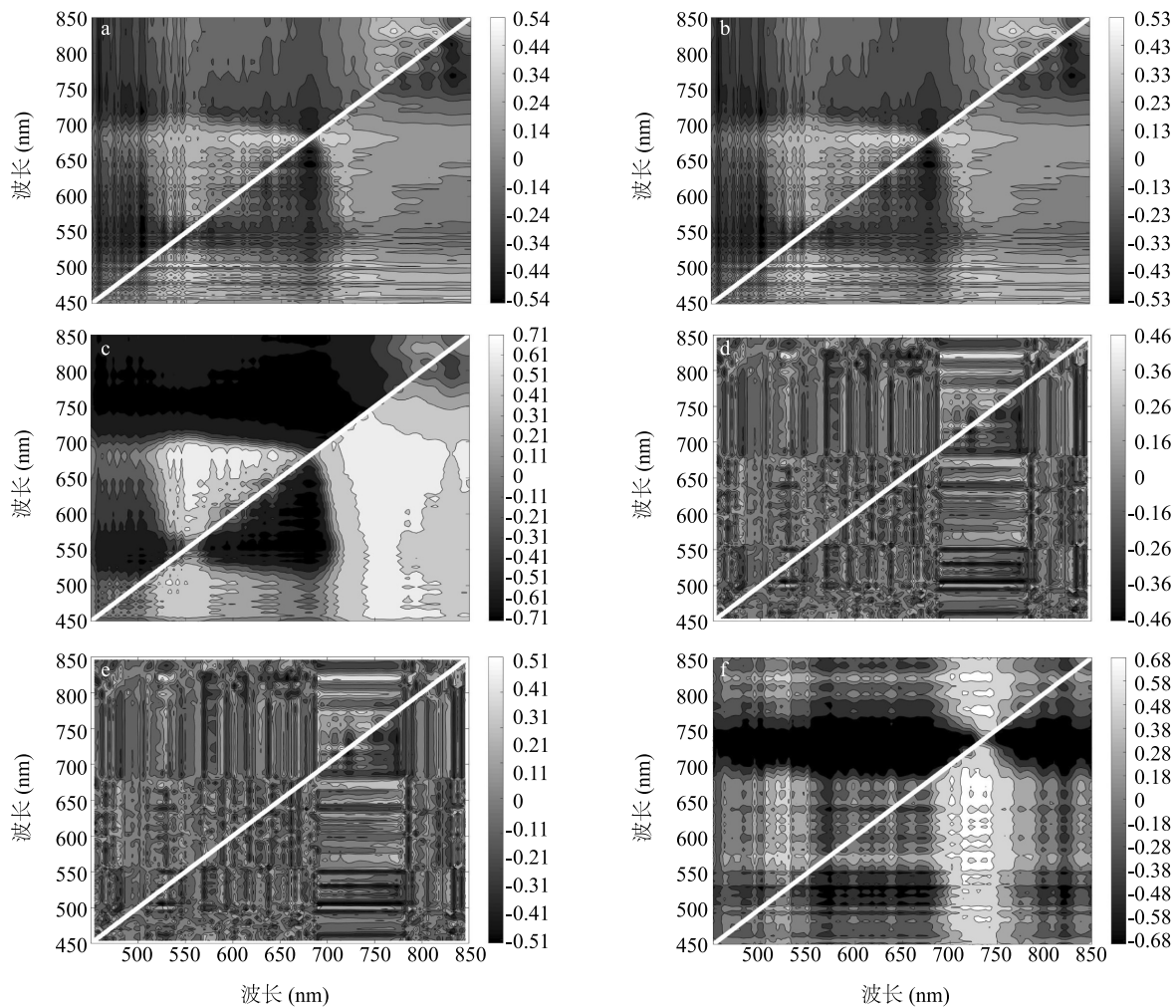
PS:原始光谱;FDS:一阶导数光谱;*RSI*:比值光谱指数;*NDSI*:归一化光谱指数;*DSI*:差值光谱指数;*Chl*:用来表示玉米叶片叶绿素含量的 Chl 值; $R^2$ :决定系数;*MRE*:平均相对误差。\*表示在 0.05 水平显著相关;\*\*表示在 0.01 水平显著相关。

