

基于多时相无人机遥感生育时期优选的冬小麦估产

王晶晶¹ 李长硕¹ 卓越¹ 檀海斌² 侯永胜³ 严海军^{1,4}

(1. 中国农业大学水利与土木工程学院, 北京 100083; 2. 国家半干旱农业工程技术研究中心, 石家庄 050051;
3. 中国农业机械化科学研究院集团有限公司土壤植物机器系统技术国家重点实验室, 北京 100083;
4. 农业节水与水资源教育部工程研究中心, 北京 100083)

摘要: 为确定无人机遥感产量估算的最优生育时期及采集次数,以砂土种植冬小麦为研究对象,设置了4组灌水(36个样区)与5组施氮(15个样区)处理,采集了起身期至灌浆后期的8次遥感数据。采用偏最小二乘法(PLS)、随机森林(RF)和套索(LASSO)算法构建了单生育时期产量估算模型。根据提出的最优模型,利用三次B样条曲线和复合梯形公式,建立了5种特定生育阶段日植被指数积分的产量估算方案。结果表明,不同生育时期的冬小麦产量估算模型精度差异显著,随冬小麦生长精度总体呈递增趋势。单生育时期中,PLS、RF和LASSO模型的最优生育时期分别为灌浆前期、灌浆前期和灌浆后期。除拔节前期外,RF模型的产量估算精度均优于PLS和LASSO。冬小麦多生育时期的产量估算精度优于单生育时期,从起身期至灌浆后期的8次遥感产量估算精度最高(决定系数 R^2 为0.96,标准均方根误差(NRMSE)为5.39%),而起身期至开花期的6次遥感产量估算精度亦达到极好(NRMSE为9.16%),可减少遥感采集次数,提前预测产量。研究结果对采用无人机遥感进行冬小麦产量预测和精度提升具有重要意义。

关键词: 无人机; 冬小麦; 产量估算; 多时相; 最优生育时期; 三次B样条曲线

中图分类号: S24; S27 **文献标识码:** A **文章编号:** 1000-1298(2022)09-0197-10

OSID:



Yield Estimation of Winter Wheat Based on Optimization of Growth Stages by Multi-temporal UAV Remote Sensing

WANG Jingjing¹ LI Changshuo¹ ZHUO Yue¹ TAN Haibin² HOU Yongsheng³ YAN Haijun^{1,4}

(1. College of Water Resources and Civil Engineering, China Agricultural University, Beijing 100083, China
2. The Semi-arid Agriculture Engineering and Technology Research Center of P. R. China, Shijiazhuang 050051, China
3. State Key Laboratory of Soil-Plant-Machine System Technology,
Chinese Academy of Agricultural Mechanization Sciences Group Co., Ltd., Beijing 100083, China
4. Engineering Research Center of Agricultural Water-Saving and Water Resources, Ministry of Education, Beijing 100083, China)

Abstract: With the development of unmanned aerial vehicle (UAV) and remote sensing technology, crop yield estimation through rapid acquisition of multi-temporal and high-resolution remote sensing images at field scale has become a research hotspot. In order to determine the optimal growth stage and sampling times for winter wheat yield estimation by UAV multispectral remote sensing, a field experiment on winter wheat in sandy soil was conducted, which was divided into four groups (36 management zones) by irrigation level and five groups (15 management zones) by nitrogen application level. Then the multi-spectral remote sensing images of eight growth stages for winter wheat from rising to late filling were collected by the UAV platform. Additionally, partial least squares (PLS), random forest (RF), least absolute shrinkage and selection operator (LASSO) were used to establish the yield prediction model of winter wheat at each growth stage. Based on the optimal model selected, five yield estimation schemes for the vegetation indices integration during specific growth periods were developed by the cubic B-spline curve and compound trapezoidal formula. The results showed that significant differences were found for estimation accuracy at different growth stages, which was increased with the growth of winter wheat. In

收稿日期: 2022-04-07 修回日期: 2022-06-16

基金项目: 河北省重点研发计划项目(21327002D、20327003D)和国家重点研发计划项目(2017YFD0201502)

作者简介: 王晶晶(1995—),女,博士生,主要从事精准灌溉与信息化技术研究,E-mail: wangjingjing9@cau.edu.cn

通信作者: 严海军(1974—),男,教授,博士生导师,主要从事节水灌溉技术与装备研究,E-mail: yanhj@cau.edu.cn

single growth period, the optimal growth periods of PLS, RF and LASSO models were early filling, early filling and late filling, respectively. Compared with PLS and LASSO models, RF had the best precision in estimating winter wheat yield except early joint stage. The accuracy of yield estimation in the multi-growth stages was better than that in a single one. The optimal yield estimation scheme was the vegetation indices from rising to the late filling stage for eight sampling times of remote sensing (the determination coefficient R^2 of 0.96 and the normalized root mean square error (NRMSE) of 5.39%). Meanwhile, the yield estimation scheme of six sampling times from rising to flowering stage also performed excellently (NRMSE of 9.16%), which meant that it can not only reduce sampling times and remote sensing cost, but also can predict the winter wheat yield in advance. The results were of great significance for the accurate prediction of winter wheat yield by UAV remote sensing.

Key words: UAV; winter wheat; yield estimation; multi-temporal; optimal growth stage; cubic B-spline curve

0 引言

小麦是中国三大粮食作物之一,种植面积达到 $2.338 \times 10^7 \text{ hm}^2$, 占全国粮食作物总面积的 22.0%^[1]。在小麦成熟收获前进行精准产量估算对保持粮食市场稳定、确保国家粮食安全具有重要意义。传统的粮食测产采用抽样方法,依赖大量的实地采样,成本高且费时费力,无法保证测产结果的时效性及准确性。卫星遥感可进行大面积的作物生长监测^[2-4],但是其估测精度易受云层、天气等影响,且遥感图像的时空分辨率均较低。

无人机遥感具有覆盖范围广、测量周期短、成本低和机动性强等特点,能够快速、准确获取作物表型关键信息,提供区域尺度的高时空分辨率图像。基于无人机遥感平台,利用解析图像获取植被指数,并结合机器学习算法,可对作物进行高效快捷的大面积产量估算^[5-7],以克服传统方法的不足^[8]。文献[9-11]利用无人机搭载各种传感器,采用不同算法构建产量估算模型,相比于传统回归方法,机器学习算法提高了产量估算精度,但在采用神经网络等算法建模时需要大量的训练样本,且极易出现过拟合问题。偏最小二乘法(PLS)将预测变量和观测变量投影到一个新空间,建立线性回归模型,可以避免数据非正态分布、因子不确定性和模型不能识别等潜在问题^[12]。随机森林(RF)是一种基于多决策树分类的新型机器学习算法,具有很好的抗噪性能和极强的拟合能力,且不会产生过拟合现象等^[13]。套索算法(LASSO)是一种同时进行特征选择和正则化的回归分析方法,可增强统计模型的预测准确性和可解释性^[14]。然而作物的光谱特征受生物因子和非生物因子的共同影响^[15],在不同生长阶段具有不同的冠层特点,其相应的敏感植被指数有所不同,产量估算模型精度亦有所不同。

