doi:10.6041/j.issn.1000-1298.2025.10.039

基于多时相 GF－1 影像和机器学习的冬小麦估产研究

卢必慧　李卫国　田　苗　王　晶　毛　星
（江苏省农业科学院农业信息研究所，南京 210014）

摘要：为了准确地估算区域冬小麦产量，本研究尝试利用多时相的 GF－1 WFV 影像，采用多种光谱植被指数表征冬小麦的生长特征，对多个时相的不同植被指数与冬小麦产量之间进行相关性分析，基于定量分析结果优选特征，然后利用反向传播神经网络、支持向量回归和随机森林回归 3 种机器学习算法构建冬小麦产量估测模型，并与传统线性回归估产模型进行了比较。结果显示，植被指数与产量的相关性随着生育期的推进逐渐增强，其中 3 月 26 日孕穗期的植被指数与产量相关性最高，达到 0.7。基于多时相植被指数组合的建模方式能够有效提高冬小麦产量的估测精度，采用拔节期和孕穗期的植被指数组合比单个生育期的植被指数的估产精度要高，反向传播神经网络估测模型的决定系数 R^2 提高约 0.104，支持向量回归模型 R^2 提高约 0.141，随机森林回归模型 R^2 提高约 0.107。针对孕穗期的多时相植被指数，3 种模型均有较高的估测精度，反向传播神经网络模型最高估测精度为 91.47%，R^2 为 0.859，RMSE 为 472.873 kg/hm^2，支持向量回归模型估测精度最高为 91.03%，R^2 为 0.826，RMSE 为 492.917 kg/hm^2，随机森林回归模型最高估测精度达到 92.25%，R^2 为 0.908，RMSE 为 445.874 kg/hm^2，3 种机器学习算法的最优精度模型对比多元回归估测模型，精度分别提升 7.53、7.09、8.31 个百分点。该研究为基于国产高分辨率卫星数据进行冬小麦产量估算研究提供了参考。

关键词：冬小麦；产量估测；GF－1；植被指数；机器学习
中图分类号：S127　　**文献标识码**：A　　**文章编号**：1000-1298(2025)10-0448-10　　**OSID**：

Yield Prediction of Winter Wheat Using Multi-temporal GF－1 Images and Machine Learning

LU Bihui　LI Weiguo　TIAN Miao　WANG Jing　MAO Xing
(*Institute of Agricultural Information*, *Jiangsu Academy of Agricultural Sciences*, *Nanjing* 210014, *China*)

Abstract: Crop yield estimation is crucial for formulating agricultural policies and ensuring regional and national food security. Multi-temporal vegetation indices (VIs) extracted from GF－1 WFV images were used to characterize the growth characteristics of winter wheat. A quantitative analysis of the correlation between different VIs across multiple growth periods and winter wheat yield was conducted. Based on the results of correlation analysis, vegetation index features were selected. Three machine learning algorithms, including back propagation neural network (BPNN), support vector regression (SVR), and random forest regression (RFR), were used to construct winter wheat yield prediction models, and accuracy evaluation was conducted. The results showed that the correlation between VIs and yield gradually increased as the growth period advanced, with the highest correlation (up to 0.7) between VIs and yield occurring during the booting period on March 26th. The modeling method based on the combination of multi-temporal VIs effectively improved the estimation accuracy of winter wheat yield. The combination of VIs during the jointing and booting stages produced higher estimation accuracy than the use of VIs at a single growth stage. The BPNN model estimated a determination coefficient R^2 that increased by about 0.104, the SVR model R^2 increased by about 0.141, and the RFR model R^2 increased by about 0.107. For the multi-temporal VIs during the booting stage, all three models demonstrated high accuracy. The highest estimation accuracy of the BPNN model was 91.47% with R^2 of

收稿日期：2024－06－27　修回日期：2024－08－20
基金项目：江苏省农业科技自主创新资金项目(CX(23)3006)
作者简介：卢必慧(1989—)，女，助理研究员，主要从事农情遥感监测研究，E-mail：20140029@jaas.ac.cn
通信作者：李卫国(1967—)，男，研究员，博士，主要从事农情遥感监测研究，E-mail：jaaslwg@126.com

0.859 and RMSE of 472.873 kg/hm^2, while the highest estimation accuracy of the SVR model was 91.03% with R^2 of 0.826 and RMSE of 492.917 kg/hm^2. The highest accuracy of the RFR model was 92.25% with R^2 of 0.908 and RMSE of 445.874 kg/hm^2. Compared with the traditional linear regression model, the accuracy of the optimal models based on three machine learning algorithms improved by 7.53, 7.09, and 8.31 percentage points, respectively. The research results provided a reference for winter wheat yield prediction based on multi-temporal domestic high-resolution satellite data.

Key words: winter wheat; yield prediction; GF－1; vegetation index; machine learning

0　引言

小麦是我国重要的粮食作物之一，及时准确地估算小麦产量对于农业生产管理、粮食安全评估以及农业政策制定等具有重要意义。随着遥感数据可用性的增强以及在空间、时间和光谱分辨率上的提高，遥感技术已被广泛地应用于农作物估产研究。

传统上，农作物遥感估产方法主要包括经验统计方法、半经验半机理方法以及机理方法[1]，在实际应用中各有优势和不足，表现为：①传统线性回归统计模型虽然较为简便、应用最为广泛，但由于作物产量形成通常具有复杂的非线性，而模型无法处理变量之间的非线性关系，因此往往精度不够高且时空泛化能力差[2-3]。②基于过程的作物模型可以模拟作物生长和产量形成的机理过程，与遥感信息同化后可以进行区域尺度产量预测，但模型往往需要大量输入参数且大规模运算非常耗时[4-5]。③以光能利用率模型为代表的半经验半机理模型，相较于机理模型较为简化，然而模型中的关键参数如光能利用率（Light use efficiency，LUE）和收获指数（Harvest index，HI），通常难以在时空分布中进行准确定量模拟[6-7]，但近年来利用荧光遥感如太阳诱导叶绿素荧光（Solar-induced fluorescence，SIF）来估算 LUE，以及利用无人机高光谱数据来估算作物的 HI 等方面的研究工作取得了显著进展[8-11]。基于机器学习（Machine learning，ML）的统计方法，如人工神经网络（Artificial neural networks，ANN）、支持向量机（Support vector machine，SVM）和随机森林（Random forrest，RF）等，因其能够考虑到预测因子与响应变量的非线性关系，相较于传统统计方法显示出较优的性能和更高的准确性。尤其是 RF 对数据集的适应能力较强，通过引入随机性使其不容易陷入过拟合并且具有更好的抗噪能力[12-13]，因而被广泛应用于作物产量预测研究[14-18]。

