

基于遥感和作物生长模型的多尺度冬小麦估产研究

余新华¹, 赵维清^{2*}, 朱再春², 徐保东³, 赵志展⁴

1. 北京师范大学地理科学学部, 地表过程与资源生态国家重点实验室, 北京 100875
2. 北京大学深圳研究生院, 城市规划与设计学院, 广东 深圳 518055
3. 华中农业大学资源与环境学院, 宏观农业研究院, 湖北 武汉 430070
4. 南京大学大气科学学院, 江苏 南京 210023

摘要 粮食安全是社会和谐、政治稳定和经济可持续发展的重要保障。准确预测区域乃至全球的农作物产量能够为各级政府、相关部门制定农业农村政策提供技术支持, 保障粮食安全。目前关于农作物估产的研究大多具有地域性、经验性, 过分依赖地面实测数据, 一种基于多光谱卫星遥感数据和作物生长模型估算农作物产量的模型框架 SCYM(Scalable Crop Yield Mapper)能够极大地减少模型对实测数据的依赖, 快速应用于不同空间尺度、不同种类作物的估产, 为多尺度农作物估产研究提供了一条有效的途径。以安徽省 2012 年—2018 年冬小麦为研究对象, 通过总结前人研究确定的敏感参数及其在研究区内的波动范围, 结合大量实割实测数据优化 WOFOST(World FOod STudies)模型参数; 将模拟产量、不同时段模拟叶面积指数(LAI)同遴选出的天气变量训练随机森林模型, 并以最佳观测日期组合下的 MODIS-LAI 代替对应时段的模拟 LAI 进行产量估算。结果表明: (1)模型产量估算值与站点实测值的总体相关性为 0.758(R^2 为 0.575), RMSE 为 790.92 kg·ha⁻¹。精度较高的站点主要分布在淮北平原(<1%)而高误差区域集中于皖南丘陵地带(>40%); (2)对 2012 年—2018 年全省范围进行冬小麦估产, 根据 7 年平均估产结果的空间分布, 小麦单产由北向南逐渐减少, 高值区出现在皖北的淮北平原, 低值区主要分布于皖中、皖南地区; (3)2012 年—2018 年实测单产平均值为 6 058.00 kg·ha⁻¹, SCYM 估算单产平均值为 5 984.95 kg·ha⁻¹, 且估算产量与实测产量的年际时间序列的相关性为 0.822, RMSE 为 189.96 kg·ha⁻¹, 每年估产的相对误差均不超过 6%。研究表明 SCYM 估产框架对安徽省冬小麦产量估算具有一定的可行性, 在产量预报方面效果良好。该方法能够在一定程度上改善以往估产模型存在的地域性、经验性问题, 在区域尺度的应用方面具有极大的潜力, 未来可为农业估产提供极其重要的理论依据和实用价值。

关键词 遥感; 农作物估产; WOFOST 模型; 冬小麦; 安徽省

中图分类号: TP79 **文献标识码**: A **DOI**: 10.3964/j.issn.1000-0593(2021)07-2205-07

引言

“民以食为天”, 农业是我国的立国之本、强国之基, 粮食安全始终处于国家安全战略地位。准确预测农作物产量不仅能为农业保险、土地租赁等问题提供决策支持, 保障农民切身利益; 更能及早对国家的不同作物产量作出准确预测, 从而帮助我国在国际粮食市场中占据优势地位。

20 世纪 70 年代以来, 卫星遥感技术飞速发展, 诸多专家学者研究发现, 建立基于不同光谱波段反射率特征的地表

信息与特定时间地点产量的经验模型, 在估产上可取得较高精度^[1-2]。随着计算机技术的不断发展, 作物生长模型以其较强的机理性和时间上的连续性成为产量预测研究领域的新兴技术。然而遥感手段过于依赖地面实测产量, 作物模型则局限于单点或小范围尺度上的模拟。研究表明, 将遥感数据与作物模型耦合, 利用空间上连续、时间上动态变化的卫星观测数据获取模型中较难获得的参数, 能够校正模型的产量形成过程, 使模拟结果更为准确^[3-4]。该方法需要以特定站点的土壤、气候等资料为支撑, 采取计算密集型方法同化耗费大量计算资源和时间, 给推广带来困难。