## 2.3 多因素模型构建及精度评价

### 2.3.1 连续投影算法提取多因素建模参数

连续投影算法是一种特征选择方法,通过连续投影的方式从原始光谱矩阵中选出预测响应变量的子集<sup>[20]</sup>,该算法不仅确保所筛选的光谱矢量共线性最小,还实现了对响应变量预测能力的最大化,其原理为第一次迭代时选光谱矩阵一列赋予迭代向量,此后分别计算该迭代向量对剩余向量的投影并提取最大投影的波段,提取完成后开始第二次迭代循环,直至

所筛选的波段数量等于给定数量,随着提取波段数量的增加,每循环一次即进行一次筛选波段对响应变量的预测能力计算。以预测能力最佳为原则,分别将原始光谱和一阶导数光谱中通过 0.01 相关性检验的波段反射率作为自变量,Chl 值作为响应变量,提取多因素建模参数(表 4),提取出的原始光谱、一阶导数光谱多因素建模参数分别为 14 个和 8 个,二者的多因素建模参数波长集中在 506~766 nm,即绿光波段至近红外波段。



a; PS-*RSI*; b; PS-*NDSI*; c; PS-*DSI*; d; FDS-*RSI*; e; FDS-*NDSI*; f; FDS-*DSI*。PS:原始光谱;FDS:一阶导数光谱;*RSI*:比值光谱指数;*NDSI*:归一化光谱指数;*DSI*:差值光谱指数。

图4 不同类型光谱指数与Chl值的相关性等势图

Fig.4 Equipotential diagram of correlation between different types of spectral indices and Chl value

表4 不同类型光谱多因素建模参数

Table 4 Multivariate modeling parameters based on different types of spectra

光谱类型	建模参数	参数个数
PS	$R_{506}$ 、 $R_{510}$ 、 $R_{514}$ 、 $R_{518}$ 、 $R_{522}$ 、 $R_{526}$ 、 $R_{534}$ 、 $R_{542}$ 、 $R_{578}$ 、 $R_{594}$ 、 $R_{602}$ 、 $R_{626}$ 、 $R_{702}$ 、 $R_{734}$	14
FDS	$R_{506}$ 、 $R_{518}$ 、 $R_{522}$ 、 $R_{550}$ 、 $R_{574}$ 、 $R_{618}$ 、 $R_{666}$ 、 $R_{766}$	8

PS:原始光谱;FDS:一阶导数光谱; $R_{506}$ :在波段506 nm处的反射率,其他同理。

**2.3.2 多因素模型建立及精度比较** 由于SPA所提取的建模参数在保证预测性最佳的同时还使得变量之间的共线性最低,故多因素模型的建立不再考虑建模参数的共线性问题,使用多元线性回归(MLR)法、核极限学习机(KELM)算法构建多因素模型。

核极限学习机是单隐含层前馈神经网络的一种,由Huang等<sup>[21]</sup>提出,其输入层到隐含层的权值是随机产生的,而隐含层到输出层的权值是计算产生的,随机产生的权值通过不断迭代,能得到较好的预测结果,该算法在普通极限学习机的基础上引入了“核”的概念,与普通神经网络相比,提升了学习速率和泛化性,在大样本下也能保证较快的计算速率,与随机森林回归相比,在样本数量偏少时也能快速学习到数据的规律,并且预测的结果能超越给定的训练集,但是在建模样本向高维空间映射的过程中,由于其函数参数的生成具有一定的随机性,往往需要迭代多次才能取得较好的结果。为了减小计算成本,本研究采用狮群优化算法(LSO)<sup>[22]</sup>对KELM

迭代过程进行优化,使计算过程向着  $R^2$  增大、 $MRE$  减小的方向收敛。采用径向基函数 (RBF) 作为 KELM 的核函数,狮群数量 ( $POP$ ) 以及迭代次数 ( $N$ ) 分别设置为 10、50,优化所得的正则化系数 ( $C$ )、核函数参数 ( $S$ ) 以及多因素建模结果 (表 5) 显示,与单因素模型相比,多因素模型精度有较大提升,各类多因素模型的建模  $R^2$ 、验证  $R^2$  均在 0.57 以