冬小麦产量形成过程较为复杂，与多光谱植被指数之间存在较为复杂的非线性关系，传统的统计回归模型难以处理这种非线性关系，同时大多是基于单一时相光谱植被指数对产量的拟合，虽然具有较好的拟合水平，但这种建模方式难以反映产量形成的过程。由多个生育期构成的数据集所得到的生长信息能更好地反映作物生长变化过程，有利于提高产量的估测精度[12,19]。另外，现有研究以使用中高分辨率（如 Landsat 和 Sentinel－2）以及中低分辨率（MODIS 等）遥感数据的大尺度产量估测为主[20-23]，而利用国产高分辨率卫星影像进行冬小麦产量估算的研究比较少。因此，本文以江苏省常熟市为研究区域，利用冬小麦主要生育期的 GF－1 WFV 卫星影像，通过比较不同光谱植被指数对产量的敏感性进行特征优选，基于 BP 神经网络、支持向量回归和随机森林回归 3 种机器学习算法构建冬小麦产量估测模型，以期提高估产精度，并显示国产卫星影像在农作物产量估测上的应用效果。

1　材料与方法

1.1　研究区概况

常熟市位于江苏省南部，隶属于苏州市，境内地势低平，大部分地区海拔为 3～7 m，市域总面积约为 1 276 km^2。地处温带，属亚热带季风性海洋气候，四季分明，气候温和，雨量充沛。年平均气温约 16.9℃，年平均降水量约 1 136 mm，年平均日照时数约为 1 814 h，年日照百分率约 39%。全市耕地面积约 5.85×10^4 hm^2，主要粮食作物轮作方式为冬小麦和水稻一年两熟，其中冬小麦是研究区的主要夏粮作物，种植时间为上年的 11 月上中旬至次年 6 月初，主要生育期为返青（2 月中下旬）、拔节（3 月上中旬）、孕穗（3 月下旬）、抽穗开花（4 月上中旬）和灌浆期（4 月底到 5 月中旬）。全市播种面积近年来稳定在 1.73×10^4 hm^{2}左右，在苏州市范围内体量最大，约占到 30%。

1.2　数据获取与预处理

1.2.1　遥感影像获取

本文以 GF－1 WFV 为主要数据源，查询了研究区 2023 年 3—5 月中旬冬小麦关键生育期（拔节期至灌浆期）的遥感影像，筛选高质量无云影像 6 景，影像信息见表 1。GF－1 WFV 影像数据首先采用中国资源卫星应用中心公布的辐射定标系数，利用 ENVI 5.3 软件进行辐射定标，然后利用 FLAASH 模块进行大气校正。几何校正采用同区域 Landsat－9

表 1 卫星遥感影像数据信息

Tab. 1 Data information on satellite remote sensing images

序号	卫星/传感器/轨道号	产品序列号	空间分辨率/m	成像时间	所在生育期	用途
1	GF－1/WFV	7156613	16	2023－03－09	拔节期	植被指数提取
2	GF－1/WFV	7168535		2023－03－26	孕穗期	
3	GF－1/WFV	7217230		2023－04－10	抽穗期	
4	GF－1/WFV	7248291		2023－04－27	开花期	
5	GF－1/WFV	7256627		2023－05－01	乳熟期	
6	GF－1/WFV	7278220		2023－05－13	蜡熟期	
7	Landsat－9/OLI－2/Pan	LC09_L1TP_119038_20230402_20230402_02_T1	15	2023－04－02	抽穗期	几何校正参考影像
8	Sentinel－2A/MSI	S2A_MSIL2A_20230408T023531_N0509_R089_T51STR	10/20	2023－04－08	抽穗期	冬小麦面积提取

全色波段作为参考影像，控制校正误差在 1 个像元之内。预处理后影像用于冬小麦植被指数的计算。

1.2.2 冬小麦面积提取

为了获取较为精确的冬小麦种植空间分布，利用研究区 2023 年 4 月 8 日的 Sentinel－2 卫星影像和地面采集的 153 个解译标志点（包括冬小麦 82 个）以及 225 个地面验证点（包括冬小麦 164 个），采用支持向量机分类方法提取常熟市冬小麦种植面积，并采用混淆矩阵方法对分类结果进行评价。2023 年常熟市冬小麦种植空间分布如图 1 所示，冬小麦总分类面积为 1.959×10^4 hm^2，总体精度为 95.66%，Kappa 系数 0.87，满足后续估产研究的精度要求。

1.2.3 地面测产数据

为建立和验证估产模型，基于 Sentinel－2 影像自动提取的常熟市冬小麦种植范围，结合基于 4 月 10 日的 GF－1WFV 影像提取的冬小麦 NDVI，在研究区范围内均匀布设了 40 个地面采样点，如图 1 所示。地面采样工作于 2023 年 5 月 25 至 26 日在冬小麦收获前进行。测产时，准确定位和记录采样点经、纬度坐标，在每个采样地块内选择 3 个 1 m^2 区域齐穗割取，装入采样网袋带回，后期对所有样品进行脱粒和风干后称量，并测量含水率。最后，根据测量结果计算所有样点的冬小麦平均单产。

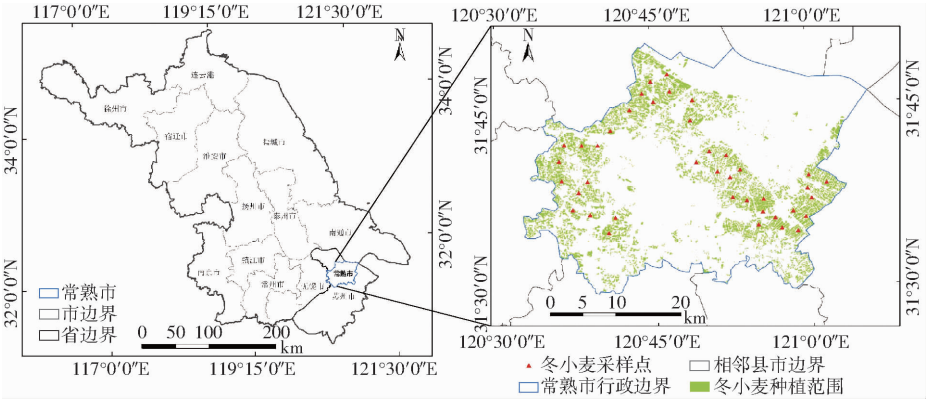


图 1 研究区地理位置及冬小麦采样点分布图

Fig. 1 Location of study area and sampling sites of winter wheat

1.3 植被指数计算

植被指数是当前基于统计模型方法的农作物监测研究中应用较为广泛的遥感变量，能够反映不同植被类型的物候差异和光谱特征[24－28]。根据冬小麦自身光谱特性，参考国内外相关的研究文献，并结合研究区特点，选取 10 种普适性较强的光谱植被指数作为构建冬小麦遥感估产模型的入选自变量，包括：①基于红波段和近红外波段的植被指数：归一化植被指数[29]（Normalized difference vegetation index，NDVI）、增强植被指数[30]（Enhanced vegetation index，EVI）、比率植被指数[31]（Ratio vegetation index，RVI）、差异植被指数[32]（Difference vegetation index，DVI）、植被衰减指数[33]（Plant senescence reflectance index，PSRI）和结构不敏感色素指数[34]（Structure insensitive pigment index，SIPI）。②优化土壤背景的红波段和近红外波段运算的植被指数：土壤校正植被指数[35]（Soil adjusted vegetation index，SAVI）和优化土壤校正植被指数[36]（Optimal soil adjusted vegetation index，OSAVI）。③基于绿波段和红或近红外波段的植被指数：绿色归一化植被