收稿日期: 2020-07-09, 修订日期: 2020-11-20

基金项目: 国家重点研发计划子课题(2017YFD0300402-6)和国家自然科学基金面上项目(41977405)资助

作者简介: 余新华, 1970 年生, 北京师范大学地理科学学部博士研究生 e-mail: yuxh1970@139.com

* 通讯作者 e-mail: iwqzhao@163.com

美国斯坦福大学的 Lobell 团队针对以上估产方法存在的问题,提出了兼具遥感数据的宏观尺度与生长模型机理性特点的 SCYM 估产框架,并应用于美国多个州的玉米估产,估计产量和实测产量的总体相关性达到 0.6^[5]。该估产框架要求输入参数少且无需大量地面实测数据的校准,为多尺度农作物估产研究提供了一种新的思路。目前这一估产框架已被用于不同产区从县级等不同空间尺度上的作物产量估算^[6-7]。

我国幅员辽阔,自然条件和社会经济条件复杂多样,地理环境和种植制度的空间差异性显著。SCYM 估产框架在我国农作物估产中的适用性目前鲜有报道。本研究旨在结合文献调查、田间实测、再分析气象资料、遥感数据、农作物机理模型和机器学习模型,利用随机森林模型代替多元线性回归模型对 SCYM 估产框架进行优化改进,并利用优化后的框架对安徽省 2012 年—2018 年冬小麦产量进行估算,分析该框架在时间、空间上的表现,为 SCYM 在我国农作物估产中的优化和应用奠定基础。

1 实验部分

1.1 研究区概况

安徽省(29°41'N—34°38'N, 114°54'E—119°37'E)地处我国南北交界处,地形种类齐全,皖西、皖南以丘陵山地为主,皖北为平原地区,总面积约 14 万 km²。该区域年均温为 15℃左右,年均降水量介于 700~1 600 mm 之间,全年光热条件良好,适宜多种农作物的生长,为粮食产业提供了优越的发展条件。

1.2 遥感数据

遥感数据来自 2012 年—2018 年 3 月—5 月 MOD15A2H8d 合成的叶面积指数产品,每年 12 景,空间分辨率为 500 m。该产品已经过辐射定标、大气校正以及几何校正等处理,在 MRT(MODIS Reprojection Tool)中完成区域拼接与投影转换。

1.3 气象数据

逐日气象数据来自中国科学院青藏高原研究所的《中国区域地面气象要素驱动数据集》^[8-9],时间分辨率为 3 h,空间分辨率为 0.1°×0.1°,包含 7 个气象要素:近地面气温、降水、气压、比湿度、10 m 风速、下行短波和长波辐射。经过格式转换与要素计算得到模型所需的气象要素。饱和水汽压差(vapor pressure deficit, VPD)由近地表 2 m 气温数据与露点温度计算得到,数据来自欧洲中期天气预报中心(European Centre for Medium-Range Weather Forecasts, ECMWF)发布的 ERA5 数据集。

1.4 冬小麦空间分布数据

研究选用安徽省冬小麦生长关键期多幅遥感影像,以县为单位,使用空间分辨率 16 m 左右的 GF1 和 Landsat8 影像,每县至少 2 幅影像,关键时相主要为农作物未播种前期上年 10 月—12 月上旬影像(北部主要为 10 月中旬影像,中南部主要为 11 月—12 月上旬影像),生长中期本年 3 月中下旬—4 月上旬影像,基于物候数据及地面调查数据选取训练样本,主要利用最大似然法(少部分县使用神经网络方法)进

行监督分类;采用 2 m 空间分辨率影像提取耕地图斑,与分类结果求交集,获得安徽省冬小麦空间分布。训练和精度验证样本为全省均匀分布的 1 198 个 200 m×200 m 样方压盖的自然地块。主要使用 2 m 分辨率 GF1 影像,补充使用 0.8 m GF2 和 BJ2 影像重采样为 2 m 影像,人工勾画地块及目视解译判读地块类型。自然地块的作物属性由人工野外调查获得。全部样方中的 90%用于训练,10%用于精度验证,采用误差矩阵评价分类精度。经验证各县的分类精度可达 85%。