为提高作物产量估算精度,相关学者开展了多时相无人机遥感的研究。ZHOU 等^[16]与韩文霆

等^[17]采用梯形公式对多时相植被指数积分,发现相比于单生育时期,采用多时相无人机遥感植被指数可提高水稻与夏玉米的产量估算精度。上述研究在建模时选择单一植被指数进行线性回归,然而不同生育时期的最优植被指数不同且存在饱和问题,同时输入参数与产量之间存在非线性关系,因此采用单个植被指数线性建模会降低模型的普适性和预测精度。程千等^[18]采用冬小麦连续 4 个生育时期遥感数据进行产量估算,结果显示多生育时期的冬小麦产量估算模型精度均高于单生育时期,但其建模方法为多个生育时期植被指数的简单合并,并不能解释小麦全生育阶段的日植被指数的动态变化趋势。综上所述,多生育时期遥感可以提高农作物产量估算精度,然而随着时间的推移,多次收集信息会增加工作量和成本,因此评估最优生育时期及采集次数至关重要。

针对多生育时期遥感信息,为解决光谱在时间序列上不均匀问题,杨雪^[19]与 ZHANG 等^[20]使用 Logistic 回归构建了水稻冠层植被指数 NDVI 随时间的动态变化趋势,但有学者认为相比于 Logistic 和 Gompertz,三次 B 样条曲线集成了三次样条和曲率最小化的优点,能够解释作物营养生长与生殖生长时期的变化趋势^[21-22],更适用于描述作物动态生长变化规律。然而不同作物的植被结构存在差异,会影响冠层光谱信息,采用样条曲线描述冬小麦逐日植被指数的规律以及对小麦产量估算精度的影响还有待研究。

为此,本文基于无人机遥感平台采集多时相冬小麦遥感影像,通过 PLS、RF 与 LASSO 构建单生育时期冬小麦产量估算模型,获得最优生育时期及算法;采用三次 B 样条曲线与牛顿复合梯形公式,利用最优算法构建 5 种特定生育阶段的产量估算方案,探究遥感采集次数对冬小麦产量估算精度的影响,以期优选出无人机遥感冬小麦产量估算的生育时期。

1 材料与方法

1.1 研究区域与试验设计

试验地点位于河北省涿州市中国农业大学教学实验场(北纬 39°27′,东经 115°21′),海拔 42 m,属于暖温带半湿润季风区,季节温差变化大且四季分明,多年平均降雨量 563.3 mm,年平均气温 11.6℃。试验地土壤类型(国际制)以砂土为主,0~80 cm 土层的砂粒、粉粒和黏粒平均质量分数为 87.92%、8.40%和 3.68%,有机质、有效磷、速效钾、硝态氮、铵态氮含量(质量比)分别为 11.72 g/kg、32.45 mg/kg、54.98 mg/kg、14.01 mg/kg、4.56 mg/kg。

研究作物冬小麦(品种为农大 212)于 2020 年 10 月 12 日播种,2021 年 6 月 8 日收获。播种量为 270 kg/hm²,行距 15 cm,播种时统一施入底肥,合计纯 N 为 54 kg/hm²、P₂O₅ 为 138 kg/hm²、K₂O 为 81 kg/hm²。在冬小麦起身期至成熟期设置施氮处理和灌水处理两个试验。施氮处理采用人工撒施尿素(总氮含量大于等于 46%),设置 5 种追施氮肥水平,即 N0(0)、N1(66.7 kg/hm², 1/3 N₃)、N2(133.4 kg/hm², 2/3 N₃)、N3(200 kg/hm²)、N4(266.7 kg/hm², 4/3 N₃)。每个水平处理重复 3 次,共 15 个样区,各样区的总灌水量均为 270 mm(含越冬灌水量 30 mm)。灌水处理试验地块呈约 60°扇形,划分成 12 个试验区,对应设置 4 个灌溉水平,即 W1(150 mm,极度亏缺)、W2(190 mm,中度亏缺)、W3(230 mm,轻度亏缺)、W4(270 mm,充分灌溉),均含越冬灌水量 30 mm。灌水试验区采用水肥一体化追施尿素(总氮含量大于等于 46%)溶液,追施氮量均为 200 kg/hm²,每个试验区内划分 3 个样区,共 36 个样区。施氮和灌水处理各样区随机布置,面积均为 6 m×6 m,统一采用一台 3 跨、半径为 140 m 的圆形喷灌机进行灌溉,其他田间管理相同,试验布置图如图 1 所示。

1.2 数据获取与处理

1.2.1 多光谱图像

使用无人机(精灵 Phantom 3 Advanced,深圳大疆创新科技有限公司)搭载多光谱相机(RedEdge, Micasense,美国)获取冬小麦多光谱图像。该相机由 5 个波段组成,波长范围分别为蓝(465~485 nm)、绿(550~570 nm)、红(663~673 nm)、红边(712~722 nm)和近红外(820~860 nm)。选择在 11:00—13:00 晴朗低风条件下飞行无人机,飞行高度为 50 m,航向及旁向的重叠度均为 75%。分别在起身期(4 月 6 日)、拔节前期(4 月 14 日)、拔节后期(4 月 22 日)、孕穗期(4 月 30 日)、抽穗期

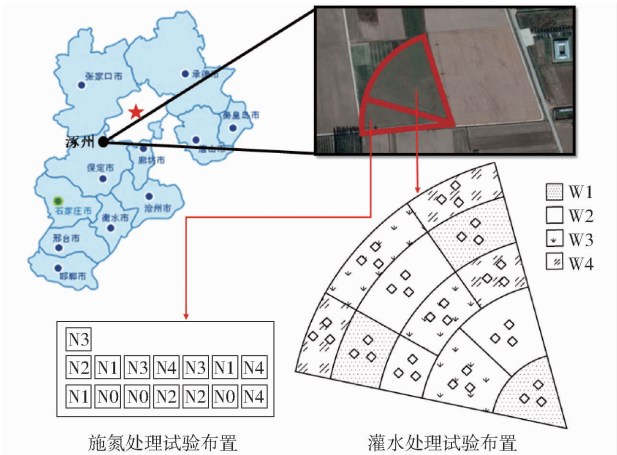


图 1 研究区域和试验设计

Fig. 1 Experimental area and design

(5 月 7 日)、开花期(5 月 11 日)、灌浆前期(5 月 18 日)、灌浆后期(5 月 27 日)进行了 8 次影像采集。使用 Pix4D Mapper 4.4 软件对多光谱图像进行几何校正和辐射校正,利用 ArcGIS 10.2 软件对拼接预处理的遥感影像进行裁剪处理,得到每个试验样区光谱图像,提取各样区光谱反射率并求得植被指数。本文选取常见的 15 种无人机遥感植被指数构建冬小麦产量估算模型,相关植被指数的计算公式如表 1 所示。