指数[37]（Green normalized difference vegetation index，GNDVI）和作物氮反射指数[38]（Nitrogen reflectance index，NRI）。

1.4　模型构建与精度评价

1.4.1　模型构建

计算各个时相的植被指数，分析各植被指数与冬小麦产量的相关性，优选相关性较高的植被指数；在特征优选的基础上，利用多元逐步回归和 3 种机器学习算法包括反向传播神经网络（Back propagation neural network，BPNN）、支持向量回归（Support vector regression，SVR）和随机森林回归（Random forest regression，RFR）来建立冬小麦产量估测模型，并对比机器学习算法模型与传统线性回归模型的估测精度。本研究的机器学习算法在 Matlab 环境下运行，其中 BPNN 模型采用 Levenberg－Marquardt（L－M）算法优化神经元和隐含层的数量来训练神经网络；SVR 模型是基于 LIBSVM3.31[39]工具包，选用径向基核函数（Radial basis function，RBF），采用网格搜索和 n-fold 交叉验证的方法进行参数寻优，以确定惩罚因子 C 和参数 σ 值。当不同的 C 和 σ 都对应最高的精度时，将参数 C 最小的作为最佳参数组合，以避免惩罚因子太高造成过拟合状态。

1.4.2　精度评价

为充分利用采样数据，采用 3 折交叉验证法来评价模型精度，将实测产量数据分为 3 组，其中有 2 组包含 13 个样本，1 组包含 14 个样本。模型训练时，将其中 2 组数据为训练集，剩余 1 组为验证集，进行 3 次迭代，保证所有样本均能参与建模和验证。模型的稳定性和预测精度采用决定系数 R^2、均方根误差（Root mean square error，RMSE）、平均绝对误差（Mean absolute error，MAE）和估算精度（Accuracy）来评价。R^2 越接近 1，表明模型拟合精度越好；RMSE 越小，Accuracy 越大，表示估算值与实测值差异越小，模型预测能力越强且精度高。

2　结果与分析

2.1　不同光谱植被指数对冬小麦产量的敏感性

通过对冬小麦关键生育期内不同时相的光谱植被指数与 40 个冬小麦测产地块的实测产量进行相关性分析，结果显示：除了 5 月 1 日和 5 月 13 日的少数植被指数外，冬小麦产量与各植被指数总体上均呈显著相关（$P<0.01$），但不同时相的植被指数与实测产量之间的相关程度有所不同。从图 2 可以看出，3 月 26 日各植被指数与实测产量相关性最高，相关系数绝对值为 0.670～0.710，其中 NDVI 相关性最高；其次是 4 月 10 日的相关性较高。2023 年常熟市冬小麦从 2 月下旬进入返青期，受气象条件和植被覆盖度影响，相关性较低；3 月上旬冬小麦进入拔节期后，植被指数与产量的相关性开始增大，随着生育期推进，生物量逐渐增加，到 3 月下旬孕穗期相关性达到最大。4 月底冬小麦进入灌浆期后，植被指数逐渐下降，与产量的相关性减小。总体来说，3 月中旬至 4 月上旬的 3 个时期植被指数对冬小麦产量敏感性较强。

同时，各个植被指数对实测产量变化也表现出不同的敏感性。以 3 月 26 日为例，可以看出，绝大多数的植被指数与实测产量之间呈显著相关，其中 NDVI、GNDVI、EVI、OSAVI 和 SAVI 与产量间相关系数均超过 0.7，表明这 5 种植被指数对产量敏感性较强。通过进一步拟合分析这 8 种植被指数与冬小麦产量之间的关系（图 3 和表 2），发现 GNDVI、RVI、NDVI、OSAVI 和 EVI 5 种植被指数拟合的模型效果较优。可见本文选取的多光谱植被指数与冬小麦产量有较强的相关性，能够用来估测冬小麦产量。

2.2　基于多元逐步回归的冬小麦产量估测

根据相关性分析结果，选取 5 个时相（0309～0501）和 5 种植被指数（NDVI、GNDVI、EVI、RVI 和 OSAVI）作为输入变量，采用逐步回归方法建立冬小麦产量估测的多元回归模型，如表 3 所示。根据 R^2 和 RMSE，模型 3 是最佳回归模型，采用三折交叉验证对模型进行精度验证。图 4 表明，冬小麦估算产量与实测产量之间具有相对较高的拟合度，R^2 为 0.668，预测误差为 766.428 kg/hm^2，预测精度为 83.94%。

2.3　基于不同机器学习算法的冬小麦产量估测

选取 5 个时相（0309、0326、0410、0427 和 0501）和 5 种植被指数（NDVI、GNDVI、EVI、RVI 和 OSAVI），利用 BPNN、SVR 和 RFR 3 种机器学习算法建立冬小麦产量估测模型。对于单时相时（如 0309、0326 等），将所选取的 5 种植被指数分别作为自变量的一个维度，形成自变量向量，即每个样本数据有 5 个属性值；对于多时相植被指数，例如 0309＋0326，则将第 2 个时相（0326）的植被指数合并到第 1 个时相（0309）的自变量向量中，形成多时相植被指数数据集；以此类推，每增加 1 个时相，自变量向量增加 5 个维度。最终建模结果如表 4 所示。

由表 4 和图 5 可知，对于单时相植被指数建模估产，从拔节期（0309）到孕穗期（0326），随着冬小麦的生长，3 种机器学习算法的产量估测精度均不断提高，其中 0326（孕穗期）估测精度最高，R^2 均超

