

DOI: 10.3724/SP.J.1006.2022.11053

基于无人机平台多模态数据融合的小麦产量估算研究

张少华¹ 段剑钊^{1,2} 贺利¹ 井宇航¹ Urs Christoph Schulthess^{2,*}
Azam Lashkari^{1,2} 郭天财¹ 王永华¹ 冯伟^{1,*}

¹ 河南农业大学农学院 / 省部共建小麦玉米作物学国家重点实验室, 河南郑州 450046; ² International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Mexico

摘要: 作物产量估测关系到人民生活质量和国家粮食安全问题, 在田块尺度下及时准确估算产量, 对于农事操作管理、收获、销售及种植计划制定均具有重要意义。选择地势起伏及空间差异较大的农田为研究区, 利用低空无人机遥感平台搭载多光谱相机、热红外相机和 RGB 相机, 同步获取小麦关键生育时期的无人机遥感影像, 并提取光谱反射率、热红外温度和数字高程信息。首先统计不同地形特征下遥感参数和生长指标的空间变异情况, 分析植被指数和温度参数与小麦产量的相关性, 然后利用多元线性回归(multiple linear regression, MLR)、偏最小二乘回归(partial least squares regression, PLSR)、支持向量机回归(support vector machine regression, SVR)和随机森林回归(random forest regression, RFR) 4 种机器学习方法以单模态数据和多模态遥感信息融合 2 种方式进行建模, 比较单模态数据和多模态数据融合的产量估测能力。结果表明, 坡度是影响作物生长和产量的重要因子, 3 个生育期内, 不同坡度等级下遥感参数差异明显, 土壤含水量、植株含水量和地上部生物量与坡度的相关性均达显著水平, 植被指数和温度参数与产量的相关性均达显著水平。依据与产量的相关性, 筛选 7 个植被指数(NDVI、GNDVI、EVI2、OSAVI、SAVI、NDRE、WDRVI)和 2 个温度参数(NRCT、CTD)作为模型输入变量, 对于单模态数据而言, 对产量的估算效应为植被指数 > 温度参数, 以灌浆期植被指数的 RFR 模型效果最好($R^2=0.724$, RMSE=614.72 kg hm⁻², MAE=478.08 kg hm⁻²); 对于双模态数据融合来说, 在植被指数基础上融入冠层温度参数表现最好, 开花期 RFR 模型效果进一步提高($R^2=0.865$, RMSE=440.73 kg hm⁻², MAE=374.86 kg hm⁻²); 在双模态数据基础上引入坡度信息进行三模态数据融合, 其产量估算效果明显优于单模态和双模态数据融合, 其中以开花期植被指数、温度参数和坡度信息融合的 RFR 估算效果最好($R^2=0.893$, RMSE=420.06 kg hm⁻², MAE=352.69 kg hm⁻²), 模型验证效果较好($R^2=0.892$, RMSE=423.55 kg hm⁻², MAE=334.43 kg hm⁻²)。可见, 在本试验条件下通过引入地形因子, 结合随机森林回归算法将多模态数据有效融合, 可充分发挥不同遥感信息源之间互补协同作用, 有效提高了产量估算模型的精度与稳定性, 为作物生长监测及产量估算提供思路参考和方法支持。

关键词: 冬小麦; 无人机; 产量估算; 地形因子; 多模态数据

Wheat yield estimation from UAV platform based on multi-modal remote sensing data fusion

ZHANG Shao-Hua¹, DUAN Jian-Zhao^{1,2}, HE Li¹, JING Yu-Hang¹, Urs Christoph Schulthess^{2,*}, Azam Lashkari^{1,2}, GUO Tian-Cai¹, WANG Yong-Hua¹, and FENG Wei^{1,*}

¹ Agronomy College of Henan Agriculture University / State Key Laboratory of Wheat and Maize Crop Science, Zhengzhou 450002, Henan, China;

本研究由国家“十三五”重点研发计划粮食丰产增效科技创新项目(2018YFD0300701), 财政部和农业农村部国家现代农业产业技术体系建设专项(CARS-03)和河南省科技攻关项目(212102110041)资助。

This study was supported by the National “13th Five-Year” Key Research and Development Program of China (2018YFD0300701), the China Agriculture Research System of MOF and MARA (CARS-03), and the Key Technologies Research & Development Program of Henan Province, China (212102110041).

* 通信作者(Corresponding authors): Urs Christoph Schulthess, E-mail: U.Schulthess@cgiar.org; 冯伟, E-mail: fengwei78@126.com

第一作者联系方式: E-mail: 15939266989@163.com

Received (收稿日期): 2021-06-02; Accepted (接受日期): 2021-10-19; Published online (网络出版日期): 2021-11-26.

URL: <https://kns.cnki.net/kcms/detail/11.1809.S.20211125.1758.002.html>

² International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Mexico

Abstract: Crop yield estimations are important for national food security, people, and the environment. Timely and accurate estimation of crop yield at the field scale is of great significance for crop management, harvest and trade. It ultimately enables farmers to optimize inputs and economic return. We selected an irrigated wheat field in a region near Kaifeng, Henan province, for this study. The terrain in that region is undulating and spatial differences. We used a low-altitude unmanned aerial vehicle (UAV) remote sensing platform equipped with a multi-spectral camera, thermal infrared camera, and RGB camera to simultaneously obtain different remote sensing parameters during the key growth stages of wheat. Based on the extracted spectral reflectivity, thermal infrared temperature, and digital elevation information, we calculated the spatial variability of remote sensing parameters, and growth indices under different terrain characteristics. We also analyzed the correlations between vegetation indices, temperature parameters and wheat yield. By means of four machine learning methods, including multiple linear regression method (MLR), partial least squares regression method (PLSR), support vector machine regression method (SVR), and random forest regression method (RFR), we compared the yield estimation capability of single-modal data versus multimodal data fusion frameworks. The results showed that slope was an important factor affecting crop growth and yield. We observed significant differences in remote sensing parameters under different slope grades. Soil water content, water content of plants, and above-ground biomass at the three growth stages were significantly correlated with slope. Most of the vegetation indices and temperature parameters of three growth stages were significantly correlated with yield as well. Based on the strength of their correlation with yield, seven vegetation indices (NDVI, GNDVI, EVI2, OSAVI, SAVI, NDRE, and WDRVI) and two temperature parameters (NRCT, CTD) were selected as the final input variables for the model. For the single-modal data framework, the model constructed with the vegetation indices was better than the yield model constructed with the temperature parameters, and the highest accuracy was obtained with a RFR model based on vegetation indices at filling stage ($R^2 = 0.724$, RMSE = 614.72 kg hm⁻², MAE = 478.08 kg hm⁻²). For the double modal data fusion approach, the highest accuracy resulted at flowering stage, using the temperature parameters combined with the vegetation indices of RFR model ($R^2=0.865$, RMSE=440.73 kg hm⁻², MAE=374.86 kg hm⁻²). Even higher accuracies were obtained, using the multimodal data fusion approach with a RFR model based on vegetation indices, temperature parameters and slope information at flowering stage ($R^2 = 0.893$, RMSE = 420.06 kg hm⁻², MAE = 352.69 kg hm⁻²), and the highest validation model ($R^2 = 0.892$, RMSE = 423.55 kg hm⁻², MAE = 334.43 kg hm⁻²) for fusion of the flowering stage. The results revealed that by using a multimodal data fusion framework of terrain factors combined with RFR, we can fully exploit the complementary and synergistic roles of different remote sensing information sources. This effectively improves the accuracy and stability of the yield estimation model, and provides a reference and support for crop growth monitoring and yield estimation.

Keywords: winter wheat; unmanned aerial vehicle (UAV); yield estimation; terrain factor; multimodal data

小麦作为我国主要的粮食作物之一,是北方居民的最重要口粮品种。在我国人口压力大与耕地面积不足的大背景下,小麦精确估产对国民经济发展、粮食政策制定和粮食市场调节等均具有重要意义^[1-2]。传统的田间调查估产方式因周期长、效率低、代表性差而表现不佳,严重影响了各级政府及农技部门对农业生产的指导及政策的制定^[3-4]。随着遥感技术的快速发展,新型遥感估产技术应运而生,以其“快速、高效和便捷”特点备受关注,通常遥感估产主要通过地面尺度和卫星尺度,近地遥感技术检测精度高,但因监测范围有限,估产效率大大降低;而卫星遥感适用于大面积作物监测,但运行周期长、分辨率较为粗糙,导致监测精度有待提高^[5]。近年来,无人机低空遥感技术快速发展,因其高分辨率、高灵活性和高准确性使得精准估产成为可能^[6-7]。无人机遥感平台可以自由搭载高光谱相机、多光谱相机、数码相机、热红外相机和激光雷达等传感器,从不同传感器上可以实时高频次获取光谱、纹理、热红外、植被盖度、冠层结构以及高程等数据信息。

因此,利用低空无人机遥感技术可以实时快速精准获得作物生长状态及产量状况,为作物种植计划制定及栽培调控提供信息支持^[8-12]。

前人关于无人机遥感单模态数据的产量估算主要通过植被指数实现,朱婉雪等^[13]采集多光谱影像构建了不同植被指数与小麦产量的线性模型;程千等^[14]利用多时相植被指数通过偏最小二乘法、支持向量机和随机森林回归模型对小麦产量进行估算。由于小麦生长和产量形成是多种因素共同作用的结果,利用多类型数据源融合进行作物产量估算是提高估测精度的一种很好的方法。刘畅等^[15]、Fu等^[16]利用光谱特征和多尺度纹理特征相结合较好地估算了小麦产量和生物量。Elsayed等^[17]利用偏最小二乘法将光谱指数、温度参数和植株含水量等数据融合,小麦产量的估测效果进一步得到改善,Rischbeck等^[18]利用偏最小二乘法将光谱信息、温度参数以及株高相结合,有效提高了大麦产量的估测精度。Maimaitijiang等^[19]利用RGB信息、光谱反射率及温度参数等多模态数据,结合深度神经网络方法准确

估测了大豆产量。综上,不同遥感数据源信息表征作物生长不同特征,多源信息综合使用、互相补充可增强估算模型的精度和准确性。

农业生产上,农田地形异质性使得水分、养分和物质能量等分布不均匀,严重影响作物生长发育^[20]。地形起伏主要影响土壤水分的分布,进而引发养分的再分配,造成植株水分和养分供应的时空差异;此外,不同地形会造成作物冠层温度和湿度差异,同样会影响作物群体的生长发育^[21-22]。因此,利用遥感技术监测作物长势和估产时,如能引入地形因子或许可以提高模型准确性。张新乐等^[23]利用植被指数结合地形因子,构建逐步多元回归和神经网络多层感知反演模型,有效提升了对大豆生物量的预测。侯淑涛等^[24]分析地形对大豆产量的影响,运用作物生长模型 DSSAT 模拟不同地形上的产量变异具有可行性,这说明将地形因子作为一个有效变量来建模以提高模型精度是可行的,但目前将地形因子、植被指数和温度参数等多源数据融合进行作物产量预测研究还相对较少,特别是在小麦作物上研究更少。