1.2.2 产量

冬小麦成熟时,在每个样区内选取 3 个代表性的 1 m×1 m 样方,对小麦脱粒后进行称量。通过谷物水分测定仪(LDS-1G 型)测量冬小麦籽粒含水率,并将其换算为含水率 13% 的产量,3 个样方产量平均值即为该样区产量实测值。施氮和灌水试验的产量数据如表 2 所示,51 个样区冬小麦产量均值为 5 675 kg/hm²,各处理间产量差异显著($P<0.01$)。

1.3 建模方法

首先将本文计算的 15 种植被指数与冬小麦产量采用 SPSS 20.0 进行相关性分析,然后选取显著相关($P<0.01$)的植被指数,作为产量估算模型的输入参数,使用机器学习算法构建单生育时期以及多生育时期产量估算模型。从 51 组样本数据中随机选取 35 组作为模型训练集,剩余 16 组作为验证集,通过 Python 软件编写回归程序。

1.3.1 单生育时期产量估算模型

为探究产量估算模型的最优生育时期和算法,分别采用 PLS、RF、LASSO 算法构建无人机遥感采集的冬小麦 8 个单生育时期产量估算模型。

1.3.2 多生育时期产量估算模型

为分析多生育时期光谱采集次数对产量估算精度的影响,采用三次 B 样条拟合多时相植被指数,将采集时间点的植被指数数据作为节点,应用广义

表 1 植被指数计算公式
Tab.1 Equations for vegetation indices

植被指数	计算公式	文献序号
比值植被指数 (Ratio vegetation index, RVI)	$RVI = \frac{NIR}{R}$	[23]
归一化植被指数 (Normalized difference vegetation index, NDVI)	$NDVI = \frac{NIR - R}{NIR + R}$	[24]
归一化差值红边指数 (Normalized difference red-edge, NDRE)	$NDRE = \frac{NIR - RE}{NIR + RE}$	[25]
增强植被指数 (Enhanced vegetation index, EVI2)	$EVI2 = \frac{2.5(NIR - R)}{NIR + 2.4R + 1}$	[26]
优化土壤调整植被指数 (Optimized soil adjusted vegetation index, OSAVI)	$OSAVI = \frac{1.16(NIR - R)}{NIR + R + 0.16}$	[27]
差值植被指数 (Difference vegetation index, DVI)	$DVI = NIR - R$	[28]
改进叶绿素吸收指数 (Modified chlorophyll absorption in reflectance index, MCARI)	$MCARI = \frac{NIR - RE - 0.2(NIR - G)}{NIR/RE}$	[29]
绿色归一化植被指数 (Green normalized difference vegetation index, GNDVI)	$GNDVI = \frac{NIR - G}{NIR + G}$	[30]
红边叶绿素指数 (Red-edge chlorophyll index, CIRE)	$CIRE = \frac{NIR}{RE} - 1$	[31]
增强植被指数 (Enhanced vegetation index, EVI)	$EVI = \frac{2.5(NIR - R)}{NIR + 6R - 7.5B + 1}$	[32]
优化植被指数 (Optimized vegetation index, Vplot)	$Vplot = 100(\ln NIR - \ln RE)$	[25]
垂直植被指数 (Perpendicular vegetation index, PVI)	$PVI = \frac{2.5(NIR - 10.489R - 6.604)}{(1 + 10.489^2)^{0.5}}$	[28]
绿色优化土壤调整植被指数 (Green optimal soil adjusted vegetation index, GOSAVI)	$GOSAVI = \frac{1.16(NIR - G)}{NIR + G + 0.16}$	[25]
改进比值植被指数 (Modified simple ratio, MSR)	$MSR = \frac{NIR/R - 1}{(NIR/R + 1)^{0.5}}$	[33]
改良土壤调整植被指数 (Modified soil adjusted vegetation index, MSAVI)	$MSAVI = 0.5\{2NIR + 1 - [(2NIR + 1)^2 - 8(NIR - R)]^{0.5}\}$	[34]

注: NIR 、 R 、 RE 、 G 、 B 分别表示近红外波段反射率、红光波段反射率、红边波段反射率、绿光波段反射率、蓝光波段反射率。

表 2 不同施氮和灌水处理的冬小麦产量
Tab.2 Winter wheat yield of different fertilization and irrigation treatments

施氮试验		灌水试验	
处理	产量/(kg·hm ⁻²)	处理	产量/(kg·hm ⁻²)
N0	(4 593 ± 145) ^c	W1	(2 556 ± 66) ^d
N1	(5 197 ± 167) ^d	W2	(4 968 ± 377) ^c
N2	(6 074 ± 645) ^c	W3	(6 111 ± 545) ^b
N3	(6 910 ± 76) ^b	W4	(7 039 ± 104) ^a
N4	(7 629 ± 99) ^a		
均值	5 675		

注: 同列数据后不同小写字母代表处理在 $P < 0.01$ 水平上差异显著。

交叉验证防止样条参数过拟合, 寻找最优平滑水平, 得到日植被指数, 并通过复合梯形公式进行时间积分, 复合梯形公式为

$$VI = \int_1^n f(t_i) dt = \frac{1}{2} \left[f(t_1) + 2 \sum_{i=2}^{n-1} f(t_i) + f(t_n) \right]$$

式中 VI ——植被指数积分值
 $f(t_1)$ 、 $f(t_i)$ 、 $f(t_n)$ ——采集时间为 t_1 、 t_i 、 t_n 的植被指数

采用单生育时期产量估算模型的最优算法构建多生育时期产量估算模型, 输入参数为各植被指数积分值, 输出参数为产量。根据计算的日植被指数, 定义 5 种特定生育阶段的日植被指数积分方案, 其中方案 1 为采集的 8 个生育时期遥感数据, 方案 2 至方案 4 依次向前推进一个生育时期, 方案 5 为从 8 个生育时期中间隔挑选 5 个。其中 4 月 6 日 (第 1 天) 开始采集数据, 5 月 27 日 (第 52 天) 结束采集, 具体方案如表 3 所示。

1.4 精度评价

植被指数拟合值与实测值的吻合程度采用相对误差 (RE) 进行评价, RE 越接近 0, 表明拟合效果越好。采用决定系数 (Coefficient of determination, R^2)、平均绝对误差 (Mean absolute error, MAE)、均方根误差 (Root mean square error, RMSE)、标准均方根误差 (Normalized root mean square error, NRMSE) 来评价产量估算结果的准确性。 R^2 越接近于 1, MAE、RMSE、NRMSE 越接近 0, 即表示模型预测精度越高; 根据 NRMSE 标准分为 4 个等级, 即极好