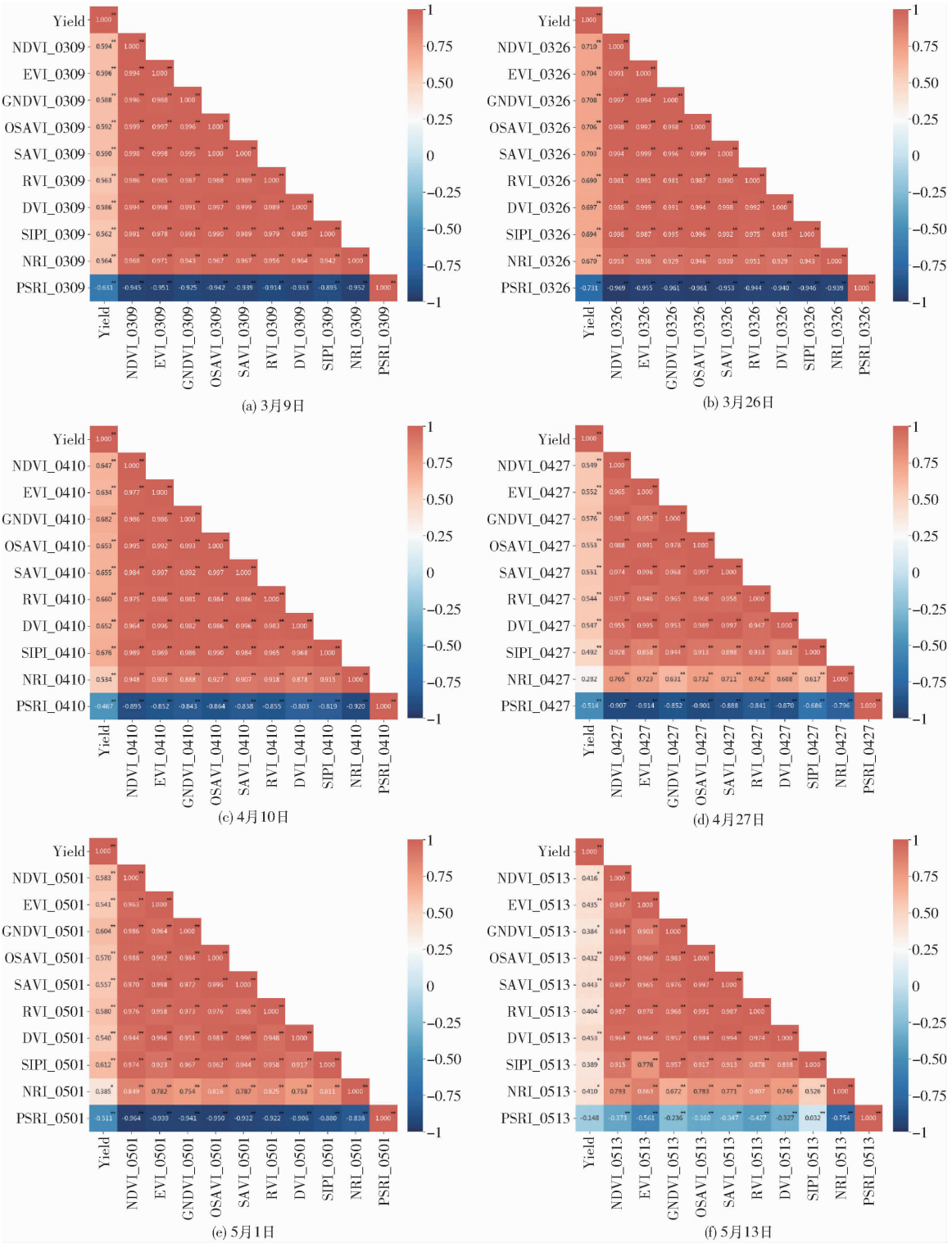


图 2 不同时相植被指数之间以及植被指数与产量的相关系数

Fig. 2 Correlation coefficients between vegetation indices and yield at different growth stages

过 0.8，RFR 模型 R^2 最高为 0.899，RMSE 为 441.877 kg/hm^2；随着生育期发展，从开花期(0410)到灌浆期(0427 和 0501)，冬小麦产量估测精度逐渐下降，这与植被指数和产量相关性的变化规律基本一致，其中 0501 估测精度最低，3 种模型 R^2 分别为 0.640、0.641 和 0.686。

对于多生育期植被指数建模，使用多时相植被指数(如生育期 0309 + 0326)比单个时相(0309 或 0326)植被指数估产精度高，BPNN 模型估测精度 R^2 分别提高约 0.104、0.010；SVR 模型 R^2 分别提高约 0.141、0.002；RFR 模型 R^2 分别提高约 0.107、0.008。生育期(0309 + 0326 + 0410)的多时相植被指数对比单个时相(0309 或 0410)的植被指数的估产精度均有提高，BPNN 模型 R^2 分别提高约 0.127、

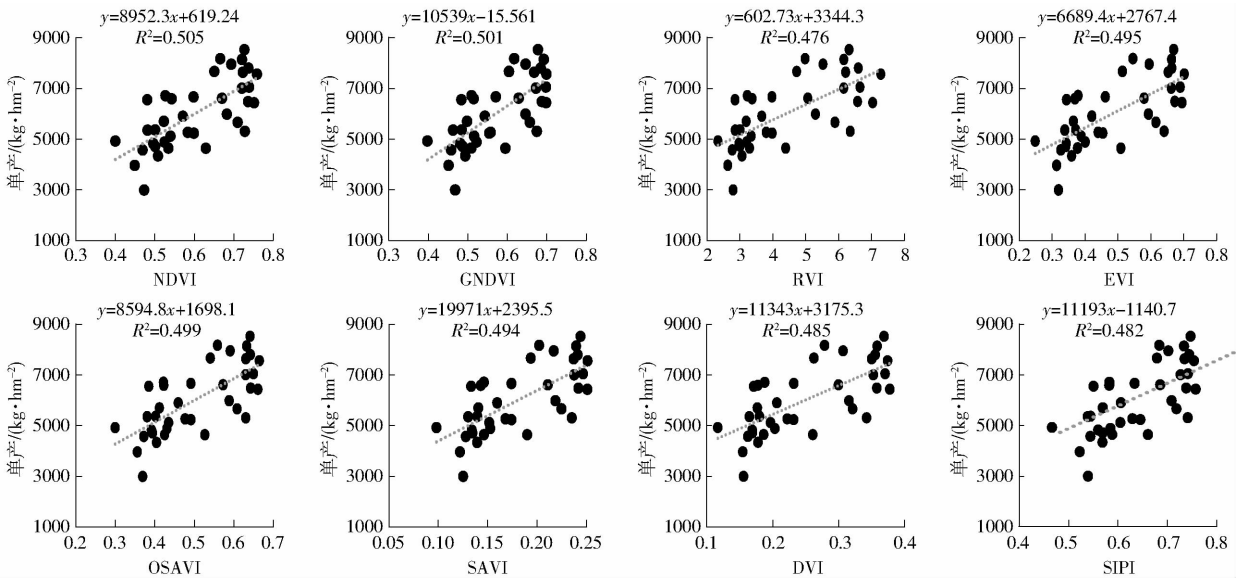


图 3　冬小麦孕穗期(3 月 26 日)植被指数与实测单产的拟合分析

Fig. 3　Fitting analysis of vegetation indices and winter wheat yield during booting stage

表 2　各时期不同植被指数与产量决定系数

Tab. 2　Coefficients of determination between vegetation indices and yield at different growth stages