为进一步提高小麦产量的估算精度,本研究使用无人机遥感影像获取播种前的田块数字高程信息 (DEM),并在小麦生长关键时期获得红外热成像和光谱图像,进而提取出与小麦长势相关的植被指数、冠层温度和坡度等信息。利用不同建模算法 (MLR、PLSR、SVR 和 RFR) 构建基于植被指数、温度参数及地形因子相融合的小麦产量估算模型,比较不同算法和模态数据间产量估算效果,优化基于多模态数据融合的小麦产量估算模型及方法,为小

麦产量的精确估算提供方法依据和技术参考。

1 材料与方法

1.1 研究区概况

本研究冬小麦试验区位于河南省开封市杞县 (34°41'11"N, 114°48'34"E), 该区域属于暖温带大陆性季风气候, 年平均气温 14.5°C, 年均降雨量 673 mm。试验田采用统一的种植方式、水肥管理及病虫害防治, 土壤为潮土类型, 地势起伏明显, 田块中坡顶与坡底的最大落差为 6.65 m, 田块内部土壤水分的空间分布差异明显, 孕穗期土壤平均含水量在 2°~4°坡度为 0.16、4°~6°为 0.13 和 6°~8°为 0.12, 开花期土壤平均含水量在 2°~4°坡度为 0.17、4°~6°为 0.15 和 6°~8°为 0.11, 灌浆期土壤平均含水量在 2°~4°坡度为 0.14、4°~6°为 0.10 和 6°~8°为 0.07。

1.2 地面数据获取

于 2020 年冬小麦孕穗期 (4 月 7 日)、开花期 (4 月 26 日) 和灌浆期 (5 月 12 日) 对冬小麦取植株样, 每个样点取 0.2 m² 小麦植株, 称取鲜重后, 105°C 杀青 30 min, 80°C 下烘干至恒重, 计算植株含水量。以上 3 个取样日期均避开降雨和灌水, 以避免温度测定随机性的影响。土壤含水率的测定: 用土钻取 0~20 cm, 20~40 cm 土层土样装入铝盒, 采用传统的土壤干燥法在 105°C 下烘干至恒重, 计算土壤样品的含水量。

于小麦成熟期测产, 为保证采集的数据均匀且具代表性, 依据地块位置及地形分布, 选取测产样点如图 1 所示, 并用 GPS 记录采样点的中心坐标, 样点测产面积为 4 m², 采集 60 个样点。

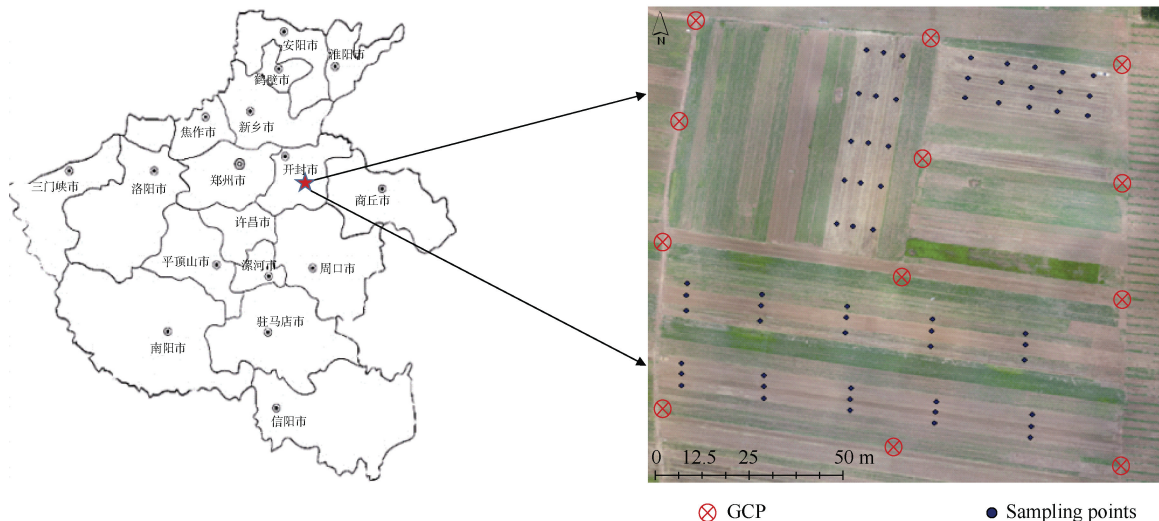


图 1 研究区位置及样点分布

Fig. 1 Location of the research areas and experiment designs

1.3 无人机平台数据的获取及处理

在小麦植株取样的同时获取无人机平台的多光谱和热红外数据。使用 Matrice600PRO 大疆六旋翼无人机与 Ronin 系列云台, 搭载 2064 像素×1544 像素 K6 机载模块化多光谱相机, 该传感器包含蓝光(450)、绿光(550)、红光(685)、红边(725)、近红外(780) 5 个多光谱通道和 1 个 4384 像素×3288 像素 RGB 原色通道。同时该平台搭载 640 像素×512 像素 Thermal Capture 2.0 热成像仪获取热红外图像。选择在 11:00—13:00, 天气晴朗, 无风无云的环境条件下进行无人机作业, 以避免环境条件的影响。飞行高度

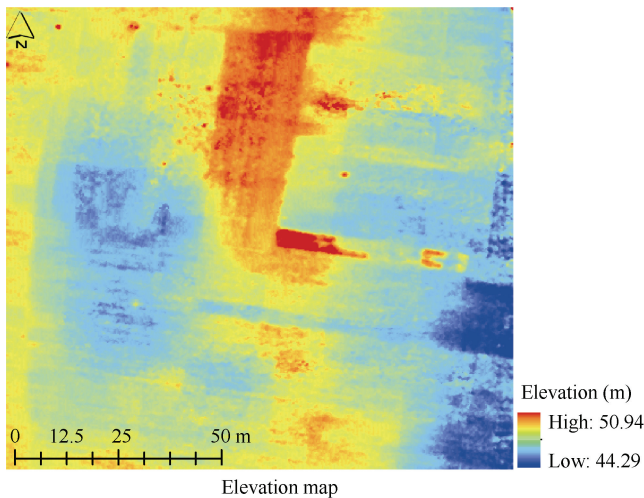


图 2 研究区地形特征

Fig. 2 Terrain features of study fields

1.5 无人机数据处理

1.5.1 热红外影像处理 将获取的小麦各生育时期无人机热红外图像采用 ThermoViewer 软件进行图像筛选与饱和度调整等预处理, 将处理好的图像导入 Pix4Dmapper 中进行图形拼接, 在 ArcGIS 中利用 RGB 图像作为基准影像对热红外图像进行配准, 利用测温枪获取 4 块靶标布及近似黑体的水体实际温度对热红外图像进行校准, 同步使用矢量文件提取试验小区的冠层温度, 如图 3 所示, 统计各样点小麦冠层温度。

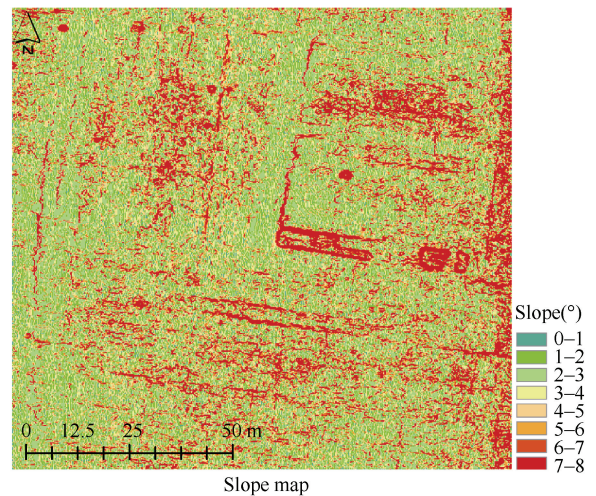
1.5.2 多光谱影像处理 将获取的小麦各生育时期无人机多光谱数据采用 Pix4Dmapper 软件进行图像拼接以得到试验区的正射影像, 利用 ArcGIS 软件进行辐射校正和几何校正, 制作试验小区 shp 矢量文件和 buffer 缓冲区, 然后计算植被指数并提取各样点的多光谱植被指数。

1.5.3 植被指数和冠层温度相关指数的选取 用

90 m, 拍摄方式为等时触发, 航向重叠度 75%, 旁向重叠度 70%。拍摄前, 使用手持式 ASD 对铺设反射率为 5%、20%、40%和 70%的靶标布进行校准。

1.4 地形数据获取

小麦整地前使用精度为 1 cm 的华策实时动态定位(real-time kinematic, RTK)对研究区域进行打点定位, 得到 12 个经纬度坐标和高程信息。通过 Photoscan 将控制点坐标导入无人机 RGB 影像中进行拼接, 生成 RGB 正射图像与高精度的数字高程模型(DEM), 并在 ArcGIS 软件中使用 Spatial Analyst 工具提取坡度因子, 获取研究区地形特征如图 2 所示。



于作物产量估算的植被指数种类较多, 依据前人结果筛选出适用于无人机传感器通道的 10 种植被指数和 2 种冠层温度相关的参数(表 1), 以用于构建小麦产量估算模型。

1.6 模型构建及精度验证

选用每个生育时期 70%样本量用于建模($n_1=42$), 剩余 30%样本量用于验证($n_2=18$)。采用决定系数 R^2 , 均方根误差 RMSE 和平均绝对误差 MAE 评价产量估算模型的效果。通常 R^2 越大, RMSE 和 MAE 越小, 模型的拟合效果与精度越好。

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

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}} \quad (2)$$

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} \quad (3)$$

式中, i 表示第 i 个样本点数据, y_i 为第 i 个样本点的冬小麦表型实测值; x_i 为模型算出的第 i 个样本点的作物表型预测值, \bar{y} 为实测平均值, \bar{x} 为模型估测平均值。