表 3 5 种多时相植被指数积分方案

Tab.3 Five multi-temporal vegetation indices integration schemes

方案	无人机遥感采集次数	采集起止时间
1	8 次(起身期、拔节前期、拔节后期、孕穗期、抽穗期、开花期、灌浆前期、灌浆后期)	4 月 6 日—5 月 27 日
2	7 次(起身期、拔节前期、拔节后期、孕穗期、抽穗期、开花期、灌浆前期)	4 月 6 日—5 月 18 日
3	6 次(起身期、拔节前期、拔节后期、孕穗期、抽穗期、开花期)	4 月 6 日—5 月 11 日
4	5 次(起身期、拔节前期、拔节后期、孕穗期、抽穗期)	4 月 6 日—5 月 7 日
5	5 次(起身期、拔节前期、孕穗期、开花期、灌浆后期)	4 月 6 日—5 月 27 日

($\leq 10\%$)、较好($(10\%, 20\%]$)、一般($(20\%, 30\%]$)、较差($>30\%$)^[35]。

2 结果与分析

2.1 植被指数与产量相关性分析

冬小麦各生育时期的植被指数与产量相关性分析结果如表 4 所示。可以看出,大部分植被指数与产量的相关系数随着冬小麦的生长呈增大趋势,在灌浆后期达到最大值,其中 NDVI、NDRE、OSAVI、GNDVI、PVI、GOSAVI 的相关系数均达到 0.90 以上。在 8 个生育时期内,EVI、DVI、MCARI 及 EVI2 与产量未全部达到显著性水平,因此选取 RVI、NDVI、NDRE、OSAVI、GNDVI、CIRE、Vplot、PVI、GOSAVI、MSR 和 MSAVI 共 11 种与产量显著相关($P<0.01$)的植被指数作为冬小麦产量估算模型的输入参数。

表 4 植被指数与产量相关系数

Tab.4 Correlation coefficients between vegetation indices and yield

生育时期	RVI	NDVI	NDRE	EVI	OSAVI	DVI	MCARI	GNDVI	CIRE	EVI2	Vplot	PVI	GOSAVI	MSR	MSAVI
起身期	0.81**	0.85**	0.81**	0.78**	0.84**	0.80**	0.01	0.82**	0.81**	0.72**	0.81**	0.84**	0.82**	0.83**	0.83**
拔节前期	0.74**	0.76**	0.86**	0.69**	0.78**	0.72**	0.12	0.72**	0.85**	0.69**	0.86**	0.74**	0.77**	0.75**	0.76**
拔节后期	0.79**	0.81**	0.77**	0.34*	0.68**	0.41*	-0.19	0.83**	0.77**	0.43**	0.77**	0.79**	0.74**	0.80**	0.58**
孕穗期	0.67**	0.64**	0.78**	0.12	0.63**	0.22	0.07	0.73**	0.78**	0.32	0.78**	0.65**	0.57**	0.67**	0.61**
抽穗期	0.82**	0.81**	0.92**	0.89**	0.88**	0.90**	-0.23	0.88**	0.90**	0.81**	0.91**	0.82**	0.90**	0.83**	0.89**
开花期	0.79**	0.84**	0.91**	0.58**	0.82**	0.66**	-0.34*	0.86**	0.90**	0.76**	0.91**	0.84**	0.82**	0.81**	0.79**
灌浆前期	0.77**	0.90**	0.95**	0.82**	0.89**	0.80**	0.47	0.93**	0.89**	0.83**	0.94**	0.88**	0.90**	0.83**	0.84**
灌浆后期	0.75**	0.92**	0.91**	0.84**	0.91**	0.82**	0.69	0.92**	0.81**	0.86**	0.88**	0.90**	0.90**	0.84**	0.86**

注:**表示 $0.001\leq P<0.01$; *表示 $0.01\leq P<0.05$ 。

2.2 基于单生育时期的产量估算精度

采用 PLS、RF 和 LASSO 构建的冬小麦单生育时期产量估算模型训练集结果如表 5 所示。可以看出,在起身期时 3 种模型的 R^2 较低,均未超过 0.40, MAE、RMSE 也均在 1 000 kg/hm² 以上,NRMSE 均大于 20%,产量估算精度较差,不具备实用意义。3 种模型的产量估算精度在拔节前期与拔节后期均有所提升。在孕穗期和抽穗期,PLS 和 LASSO 模型的精度接近,而 RF 模型的精度优于 PLS 和 LASSO,但优势不明显。开花期之后,3 种模型的精度进一步提升,其中 PLS 和 RF 模型均在灌浆前期取得最佳值,相应的 R^2 分别为 0.81、0.88,MAE 分别为 509、

表 5 单生育时期产量估算模型训练集结果

Tab.5 Results of training set of yield prediction model for single growth stage

生育时期	PLS				RF				LASSO			
	R^2	MAE/ (kg·hm ⁻²)	RMSE/ (kg·hm ⁻²)	NRMSE/ %	R^2	MAE/ (kg·hm ⁻²)	RMSE/ (kg·hm ⁻²)	NRMSE/ %	R^2	MAE/ (kg·hm ⁻²)	RMSE/ (kg·hm ⁻²)	NRMSE/ %
起身期	0.29	1 116	1 298	23.43	0.39	1 036	1 197	21.58	0.28	1 098	1 305	23.54
拔节前期	0.42	898	1 164	21.00	0.46	901	1 123	20.26	0.52	820	1 054	19.01
拔节后期	0.49	945	1 096	19.77	0.56	893	1 020	18.40	0.46	1 025	1 126	20.31
孕穗期	0.47	1 042	1 114	20.09	0.70	673	832	15.01	0.54	933	1 041	18.78
抽穗期	0.64	731	917	16.54	0.70	652	832	15.00	0.62	759	943	17.01
开花期	0.71	713	829	14.96	0.75	623	770	13.88	0.73	650	803	14.49
灌浆前期	0.81	509	658	11.87	0.88	454	527	9.51	0.81	542	669	12.09
灌浆后期	0.78	582	715	12.91	0.87	440	542	9.77	0.81	570	666	12.01

454 kg/hm², RMSE 分别为 658、527 kg/hm², 但此时 PLS 模型的 NRMSE 为 11.87%, 而 RF 的 NRMSE 为 9.51%, 达到了极好水平。LASSO 模型的精度在灌浆后期达到了最佳值, R^2 为 0.81, MAE、RMSE 分别为 570、666 kg/hm², NRMSE 为 12.01%。综上所述, 单生育时期产量估算模型精度随着冬小麦的生长总体呈递增趋势; 除拔节前期外, 3 种算法中 RF 的产量估算效果最好。

图 2 为单生育时期冬小麦产量估算模型的验证

集结果, 可以看出验证集的产量估算效果和训练集保持一致, 即从起身期到灌浆后期, 3 种模型的预测精度逐渐提高, 其中 PLS 和 RF 在灌浆前期预测效果最好, 验证集 R^2 分别为 0.90、0.91, NRMSE 分别为 10.26%、9.91%。而 LASSO 模型在灌浆后期预测效果最好, 验证集 R^2 为 0.87, NRMSE 为 12.00%。比较各生育时期 3 种算法的产量估算精度, 最终选择 RF 作为冬小麦单生育时期产量反演的最优算法, 并用于构建多生育时期冬小麦产量估算模型。