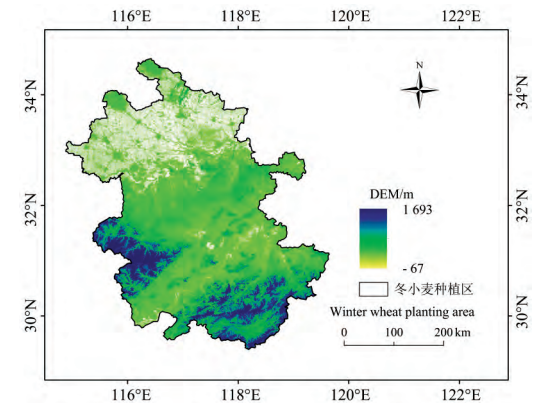


图 1 研究区概况

Fig 1 Overview of the study area

1.5 实割实测产量数据

实割实测产量数据来自国家统计局,为地块尺度数据,处理成村级尺度的冬小麦单位面积产量。2012 年—2018 年间,全省共包含 1 265 个可用的村级站点/年组合,其中皖北 650 个,皖中 506 个,皖南 109 个。

1.6 SCYM 估产框架

SCYM 是 Lobell 团队提出的一种新的多尺度农作物估产框架,旨在利用作物生理知识去解释卫星观测结果,可快速应用于不同田块、不同种类作物的估产,估产精度高并且可扩展性强。基本流程如下:(1)针对一系列合理的管理措施和环境条件进行作物生长模拟。(2)训练模拟叶面积指数(leaf area index, LAI)、季节性天气变量与模拟产量的随机森林模型。(3)基于遥感观测 LAI 和随机森林模型进行产量估算。

1.7 WOFOST 作物生长模型

WOFOST 模型由荷兰 Wageningen 农业大学和世界粮食研究中心(CWFS)共同开发研制,以作物品种、土壤属性、气候资料及田间管理措施为驱动,逐日模拟作物出苗到成熟的生理过程,目前在诸多国家和地区得到了广泛应用^[10-12]。LAI 是模型中最重要的变量之一,其影响着作物的干物质积累,进而决定了最终产量,并将贮藏器官干重(WSO)的最大值作为冬小麦模拟产量指标。

1.8 随机森林

随机森林由多棵决策树组成,本质是将若干个弱分类器集成为强分类器的集成算法。利用 bootstrap 有放回地抽取与样本集相同大小的训练集,构建分类回归树,针对每棵树的分类结果,采取投票(取众数)或算术平均的方法得到分类

或回归结果，误差则基于未被抽取的样本组成的数据集来计算。本研究中将季节性天气变量（播种—越冬期积温、全生育期降水量、5 月太阳辐射、12 月—2 月平均气温）及不同观测日期下的模拟 LAI 作为随机森林的特征变量集合，模拟产量作为输出结果训练模型。

1.9 SCYM 在安徽地区冬小麦的应用

1.9.1 WOFOST 模型参数优化

WOFOST 模型中作物品种、土壤属性、田间管理等参数同天气变化一起影响着最终产量的形成，模型模拟结果对不同参数变化的响应程度不同。搜集了 WOFOST 模型模拟

冬小麦生育动态相关的大量文献，在这些文献中完成了参数敏感性分析以及本地化率定工作，并提供了全部参数的取值范围(表 1)。

鉴于田间实验数据难以获取这一重要原因，基于大量前人研究成果^[10]，挑选出高敏感性的生长参数：从出苗到开花的积温 TSUM1、叶片衰老系数 SPAN、从出苗到开花的比叶面积 SLATB2 以及播种日期 IDSOW，综合考虑安徽省气象条件和作物生长物候期，分别在其适宜范围内取等间隔的 10 个值，对四个参数的取值排列组合，以各站点气象数据为驱动，分别在皖北、皖中、皖南区域进行模拟。

表 1 WOFOST 模型中冬小麦主要参数标定范围
Table 1 Range of calibrated values of maincrop parameters of WOFOST model