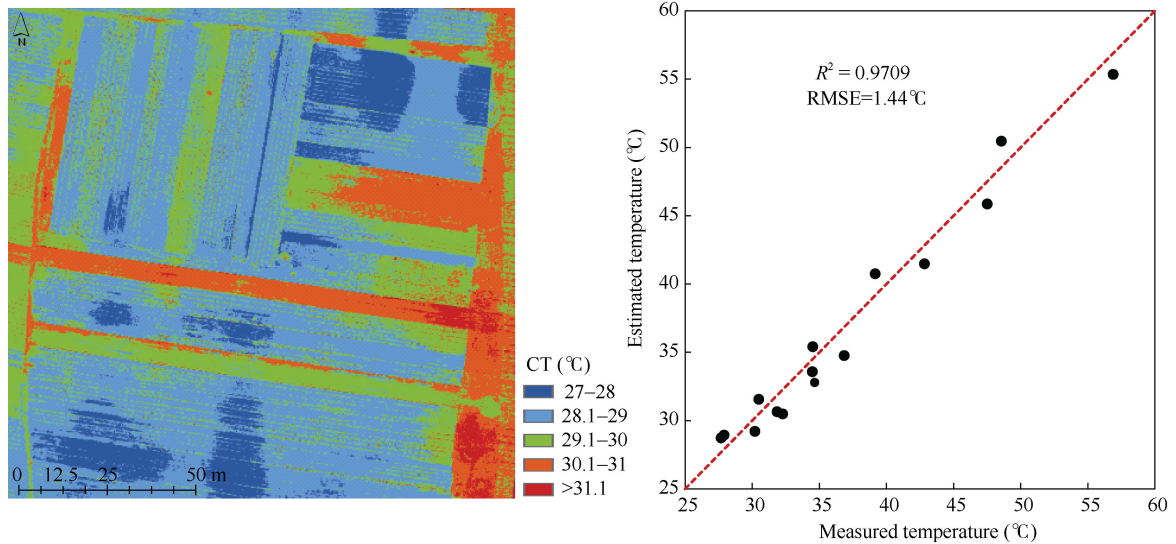


图 3 开花期冠层温度分布图及实测和预测的冠层温度
Fig. 3 Canopy temperature distribution map measured and estimated crop canopy temperature at flowering stage

表 1 本文采用的光谱指数和温度参数
Table 1 Spectral indices and temperature parameter in this study

参数 Index	公式 Formula	参考文献 Reference
归一化植被指数 Normalized difference vegetation index (NDVI)	$\frac{(R_{nir} - R_{red})}{(R_{nir} + R_{red})}$	[25]
归一化绿度植被指数 Green normalized difference vegetation index (GNDVI)	$\frac{(R_{nir} - R_{green})}{(R_{nir} + R_{green})}$	[26]
双波段增强植被指数 Two-band enhanced vegetation index (EVI2)	$2.5 \times \frac{(R_{nir} - R_{red})}{(R_{nir} + 2.4 \times R_{red} + 1)}$	[27]
优化土壤调整植被指数 Optimized soil adjusted vegetation index (OSAVI)	$1.16 \times \frac{(R_{nir} - R_{red})}{(R_{nir} + R_{red} + 0.16)}$	[28]
土壤调整植被指数 Soil adjusted vegetation index (SAVI)	$1.5 \times \frac{(R_{nir} - R_{red})}{(R_{nir} + R_{red} + 0.5)}$	[29]
红边归一化植被指数 Red edge normalized index (NDRE)	$\frac{(R_{nir} - R_{re})}{(R_{nir} + R_{re})}$	[30]
宽动态植被指数 Wide dynamic range vegetation index (WDRVI)	$\frac{(0.1 \times R_{nir} - R_{red})}{(0.1 \times R_{nir} + R_{red})}$	[31]
改善简单比率植被指数 Modified simple ratio index (MSR)	$\frac{\left(\frac{R_{nir}}{R_{red}} - 1\right)}{\left(\sqrt{\frac{R_{nir}}{R_{red}}} + 1\right)}$	[32]
改良叶绿素吸收率指数 Modified chlorophyll absorption ratio index (MCARI)	$\frac{[R_{re} - R_{red} - 0.2(R_{re} - R_{green})]}{\left(\frac{R_{re}}{R_{red}}\right)}$	[33]

(续表 1)

参数 Index	公式 Formulas	参考文献 References
红波段比值植被指数 Red ratio vegetation index (RVI _{red})	$\frac{R_{nir}}{R_{red}}$	[34]
标准化相对冠层温度 Normalized relative canopy temperature (NRCT)	$\frac{(T_c - T_{min})}{(T_{max} + T_{min})}$	[35]
冠-气温差 Canopy temperature depression (CTD)	$T_c - T_a$	[36]

R_{green} 为绿波段反射率; R_{red} 为红波段反射率; R_{re} 为红边波段反射率; R_{nir} 为近红外波段反射率; T_c 为冠层温度; T_{min} 为最低温度; T_{max} 为最高温度; T_a 为大气温度。

R_{green} is green band; R_{red} is red band; R_{re} is red edge band; R_{nir} is near infrared band; T_c is canopy temperature; T_{min} is the lowest temperature measured; T_{max} is the highest temperature measured; T_a is the air temperature.

2 结果与分析

2.1 不同地形特征下遥感影像参数的统计描述

为了准确评价地形因素对所获取遥感参数的影响, 依据研究区域的地形地貌特点, 将坡度在 $2^\circ \sim 8^\circ$ 之间的像元以 2° 为间隔分割成 3 个梯度 ($2^\circ \sim 4^\circ$ 、 $4^\circ \sim 6^\circ$ 和 $6^\circ \sim 8^\circ$)。统计每个坡度等级范围内归一化绿度植被指数 GNDVI 和冠层温度 CT 数据的变化情况 (表 2)。整体上, 所调查田块区域的地形坡度在 $2^\circ \sim 8^\circ$ 之间, 海拔集中在 44~60 m 之间, 随着坡度的增加, GNDVI 逐渐降低, 小麦生长状况变差, 而冠层温度逐渐增加。从孕穗期到灌浆期, GNDVI 先增加而后降低, 总体变化幅度为 0.518~0.837, 坡度增加时, GNDVI 变异度逐渐加大, 尤其在孕穗期和灌浆期的最大坡度时变异更大; 开花期植株生长旺盛, 盖度大, 绿度高。冠层温度随着生育时期后移逐渐升高, 整体变化幅度为 $25.46 \sim 33.97^\circ\text{C}$, 在坡度较高时, 温度变异度较大。可见, 坡度引起降雨和灌水径流, 影

响水分和养分的田间分布及再分配, 进而导致小麦长势的空间差异分布。

2.2 坡度因子与小麦生长参数间相关性

根据试验区不同地形条件选取采样点, 测定不同地形采样点的土壤含水量、植株含水量和生物量。将地形坡度与小麦植株含水量和生物量分别进行相关分析 (表 3), 坡度与小麦生长参数均呈显著负相关。生长参数间比较, 土壤含水量相关性最高, 其次为植株含水量, 地上部生物量相对较低。不同生育时期也存在差异, 灌浆期最高, 其次为开花期, 孕穗期最低。在灌浆期, 土壤含水量、植株含水量和地上部生物量与坡度的相关系数 r 分别为 -0.731 、 -0.521 和 -0.434 , 相关性均达极显著水平; 在孕穗期相关系数 r 分别为 -0.605 、 -0.450 和 -0.332 , 相关性依然达到显著水平。土壤含水量、植株含水量和地上部生物量与坡度的相关性随着生育时期推进呈逐渐增强的趋势, 说明地形是造成该区域小麦产量差异的重要原因之一。

表 2 不同坡度等级下遥感影像参数的描述性统计

Table 2 Descriptive statistics of images remote sensing indices on slopes

生育期 Growth stage	坡度 Slope ($^\circ$)	GNDVI				CT ($^\circ\text{C}$)			
		最大值 Max.	最小值 Min.	均值 Mean	标准差 SD	最大值 Max.	最小值 Min.	均值 Mean	标准差 SD
孕穗期 Booting stage	2~4	0.805	0.731	0.778	0.051	26.62	25.46	26.24	0.44
	4~6	0.723	0.602	0.637	0.049	28.24	27.56	27.42	0.34
开花期 Flowering stage	6~8	0.671	0.526	0.594	0.087	28.51	27.84	28.26	0.61
	2~4	0.837	0.758	0.809	0.047	28.64	27.14	27.84	0.46
	4~6	0.772	0.652	0.725	0.054	29.85	28.37	29.02	0.51
灌浆期 Filling stage	6~8	0.713	0.608	0.634	0.069	30.64	29.02	29.79	0.58
	2~4	0.796	0.719	0.747	0.053	31.05	29.46	30.54	0.52
	4~6	0.705	0.564	0.598	0.064	31.46	29.85	31.12	0.56
	6~8	0.669	0.518	0.573	0.081	33.97	30.31	32.85	0.78

GNDVI: 归一化绿度植被指数; CT: 冠层温度。

GNDVI: green normalized difference vegetation index; CT: canopy temperature.

表 3 不同生育期农学指标与坡度相关系数

Table 3 Correlation between growth parameters and slope at different growth stages

生长参数 Growth parameter	孕穗期 Booting stage	开花期 Flowering stage	灌浆期 Filling stage
土壤含水量 Soil moisture content (%)	-0.605**	-0.673**	-0.731**
植株含水量 Plant water content (%)	-0.450**	-0.503**	-0.521**
地上部生物量 Above-ground biomass (kg hm ⁻²)	-0.332*	-0.406**	-0.434**

表示 0.01 水平显著; *表示 0.05 水平显著。: $P < 0.01$; *: $P < 0.05$.