多光谱	冬小麦不同生长时期					
植被指数	0309	0326	0410	0427	0501	0513
NDVI	0. 352	0. 505	0. 419	0. 301	0. 340	0. 173
GNDVI	0. 346	0. 501	0. 465	0. 332	0. 365	0. 147
RVI	0. 373	0. 476	0. 438	0. 306	0. 357	0. 185
EVI	0. 356	0. 495	0. 401	0. 305	0. 293	0. 189
OSAVI	0. 351	0. 499	0. 427	0. 306	0. 325	0. 187
SAVI	0. 349	0. 494	0. 429	0. 304	0. 310	0. 180
DVI	0. 376	0. 485	0. 426	0. 301	0. 302	0. 222
SIPI	0. 316	0. 482	0. 457	0. 242	0. 375	0. 151
NRI	0. 460	0. 449	0. 286	0. 079	0. 148	0. 168
PSRI	0. 398	0. 452	0. 219	0. 264	0. 260	0. 022

0. 108，SVR 模型 R^2 分别提高约 0. 133、0. 056，RFR 模型 R^2 分别提高约 0. 031、0. 021；但对比单时相(0326)植被指数的估产精度有所下降，其中 SVR 模型 R^2 下降约 0. 006，RFR 模型 R^2 下降约 0. 067。同样地，生育期(0309 + 0326 + 0410 + 0427)对比单个时相的植被指数的估测精度均有不同程度地下降。由此可以看出，采用灌浆期(0427)之前的多时相植被指数能够较好地提高估产精度，孕穗期(0326)的植被指数与产量相关性更好，估测精度更高。

基于本研究的样本数据集，对比这 3 种机器学习算法，当时相增加、数据属性值增多时，BPNN 和 SVR 模型估测能力低于 RFR 模型，并且 RFR 模型针对整个生育期的不同时相植被指数组合的数据集，总体上均表现出较高的估测精度。BPNN 模型针对生育期(0309 + 0326 + 0410)数据集估测精度最高，达到 91. 47%，R^2 为 0. 859，RMSE 为 472. 873 kg/hm^2；而 SVR 和 RFR 模型针对生育期(0309 + 0326)数据集估测精度最高，其中 SVR 模型预测精度为 91. 03%，R^2 为 0. 826，RMSE 为 492. 917 kg/hm^2，RFR 模型预测精度为 92. 25%，R^2 为 0. 908，RMSE 为 445. 874 kg/hm^2。将 3 种机器学习算法的最优精度模型与基于多元逐步回归的估测模型进行对比，估测精度分别提升 7. 53、7. 09、8. 31

表 3　不同线性回归估产模型方程

Tab. 3　Linear regression equations for different yield estimation models

估测模型	预测变量	模型方程	R^2	RMSE/(kg·hm^{-2})
模型 1	NDVI_0326	$y = 8\,952.287x_1 + 619.235$	0. 491	934. 689
模型 2	NDVI_0326，RVI_0309	$y = 19\,770.585x_1 - 1\,156.362x_2 - 2\,402.285$	0. 578	839. 367
模型 3	NDVI_0326，RVI_0309，EVI_0326	$y = -5\,092.075x_1 - 1\,980.161x_2 + 24\,783.924x_3 + 3\,042.972$	0. 646	767. 190
模型 4	RVI_0309，EVI_0326	$y = -1\,895.423x_2 + 20\,363.734x_3 + 1\,863.402$	0. 638	769. 836

注：表中 NDVI_0326、RVI_0309、EVI_0326 分别用变量 x_1、x_2、x_3 表示。

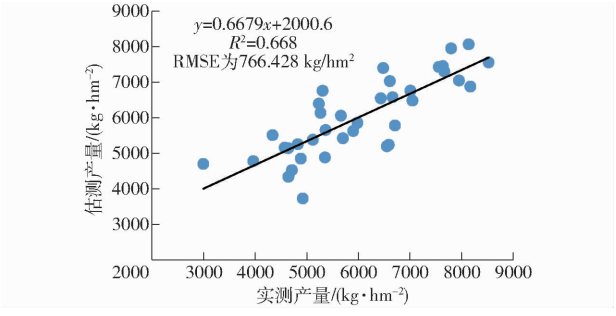


图 4　冬小麦估测与实测产量拟合精度验证

Fig. 4　Validation of fitting accuracy for winter wheat estimated and measured yields

表 4 不同机器学习算法估产模型统计结果

Tab. 4 Statistical results of yield prediction models using different machine learning methods

输入特征（生育期）	BPNN			SVR			RFR		
	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})
0309	0.731	576.905	453.534	0.685	594.469	511.252	0.801	479.261	389.589
0326	0.825	473.772	388.943	0.824	567.782	429.923	0.899	441.877	391.500
0309+0326	0.835	528.118	444.810	0.826	492.917	422.479	0.908	445.874	374.263
0410	0.750	558.553	427.600	0.762	632.285	536.698	0.811	486.611	392.580
0309+0410	0.719	589.685	486.075	0.742	539.787	385.474	0.769	543.107	455.407
0326+0410	0.849	557.346	494.595	0.808	473.361	375.541	0.839	490.004	398.532
0309+0326+0410	0.859	472.783	376.941	0.818	462.786	425.388	0.832	424.666	343.726
0427	0.677	626.039	508.036	0.664	687.763	524.775	0.769	527.932	451.482
0309+0427	0.750	577.502	485.539	0.774	555.211	468.360	0.758	527.964	417.225
0326+0427	0.730	561.374	456.756	0.711	533.119	424.991	0.778	585.936	491.321
0410+0427	0.691	566.666	502.999	0.634	579.661	471.255	0.679	576.966	504.645
0309+0326+0427	0.790	497.733	459.350	0.796	484.163	416.056	0.802	479.828	431.887
0309+0410+0427	0.799	561.255	464.797	0.750	605.060	519.472	0.772	454.926	422.833
0326+0410+0427	0.730	557.023	398.739	0.752	531.678	440.475	0.795	518.897	441.404
0309+0326+0410+0427	0.711	563.284	467.600	0.792	546.464	467.567	0.835	537.834	462.620
0501	0.640	641.833	501.580	0.641	684.646	565.996	0.686	593.778	527.644
0309+0501	0.738	562.077	418.820	0.689	664.114	546.967	0.756	561.098	429.553
0326+0501	0.693	594.729	475.132	0.714	567.412	480.638	0.725	488.589	424.339
0410+0501	0.672	625.547	561.264	0.676	614.443	564.391	0.685	595.822	513.166
0427+0501	0.673	569.924	480.198	0.605	564.185	455.536	0.749	523.983	439.609
0309+0326+0501	0.770	590.380	453.215	0.775	577.149	506.155	0.779	538.221	480.349
0309+0410+0501	0.769	572.254	427.955	0.681	582.116	479.827	0.789	530.192	440.533
0326+0410+0501	0.742	547.976	494.888	0.752	549.666	453.924	0.786	538.673	445.551
0309+0427+0501	0.747	591.275	482.955	0.716	612.607	535.260	0.742	592.775	483.493
0326+0427+0501	0.770	503.631	384.875	0.793	523.591	409.712	0.791	542.280	482.873
0410+0427+0501	0.619	645.399	541.114	0.604	685.240	528.607	0.721	524.540	438.521
0309+0326+0410+0501	0.791	577.393	484.411	0.798	564.230	425.222	0.797	526.423	438.258
0309+0326+0427+0501	0.793	527.679	400.710	0.759	577.062	469.206	0.763	544.907	479.739
0309+0410+0427+0501	0.644	608.440	545.198	0.621	668.174	540.128	0.706	565.530	467.500
0326+0410+0427+0501	0.672	585.373	471.280	0.619	643.053	581.097	0.719	520.903	428.943
0309+0326+0410+0427+0501	0.665	572.276	484.198	0.652	573.014	462.508	0.670	571.839	440.487