参数名称	定义	单位	取值范围
TSUM1	出苗至开花有效积温	℃ · d	600.0~1 500.0
TSUM2	开花至成熟有效积温	℃ · d	500.0~1 200.0
IDSOW	播种日期	d	273~313
TDWI	初始生物量	kg · hm ⁻²	100~210
SPAN	叶片衰老系数	d	21~41
LAIEM	出苗时叶面积指数	hm ⁻² · hm ⁻²	0.136 5
CVL	叶生长同化物转化效率	kg · kg ⁻¹	0.485~0.685
CVO	贮存器官生长同化物转化效率	kg · kg ⁻¹	0.729~0.81
CVR	根同化物转化效率	kg · kg ⁻¹	0.694
CVS	茎同化物转化效率	kg · kg ⁻¹	0.662~0.762
EFFTB	对应生育期单片光能利用效率	kg · hm ⁻² · h ⁻¹ · J ⁻¹ · m ⁻² · s ⁻¹	0.35~0.5
AMAX01(DVS=0.00)	叶片最大 CO ₂ 同化速率	kg · hm ⁻²	35.83
AMAX02(DVS=1.00)			38.83
AMAX03(DVS=1.30)			30~35.83
AMAX04(DVS=2.00)			4.48
SLA01(DVS=0.00)	比叶面积	hm ² · kg ⁻¹	0.001 25~0.002 12
SLA02(DVS=0.50)			0.001 35~0.002 75
SLA03(DVS=2.00)			0.0~0.001 5

注：在 WOFOST 模型中物候用无量纲状态变量 DVS 来描述，对于大多数一年生作物，DVS 在出苗期设为 0，开花期设为 1，成熟期设为 2
Note: In WOFOST phenology is described by the dimensionless state variable DVS. For most annual crops, DVS is set to 0 at seedling emergence, 1 at flowering and 2 at maturity

1.9.2 产量预测模型建立

基于前人研究成果，结合安徽省冬小麦的各生育阶段，设计出以下指标作为待选变量(表 2)。

利用安徽省每年 10 月至次年 6 月的再分析资料，结合冬小麦生育期计算出以上指标。利用待选指标对相应站点/年的模拟产量逐步回归，最终选取播种—越冬期积温、全生育期降水量、5 月太阳辐射及 12 月—2 月平均气温作为构建产量预测模型的天气变量。

将返青至成熟期(3 月—5 月)按 8 d 为单位分段，提取每个时段的最大模拟 LAI，选择以 8 d 为单位是便于后续将模拟 LAI 与 LAI 遥感观测值一一对应。将上文选取的季节性天气变量与两个不同时段 LAI 值排列组合，作为随机森林的输入特征变量，并将不同日期组合训练成的随机森林模型储存起来，供后续使用。

1.9.3 产量估算

根据 MOD15A2H 产品的质量控制文件及冬小麦空间分

布数据筛选出总体质量好且属于冬小麦种植区的像元。产量验证数据为村级尺度，而 MODIS 产品分辨率为 500 m×500 m，需要对 LAI 数据进行均值滤波处理。付立哲等研究表明，由于农作物种植范围相对较小，MODISLAI 数据中不可避免地会出现大量混合像元，LAI 被低估的现象普遍存在^[13]。利用 2012 年—2018 年 3 月—5 月均值滤波后的 LAI 数据与相应时段的模拟 LAI 进行线性拟合，以修正遥感观测值。

对不同时相 LAI 组合下模型的解释能力按照从高到低排序，提取出每个站点/年所对应像元的最佳可用观测日期组合。将每个站点/年的季节性天气变量及两个时段的 LAI 修正值代入到训练好的随机森林模型中预测产量。

2 结果与讨论

2.1 模型模拟与回归校准

将全部参数组合下各站点/年的模拟产量与地面实测单

表 2 待选择的季节性天气变量
Table 2 Seasonal meteorological weather indicators to be selected

天气变量	待选指标
平均气温/℃	12 月平均气温 1 月平均气温 2 月平均气温
平均最低气温/℃	12 月—2 月平均气温 3 月—5 月平均气温 12 月平均最低气温 1 月平均最低气温 2 月平均最低气温
降水量/mm	12 月—2 月平均最低气温 越冬前降水量 返青期降水量 拔节期降水量 抽穗扬花期降水量 灌浆期降水量 全生育期总降水量
太阳辐射/(kJ·m ⁻²)	灌浆期太阳辐射 3 月太阳辐射 4 月太阳辐射 5 月太阳辐射
饱和水汽压差/kPa	3 月—5 月平均太阳辐射 3 月平均饱和水汽压差 4 月平均饱和水汽压差 5 月平均饱和水汽压差
有效活动积温/℃	3 月—5 月平均饱和水汽压差 播种—越冬期积温 返青—成熟期积温 全生育期积温 越冬期负积温