2.3 植被指数和温度参数与产量间相关性分析

将选取的植被指数、冠层温度参数分别与冬小麦产量进行相关分析(表 4)。除 RVI 与产量相关性达显著水平外($P < 0.05$), 其他遥感参数均表现为极显著相关水平($P < 0.01$)。从生育时期上看, 随着生育进程的推进相关性逐渐增强, 在灌浆期有 4 个植被指数的相关系数高于 0.80, GNDVI 表现最高(0.821); 开花期的相关系数范围为 0.614~0.770, NDRE 最优(0.770); 而孕穗期的相关系数范围为 0.280~0.668, SAVI 的相关系数最高(0.668)。不同植被指数在不同时期的相关系数存在较大差异, 整体而言, NDVI、GNDVI、EVI2、OSAVI 和 SAVI 在 3 个生育时期的表现均较好, 其次为 NDRE 和 WDRVI, 在不同时期的相关系数也均维持较高水平。两个温度参数 NRCT 和 CTD 与产量间相关性在开花期最高(r 为 -0.758 和 0.669), 与植被指数的相关系数表现接近; 而在孕穗期和灌浆期的相关系数相对较差, 孕穗期内温度参数的相关系数 r 分别为 -0.487 和 0.443, 整体低于植被指数, 而在灌浆期, 相关系数 r 分别为

-0.590 和 0.467, 与植被指数的差距较大。

2.4 基于单模态数据的产量估算模型构建

为了评价单模态数据估算产量的能力, 采用不同的建模方法比较分析了植被指数和温度参数在不同模型估算中的效果。根据与产量的相关性分别选取 7 个植被指数和 2 个冠层温度参数, 输入模型的因子分别为 NDVI、GNDVI、EVI2、OSAVI、SAVI、NDRE、WDRVI 和 NRCT、CTD, 选用 MLR、PLSR、SVR 和 RFR 共 4 种建模方法估算不同生育时期的冬小麦产量(表 5)。由表 5 可知, 基于植被指数估算产量时, 随着生育时期的推进, 不同算法的估测效果均逐渐变好, 灌浆期最好(平均 $R^2=0.681$ 、 $RMSE=628.92$ kg hm⁻²、 $MAE=497.49$ kg hm⁻²), 其次为开花期(平均 $R^2=0.668$ 、 $RMSE=653.93$ kg hm⁻²、 $MAE=550.56$ kg hm⁻²), 以上两时期较好结果与植被指数和产量间相关性是一致的。不同建模算法间比较, RFR 模型估算效果最佳, 其次为 SVR, 而 MLR 和 PLSR 模型相对较差。从不同算法在各个时期的最优化效果看, RFR 模型在灌浆期和开花期最佳

表 4 不同生育期植被指数和温度参数与产量相关系数

Table 4 Correlation between vegetation indices, temperature parameters and yield at different growth stages

指数 Index	孕穗期 Booting stage	开花期 Flowering stage	灌浆期 Filling stage
NDVI	0.587**	0.718**	0.799**
GNDVI	0.616**	0.756**	0.821**
EVI2	0.663**	0.732**	0.806**
OSAVI	0.653**	0.741**	0.820**
SAVI	0.668**	0.735**	0.809**
NDRE	0.666**	0.770**	0.743**
WDRVI	0.567**	0.705**	0.796**
MSR	0.390**	0.689**	0.688**
MCARI	0.475**	0.614**	0.714**
RVI	0.280*	0.664**	0.752**
NRCT	-0.487**	-0.758**	-0.590**
CTD	0.443**	0.669**	0.467**

**表示 0.01 水平显著; *表示 0.05 水平显著。缩写同表 1。

** : $P < 0.01$; * : $P < 0.05$. Abbreviations are the same as those given in Table 1.

($R^2=0.711\sim 0.724$ 、 $RMSE=614.72\sim 645.85\text{ kg hm}^{-2}$ 、 $MAE=478.08\sim 527.65\text{ kg hm}^{-2}$), 其次为 SVR 模型 ($R^2=0.673\sim 0.680$ 、 $RMSE=618.41\sim 648.37\text{ kg hm}^{-2}$ 、 $MAE=486.64\sim 532.46\text{ kg hm}^{-2}$)。相比于植被指数, 基于温度参数的产量估算效果相对较差, 其估算效果随着生育时期推进, R^2 呈先升高后降低的趋势, $RMSE$ 和 MAE 则相反, 其中, 以开花期的估算效果最佳 ($R^2=0.653$ 、 $RMSE=666.84\text{ kg hm}^{-2}$ 、 $MAE=584.2\text{ kg hm}^{-2}$), 而灌浆期和孕穗期均较差。以温度参数为输入变量的不同估算方法间的差异与植被指数基本一致, 具体为 $RFR > SVR > MLR$ 和 $PLSR$, RFR 和 SVR 模型在开花期表现较好 ($R^2=0.659\sim 0.707$ 、 $RMSE=642.19\sim 668.87\text{ kg hm}^{-2}$ 、 $MAE=530.89\sim 581.64\text{ kg hm}^{-2}$)。

为验证单模态数据估算小麦产量的效果, 对训

练集表现较好的模型利用验证集数据进行检验, 并做实测值与预测值之间关系图(图 4)。图 4-a~c 是基于冬小麦孕穗期至灌浆期植被指数建立的最佳 RFR 模型验证产量预测值与实测值间的拟合图, 图 4-d~f 是基于冬小麦孕穗期至灌浆期温度参数建立的最佳 RFR 模型验证产量预测值与实测值间的拟合图。不同估算方法的验证集精度同训练集基本一致, 对于植被指数而言, RFR 模型在开花期的表现为 $R^2=0.707$ 、 $RMSE=699.72\text{ kg hm}^{-2}$ 和 $MAE=642.01\text{ kg hm}^{-2}$, 在灌浆期的结果为 $R^2=0.720$ 、 $RMSE=603.60\text{ kg hm}^{-2}$ 和 $MAE=506.73\text{ kg hm}^{-2}$; 对于温度参数来说, 在开花期 RFR 的检验结果为 $R^2=0.689$ 、 $RMSE=647.26\text{ kg hm}^{-2}$ 和 $MAE=597.52\text{ kg hm}^{-2}$, 这表明基于单模态数据的 RFR 模型可以对小麦产量进行估算。

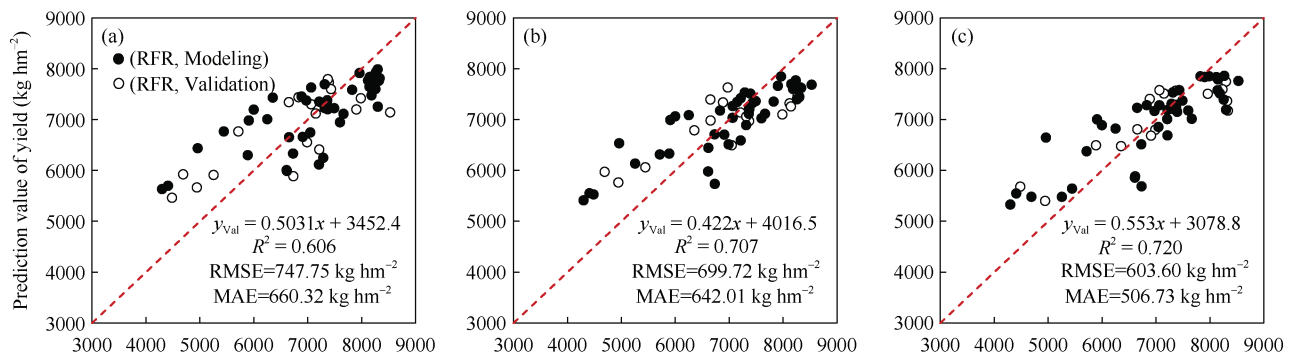
表 5 基于单模态数据的小麦产量估算精度

Table 5 Yield estimation of wheat based on single modal data

参数 Index	回归模型 Regression model	孕穗期 Booting stage			开花期 Flowering stage			灌浆期 Filling stage		
		R^2	RMSE (kg hm^{-2})	MAE (kg hm^{-2})	R^2	RMSE (kg hm^{-2})	MAE (kg hm^{-2})	R^2	RMSE (kg hm^{-2})	MAE (kg hm^{-2})
VI	MLR	0.496	760.76	648.32	0.663	659.23	567.27	0.672	638.43	502.93
	PLSR	0.508	765.69	635.14	0.626	662.28	574.84	0.648	644.11	523.51
	SVR	0.521	755.52	597.46	0.673	648.37	532.46	0.680	618.41	486.64
	RFR	0.564	700.62	576.70	0.711	645.85	527.65	0.724	614.72	478.08
TH	MLR	0.340	902.84	728.61	0.637	676.12	608.38	0.391	859.12	654.23
	PLSR	0.349	883.15	694.48	0.609	680.17	615.89	0.368	880.24	676.27
	SVR	0.361	876.73	671.52	0.659	668.87	581.64	0.463	807.35	594.65
	RFR	0.417	839.74	650.37	0.707	642.19	530.89	0.540	788.59	563.06

VI 为植被指数, TH 为温度参数。MLR 为多元线性回归, PLSR 为偏最小二乘回归, SVR 为支持向量机回归, RFR 为随机森林回归。 R^2 为决定系数, $RMSE$ 为均方根误差, MAE 为平均绝对误差。

VI: vegetation index; TH: temperature parameter. MLR: multiple linear regression; PLSR: partial least squares regression; SVR: support vector machine regression; RFR: random forest regression. R^2 : the coefficient of determination; $RMSE$: root mean square error; MAE is mean absolute error.



(图 4)

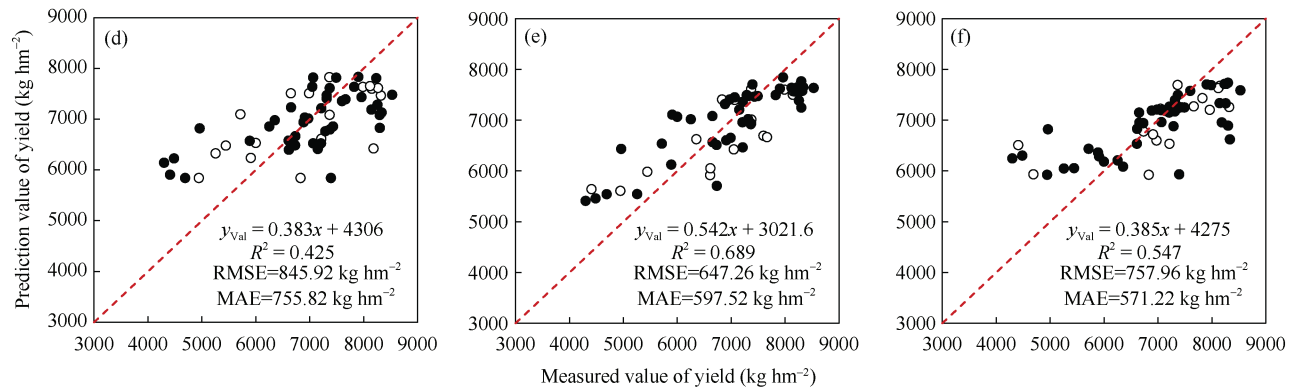


图 4 基于单模态冬小麦产量实测值与预测值间关系

Fig. 4 Relationship between the measured and predicted yields of winter wheat

a-c 是基于冬小麦孕穗期、开花期、灌浆期植被指数建立的最佳 RFR 模型验证产量预测值与实测值间的散点图; d-f 是基于冬小麦孕穗期、开花期、灌浆期温度参数建立的最佳 RFR 模型验证产量预测值与实测值间的散点图。

a-c are the scatter plots between the predicted yield and the measured value based on the best RFR model established based on the vegetation indices in the booting stage, flowering stage, and the filling stage of winter wheat; d-f are the scatter plots between the predicted yield and the measured value based on the best RFR model established based on the temperature parameters in the booting stage, flowering stage, and the filling stage of winter wheat.