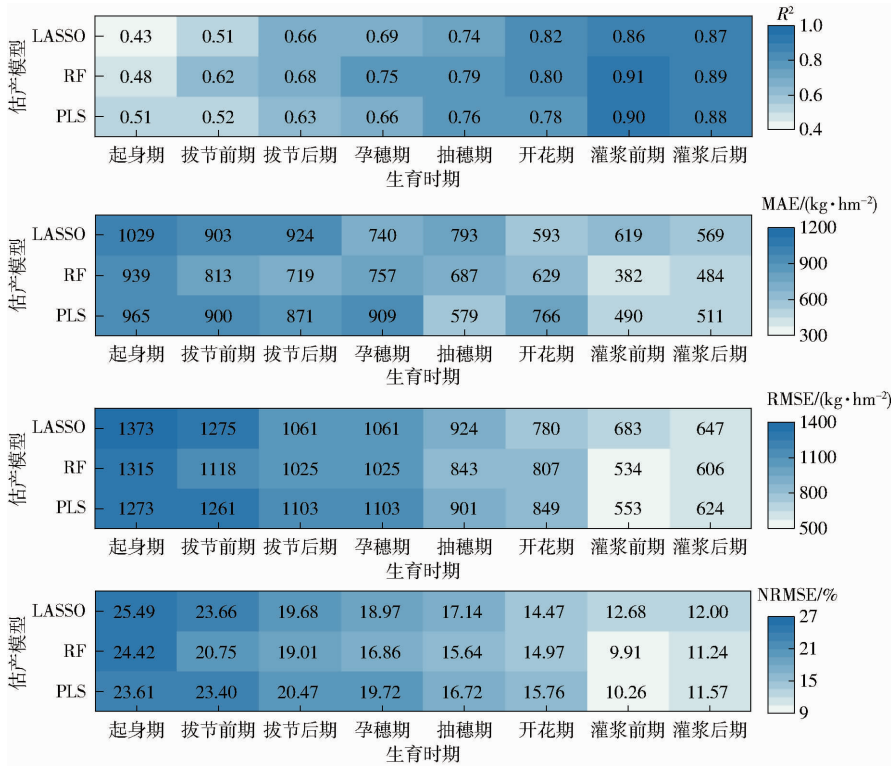


图 2 冬小麦单生育时期产量估算模型验证集精度

Fig. 2 Verification accuracy of yield of winter wheat for single growth stage

2.3 基于多生育时期的产量估算精度

2.3.1 植被指数样条曲线拟合结果

选出与冬小麦产量显著相关的 11 种植被指数, 采用三次 B 样条曲线进行拟合, 可以得到植被指数在冬小麦起身期至灌浆后期的变化趋势。由于方案 2、3、4 植被指数变化规律与方案 1 一致, 受篇幅所限, 本节比较了方案 1 与方案 5 的拟合结果。图 3 仅列出有代表性的施氮试验 N3 处理和灌水试验 W4 处理拟合结果。可以看出, 冬小麦起身期植被覆盖度较小, 冠层光谱影像中包含大量的土壤背景信息, 植被指数较小; 拔节期至孕穗期冬小麦快速生长, 植被指数呈快速增长趋势; 孕穗期至开花期植被指数变化较稳定, 并在抽穗期前后达到峰值; 开花期至灌浆后期为小麦生殖生长阶段, 植被指数呈现略微下降趋势。

图 4 为方案 1 与方案 5 中施氮处理 N3、灌水处

理 W4 的 11 种植被指数拟合值与实测值的相对误差。可以看出, 方案 1 中 N3 处理的 RE 在拔节前期达到最大值 9.36%, W4 处理的 RE 在灌浆前期达到最大值 11.52%; 方案 5 中两种处理的 RE 均在拔节前期达到最大值, 分别为 10.41%、12.23%。两种方案不同生育时期的 RE 均小于 15%, 表明采用三次 B 样条曲线拟合计算的日植被指数, 可用于构建多时相植被指数积分的冬小麦产量估算模型。

2.3.2 积分预测模型精度

根据样条曲线拟合结果, 将日植被指数按照设定的 5 种方案采用复合梯形公式进行积分, 采用单生育时期优选的 RF 算法构建模型, 模型的训练集与验证集的结果如图 5 所示。5 种方案的预测值均与实测值呈良好线性相关, 均匀分布在 1:1 线两侧。相比于单生育时期, 方案 1、2、3、5 的 R^2 、MAE、RMSE、NRMSE 有明显改善, 均优于单生育时期产量