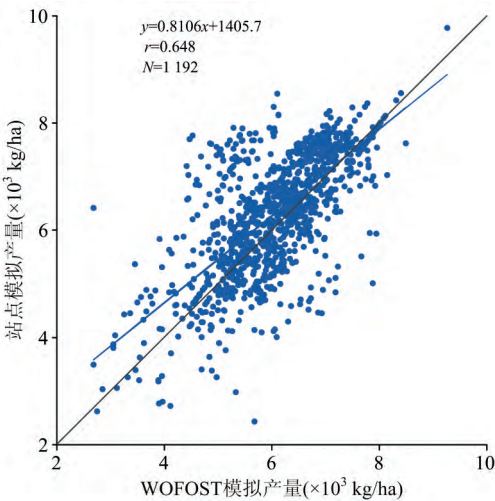


图 2 WOFOST 模拟产量与站点实测产量的比较
Fig 2 Comparison of WOFOST simulated yield with site-measured yield

产线性拟合，剔除异常值点，确定拟合优度最高的参数组合。最终 WOFOST 模型估算产量与站点实测产量的 R^2 达到

0.421，皖北地区为 0.452，皖中地区为 0.376，皖南地区为 0.668。随机抽取 50 次最优参数组合下的模拟结果，绘制冬小麦生育期 LAI 及 WSO 时序曲线[图 3(a)和(b)]。可以看出 LAI 最大值普遍在 3~7 之间，冬小麦单产在 5 000~8 000 kg·ha⁻¹ 波动。在不同种植场景下，安徽省冬小麦的生长情况呈现出广泛的变异性。

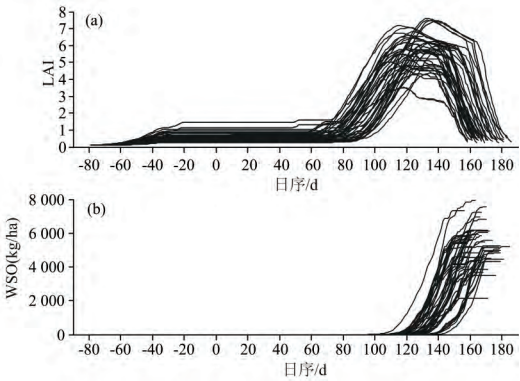


图 3 50 次随机模拟的 LAI(a)和 WSO(b)时间序列曲线
Fig 3 LAI (a) and WSO (b) daily outputs from 50 random simulations of the WOFOST

图 4 表示不同时段 LAI 组合作为特征变量时模型的解释能力。可以看出，基于 LAI 模拟值和季节性天气变量的产量预测模型对模拟结果变异性的解释能力取决于 LAI 的具体观测日期，但总体上决定系数较高，均在 0.68 以上。相较于前期(3 月初)，在冬小麦发育后期(4 月中旬以后)进行 LAI 观测时模型的解释能力更强。在冬小麦播种来年的 5 月 9 日至 16 日、5 月 17 日至 5 月 24 日之间进行 LAI 观测，可以获得该区域最准确的产量估计值。

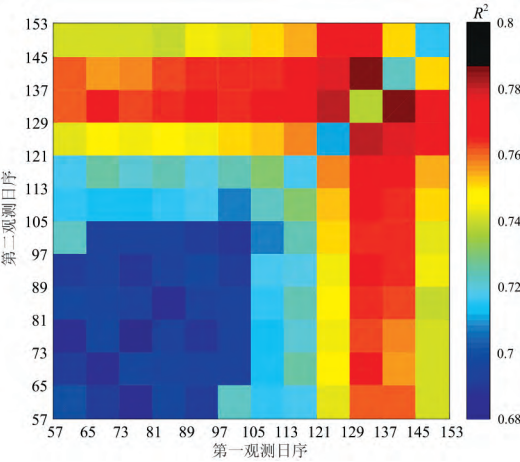


图 4 基于 LAI 模拟值和季节性天气变量的随机森林模型的决定系数(R^2)
Fig 4 The coefficient of determination (R^2) for the Random Forests models based on simulated LAI and monthly weather

2.2 SCYM 估产框架精度分析
2.2.1 站点尺度估产精度分析

将全部站点/年组合视为一个整体时, SCYM 的产量估算值与实测值总体相关性为 0.758 (R^2 为 0.575) (图 5), RMSE 为 $790.92 \text{ kg} \cdot \text{ha}^{-1}$ 。其中皖北地区冬小麦产量预测效果最佳, 相关系数可达 0.604 (R^2 为 0.365), RMSE 为 $807.77 \text{ kg} \cdot \text{ha}^{-1}$, 皖南地区预测效果不佳, 相关系数仅为 0.210 (R^2 为 0.044), RMSE 为 $786.56 \text{ kg} \cdot \text{ha}^{-1}$ 。