2.5 基于多模态数据的产量估算模型构建

利用植被指数(VI)、温度参数(TH)和坡度因子(TF), 通过 4 种建模方法将不同类型自变量参数组合进行产量估算(表 6)。如表 6 所示, 整体而言, 无论将两类传感器获取的遥感参数结合, 或者 3 类传感器参数结合作为自变量, 不同建模方法估算结果的趋势是一致的, 即模型精度表现为 RFR > SVR > MLR 和 PLSR, RFR 和 SVR 之间差异较小, PLSR 和 MLR 模型的质量和精度均相对较低, 且明显小于 RFR 和 SVR 模型。对于不同模态数据结合而言, 模型估算效果在不同生育期间存在差异, 以开花期和灌浆期较好, 平均 R^2 分别为 0.757 和 0.710, 平均 RMSE 分别为 $559.25 \text{ kg hm}^{-2}$ 和 $611.29 \text{ kg hm}^{-2}$, 平均 MAE 分别为 $472.08 \text{ kg hm}^{-2}$ 和 $498.37 \text{ kg hm}^{-2}$, 而孕穗期明显变差, 平均 R^2 为 0.593, 平均 RMSE 为 $722.39 \text{ kg hm}^{-2}$, 平均 MAE 为 $590.01 \text{ kg hm}^{-2}$ 。

将植被指数与坡度结合作为自变量, 产量估算效果与单模态植被指数模型表现一致, 随着生育时期推进, 估算精度逐渐增加, 灌浆期的 RFR 模型估算结果最好 ($R^2=0.786$ 、 $\text{RMSE}=561.47 \text{ kg hm}^{-2}$ 、 $\text{MAE}=459.64 \text{ kg hm}^{-2}$), 相比于仅采用植被指数构建的 RFR 模型 R^2 提高 8.56%, RMSE 降低 8.66%, MAE 降低 3.85%。

将植被指数与温度参数结合作为自变量, 从 4 种模型产量估算的平均效果看, 以开花期最佳 ($R^2=0.820$ 、 $\text{RMSE}=500.17 \text{ kg hm}^{-2}$ 、 $\text{MAE}=423.64 \text{ kg hm}^{-2}$), 其次为灌浆期 ($R^2=0.790$ 、 $\text{RMSE}=564.54 \text{ kg}$

hm^{-2} 、 $\text{MAE}=462.1 \text{ kg hm}^{-2}$), 孕穗期最低 ($R^2=0.654$ 、 $\text{RMSE}=711.34 \text{ kg hm}^{-2}$ 、 $\text{MAE}=593.73 \text{ kg hm}^{-2}$)。其中, RFR 模型在开花期估算效果最好 ($R^2=0.865$ 、 $\text{RMSE}=440.73 \text{ kg hm}^{-2}$ 、 $\text{MAE}=374.86 \text{ kg hm}^{-2}$), 与仅利用植被指数构建的 RFR 模型相比, R^2 提高 21.66%, RMSE 降低 31.76%, MAE 降低 28.96%。

将温度参数与坡度结合作为自变量时, 由于参数较少, 产量估算效果较差, 开花期效果最好, 4 种模型的平均 R^2 仅为 0.677, RMSE 为 $666.22 \text{ kg hm}^{-2}$, MAE 为 $555.76 \text{ kg hm}^{-2}$, 孕穗期最差, 其中 RFR 模型在开花期估算效果最好 ($R^2=0.727$ 、 $\text{RMSE}=640.49 \text{ kg hm}^{-2}$ 、 $\text{MAE}=518.79 \text{ kg hm}^{-2}$), 比仅采用温度参数构建的产量模型 R^2 提高 2.25%, RMSE 降低 0.83%, MAE 降低 2.26%, 模型效果提升有限。

由以上结果可知, 以植被指数为基础, 将温度和坡度因子引入结合, 模型精度得到大幅度提升, 因此可将这 3 种数据源融合作为自变量预测小麦产量。从 4 种建模方法的平均估算效果看, 不同生育时期比较具体为开花期 ($R^2=0.835$ 、 $\text{RMSE}=459.48 \text{ kg hm}^{-2}$ 、 $\text{MAE}=379.96 \text{ kg hm}^{-2}$) > 灌浆期 ($R^2=0.802$ 、 $\text{RMSE}=507 \text{ kg hm}^{-2}$ 、 $\text{MAE}=432.28 \text{ kg hm}^{-2}$) > 孕穗期 ($R^2=0.686$ 、 $\text{RMSE}=626.28 \text{ kg hm}^{-2}$ 、 $\text{MAE}=500.13 \text{ kg hm}^{-2}$)。其中, RFR 模型在开花期的估算表现最佳 ($R^2=0.893$ 、 $\text{RMSE}=420.06 \text{ kg hm}^{-2}$ 、 $\text{MAE}=352.69 \text{ kg hm}^{-2}$), 该模型的关键参数具体为: 优化决策树的数量为 500、内部节点随机采样的个数为 3、进行 1000 次迭代和基础学习。以 3 种数据源融合为输入变量的

开花期 RFR 模型结果与仅利用植被指数相比, R^2 提高 25.6%, RMSE 降低 34.96%, MAE 降低 33.16%。此外, SVR 模型在开花期也取得了较好结果, R^2 达 0.824, RMSE 为 $451.26 \text{ kg hm}^{-2}$, MAE 为 $379.65 \text{ kg hm}^{-2}$ 。

为进一步验证基于多模态数据估算小麦产量的效果, 利用验证集数据进行检验, 并做实测值与预测值之间关系图(图 5)。图 5-a~c 是基于冬小麦孕穗期至灌浆期植被指数结合温度参数建立的最佳 RFR 模型验证产量预测值与实测值间的拟合图, 图 5-d~f 是基于冬小麦孕穗期至灌浆期植被指数、温度参数

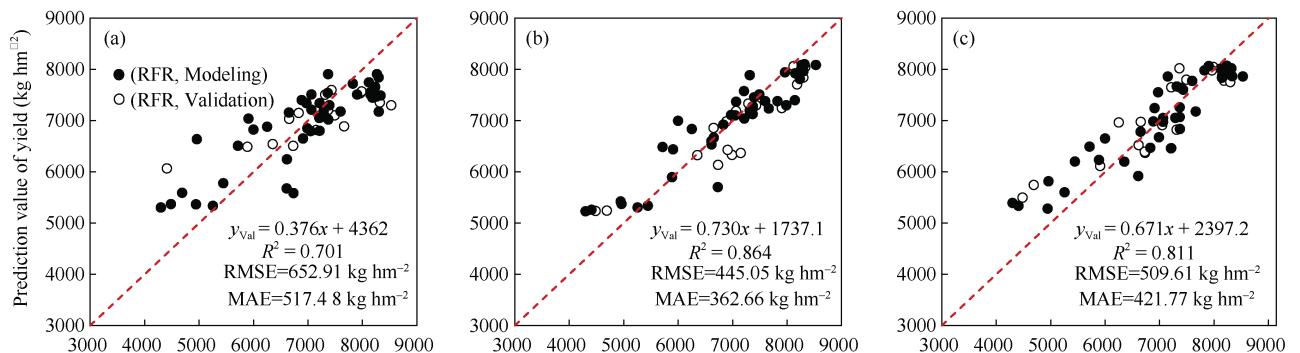
与坡度建立的最佳 RFR 模型验证产量预测值与实测值间的拟合图。不同产量估算模型间比较, 验证集同训练集的表现基本一致, 基于双模态数据融合的 RFR 模型在开花期的检验效果最好 ($R^2=0.864$ 、 $\text{RMSE}=445.05 \text{ kg hm}^{-2}$ 、 $\text{MAE}=362.66 \text{ kg hm}^{-2}$), 基于三模态数据融合的 RFR 模型依然在开花期表现最好 ($R^2=0.892$ 、 $\text{RMSE}=423.55 \text{ kg hm}^{-2}$ 、 $\text{MAE}=334.43 \text{ kg hm}^{-2}$)。遥感信息源趋于丰富, 估算模型的稳定性增强。可见, 将不同模态的遥感数据融合可进一步提高产量估算效果及模型稳定性。

表 6 基于多模态数据融合的小麦产量估算
Table 6 Yield estimation of wheat based on multimodal data fusion

参数 Index	回归模型 Regression model	孕穗期 Booting stage			开花期 Flowering stage			灌浆期 Filling stage		
		R^2	RMSE (kg hm^{-2})	MAE (kg hm^{-2})	R^2	RMSE (kg hm^{-2})	MAE (kg hm^{-2})	R^2	RMSE (kg hm^{-2})	MAE (kg hm^{-2})
VI+TF	MLR	0.544	766.67	642.58	0.677	615.09	546.26	0.727	625.12	497.92
	PLSR	0.513	722.35	615.79	0.635	654.25	559.85	0.711	650.38	513.67
	SVR	0.604	693.45	567.62	0.729	598.46	514.24	0.769	584.27	481.37
	RFR	0.633	671.14	541.46	0.746	576.69	495.46	0.786	561.47	459.64
VI+TH	MLR	0.644	743.29	634.84	0.801	521.62	439.67	0.780	596.92	482.47
	PLSR	0.609	749.22	629.75	0.790	542.45	456.87	0.752	606.89	489.67
	SVR	0.663	709.49	574.57	0.822	495.88	423.15	0.795	573.61	474.35
	RFR	0.699	643.36	535.74	0.865	440.73	374.86	0.831	480.73	401.91
TF+TH	MLR	0.371	863.76	705.27	0.652	671.64	571.29	0.417	816.05	639.37
	PLSR	0.403	854.87	685.23	0.622	694.52	586.13	0.432	832.48	656.74
	SVR	0.495	838.46	664.67	0.708	658.24	546.82	0.562	727.13	587.68
	RFR	0.561	797.06	642.13	0.727	640.49	518.79	0.594	697.58	559.98
VI+TF+TH	MLR	0.660	656.07	527.41	0.817	460.94	375.15	0.787	507.92	447.26
	PLSR	0.654	637.50	514.79	0.805	505.66	412.35	0.753	585.85	479.84
	SVR	0.702	600.37	487.37	0.824	451.26	379.65	0.813	476.67	421.31
	RFR	0.729	611.19	470.94	0.893	420.06	352.69	0.856	457.56	380.71

VI 为植被指数, TH 为温度参数, TF 为地形因子。其余缩写同表 5。

VI: vegetation index; TH: temperature parameter; TF: terrain factor. Other abbreviations are the same as those given in Table 5.