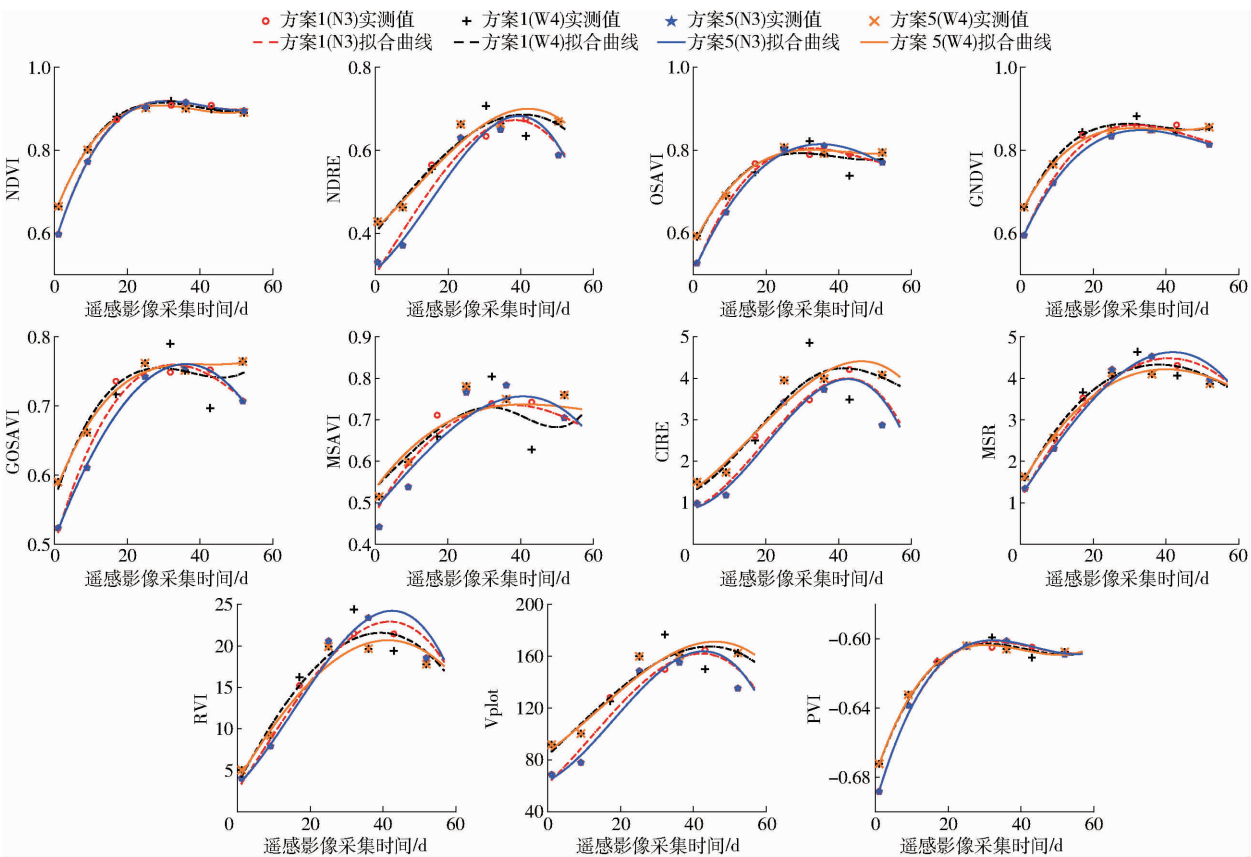


图 3 方案 1 与方案 5 中施氮处理 N3、灌水处理 W4 植被指数样条曲线拟合结果

Fig. 3 Vegetation indices fitted by spline curve of nitrogen treatment N3 and irrigation treatment W4 in scheme 1 and scheme 5

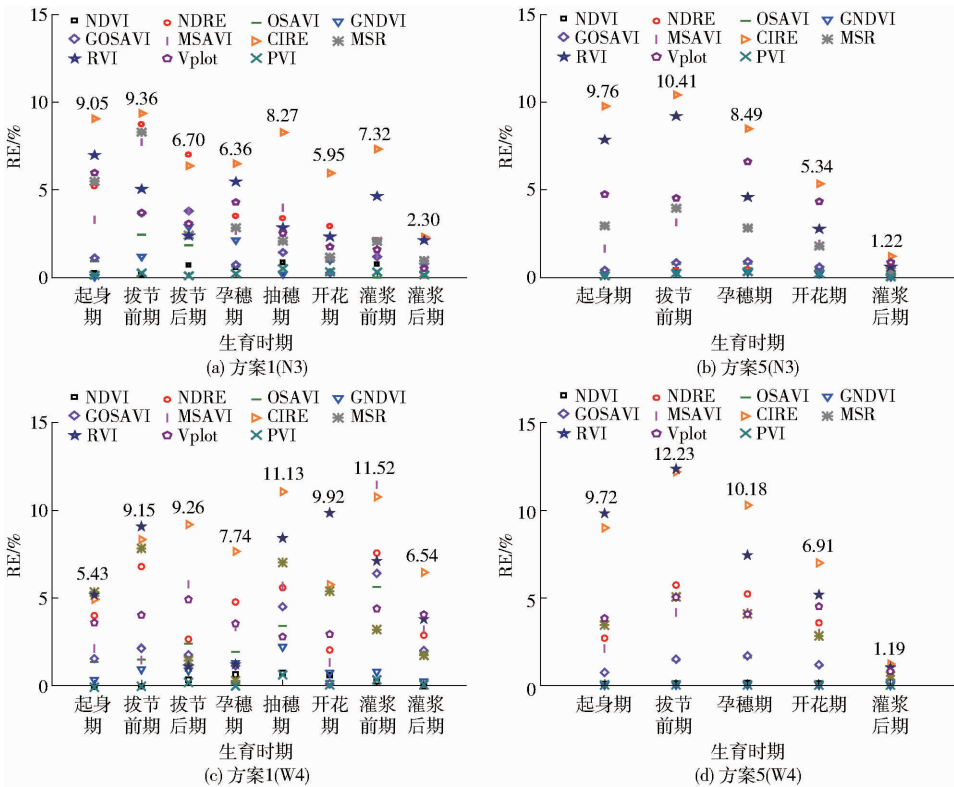


图 4 方案 1 与方案 5 中施氮处理 N3、灌水处理 W4 植被指数拟合值与实测值的相对误差

Fig. 4 Relative errors of vegetation indices of nitrogen treatment N3 and irrigation treatment W4 in scheme 1 and scheme 5

估算模型精度。与起身期至灌浆后期 8 次遥感(方案 1)相比,起身期至灌浆后期 5 次遥感(方案 5)的

训练集模型精度 R^2 从 0.96 降为 0.94,MAE、RMSE 分别从 212、299 kg/hm^2 增至 318、378 kg/hm^2 ,

NRMSE 由 5.39% 增至 6.82%, 可见方案 1 与方案 5 的模型预测效果均为极好, 验证集各参数结果与训练集结果一致, 表明适当减少采集次数并没有明显降低模型预测精度。