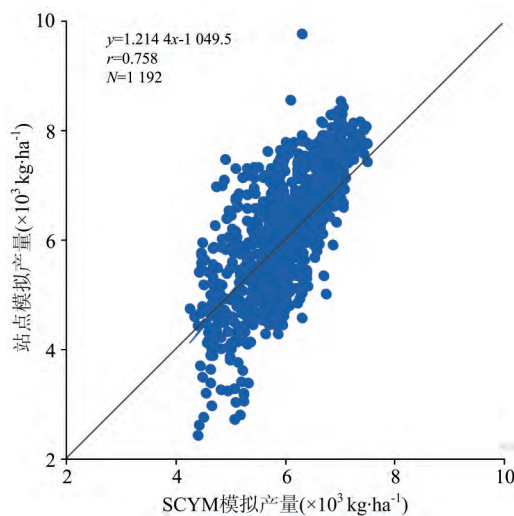


图 5 SCYM 估算产量与站点实测产量的比较

Fig 5 Comparison of SCYM estimated yield with site-measured yield

从 SCYM 估产的空间分布差异来看(图 6), 估产精度的空间异质性显著, 估产精度较高的站点主要分布在淮北平原, 相对误差普遍低于 1%, 而高误差区域集中于皖南丘陵地带, 最大相对误差为 60.6%。可能的原因是皖北种植规模较大, 产量验证点分布较为密集, 而皖南多为山地丘陵, 站点分布稀疏, 为 SCYM 框架解释低产信息带来了一定的困难。

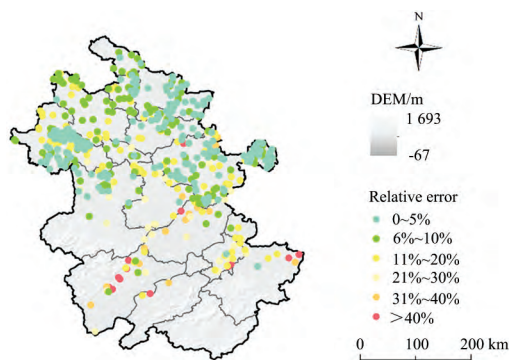


图 6 SCYM 估算产量与站点实测产量的相对误差空间分布

Fig 6 Spatial distribution of relative error between the SCYM estimated yield and the site-measured yield

2.2.2 安徽省冬小麦估产精度分析

以季节性天气变量及每个像元最佳观测日期组合下的 LAI 为驱动, 对 2012 年—2018 年安徽省冬小麦的产量进行预测, 多年平均产量分布情况如图 7。图中白色为非耕地区

域, 小麦单产在 $4\,531.91 \sim 6\,997.58 \text{ kg} \cdot \text{ha}^{-1}$ 。总体而言, 小麦单产由北向南逐渐减少, 高值区出现在皖北的淮北平原, 低值区则在皖中、皖南均有分布。

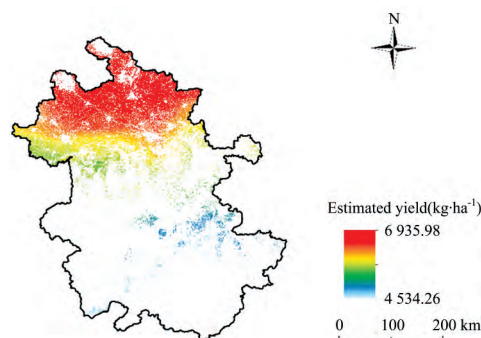


图 7 2012 年—2018 年 SCYM 估算产量平均值

Fig 7 Average SCYM estimated yield in 2012—2018

将 SCYM 估产结果与站点实测数据逐年比较(图 8), 2014 年后估算产量持续下降, 分析七年的天气变量发现, 2014 年后 5 月太阳辐射总量下降 11.92%, 全生育期降水量增加 22.94%, 多雨寡照的气候条件对冬小麦灌浆十分不利, 从而导致冬小麦减产。自 2015 年开始, 播种—越冬期积温、12 月—2 月平均气温不断升高, 研究发现, 气候变暖导致小麦发育前期旺长, 若气温回升而不及时调控, 后期易早衰、倒伏, 从而产量大幅下降。但总体来看, 2012 年—2018 年实测单产平均值为 $6\,058.00 \text{ kg} \cdot \text{ha}^{-1}$, SCYM 估算单产平均值为 $5\,984.95 \text{ kg} \cdot \text{ha}^{-1}$, 两条时序曲线的相关性达到 0.822, RMSE 为 $189.96 \text{ kg} \cdot \text{ha}^{-1}$, 每年估产的相对误差均不超过 6%, 说明该方法具有一定的可行性, 可为安徽冬小麦的产量估算提供技术支持。