(图 5)

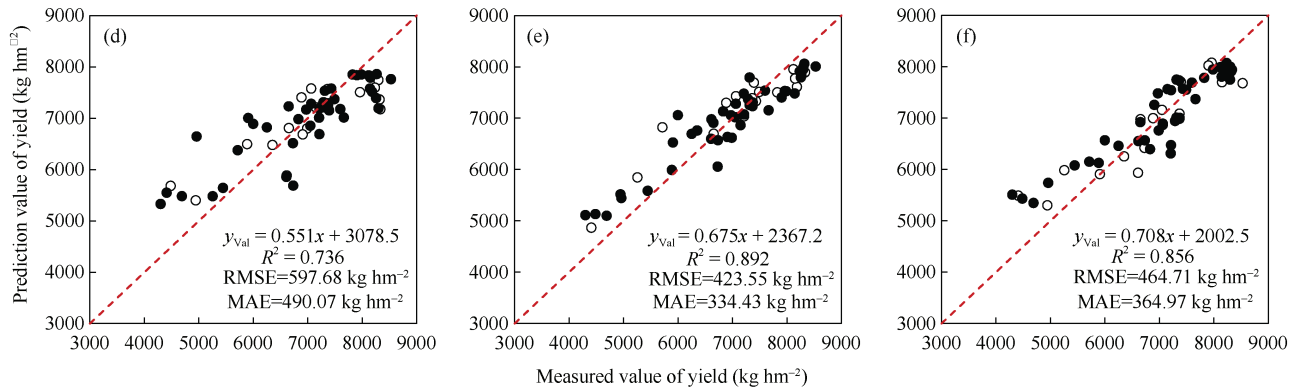


图 5 基于多模态冬小麦产量实测值与预测值间关系

Fig. 5 Relationship between measured and predicted winter wheat yield of winter wheat

a-c 是基于冬小麦孕穗期、开花期、灌浆期植被指数结合温度参数建立的最佳 RFR 模型验证产量预测值与实测值间的散点图; d-f 是基于冬小麦孕穗期、开花期、灌浆期植被指数、温度参数与坡度建立的最佳 RFR 模型验证产量预测值与实测值间的散点图。

a-c are the scatter plots between the predicted yield and the measured value based on the best RFR model established based on the vegetation indices and temperature parameters in the booting stage, flowering stage, and the filling stage of winter wheat; d-f are the scatter plots between the predicted yield and the measured value based on the best RFR model established based on the vegetation indices, temperature parameters and slope in the booting stage, flowering stage, and the filling stage of winter wheat.

3 讨论

3.1 微地形因子对作物产量的影响

坡度表示局部地表表面的倾斜程度, 坡度大小直接影响着地表物质流动与能量转换的规模与强度, 是制约生产力空间布局的重要因子^[37]。本研究区的地形坡度集中在 $2^{\circ}\sim 8^{\circ}$, 地形起伏明显, 进而影响了土壤水分分布和植株含水量以及生物量形成 ($P < 0.05$), 导致作物在不同坡度下的生长状况及产量差异显著。前人研究表明, 地形状况显著影响植被指数, 导致不同坡度及坡向下植被指数与实际情况产生较大偏差, 并且这种地形效应随着坡度的增大而增加^[38-39]。另一方面, 地形属性如高程或坡度影响着生态系统结构和组成, 进而引起太阳的有效光辐射以及降水的空间再分布, 进而影响冠层温度^[40]。前人通过获取中小尺度范围的高精度地形因子反演作物生物量, 取得了较好效果, 这与地形因子与植株生长密切相关, 并最终影响产量的形成有关^[41-42]。本试验在不同坡度下, 孕穗期到灌浆期在最大坡度与最小坡度下 GNDVI 差值为 $0.229\sim 0.279$, 冠层温度差值为 $3.05\sim 4.51^{\circ}\text{C}$, 表明作物长势的空间差异明显。前人主要是通过遥感的技术手段获取植被冠层光谱信息进行产量估测, 但在地形复杂的地区, 仅仅依靠植被冠层光谱信息估测作物产量的误差通常较大。因此, 本研究在植被指数和冠层温度结合的同时, 引入地形特征数据参与估算建模, 在一定程度上缓解了地形效应造成的数据偏差及饱和问题,

所构建多模态数据融合的产量模型在复杂地形条件下更具适用性和稳定性。可见, 在作物遥感估产中, 尤其在那些地势复杂地区要重视地形因子对产量估算结果的影响, 充分利用地形因子的信息补充及贡献, 最大限度提升产量估算精度。

3.2 基于多模态遥感数据的产量估算

小麦从孕穗期至灌浆期, 是从营养生长与生殖生长并进阶段逐渐转向生殖生长为主的阶段, 该时期是产量形成的重要阶段。随着生育时期的推进, 植被指数与产量的相关性逐渐增强, 但在种植密度较大和产量水平较高时植被指数存在饱和现象, 仅仅利用植被指数估算产量的精度还有待进一步提高^[43], 本研究中在孕穗期估算效果最差, 灌浆期最好 ($R^2=0.724$, $\text{MAE}=478.08 \text{ kg hm}^{-2}$)。此外, 光谱数据主要反应光与植物或者土壤的相互作用, 尽管在很大程度上反映作物长势状况, 但是光谱数据依然存在信息表征的内在局限性, 其与蒸腾作用并没有直接关系, 而冠层温度和冠层蒸腾作用存在密切联系, 利用冠层温度信息可以在一定程度上评价作物产量状况, 本文中其与产量的相关性均达显著水平, 最优化的相关系数达 -0.758 。Guo 等^[44]与 Hou 等^[45]研究证明冠层温度可反应作物的生长状况, 尤其当胁迫发生时, 冠层温度变化较为敏感, 其与植株含水量、土壤含水量和色素浓度等密切相关。同时, 冠层温度也能够反应作物气孔导度的动态变化, 前人利用温度参数预测作物产量开展了较多研究, 且取得较好效果^[17-19]。本文利用冠层温度信息对小

麦产量估算也取得较好效果($R^2=0.707$ 、 $MAE=530.89$ kg hm^{-2}), 但较植被指数稍差。进一步将冠层温度与光谱信息结合, 预测精度均有所提高, 这与前人的研究结果一致。可见, 在以后评价作物生长及产量估算时, 冠层温度信息可作为冠层光谱信息的有益补充^[46-47]。

无人机平台搭载的不同传感器可以获得多源数据, 这些多模态遥感信息融合能够显著提高对作物长势的监测及产量的估算。将植被指数与冠层温度参数结合, 其产量估算效果明显优于单模态数据估算^[17-19], 本研究也证实了以上 2 种类型数据结合好于单一类型数据, 其开花期 RFR 模型决定系数比仅利用植被指数提高 21.66%。由于先天地形及家庭联产承包导致的微地形变化在生产上较为常见, 坡度的变化直接影响降雨及灌水在田间的差异分布, 将地形因子融入多模态数据中, 有效提高了作物产量的估测精度^[23-24]。Rischbeck 等^[18]与 Maimaitijiang 等^[19]也证明了多类型数据源结合对作物产量的估算具有较好效果, 较单模态数据模型精度有较大程度提升。本研究结果显示, 将多光谱相机的冠层光谱信息、热红外相机的冠层温度信息以及数字高程信息有机结合, 产量估算精度较两种传感器信息融合得到进一步提高。本文中植被指数与温度结合的开花期 RFR 模型($R^2=0.865$ 、 $MAE=374.86$ kg hm^{-2}), 而在此基础上将坡度信息融入, RFR 模型效果提高($R^2=0.893$ 、 $MAE=352.69$ kg hm^{-2})。这表明冠层光谱、温度和数字高程信息在作物产量预测中可以信息互补, 能够很好地估算小麦产量。比较而言, 3 个生育期间产量预测效果存在较大差异, 本研究中以开花期最好, 其次为灌浆期, 而孕穗期最差。其原因为孕穗期正值小麦穗花发育关键期, 营养生长与生殖生长并进, 此时小花小穗发育受外界环境影响较大, 遥感监测该期小麦长势进而预测产量具有很大的不确定性; 至开花期, 小麦穗数确定, 此时小花受精结实对外界条件非常敏感, 若开花期生长良好, 也可为随后的籽粒灌浆提供更充分的物质基础, 该期遥感监测小麦长势, 对穗数和粒数的解释性较好, 同时对籽粒灌浆也具有较弱的预测性^[48]; 籽粒灌浆期是粒重形成的关键时期, 叶片叶绿素含量下降, 温度差异变小, 此时生长状况对穗粒数反映不足, 利用热红外仪感知胁迫状况效果较差, 这些导致了灌浆期遥感估算小麦产量的精度下降, 这与前人利用温度信息估算小麦产量的结果一致^[49-50]。可见,

开花期承前启后, 即反映作物前期生长基础, 又可预测后期生长趋势, 在本试验条件下开花期可作为小麦产量估算的关键时期, 为种植管理及决策提供参考依据。

3.3 基于不同建模算法的产量估算

随着计算机建模科学的快速发展, 机器学习技术在作物监测及分类中得到了广泛应用。但由于不同的建模方法, 其原理不尽相同, 监测目标物属性也有区别, 导致所得到的结果可能存在很大差异^[51-52]。为了探究不同的回归方法对产量的估算效果, 本研究比较了 MLR、PLSR、SVR 和 RFR 四种建模方法估算小麦产量的潜力。基于 MLR 模型可以有效解释自变量与因变量之间的线性关系, 但是在产量估算当中受到很大的限制, 最典型的就是自变量之间的多重共线性问题, 这种变量之间的多重相关性会严重破坏产量估计, 扩大模型误差, 验证效果不佳并破坏模型的稳定性^[53]。PLSR 模型在 MLR 模型的基础上, 加入典型相关分析与主成分分析, 对系统中的数据信息进行分解和筛选, 提取对产量的解释性最强的综合变量, 从而更好地克服变量多重相关性在建模中的不良作用, 同时 PLSR 模型在处理复杂多维度的关系与辨识系统中的信息与噪声时仍存在短板^[54]。本文利用 MLR 与 PLSR 算法的最高估算的 R^2 仅为 0.817 和 0.805, RMSE 分别为 460.94 kg hm^{-2} 和 505.66 kg hm^{-2} , MAE 分别为 375.15 kg hm^{-2} 和 412.35 kg hm^{-2} , 在处理多元复杂因子时表现出较大的局限性。SVR 模型和 RFR 模型在处理这些问题时表现出很好的适应性, 使产量模型的评价指标进一步提高, 估算能力增强, 但 SVR 模型的核心问题是核函数确定及相关参数的选择, 构建的产量模型在一定程度上受到限制^[55], SVR 模型最高 R^2 为 0.824, RMSE 为 451.26 kg hm^{-2} , MAE 为 379.65 kg hm^{-2} 。RFR 具有决策树系统的优点, 能够处理具有高维特征的样本, 无需降维就能评估各个特征在分类问题上的重要性, 并且能够有效运行大数据集, 具有较高的准确性和稳定性, 更适合解决反演问题^[56-58]。在本试验条件下, RFR 模型估算最高 R^2 为 0.893, RMSE 为 420.06 kg hm^{-2} , MAE 为 352.69 kg hm^{-2} , 说明 RFR 模型在融合多模态数据方面具有较大优势, 可作为小麦产量估算的有效途径, 为农艺调控及政策指导提供参考。