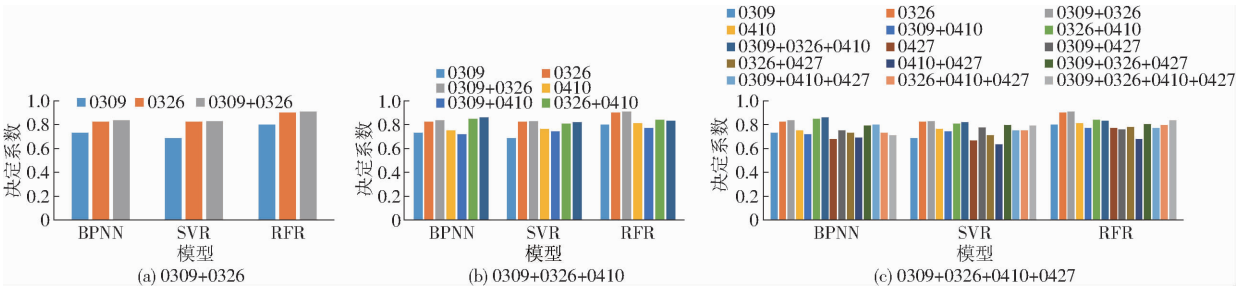


图 5 针对不同时相植被指数组合的回归模型产量估测决定系数

Fig. 5 R^2 of regression models using combinations of vegetation indices across different growth periods

个百分点。基于机器学习的各自最优精度模型所预测的冬小麦产量空间分布如图 6 所示。

3 讨论

（1）针对基于单一时相光谱植被指数的传统统计回归模型无法处理产量与多光谱植被指数之间的非线性关系，估产建模方式难以反映产量形成的过程，本文基于多时相 GF－1 卫星影像，选取了 10 种植被指数，分析了不同时相的各植被指数与产量的定量关系，通过相关性分析优选模型的输入特征，利

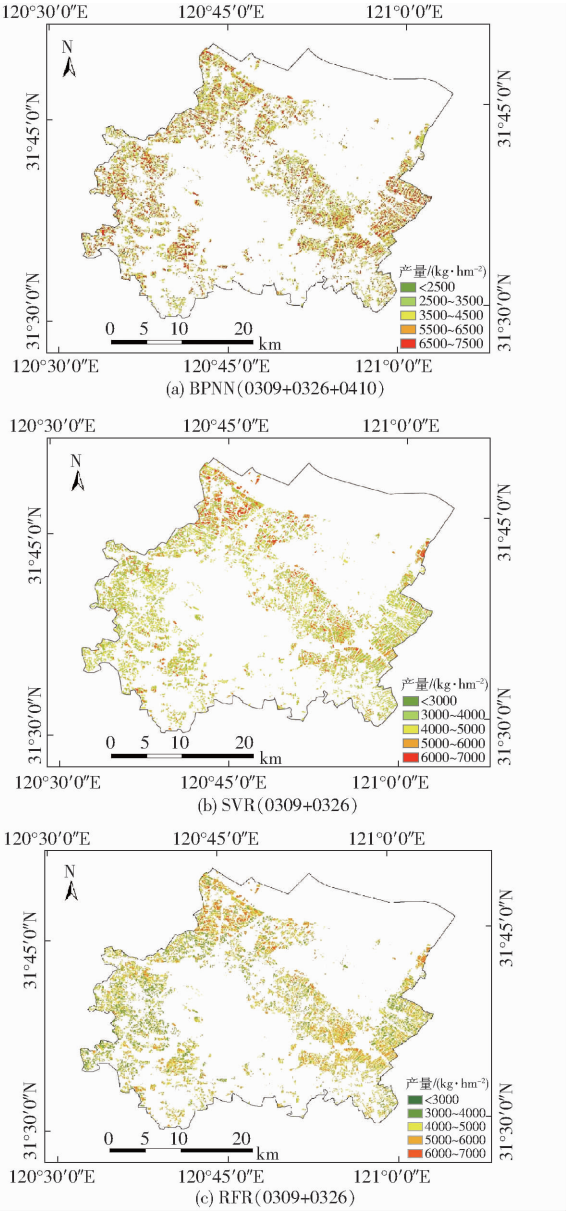


图 6　基于 3 种机器学习算法预测的冬小麦产量空间分布

Fig. 6　Spatial distribution of winter wheat yield predicted by three machine learning methods

用 BPNN、SVR 和 RFR 3 种机器学习算法，建立了不同形式的多时相植被指数组合的产量估测模型。通过将新生育期的植被指数作为新的属性值，添加到上一个生育期的数据集中，形成新的多时相植被指数数据集，以增加样本数据属性值的维度，提升产量估测模型的训练，从而提高了冬小麦产量估测精度，这与以往的研究结论一致[12,19]，同时对比传统的线性回归模型，估产精度最高提升 8.31 个百分点。

（2）根据本研究的样本数据集，所建立的不同形式的多时相植被指数产量估测模型，对比发现，针对孕穗期的多时相植被指数的模型估测精度较高，这与以往的冬小麦估产研究所使用的最佳生育期模型（如开花期和灌浆期）有所不同[12,26,40－42]，这种不一致性可能是与研究区的气候和环境条件差异、冬小麦种植品种和管理方式的不同以及所采用的遥感数据源和建模方法等均有一定的关系。尤其是气候环境条件和种植管理方式的不同会导致冬小麦生长动态上的不同和物候上的空间差异，导致某个生育期的数据对产量预测更加关键。就本研究区而言，位于苏南地区的常熟市冬小麦的物候期相较于苏北地区提前，对比中国北方地区冬小麦的生育期物候差异更为明显。此外，所使用的遥感数据源和方法也各不相同，以上这些因素均在不同程度上影响到最终的估产模型。