上,高于单因素的  $R^2$  最大值 0.52。就建模方法而言,MLR 算法的建模效果要低于 KELM 算法;从光谱类型来看,原始光谱的模型精度优于一阶导数光谱;综合比较可得,原始光谱的 LSO-KELM 模型是所有方法中的最优模型,其建模、验证  $R^2$  分别为 0.73、0.70,相应  $MRE$  处于较低水平,为 3.56% 和 3.53%。

表 5 不同光谱类型下多因素回归模型

Table 5 Multiple factors regression models based on different types of spectra

模型类型	模型表达式	建模精度		验证精度	
		$R^2$	$MRE$ (%)	$R^2$	$MRE$ (%)
PS-MLR	$Chl = 58.97 + 4.59R_{506} - 4.69R_{510} + 4.38R_{514} - 3.04R_{518} - 3.59R_{522} + 4.87R_{526} - 3.49R_{534} - 3.42R_{542} + 8.58R_{578} + 2.89R_{594} - 5.16R_{602} - 5.48R_{626} + 4.49R_{702} - 0.95R_{734}$	0.71 **	3.62	0.69 **	3.74
PS-LSO-KELM	$C = 0.91; S = 0.60$	0.73 **	3.56	0.70 **	3.53
FDS-MLR	$Chl = 58.09 - 4.05R_{506} - 6.44R_{518} + 2.22R_{522} - 4.89R_{550} + 14.21R_{574} + 4.58R_{618} + 10.18R_{666} - 4.64R_{766}$	0.60 **	3.95	0.58 **	4.08
FDS-LSO-KELM	$C = 2.18; S = 5.71$	0.70 **	3.71	0.68 **	3.90

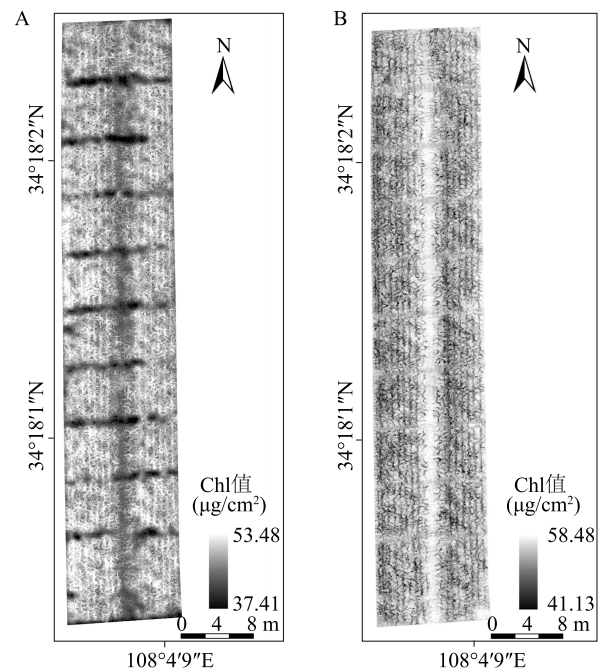
PS:原始光谱;FDS:一阶导数光谱;MLR:多元线性回归算法;LSO-KELM:狮群算法优化的核极限学习机算法; $Chl$ :表示玉米叶片叶绿素含量的  $Chl$  值; $R_{506}$  表示在波段 506 nm 处的反射率,其他同理; $R^2$ :决定系数; $MRE$ :平均相对误差; $C$ :正则化系数; $S$ :核函数参数。\*\*表示在 0.01 水平显著相关。