本文在应用不同模型算法协同植被指数、温度和地形多源数据估算小麦产量时, 并没有考虑不同

数据源参数对模型的贡献度,如何利用不同算法确定不同数据源参数的权重进一步提升模型精度还有待进一步研究。此外,产量是由多因素环境因子与自身条件共同作用下形成的,是一个复杂的生物学及其环境适应的过程,也应当考虑冠层结构及纹理信息等对产量估算的贡献和作用,深入探讨多源遥感信息协同下产量模型的估算精度。本试验结果是依据小麦生长前期的生长状况进而估算成熟期产量,生长后期一喷三防及管理措施到位,小麦灌浆未遭受病虫及干旱胁迫的影响,估算效果较好。由于小麦生育后期灾害较多,若防控不及时易发生如病虫等灾害,这对产量估算影响较大,往往导致估产结果偏高。为此,应考虑极端天气及病虫等灾害因子对产量的影响,融入灾害因子模块,进而修正产量估算结果。当然,本文仅涉及一年单点试验数据,确立的 RFR 最佳产量估算还需要经过不同年份和不同区域检验和验证,以增强产量估算模型的可靠性和适用性。同时,生产条件多样,随着特征变量的增加,样本数量的扩大,尤其是土壤、灌溉、施肥等时空变异及其各变量之间复杂性增大,应当考虑深度学习等模型算法的尝试与应用。相较于近地遥感,低空无人机遥感具有成本低、采集速度快及监测面积大等优点,尤其随着土地流转进程加快,农业新型经营主体队伍不断壮大,土地集中规模化经营的背景下,无人机遥感将越来越显示出巨大的应用潜力及实用价值。因此,下一步的研究重点是如何把低空无人机遥感的机动性和近地遥感的细致性以及卫星遥感的宏观性相结合来进行不同尺度的作物产量估算。

4 结论

坡度影响降雨和灌水的田间径流分布,导致水肥二次分配,是造成产量空间差异的重要因素。本文在不同生育时期条件下,利用 MLR、PLSR、SVR 和 RFR 算法将不同模态数据融合进行小麦产量估算,多模态数据融合优于单模态数据。对于单模态的植被指数与温度参数而言,对产量的作用顺序为植被指数>温度参数,以灌浆期植被指数的 RFR 模型效果最好;对于多模态数据来说,RFR 模型将植被指数(NDVI、GNDVI、EVI2、OSAVI、SAVI、NDRE、WDRVI)、温度参数(NRCT、CTD)和坡度信息进行融合,估算效果最佳,且以开花期最为有效。在本试验条件下,随机森林方法可以提升经典机器学习模型对作物产量的估计精度和稳定性,并且通过多模

态遥感数据融合增强模型的准确度和适用性。

References

- [1] He L, Asseng S, Zhao G, Wu D R, Yang X Y, Zhuang W, Jin N, Yu Q. Impacts of recent climate warming, cultivar changes, and crop management on winter wheat phenology across the Loess Plateau of China. *Agric Forest Meteorol*, 2015, 200: 135–143.
- [2] Mueller N, Gerber J S, Johnston M, Ray D K, Ramankutty N, Foley J A. Closing yield gaps through nutrient and water management. *Nature*, 2012, 490: 254–257.
- [3] 杨绚, 汤绪, 陈葆德, 田展, 赵思健. 利用 CMIP5 多模式集合模拟气候变化对中国小麦产量的影响. *中国农业科学*, 2014, 47: 3009–3024.
Yang X, Tang X, Chen B D, Tian Z, Zhao S J. Impacts of climate change on wheat yield in China simulated by CMIP5 multi-model ensemble projections. *Sci Agric Sin*, 2014, 47: 3009–3024 (in Chinese with English abstract).
- [4] Reshef I B, Vermote E, Lindeman M, Justice C. A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sens Environ*, 2010, 114: 1312–1323.
- [5] Zaman-Allah M, Vergara O, Araus J L, Tarekegne A, Magorokosho C, Zarco-Tejada P J, Hornero A, Albà A H, Das B, Craufurd P, Olsen M, Prasanna B M, Cairns J. Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. *Plant Methods*, 2015, 11: 35.
- [6] Wang L G, Tian Y G, Yao X, Zhu Y, Cao W X. Predicting grain yield and protein content in wheat by fusing multisensory and multi-temporal remote-sensing images. *Field Crops Res*, 2014, 164: 178–188.
- [7] 纪景纯, 赵原, 邹晓娟, 宣可凡, 王伟鹏, 刘建立, 李晓鹏. 无人机遥感在农田信息监测中的应用进展. *土壤学报*, 2019, 56: 773–784.
Ji J C, Zhao Y, Zou X J, Xuan K F, Wang W P, Liu J L, Li X P. Advancement in application of UAV remote sensing to monitoring of farmlands. *Acta Pedol Sin*, 2019, 56: 773–784 (in Chinese with English abstract).
- [8] 胡健波, 张健. 无人机遥感在生态学中的应用进展. *生态学报*, 2018, 38: 20–30.
Hu J B, Zhang J. Unmanned aerial vehicle remote sensing in ecology: advances and prospects. *Acta Ecol Sin*, 2018, 38: 20–30 (in Chinese with English abstract).
- [9] 孙刚, 黄文江, 陈鹏飞, 高帅, 王秀. 轻小型无人机多光谱遥感技术应用进展. *农业机械学报*, 2018, 49(3): 1–17.
Sun G, Huang W J, Chen P F, Gao S, Wang X. Advances in UAV-based multispectral remote sensing applications. *Trans CSAM*, 2018, 49(3): 1–17 (in Chinese with English abstract).
- [10] 裴浩杰, 冯海宽, 李长春, 金秀良, 李振海, 杨贵军. 基于综合指标的冬小麦长势无人机遥感监测. *农业工程学报*, 2017, 33(20): 74–82.
Pei H J, Feng H K, Li C C, Jin X L, Li Z H, Yang Z J. Remote sensing monitoring of winter wheat growth with UAV based on comprehensive index. *Trans CSAE*, 2017, 33(20): 74–82 (in Chinese with English abstract).
- [11] Navarro A, Young M, Allan B, Carnell P, Macreadie P,