（3）基于 3 种机器学习算法所建立的模型中，RFR 模型表现出较高的估测精度，具有更好的抗噪能力。但是考虑到本文仅用了一年的观测数据，训练样本有限，可能会导致模型的不稳定性和泛化能力不足等问题，在后续的工作中应获取多年的地面观测以提高模型的稳定性；同时，当样本数据量过大或特征过多时，模型的预测能力将下降，或将考虑训练性能更为优越的机器学习模型（如基于梯度提升的决策树模型 LightGBM 和 XGBoost）或者深度学习模型。在模型的输入特征上，本文只选取了植被指数，未能考虑到环境变量（如气象数据和土壤属性等）。另外，与基于反射率的植被指数相比，一些能潜在提升产量预测结果的卫星数据，例如 SIF 已被认为是估计光合作用的一个新的优势指标，并被用于区域性的作物产量预测。因此，今后的研究将考虑构建融合多源数据的冬小麦产量预测模型，以进一步提高预测精度。

4　结论

（1）随着冬小麦生育期的推进，基于多时相 GF－1 影像数据的不同植被指数与产量的相关性不断增强，其中 3 月 26 日孕穗期植被指数与产量相关性最高，多个植被指数与产量相关系数达到 0.7，基于此生育期植被指数建立的冬小麦产量估测模型的精度也相对较高。

（2）采用多个植被指数作为机器学习模型的输入参数，分别以单一时相、任意 2 个时相以及多个时相的植被指数组合作为自变量向量，构建产量估测模型。结果表明，基于多时相特征参数的输入能够有效提高冬小麦产量的估测精度，采用拔节期和孕穗期的植被指数组合比采用单个生育期的植被指数的估产精度要高，BPNN 模型估测决定系数 R^2 提高约 0.104，SVR 模型 R^2 提高约 0.141，RFR 模型 R^2 提高约 0.107。

（3）对比基于 BPNN、SVR 和 RFR 3 种机器学习

方法所建立的产量估测模型,针对孕穗期的多时相植被指数,3 种模型均有较高的估测精度:BPNN 模型估测精度最高为 91.47%,R^2 为 0.859,RMSE 为 472.873 kg/hm^2;SVR 模型估测精度最高为 91.03%,R^2 为 0.826,RMSE 为 492.917 kg/hm^2;RFR 模型估测精度最高为 92.25%,R^2 为 0.908,RMSE 为 445.874 kg/hm^2;其中 RFR 模型具有更好的稳定性。3 种机器学习算法的最优精度模型对比传统线性回归模型,精度分别提升 7.53、7.09、8.31 个百分点。

参 考 文 献

[1] 赵春江. 农业遥感研究与应用进展[J]. 农业机械学报, 2014, 45(12): 277 - 293.
ZHAO Chunjiang. Advances of research and application in remote sensing for agriculture[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(12): 277 - 293. (in Chinese)

[2] GROTEN S M E. NDVI-crop monitoring and early yield assessment of Burkina Faso[J]. International Journal of Remote Sensing, 1993, 14(8): 1495 - 1515.

[3] 黎锐, 李存军, 徐新刚, 等. 基于支持向量回归(SVR)和多时相遥感数据的冬小麦估产[J]. 农业工程学报, 2009, 25(7): 114 - 117.
LI Rui, LI Cunjun, XU Xin'gang, et al. Winter wheat yield estimation based on support vector machine regression and multi-temporal remote sensing data[J]. Transactions of the CSAE, 2009, 25(7): 114 - 117. (in Chinese)

[4] CAMMARANO D, BASSO B, HOLLAND J, et al. Modeling spatial and temporal optimal N fertilizer rates to reduce nitrate leaching while improving grain yield and quality in malting barley[J]. Computer and Electronics in Agriculture, 2021, 182: 105997.

[5] LOBELL D B, BURKE M B. On the use of statistical models to predict crop yield responses to climate change[J]. Agricultural and Forest Meteorology, 2010, 150: 1443 - 1452.

[6] TURNER D P, GOWER S T, COHEN W B, et al. Effects of spatial variability in light use efficiency on satellite-based NPP monitoring[J]. Remote Sensing of Environment, 2002, 80(3): 397 - 405.

[7] 欧阳宏达, 赵书河, 张新明. 基于多源遥感数据和 EC - LUE 模型的冬小麦产量估算[J]. 中国农业信息, 2023, 35(1): 27 - 42.
OUYANG Hongda, ZHAO Shuhe, ZHANG Ximming. Winter wheat yield estimation based on multi-source remote sensing data and EC - LUE model[J]. China Agricultural Informatics, 2023, 35(1): 27 - 42. (in Chinese)

[8] GUAN K, WU J, KIMBALL J S, et al. The shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating large-scale plant photosynthesis[J]. Remote Sensing of Environment, 2017, 199: 333 - 349.

[9] 郭建茂, 郭彩云, 李羚, 等. 基于地面通量观测与 MODIS 数据估算玉米光能利用率[J]. 气象科学, 2023, 43(2): 186 - 195.
GUO Jianmao, GUO Caiyun, LI Ling, et al. Retrieval of light use efficiency of maize based on surface flux observation and MODIS data[J]. Journal of the Meteorological Sciences, 2023, 43(2): 186 - 195. (in Chinese)

[10] REN J, ZHANG N, LIU X, et al. Dynamic harvest index of winter wheat based on UAV hyperspectral remote sensing considering crop aboveground biomass change and the grain filling process[J]. Remote Sensing, 2022, 14: 1955.

[11] 任建强, 张宁丹, 刘杏认, 等. 基于哨兵-2A 模拟反射率及其影像的冬小麦收获指数估算[J]. 农业机械学报, 2022, 53(12): 231 - 243.
REN Jianqiang, ZHANG Ningdan, LIU Xingren, et al. Estimation of harvest index of winter wheat based on simulated Sentinel - 2A reflectance data and its real remote sensing imagery[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(12): 231 - 243. (in Chinese)

[12] 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): 161 - 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): 161 - 167. (in Chinese)

[13] 林滢, 邵怀勇. 基于随机森林算法的河南省冬小麦产量预测最佳时间窗和影响因子研究[J]. 麦类作物学报, 2020, 40(7): 874 - 880.
LIN Ying, SHAO Huaiyong. Study on optimal time and influencing factors for winter wheat yield prediction in Henan based on random forest algorithm[J]. Journal of Triticeae Crops, 2020, 40(7): 874 - 880. (in Chinese)

[14] 杨北萍, 陈圣波, 于海洋, 等. 基于随机森林回归方法的水稻产量遥感估算[J]. 中国农业大学学报, 2020, 25(6): 26 - 34.
YANG Beiping, CHEN Shengbo, YU Haiyang, et al. Remote sensing estimation of rice yield based on random forest regression method[J]. Journal of China Agricultural University, 2020, 25(6): 26 - 34. (in Chinese)

[15] 曾妍, 王迪, 赵晓娟. 基于支持向量回归的关中平原冬小麦估产研究[J]. 中国农业信息, 2019, 31(6): 10 - 20.
ZENG Yan, WANG Di, ZHAO Xiaojuan. Study on yield prediction of winter wheat in Guanzhong Plain based on SVR[J]. China Agricultural Informatics, 2019, 31(6): 10 - 20. (in Chinese)

[16] PANTAZI X E, MOSHOU D, ALEXANDRIDIS T, et al. Wheat yield prediction using machine learning and advanced sensing techniques[J]. Computers and Electronics in Agriculture, 2016, 121(2): 57 - 65.