## 2.4 模型精度比较

基于单因素、多因素模型中的最优模型,分别进行高光谱遥感影像  $Chl$  值反演估算,同时将反演得到的预测值与实地采样得到的实测值进行拟合,实现模型精度比较,1:1 线设置为虚线。图 5 显示,PS- $DSI$  模型和 PS-LSO-KELM 模型下的  $Chl$  值反演图基本符合小区排列,但是外在表现有所差异。与 PS- $DSI$  模型相比,PS-LSO-KELM 模型的反演图不仅体现了玉米冠层  $Chl$  值的空间分布,还体现了玉米冠层的形态特征, $Chl$  值的位置信息更加准确,蕴含的信息更多,可见在  $Chl$  值的遥感反演估算方面,机器学习算法优于传统回归算法。图 6 显示,PS-LSO-KELM 模型的预测值比 PS- $DSI$  模型更加贴拟合线,同时拟合线与 1:1 线差距不大,故模型精度更高,其拟合方程  $R^2$  为 0.70,因此,PS-LSO-KELM 模型是此次研究中的最优模型。

## 3 讨论

玉米冠层  $Chl$  值与 S185 高光谱相机的原始光谱主要在绿光波段和近红外波段处通过 0.01 相关性检验,与一阶导数光谱则主要在近红外波段处通过 0.01 相关性检验,这与王婷婷等<sup>[6]</sup>等的研究结果基本一致。构造的光谱指数中,2 种光谱下均为  $DSI$



A: 基于原始光谱的差值光谱指数模型; B: 基于原始光谱的狮群算法优化的核极限学习机算法模型。 $Chl$  值: 叶绿素含量。

图 5 不同类型最优模型下玉米冠层  $Chl$  值空间分布图

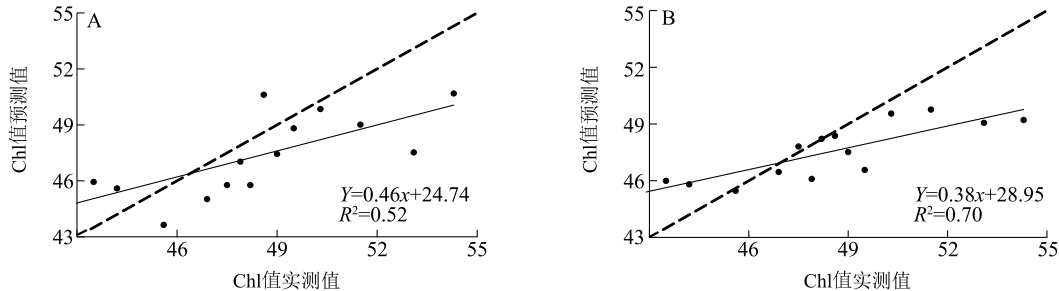
Fig.5 Spatial distribution of  $Chl$  values in maize canopy under different types of optimal models

的相关性最好,这与前人的研究结果<sup>[12,23-24]</sup>不一致,



一方面是因为光谱指数本质上是一种消除背景噪声、增强反演特征的算法,故该算法对仪器、作物种类以及光谱预处理方式等具有一定依赖性;另一方面是尺度效应,前人的光谱曲线多是地面遥感获取,属于叶片尺度,本研究的光谱曲线来自于低空无人

机影像,属于冠层尺度。这两方面原因都会造成光谱指数对 Chl 值的敏感性差异。未来的研究应该增加关中地区冠层玉米叶片的采样数量,同时还要分多个生育期进行试验,基于 S185 高光谱相机确立不同生育期的最优光谱指数。



A: 基于原始光谱的差值光谱指数模型; B: 基于原始光谱的狮群算法优化的核极限学习机算法模型。x: Chl 值的实测值; Y: Chl 值的预测值;  $R^2$ : 决定系数。

图6 不同类型最优模型预测值与实测值分布

Fig.6 Distribution of predicted and measured values of different types of optimal models

SPA 算法属于特征选择的一种,与基于相关性的特征选择算法(FCBF)相比,SPA 所筛选的自变量数量可控,所建立模型相对简单,与传统的特征提取算法如主成分分析(PCA)等相比,SPA 提取的自变量可解释性高,所建立的模型具有一定的物理化学意义。