- Ierodiaconou D. The application of unmanned aerial vehicles (UAVs) to estimate aboveground biomass of mangrove ecosystems. *Remote Sens Environ*, 2020, 242: 111747.
- [12] Zhang C H, Kovacs J M. The application of small unmanned aerial systems for precision agriculture: a review. *Precis Agric*, 2012, 13: 693–712.
- [13] 朱婉雪, 李仕冀, 张旭博, 李洋, 孙志刚. 基于无人机遥感植被指数优选的田块尺度冬小麦估产. *农业工程学报*, 2018, 34(11): 78–86.
Zhu W X, Li S J, Zhang X B, Li Y, Sun Z G. Estimation of winter wheat yield using optimal vegetation indices from unmanned aerial vehicle remote sensing. *Trans CSAE*, 2018, 34(11): 78–86 (in Chinese with English abstract).
- [14] 程千, 徐洪刚, 曹引波, 段福义, 陈震. 基于无人机多时相植被指数的冬小麦产量估测. *农业机械学报*, 2021, 52(3): 160–167.
Cheng Q, Xu H G, Cao Y B, Duan F Y, Chen Z. Grain yield prediction of winter wheat using multi-temporal UAV based multispectral vegetation index. *Trans CSAM*, 2021, 52(3): 160–167 (in Chinese with English abstract).
- [15] 刘畅, 杨贵军, 李振海, 汤伏全, 王建雯, 张春兰, 张丽妍. 融合无人机光谱信息与纹理信息的冬小麦生物量估测. *中国农业科学*, 2018, 51: 3060–3073.
Liu C, Yang G J, Li Z H, Tang F Q, Wang J W, Zhang C L, Zhang L Y. Biomass estimation in winter wheat by UAV spectral information and texture information fusion. *Sci Agric Sin*, 2018, 51: 3060–3073 (in Chinese with English abstract).
- [16] Fu Y Y, Yang G J, Song X Y, Li Z H, Xu X G, Feng H K, Zhao C J. Improved estimation of winter wheat aboveground biomass using multiscale textures extracted from UAV-based digital images and hyperspectral feature analysis. *Remote Sens*, 2021, 13: 581.
- [17] Elsayed S, Elhoweity M, Ibrahim H H, Dewir Y H, Migdadi H M, Schmidhalter U. Thermal imaging and passive reflectance sensing to estimate the water status and grain yield of wheat under different irrigation regimes. *Agric Water Manage*, 2017, 189: 98–110.
- [18] Rischbeck P, Elsayed S, Mistele B, Barmer G, Heil K, Schmidhalter U. Data fusion of spectral, thermal and canopy height parameters for improved yield prediction of drought stressed spring barley. *Eur J Agron*, 2016, 78: 44–59.
- [19] Maimaitijiang M, Sagan V, Sidike P, Hartling S, Esposito F, Fritschi F B. Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sens Environ*, 2020, 237: 111599.
- [20] Riihimäki H, Heiskanen J, Luoto M. The effect of topography on arctic-alpine aboveground biomass and NDVI patterns. *Int J Appl Earth Obs*, 2017, 56: 44–53.
- [21] Nie X D, Guo W, Huang B, Zhuo M N, Li D Q, Li Z W, Yuan Z J. Effects of soil properties, topography and landform on the understory biomass of a pine forest in a subtropical hilly region. *Catena* (Giessen), 2019, 176: 104–111.
- [22] Yang Q Y, Jiang Z C, Li W J, Li H. Prediction of soil organic matter in peak-cluster depression region using kriging and terrain indices. *Soil Tillage Res*, 2014, 144: 126–132.
- [23] 张新乐, 徐梦园, 刘焕军, 孟令华, 邱政超, 潘越, 谢雅慧. 引入地形因子的黑土区大豆干生物量遥感反演模型及验证. *农业工程学报*, 2017, 33(16): 168–173.
Zhang X L, Xu M Y, Liu H J, Meng L H, Qiu Z C, Pan Y, Xie Y H. Remote sensing inversion models and validation of aboveground biomass in soybean with introduction of terrain factors in black soil area. *Trans CSAE*, 2017, 33(16): 168–173 (in Chinese with English abstract).
- [24] 侯淑涛, 崔杨, 孟令华, 武丹茜, 钱蕾, 鲍依临, 叶强, 刘焕军. 地形对漫川漫岗黑土区大豆产量的影响. *农业工程学报*, 2020, 36(8): 88–95.
Hou S T, Cui Y, Meng L H, Wu D Q, Qian L, Bao Y L, Ye Q, Liu H J. Effects of terrain on soybean yields in rolling hilly black soil areas. *Trans CSAE*, 2020, 36(8): 88–95 (in Chinese with English abstract).
- [25] Qi J, Kerr Y, Moran M, Weltz M, Huete A, Sorooshian S, Bryant R. Leaf area index estimates using remotely sensed data and BRDF models in a semiarid region. *Remote Sens Environ*, 2000, 73: 18–30.
- [26] Anatoly A G, Mark N M. Remote sensing of chlorophyll concentration in higher plant leaves. *Adv Space Res*, 1998, 22: 689–692.
- [27] Jiang Z, Huete A R, Didan K, Miurab T. Development of a two-band enhanced vegetation index without a blue band. *Remote Sense Environ*, 2008, 112: 3833–3845.
- [28] Geneviève R, Michael S, Frédéric B. Optimization of soil-adjusted vegetation indices. *Remote Sense Environ*, 1996, 55: 95–107.
- [29] Huete A R. Soil-adjust vegetation index (SAVI). *Remote Sense Environ*, 1988, 25: 295–309.
- [30] Githlson A, Merzlyak M N. Spectral reflectance changes associated with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. leaves. Spectral features and relation to chlorophyll estimation. *J Plant Physiol*, 1994, 143: 286–292.
- [31] Gitelson A A. Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. *J Plant Physiol*, 2004, 161: 165–173.
- [32] Chen J M. Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Can J Remote Sens*, 2014, 22: 229–242.
- [33] Daughtry C S T, Walthall C L, Kim M S, Colstoun E B D, McMurtrey J E. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sense Environ*, 2000, 74: 229–239.
- [34] Tucker C J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sense Environ*, 1979, 8: 127–150.
- [35] Elsayed S, Rischbeck P, Schmidhalter U. Comparing the performance of active and passive reflectance sensors to assess the normalized relative canopy temperature and grain yield of drought-stressed barley cultivars. *Field Crops Res*, 2015, 177: 148–160.
- [36] Purushothaman R, Thudi M, Krishnamurthy L, Upadhyaya H D, Kashiwagib J, Gowda C L L, Varshney R K. Association of mid-reproductive stage canopy temperature depression with the molecular markers and grain yields of chickpea (*Cicer arietinum* L.) germplasm under terminal drought. *Field Crops Res*, 2015, 174: 1–11.
- [37] Valtera M, Šamonil, Svoboda M, Janda P. Effects of topography

- and forest stand dynamics on soil morphology in three natural *Picea abies* mountain forests. *Plant Soil*, 2015, 392: 57–69.
- [38] 马雨阳, 官海翔, 杨昊轩, 邵帅, 邵逸群, 刘焕军. 利用地形、土壤和作物信息辅助提高东北漫岗地数字高程模型精度的新方法. *中国农业科学*, 2021, 54: 1715–1727.
Ma Y Y, Guan H X, Yang H X, Shao S, Shao Y Q, Liu H J. A new method to improve the accuracy of digital elevation model in northeast China by using terrain, soil and crop information. *Sci Agric Sin*, 2021, 54: 1715–1727 (in Chinese with English abstract).
- [39] 朱高龙, 柳艺博, 居为民, 陈镜明. 4 种常用植被指数的地形效应评估. *遥感学报*, 2013, 17: 210–234.
Zhu G L, Liu Y B, Ju W M, Chen J M. Evaluation of topographic effects on four commonly used vegetation indices. *J Remote Sens*, 2013, 17: 210–234 (in Chinese with English abstract).
- [40] 刘海丰, 薛达元, 桑卫国. 地形因子对暖温带森林群落物种丰富度-地上生物量关系的影响. *生态环境学报*, 2012, 21: 1403–1407.
Liu H F, Xue D Y, Sang W G. Effect of topographic factors on the relationship between species richness and aboveground biomass in a warm temperate forest. *Ecol Environ Sci*, 2012, 21: 1403–1407 (in Chinese with English abstract).
- [41] Takyu M, Aiba S I, Kitayama K. Changes in biomass, productivity and decomposition along topographical gradients under different geological conditions in tropical lower montane forests on mount Kinabalu, Borneo. *Oecologia*, 2003, 134: 397–404.
- [42] 杨仕曦, 吕广斌, 黄云, 向华辉, 王正银. 九龙坡花椒种植区地形、土壤肥力与花椒产量的关系. *中国生态农业学报*, 2019, 27: 1823–1832.
Yang S X, Lyu G B, Huang Y, Xiang H H, Wang Z Y. Relationships of *Zanthoxylum bungeanum* yield with topography and soil fertility in Jiulongpo area. *Chin J Eco-Agric*, 2019, 27: 1823–1832 (in Chinese with English abstract).
- [43] Kross A, Mcnairn H, Lapen D, Sunohara M, Champagne C. Assessment of rapid eye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *Int J Appl Earth Obs*, 2015, 34: 235–248.
- [44] Guo J X, Tian G L, Zhou Y, Wang M, Ling N, Shen Q R, Guo S W. Evaluation of the grain yield and nitrogen nutrient status of wheat (*Triticum aestivum* L.) using thermal imaging. *Field Crops Res*, 2016, 196: 463–472.
- [45] Hou A, Chen P, Alloatti J, Li D, Mozzoni L, Zhang B, Shi A. Genetic variability of seed sugar content in worldwide soybean germplasm collections. *Crop Sci*, 2009, 49: 903–912.
- [46] Virlet N, Lebourgeois V, Martinez S, Costes E, Labbe S, Regnard J L. Stress indicators based on airborne thermal imagery for field phenotyping a heterogeneous tree population for response to water constraints. *J Exp Bot*, 2014, 65: 5429–5442.
- [47] Du W Y, Zhang L D, Hu Z F, Shamaila Z, Zeng A J, Song J L, Liu Y J, Wolfram S, Joachim M, He X K. Utilization of thermal infrared image for inversion of winter wheat yield and biomass. *Spectr Spectr Anal*, 2011, 31: 1476–1480.
- [48] González F G, Miralles D J, Slafer G A. Wheat floret survival as related to pre-anthesis spike growth. *J Exp Bot*, 2011, 62: 4889–4901.
- [49] Laurent A, Loyce C, Makowski D, Pelzer E. Using site-specific data to estimate energy crop yield. *Environ Modell Softw*, 2015, 74: 104–113.
- [50] Yue J B, Yang G J, Li C C, Li Z H, Wang Y J, Feng H K, Xu B. Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models. *Remote Sens*, 2017, 9: 708.
- [51] 陶惠林, 徐良骥, 冯海宽, 杨贵军, 杨小冬, 苗梦珂, 代阳. 基于无人机数码影像的冬小麦株高和生物量估算. *农业工程学报*, 2019, 35(19): 107–116.
Tao H L, Xu L J, Feng H K, Yang G J, Yang X D, Miao M K, Dai Y. Estimation of plant height and biomass of winter wheat based on UAV digital image. *Trans CSAE*, 2019, 35(19): 107–116 (in Chinese with English abstract).
- [52] 潘治利, 祁萌, 魏春阳, 李锋, 张仕祥, 王建伟, 过伟民, 艾志录. 基于图像处理和支向量机的初烤烟叶颜色特征区域分类. *作物学报*, 2012, 38: 374–379.
Pan Z L, Qi M, Wei C Y, Li F, Zhang S X, Wang J W, Guo W M, Ai Z L. Color region classification of flue-cured tobacco leaves based on the image. *Acta Agron Sin*, 2012, 38: 374–379 (in Chinese with English abstract).
- [53] Xu J X, Ma J, Tang Y N, Wu W X, Shao J H, Wu W B, Wei S Y, Liu Y F, Wang Y C, Guo H Q. Estimation of sugarcane yield using a machine learning approach based on UAV-LiDAR data. *Remote Sens*, 2020, 12: 2823.
- [54] Nawar S, Buddenbaum H, Hill J, Kozak J. Modeling and mapping of soil salinity with reflectance spectroscopy and landsat data using two quantitative methods (PLSR and MARS). *Remote Sens*, 2014, 6: 10813–10834.
- [55] Liu H Z, Shi T Z, Chen Y Y, Wang J J, Fei T, Wu G F. Improving spectral estimation of soil organic carbon content through semi-supervised regression. *Remote Sens*, 2017, 9: 29.
- [56] Aghighi H, Azadbakht M, Ashourloo D, Shahrabi H S, Radiom S. Machine learning regression techniques for the silage maize yield prediction using time-series images of Landsat 8 OLI. *IEEE J-Selec Topic Appl Earth Observ Remote Sens*, 2018, 11: 4563–4577.
- [57] 杨北萍, 陈圣波, 于海洋, 安秦. 基于随机森林回归方法的水稻产量遥感估算. *中国农业学报*, 2020, 25(6): 26–34.
Yang B P, Chen S B, Yu H Y, An Q. Remote sensing estimation of rice yield based on random forest regression method. *J China Agric Univ*, 2020, 25(6): 26–34 (in Chinese with English abstract).
- [58] Jeong J H, Resop J P, Mueller N D, Fleisher D H, Yun K, Butler E E, Timlin D J, Shim K M, Gerber J S, Reddy V R, Kim S H. Random forests for global and regional crop yield predictions. *PLoS One*, 2016, 11: e0156571.