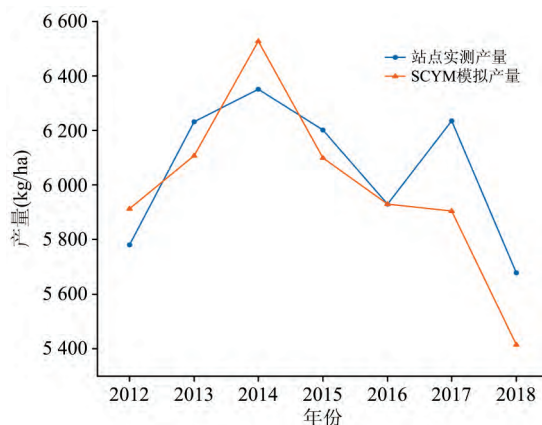


图 8 安徽省冬小麦单产变化趋势

Fig 8 Trend of winter wheat yield in Anhui Province

3 结论

以安徽省冬小麦为研究对象, 通过总结前人研究确定的敏感参数及其在研究区内的波动范围, 结合大量实割实测数据优化 WOFOST 模型参数, 在 SCYM 基本框架基础上, 利用随机森林模型代替多元线性回归模型, 在站点/年尺度及

全省冬小麦种植空间分别进行产量估算,均取得良好效果。

SCYM 估产框架仅需要输入气象数据及遥感观测值便可快速应用于不同种植场景,实现不同区域、不同作物的产量估计。该方法能够在一定程度上改善当前估产模型存在的地域性、经验性问题,为农业估产提供重要的理论依据和实用价值。

但是对于 SCYM 的研究尚在探索阶段,距离真正投入使用还有待于更加深入的研究:

(1)由于缺乏专业农业气象试验站点数据,本研究确定的最优模型参数在其他地区可能不再适用。未来将会注重田间数据的采集工作,加强作物生长模拟方面的研究。

(2)仅使用 MODIS LAI 数据进行估产,今后引入多源遥感数据,有望进一步提高 SCYM 的估产精度和适用性。

(3)本研究的估产框架理论上可用于我国绝大多数农作物的多尺度估产。后续将逐步评估 SCYM 估产框架对我国其他主要粮食作物的多尺度估产效果。

References

- [1] REN Jian-qiang, CHEN Zhong-xin, ZHOU Qing-bo(任建强, 陈仲新, 周清波, 等). Journal of Remote Sensing(遥感学报), 2015, 19(4): 568.
- [2] GONG Yan, XIAO Jie, HOU Jin-yu, et al(龚 龔, 肖 洁, 候金雨, 等). Journal of Geomatics(测绘地理信息), 2017, 42(6): 40.
- [3] BAO Shan-ning, CAO Chun-xiang, HUANG Jian-xi, et al(包姍宁, 曹春香, 黄健熙, 等). Journal of Geo-Information Science(地球信息科学学报), 2015, 17(7): 871.
- [4] Jin H, Li A, Wang J, et al. European Journal of Agronomy, 2016, 78: 1.
- [5] Lobell D B, Thau D, Seifert C, et al. Remote Sensing of Environment, 2015, 164: 324.
- [6] Azzari G, Jain M, Lobell D B. Remote Sensing of Environment, 2017, 202: 129.
- [7] Burke M, Lobell D B. Proceedings of the National Academy of Sciences, 2017, 114(9): 2189.
- [8] Yang K, He J, Tang W J, et al. Agricultural and Forest Meteorology, 2009, 150(1): 38.
- [9] He J, Yang K, Tang W, et al. Scientific Data, 2020, 7(1): 1.
- [10] HUANG Jian-xi, JIA Shi-ling, MA Hong-yuan, et al(黄健熙, 贾世灵, 马鸿元, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2017, 33(10): 230.
- [11] Ceglar A, Van d W R, De Wit A, et al. Agricultural Systems, 2018, 168: 168.
- [12] LONG Ze-hao, QIN Qi-ming, ZHANG Tian-yuan, et al(龙泽昊, 秦其明, 张添源, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2020, 40(3): 898.
- [13] FU Li-zhe, QU Yong-hua, WANG Jin-di(付立哲, 屈永华, 王锦地). Journal of Remote Sensing(遥感学报), 2017, 21(2): 206.