在处理非线性回归问题时,机器学习算法往往比传统回归算法效果更佳<sup>[25]</sup>。原因在于机器学习算法是通过决策树或模拟人类神经元工作的方式来学习数据本身的规律,相较于单一的传统回归算法,机器学习算法体现了群体的特点。核函数的选取以及参数的不同会影响到 KELM 对 Chl 值的拟合结果,KELM 的核函数包括径向基核、线性核、小波核等,由于不同核函数的侧重点不同,也有学者提出双核函数的想法,充实机器学习算法的实用性<sup>[26]</sup>。麻雀搜索算法、灰狼算法等群体智能算法能快速确定 KELM 的最佳参数,减小 KELM 的运行成本,本研究采用狮群算法优化 KELM,以较少的迭代次数取得较好的玉米冠层 Chl 值拟合结果。

在本研究构建的所有模型中,多因素模型拟合效果优于单因素模型,机器学习算法优于传统回归算法,这与前人的研究结果<sup>[27-28]</sup>一致。实际应用时,通过无人机搭载 S185 高光谱相机获取目标区域抽雄期玉米高光谱遥感影像,从中提取模型参数,应用于已经训练好的 KELM 模型中即可得到玉米冠

层叶片 Chl 值的空间分布图。采样点的 Chl 值高,说明玉米长势较好,无需追肥;采样点的 Chl 值低,说明玉米长势较差,需要适量追肥。但是此次研究的区域局限在关中地区,研究时期限制在抽雄期,构建的模型是否适用于其他地区和其他生育期的玉米还有待测试。故未来应该在不同地区、不同生育期背景下开展低空无人机搭载 S185 高光谱相机反演玉米冠层 Chl 值的研究,在更多样本的支持下,确立相应的敏感波段,以期未来可以通过无人机高光谱技术监测大区域的玉米 Chl 值。

## 4 结论

本研究以陕西省关中地区抽雄期玉米为研究对象,基于玉米叶片 Chl 值和 S185 高光谱影像,构建反演玉米叶片 Chl 值的单因素模型、多因素模型以及狮群算法优化的核极限学习机模型并填图,得到如下结论:第一,玉米抽雄期冠层原始光谱反演 Chl 值的敏感波段主要集中于绿光波段至近红外波段,一阶导数光谱反演 Chl 值的敏感波段主要集中于近红外波段,二者的单因素最优建模参数均为  $DSI$ ,相关系数分别为 0.71、0.68。第二,连续投影算法降维效果较好,从原始光谱和一阶导数光谱中分别提取了 14 个、8 个建模参数,该算法未对光谱作修改,所提取的建模参数可解释性较好。第三,从填图效果以及预测值验证结果来看,PS-LSO-KELM 模型是