[17] GUO Y H, ZHANG X, CHEN X Z, et al. Integrated UAV-based multi-source data for predicting maize grain yield using machine learning approaches[J]. Remote sensing, 2022, 14: 6290.

[18] 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）
[19] WANG L，TIAN Y，YAO X，et al. Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images[J]. Field Crops Research，2014，164：178－188.
[20] LOPRESTI M F，BELLA C M D，DEGIOANNI A J，et al. Relationship between MODIS－NDVI data and wheat yield[J]. Information Processing in Agriculture，2015，2(2)：73－84.
[21] 王恺宁，王修信. 多植被指数组合的冬小麦遥感估产方法研究[J]. 干旱区资源与环境，2017，31(7)：44－49.
WANG Kaining，WANG Xiuxin. Research on winter wheat yield estimation with the multiply remote sensing vegetation index combination[J]. Journal of Arid Land Resources and Environment，2017，31(7)：44－49.（in Chinese）
[22] LI L C，WANG B，FENG P Y，et al. Developing machine learning models with multi-source environmental data to predict wheat yield in China[J]. Computers and Electronics in Agriculture，2022，194：106790.
[23] 郭瑞，朱秀芳，李石波，等. 山东省冬小麦单产监测与预报方法研究[J]. 农业机械学报，2020，51(7)：156－163.
GUO Rui，ZHU Xiufang，LI Shibo，et al. Monitoring and forecasting method of winter wheat yield in Shandong Province[J]. Transactions of the Chinese Society for Agricultural Machinery，2020，51(7)：156－163.（in Chinese）
[24] RANJAY S，LIPING D，EUGENE G Y，et al. Regression model to estimate flood impact on corn yield using MODIS NDVI and USDA cropland data layer[J]. Journal of Integrative Agriculture，2017，16(2)：398－407.
[25] XUE J，SU B. Significant remote sensing vegetation indices：a review of developments and applications[J]. Journal of Sensors，2017，2017：1－17.
[26] 谭昌伟，杜颖，童璐，等. 基于开花期卫星遥感数据的大田小麦估产方法比较[J]. 中国农业科学，2017，50(16)：3101－3109.
TAN Changwei，DU Ying，TONG Lu，et al. Comparison of the methods for predicting wheat yield based on satellite remote sensing data at anthesis[J]. Scientia Agricultura Sinica，2017，50(16)：3101－3109.（in Chinese）
[27] 任建强，陈仲新，周清波，等. MODIS 植被指数的美国玉米单产遥感估测[J]. 遥感学报，2015，19(4)：568－577.
REN Jianqiang，CHEN Zhongxin，ZHOU Qingbo，et al. MODIS vegetation index data used for estimating corn yield in USA[J]. Journal of Remote Sensing，2015，19(4)：568－577.（in Chinese）
[28] 欧阳玲，毛德华，王宗明，等. 基于 GF－1 与 Landsat8 OLI 影像的作物种植结构与产量分析[J]. 农业工程学报，2017，33(11)：147－156.
OUYANG Ling，MAO Dehua，WANG Zongming，et al. Analysis crops planting structure and yield based on GF－1 and Landsat8 OLI images[J]. Transactions of the CSAE，2017，33(11)：147－156.（in Chinese）
[29] SERRANO J，FILELLA I，PENUELAS J. Remote sensing of biomass and yield of winter wheat under different nitrogen supplies[J]. Crop Science，2000，40：723－731.
[30] 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)：457－465.
[31] PEARSON R L，MILLER D L. Remote mapping of standing crop biomass for estimation of the productivity of the short grass prairie[C]// Proceedings of the Eighth International Symposium on Remote Sensing of Environment. Michigan：Ann Arbor，1972：1357－1381.
[32] JORDAN C F. Derivation of leaf-area index from quality of light on the forest floor[J]. Ecology，1969，50(4)：663－666.
[33] SIMS D A，GAMON J A. Relationships between leaf pigment content and spectral reflectance across a wide range of species，leaf structures and developmental stages[J]. Remote Sensing of Environment，2002，81：337－354.
[34] PENUELAS J，BARET F，FILLELLA I. Semi-empirical indices to assess carotenoids /chlorophyll a ratio from leaf spectral reflectance[J]. Photosynthetica，1995，31：221－230.
[35] HUETE A R. A soil-adjusted vegetation index (SAVI)[J]. Remote sensing of Environment，1988，25：295－309.
[36] RONDEAUX G，STEVEN M D，BARET F. Optimization of soil adjusted vegetation indices[J]. Remote Sensing of Environment，1996，55：95－107.
[37] DAUGHTRY C S T，WALTHALL C L，KIM M S，et al. Estimating corn foliar chlorophyll content from leaf and canopy reflectance[J]. Remote Sensing of Environment，2000，74：229－239.
[38] SCHLEICHER T D，BAUSCH W C，DELGADO J A，et al. Evaluation and refinement of the nitrogen reflectance index (NRI) for site-specific fertilizer management[C]// ASAE Annual International Meeting Report，2001.
[39] CHANG C C，LIN C J. LIBSVM：a library for support vector machines[J]. ACM Transactions on Intelligent System and Technology，2011，2(3)：27.
[40] 王利军，郭燕，贺佳，等. 基于地块单元的冬小麦遥感估产方法研究[J]. 中国农业资源与区划，2021，42(7)：243－253.
WANG Lijun，GUO Yan，HE Jia，et al. Approach for winter wheat yield estimation with remote sensing image at field scale[J]. Chinese Journal of Agricultural Resource and Regional Planning，2021，42(7)：243－253.（in Chinese）
[41] 刘新杰，魏云霞，焦全军，等. 基于时序定量遥感的冬小麦长势监测与估产研究[J]. 遥感技术与应用，2019，34(4)：756－765.
LIU Xinjie，WEI Yunxia，JIAO Quanjun，et al. Growth monitoring and yield prediction of winter wheat based on time-series quantitative remote sensing data[J]. Remote Sensing Technology and Application，2019，34(4)：756－765.（in Chinese）
[42] ZHU J，YIN Y，LU J，et al. The relationship between wheat yield and sun-induced chlorophyll fluorescence from continuous measurements over the growing season[J]. Remote Sensing of Environment，2023，298：113791.