Research in Crop Yield Estimation Models on Different Scales Based on Remote Sensing and Crop Growth Model

YU Xin-hua¹, ZHAO Wei-qing^{2*}, ZHU Zai-chun², XU Bao-dong³, ZHAO Zhi-zhan⁴

1. State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
2. School of Urban Planning and Design, Shenzhen 518055, China
3. College of Resources and Environmental Sciences/Macro Agriculture Research Institute Huazhong Agricultural University, Wuhan 430070, China
4. School of Atmospheric Science, Nanjing University, Nanjing 210023, China

Abstract Food security is a guarantee for social harmony, political stability and sustainable development of the economy. However, current research on crop yield estimation is mostly regional and empirical, relying too much on ground-measured data. Scalable Crop Yield Mapping (SCYM) is a satellite data based framework for estimating crop yield. It can be quickly applied to the estimated yield of different spatial scales and different types of crops without relying on measured data. This framework provides an important theoretical basis for multi-scale crop yield estimation research. We took the winter wheat of Anhui Province from 2012 to 2018 as the study object. Firstly, the sensitive parameters determined by the predecessors and their fluctuation ranges in the study area are summarized. Combined with a large amount of site data, the parameters optimization of the WOFOST model was completed. Secondly, random forest models were established based on the simulated yield, simulated leaf area index (LAI) at different periods, and selected meteorological indicators. Finally, the MODIS-LAI under the best observation date combination replaced the simulated LAI for the corresponding time periods to estimate the winter wheat yield in Anhui Province.

The main outcomes in this study are as follows: (1) The overall correlation between the estimated outputs and the measured data of the stations is 0.758 (R^2 is 0.575), and the RMSE is $790.92 \text{ kg} \cdot \text{ha}^{-1}$. The sites with higher production accuracy are mainly distributed in the Huaibei Plain ($<1\%$), while the areas with high errors are concentrated in the hilly areas of southern Anhui ($>40\%$). (2) The winter wheat yield in Anhui Province from 2012 to 2018 was estimated by SCYM. According to the spatial distribution of the 7-year average yield estimation, the yield is gradually decreasing from north to south. The high-value areas are located in the Huaibei Plain in northern Anhui, and the low-value areas are distributed in central Anhui and southern Anhui. (3) The average measured yield from 2012 to 2018 is $6\,058.00 \text{ kg} \cdot \text{ha}^{-1}$, while the average yield of the SCYM is $5\,984.95 \text{ kg} \cdot \text{ha}^{-1}$. The correlation between them in the interannual time series is 0.822, and the RMSE is $189.96 \text{ kg} \cdot \text{ha}^{-1}$. In seven years, the relative error each year does not exceed 6%. This study shows that the SCYM framework is feasible for estimating winter wheat yield in Anhui Province and has a good effect on yield forecast. This method can improve the regionality and empiricism of the previous crop yield estimation models to a certain extent. Meanwhile, it also solves the shortcomings of computationally intensive methods, which are costly and difficult to popularize. Thus, SCYM has great potential in applying of regional scales, and it will provide an extremely important theoretical basis and practical value for agricultural production in the future.

Keywords Remote sensing; Yield estimation; WOFOST; Winter wheat; Anhui Province

(Received Jul. 9, 2020; accepted Nov. 20, 2020)

* Corresponding author

敬告读者——《光谱学与光谱分析》已全文上网

从 2008 年第 7 期开始在《光谱学与光谱分析》网站(www.gpxygpx.com)“在线期刊”栏内发布《光谱学与光谱分析》期刊全文,读者可方便地免费下载摘要和 PDF 全文,欢迎浏览、检索本刊当期的全部内容;并陆续刊出自 2004 年以后出版的各期摘要和 PDF 全文内容。2009 年起《光谱学与光谱分析》每期出版日期改为每月 1 日。

《光谱学与光谱分析》期刊社