此次研究中的最优模型,其建模  $R^2$  和验证  $R^2$  分别达到了 0.73、0.70,  $MRE$  分别为 3.56%、3.53%。

## 参考文献:

- [1] 袁小康,周广胜,王秋玲,等.不同灌溉量夏玉米叶绿素含量的高光谱特征及其反演[J].生态学报,2021,41(2):543-552.
- [2] 陈澜,常庆瑞,高一帆,等.猕猴桃叶片叶绿素含量高光谱估算模型研究[J].西北农林科技大学学报(自然科学版),2020,48(6):79-89,98.
- [3] 努尔凯麦尔·木拉提,杨亚杰,帕尔哈提·阿布都克日木,等.小麦叶绿素含量测定方法比较[J].江苏农业科学,2021,49(9):156-159.
- [4] 周静,胡忠军,李培培,等.基于叶绿素a测定的分光光度法与bbe法比较:以千岛湖为例[J].水生生态学杂志,2013,34(2):80-84.
- [5] 崔小涛,常庆瑞,屈春燕,等.基于高光谱和MLSR-GA-BP神经网络模型油菜叶片SPAD值遥感估算[J].东北农业大学学报,2020,51(8):74-84.
- [6] 王婷婷,常庆瑞,刘梦云,等.基于DualEx植物多酚-叶绿素仪的冬小麦叶绿素含量高光谱估算[J].麦类作物学报,2019,39(5):595-604.
- [7] 王伟东,常庆瑞,王玉娜.基于UHD185成像光谱仪的冬小麦叶绿素监测[J].干旱区研究,2020,37(5):1362-1369.
- [8] 廖小罕,周成虎,苏奋振,等.无人机遥感众创时代[J].地球信息科学学报,2016,18(11):1439-1447.
- [9] 黄耀欢,李中华,朱海涛.作物胁迫无人机遥感监测研究评述[J].地球信息科学学报,2019,21(4):512-523.
- [10] 田明璐,班松涛,常庆瑞,等.基于低空无人机成像光谱仪影像估算棉花叶面积指数[J].农业工程学报,2016,32(21):102-108.
- [11] 陈鹏飞,梁飞.基于低空无人机影像光谱和纹理特征的棉花氮素营养诊断研究[J].中国农业科学,2019,52(13):2220-2229.
- [12] 陈晓凯,李粉玲,王玉娜,等.无人机高光谱遥感估算冬小麦叶面积指数[J].农业工程学报,2020,36(22):40-49.
- [13] 陶惠林,徐良骥,冯海宽,等.基于无人机高光谱长势指标的冬小麦长势监测[J].农业机械学报,2020,51(2):180-191.
- [14] SINGHAL G, BANSOD B, MATHEW L, et al. Estimation of leaf chlorophyll concentration in turmeric (*Curcuma longa*) using high-resolution unmanned aerial vehicle imagery based on kernel ridge regression[J]. Journal of the Indian Society of Remote Sensing, 2019, 47(7): 1111-1112.
- [15] COLORADO J D, CERA-BORNACELLI N, CALDAS J S, et al. Estimation of nitrogen in rice crops from UAV-captured images[J]. Remote Sensing, 2020, 12(20): 3396.
- [16] 陈妮娜,纪瑞鹏,贾庆宇,等.关键发育期干旱对春玉米产量及籽粒品质的影响[J].生态学杂志,2021,40(6):1687-1694.
- [17] 陈鹏飞,李刚,石雅娇,等.一款无人机高光谱传感器的验证及其在玉米叶面积指数反演中的应用[J].中国农业科学,2018,51(8):1464-1474.
- [18] GAUTAM K A, SINGH V, BECTOR V, et al. Trend analysis of vegetation indices using spectroradiometer at different growth stages of cotton[J]. Agricultural Mechanization in Asia Africa & Latin America, 2018, 49(1):63-66.
- [19] 李传华,赵军,师银芳,等.基于变异系数的植被NPP人为影响定量研究——以石羊河流域为例[J].生态学报,2016,36(13):4034-4044.
- [20] SOARES S, GOMES A, FILHO A, et al. The successive projections algorithm[J]. TrAC Trends in Analytical Chemistry, 2013, 42:84-97.
- [21] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: Theory and applications[J]. Neurocomputing, 2006, 70(1-3): 489-501.
- [22] 刘生建,杨艳,周永权.一种群体智能算法——狮群算法[J].模式识别与人工智能,2018,31(5):431-441.
- [23] 刘秀英,刘晨洲,吴娜微,等.玉米叶片花青素相对含量高光谱遥感反演[J].遥感信息,2018,33(6):1-8.
- [24] 杨婧,廖桂平,刘凡,等.基于高光谱成像技术的油菜叶片叶绿素含量预测[J].中国农业科技导报,2020,22(5):86-96.
- [25] YANN L, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [26] 聂立新,张天侠,赵波.粒子群算法优化双核支持向量机及应用[J].振动、测试与诊断,2014,34(3):565-569.
- [27] WANG L, CHANG Q R, YANG J, et al. Estimation of paddy rice leaf area index using machine learning methods based on hyperspectral data from multi-year experiments[J]. PLoS One, 2018, 13(12): e0207624.
- [28] LEE H, WANG J, LEBLON B. Using linear regression, random forests, and support vector machine with unmanned aerial vehicle multispectral images to predict canopy nitrogen weight in corn[J]. Remote Sensing, 2020, 12(13):2071.

(责任编辑:王妮)