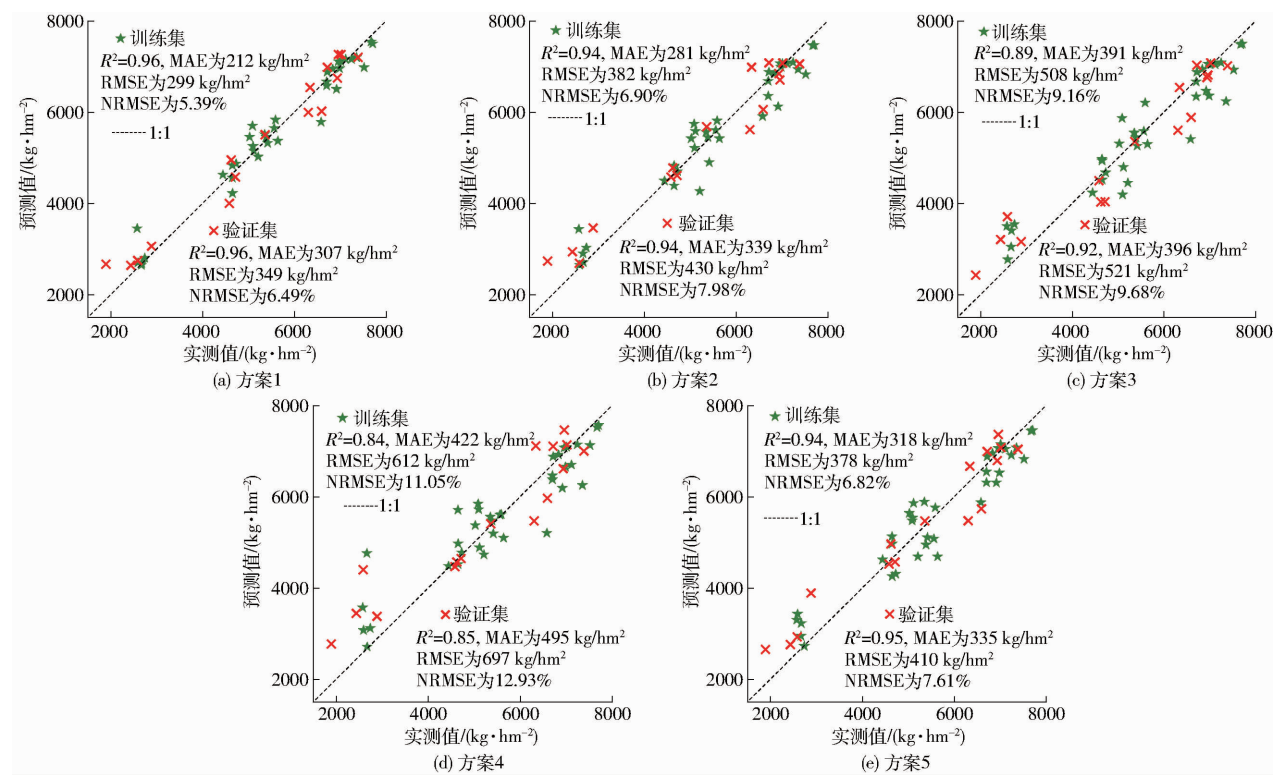


图 5 5 种积分方案产量预测结果

Fig. 5 Yield prediction results of five integral schemes

随着多光谱遥感采集生育时期前移, 方案 1、2、3、4 的训练集模型精度呈逐渐下降趋势。方案 1、2、3 的 NRMSE 均低于 10%, 而方案 4 的 NRMSE 高达 11.05%。表明采用起身期至开花期 6 次遥感(方案 3)的产量估算模型精度达到极好的水平, 可用于冬小麦产量的早期估算。

3 讨论

本研究优选的 RVI、NDVI、NDRE、OSAVI、GNDVI、CIRE、Vplot、PVI、GOSAVI、MSR 和 MSAVI 11 种植被指数与冬小麦产量呈现显著相关 ($P < 0.01$), 随着冬小麦的生长, 大部分植被指数与产量的相关性呈增大趋势。采用 PLS、RF 和 LASSO 构建单生育时期植被指数的产量估算模型精度在冬小麦营养生长阶段呈快速递增趋势, 在生殖生长阶段呈现缓慢递增趋势, 这是由于随着冬小麦生长表型差异逐渐显著。3 种模型产量估算的最优生育时期不同, PLS 和 RF 模型为灌浆前期, LASSO 模型为灌浆后期。这与文献[36-37]提出在小麦灌浆期进行产量估算有较高的预测精度相一致, 因此灌浆期可能是冬小麦产量估算的最佳时期。本研究发现除拔节前期外, RF 模型精度均高于 PLS 和 LASSO, 与文献[18]发现 RF 构建的冬小麦产量估算精度优于

PLS 和支持向量机(SVR)的结果相似, 这是由于 RF 采用多个决策树进行样本训练, 在处理复杂输入参数中优势明显^[13]。

本研究采用 RF 构建的多生育时期冬小麦产量估算模型精度高于单生育时期, 从起身期至灌浆后期 8 次遥感产量估算精度最高 ($R^2 = 0.96$, MAE 为 212 kg/hm²、RMSE 为 299 kg/hm²、NRMSE 为 5.39%)。与 8 次遥感数据相比, 从起身期至灌浆后期 5 次遥感数据的模型精度 NRMSE 为 6.82%, 表明减少遥感采集次数并未明显降低产量预测精度; 而采用起身期至开花期 6 次遥感的模型精度 NRMSE 为 9.16%, 亦达到极好效果, 这可以提前估测冬小麦产量。由于冬小麦不同生育时期的光谱反射特性会不同程度地受到外界环境影响, 需要进一步有效融合多传感器数据或集成多种机器学习算法, 以消除外界因素干扰, 提高产量估算精度。

4 结论

(1) 单生育时期产量估算模型精度随着冬小麦生育时期的发展总体呈现递增趋势。PLS、RF、LASSO 模型的最优生育时期分别为灌浆前期、灌浆前期、灌浆后期。除拔节前期外, RF 模型优于 PLS 和 LASSO。

(2)采用三次 B 样条曲线拟合获得冬小麦起身期至灌浆后期的 8 次遥感(方案 1)与 5 次遥感(方案 5)的植被指数拟合值,与实测值的 RE 均小于 15%,可用于构建植被指数积分的冬小麦产量估算模型。

(3)采用 RF 构建的多生育时期产量估算模型精度高于单生育时期,最优产量估算方案为起身期至灌浆后期 8 次遥感数据(方案 1,NRMSE 为 5.39%);与之相比,起身期至灌浆后期的 5 次遥感(方案 5)减少了采集次数,可获得极好的产量预测精度(NRMSE 为 6.82%);起身期至开花期的 6 次遥感(方案 3)亦达到极好的预测效果(NRMSE 为 9.16%)。

参 考 文 献

[1] 中华人民共和国国家统计局. 2019 年中国统计年鉴[M]. 北京: 中国统计出版社, 2020.

[2] 王来刚, 郑国清, 郭燕, 等. 融合多源时空数据的冬小麦产量预测模型研究[J]. 农业机械学报, 2022, 53(1): 198–204.
WANG Laigang, ZHENG Guoqing, GUO Yan, et al. Prediction of winter wheat yield based on fusing multi-source spatio-temporal data[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(1): 198–204. (in Chinese)

[3] 刘峻明, 周舟, 和晓彤, 等. 基于 NDWI 和卷积神经网络的冬小麦产量估测方法[J]. 农业机械学报, 2021, 52(12): 273–280.
LIU Junming, ZHOU Zhou, HE Xiaotong, et al. Winter wheat yield estimation method based on NDWI and convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(12): 273–280. (in Chinese)

[4] FAJARDO M, WHELAN B M. Within-farm wheat yield forecasting incorporating off-farm information [J]. Precision Agriculture, 2021, 22(2): 569–585.

[5] ZHANG X Y, ZHAO J M, YANG G J, et al. Establishment of plot-yield prediction models in soybean breeding programs using UAV-based hyperspectral remote sensing[J]. Remote Sensing, 2019, 11(23): 2752.

[6] MARQUES R A P, PRADO O L, ELIS G F D, et al. A random forest ranking approach to predict yield in maize with UAV-based vegetation spectral indices[J]. Computers and Electronics in Agriculture, 2020, 178: 105791.

[7] SILVA E E D, BIAO F H R, TEODORO L P R, et al. UAV-multispectral and vegetation indices in soybean grain yield prediction based on in situ observation[J]. Remote Sensing Applications: Society and Environment, 2020, 18: 100318.

[8] 彭要奇, 肖颖欣, 郑永军, 等. 无人机光谱成像技术在大田中的应用研究进展[J]. 光谱学与光谱分析, 2021, 40(5): 1356–1361.
PENG Yaoqi, XIAO Yingxin, ZHENG Yongjun, et al. Research progress in the application of UAV spectral imaging technology in field[J]. Spectroscopy and Spectral Analysis, 2021, 40(5): 1356–1361. (in Chinese)

[9] 陶惠林, 徐良骥, 冯海宽, 等. 基于无人机高光谱遥感数据的冬小麦产量估算[J]. 农业机械学报, 2020, 51(7): 146–155.
TAO Huilin, XU Liangji, FENG Haikuan, et al. Winter wheat yield estimation based on UAV hyperspectral remote sensing data [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(7): 146–155. (in Chinese)

[10] LI J T, SIVAKUMAR A N V, BHATTA M, et al. Principal variable selection to explain grain yield variation in winter wheat from features extracted from UAV imagery[J]. Plant Methods, 2019, 15:123.

[11] MAIMAITIJIAN M, SAGAN V, SIDEIKE P, et al. Soybean yield prediction from UAV using multimodal data fusion and deep learning[J]. Remote Sensing of Environment, 2020, 237: 111599.

[12] SVANTE W, MICHAEL S, LENNART E. PLS-regression; a basic tool of chemometrics[J]. Chemometrics and Intelligent Laboratory Systems, 2001, 58(2): 109.

[13] BREOMAN L. Random forests[J]. Machine Learning, 2001, 45(1): 5–32.

[14] TIBSHIRANI R. Regression shrinkage and selection via the LASSO[J]. Journal of the Royal Statistical Society. Series B (Methodological), 1996, 58(1): 267.

[15] 冯雷, 高吉兴, 何勇, 等. 波谱成像技术在作物病害信息早期检测中的研究进展[J]. 农业机械学报, 2013, 44(9): 169–176.
FENG Lei, GAO Jixing, HE Yong, et al. Research development of spectral imaging technology in early detection of botanical diseases[J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(9): 169–176. (in Chinese)

[16] ZHOU X, ZHENG H B, XU X Q, et al. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2017, 130: 246–255.

[17] 韩文霆, 彭星硕, 张立元, 等. 基于多时相无人机遥感植被指数的夏玉米产量估算[J]. 农业机械学报, 2020, 51(1): 148–155.
HAN Wenting, PENG Xingshuo, ZHANG Liyuan, et al. Summer maize yield estimation based on vegetation index derived from multi-temporal UAV remote sensing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(1): 148–155. (in Chinese)

[18] 程千, 徐洪刚, 曹引波, 等. 基于无人机多时相植被指数的冬小麦产量估测[J]. 农业机械学报, 2021, 52(3): 160–167.
CHENG Qian, XU Honggang, CAO Yinbo, et al. Grain yield prediction of winter wheat using multi-temporal UAV based on multispectral vegetation index[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(3): 160–167.

(in Chinese)

- [19] 杨雪. 稻麦两熟区冬小麦适宜氮素指标动态模型构建与追氮调控研究[D]. 南京: 南京农业大学, 2015.
YANG Xue. Study on suitable nitrogen indices dynamic models and nitrogen fertilizer regulation of winter wheat in wheat-rice cropping area[D]. Nanjing: Nanjing Agricultural University, 2015. (in Chinese)
- [20] ZHANG K, GE X K, SHEN P C, et al. Predicting rice grain yield based on dynamic changes in vegetation indexes during early to mid-growth stages[J]. *Remote Sensing*, 2019, 11(4): 387.
- [21] BRIEN C, JEWELL N, WATTS-WILLIAMS S J, et al. Smoothing and extraction of traits in the growth analysis of noninvasive phenotypic data[J]. *Plant Methods*, 2020, 16(1): 1–21.
- [22] SEBASTIAN V, TAYLOR P, CARL J B, et al. Understanding growth dynamics and yield prediction of sorghum using high temporal resolution UAV imagery time series and machine learning[J]. *Remote Sensing*, 2021, 13(9): 1763.
- [23] PEARSON R L, MILLER L D. Remote mapping of standing crop biomass for estimation of the productivity of the short grass prairie, Pawnee national grasslands, Colorado[C] // *Proceedings of the 8th International Symposium on Remote Sensing of Environment*, 1972: 1355–1379.
- [24] ROUSE J W, HAAS R H, SCHELL J A, et al. Monitoring vegetation systems in the great plains with ERTS[R]. NASA, 1974: 309.
- [25] GITELSON A A, MERZLYAK M N. Remote estimation of chlorophyll content in higher plant leaves[J]. *International Journal of Remote Sensing*, 1997, 18(12): 2691–2697.
- [26] JIANG Z, HUETE A, DIDAN K, et al. Development of a two-band enhanced vegetation index without a blue band[J]. *Remote Sensing of Environment*, 2008, 112(10): 3833–3845.
- [27] CAO Q, MIAO Y X, WANG H Y, et al. Non-destructive estimation of rice plant nitrogen status with crop circle multispectral active canopy sensor[J]. *Field Crops Research*, 2013, 153(3): 133–144.
- [28] BROGE N H, MORTENSEN J V. Deriving green crop area index and canopy chlorophyll density of winter wheat from spectral reflectance data[J]. *Remote Sensing of Environment*, 2002, 81(8): 45–47.
- [29] DAUGHTRY C S T, WALTHALL C L, KIM M S, et al. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance[J]. *Remote Sensing of Environment*, 2000, 74(2): 229–239.
- [30] GITELSON A A, KAUFMAN Y J, MERZLYAK M N. Use of a green channel in remote sensing of global vegetation from EOS–MODIS[J]. *Remote Sensing of Environment*, 1996, 58(3): 289–298.
- [31] GITELSON A A, VINA A, CIGANDA V, et al. Remote estimation of canopy chlorophyll content in crops[J]. *Geophysical Research Letters*, 2005, 32(8): L08403.
- [32] LIU H Q, HUETE A. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1995, 33(2): 457–576.
- [33] CAO Q, MIAO Y X, LI F, et al. Developing a new crop circle active canopy sensor-based precision nitrogen management strategy for winter wheat in North China Plain[J]. *Precision Agriculture*, 2017, 18(2): 2–18.
- [34] HABOUDANE D, MILLER J R, PARREY E, et al. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture[J]. *Remote Sensing of Environment*, 2004, 90(3): 337–352.
- [35] HUI X, LIN X J, ZHAO Y, et al. Assessing water distribution characteristics of a variable-rate irrigation system[J]. *Agricultural Water Management*, 2022, 206: 107276.
- [36] 朱婉雪, 李仕冀, 张旭博, 等. 基于无人机遥感植被指数优选的田块尺度冬小麦估产[J]. *农业工程学报*, 2018, 34(11): 78–86.
ZHU Wanxue, LI Shiji, ZHANG Xubo, et al. Estimation of winter wheat yield using optimal vegetation indices from unmanned aerial vehicle remote sensing[J]. *Transactions of the CSAE*, 2018, 34(11): 78–86. (in Chinese)
- [37] 陶惠林, 冯海宽, 杨贵军, 等. 基于无人机数码影像和高光谱数据的冬小麦产量估算对比[J]. *农业工程学报*, 2019, 35(27): 111–118.
TAO Huilin, FENG Haikuan, YANG Guijun, et al. Comparison of winter wheat yields estimated with UAV digital image and hyperspectral data[J]. *Transactions of the CSAE*, 2019, 35(27): 111–118. (in Chinese)