

VPM模型与转产系数结合的吉林省玉米估产

王永昊¹, 王鸣雷^{1,2}, 闫慧敏^{1,2}, 杨建宇³, 史文娇^{1,2*}

(1. 中国科学院地理科学与资源研究所陆地表层格局与模拟院重点实验室, 北京 100101;

2. 中国科学院大学资源与环境学院, 北京 100049; 3. 中国农业大学土地科学与技术学院, 北京 100193)

摘要: 准确估测农作物产量对于保障粮食安全、指导农业生产和调整生产策略等具有重要作用。为解决大范围估产参数确定难、成本高的问题, 该研究基于 2016–2021 年 Sentinel-2 遥感数据和气象数据等, 提出一种综合植被光合作用模型 (vegetation photosynthesis model, VPM) 与转产系数的产量估测方法, 对吉林省 30 个玉米主产区进行估产研究。结果表明: 1) 该研究提出的模型估产精度较传统 VPM 模型表现出更高的准确性和可靠性 (决定系数提升 0.18; 相对均方根误差降低 3.24%); 2) 研究区玉米单产范围为 7~13 t/hm², 高值区主要集中在中部地区, 并且呈现由中部向西北和东南地区递减的趋势; 3) 模型敏感性分析表明, 更精细的转产系数、更高分辨率的玉米空间分布数据和遥感数据能够有效提高模型估产精度。该研究提出的模型可为低成本、大规模、快速精确的估产工作提供解决方案, 对实施农业估产具有重要的现实意义和推广价值。

关键词: 模型; 农作物; 产量; 哨兵 2 号; 吉林省; 最大光能利用率

doi: 10.11975/j.issn.1002-6819.202404042

中图分类号: S127

文献标志码: A

文章编号: 1002-6819(2024)-20-0195-07

王永昊, 王鸣雷, 闫慧敏, 等. VPM 模型与转产系数结合的吉林省玉米估产[J]. 农业工程学报, 2024, 40(20): 195-201.

doi: 10.11975/j.issn.1002-6819.202404042 <http://www.tcsae.org>

WANG Yonghao, WANG Minglei, YAN Huimin, et al. Estimating maize yield in Jilin Province of China using VPM model combined with conversion coefficient[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(20): 195-201. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202404042 <http://www.tcsae.org>

0 引言

准确地估测农作物产量对于保障粮食安全至关重要。现有估产方法可以归纳为传统估产和基于遥感手段估产两类。其中, 传统估产方法, 如抽样调查和实地观测等, 虽然估测精度高, 但是估测过程中需要耗费大量人力和物力, 估测成本相对较高。此外, 传统方法只能对少数样本进行实测, 无法满足大范围的估产工作^[1]。随着遥感技术的不断发展, 遥感手段逐渐成为农作物估产的主流方法^[2]。目前, 基于遥感技术的农作物估产方法可归纳为统计回归法、作物生长模型法、生物量和收获指数法、机器学习及深度学习法等。其中统计回归法原理简单、效率较高, 但缺乏作物生长机理的融合^[3]; 作物生长模型法虽然与作物生长机理紧密结合, 但需要大量大田作物参数及各种影响因子对模型进行修正^[4]; 机器学习及深度学习法估产精度较高, 但其“黑箱”模型的特点使其在机理方面难以解释^[5]; 生物量和收获指数法既可以很好地与作物生长机理相结合, 也具备效率高、成

本低的特点, 在作物估产研究中被广泛应用^[6]。

在生物量和收获指数法中, 陆地生态系统总初级生产力 (gross primary production, GPP) 和陆地生态系统净初级生产力 (net primary production, NPP) 具有关键作用。近些年来, 采用以遥感方法为主的光能利用率模型, 如卡内基-艾姆斯-斯坦福方法 (Carnegie Ames Stanford approach, CASA)^[7]、净光合生产力模型 (photosynthesis, PSN)^[8] 和植被光合作用模型 (vegetation photosynthesis model, VPM)^[9] 等, 进行 GPP 和 NPP 估算, 成为陆地生态系统生产力模拟研究的一个重要发展方向。其中, 相较于 CASA 和 PSN 模型, VPM 模型结合了增强植被指数 (enhanced vegetation index, EVI) 和陆表水指数 (land surface water index, LSWI)^[9], 在农田 C3 和 C4 作物 GPP 模拟中应用更加广泛^[10], 并表现出误差低、估产精度高^[11] 的特点。

基于传统 VPM 模型的农作物估产工作中, 需要确定最大光能利用率、收获指数和植物碳素含量等关键参数。各项关键参数因作物类型、作物品种和研究地区的差异而有所不同, 如玉米的最大光能利用率在美国为 3.31 g/MJ (以 C 计, 下同), 在中国范围内则介于 2.17~2.29 g/MJ 之间^[12]; 中国不同地区的收获指数波动范围为其平均值的 20% 左右^[13]。因此, 现有研究多通过实地测量的方式获取这些关键参数, 以实现准确的农作物估产^[14]。然而, 这种数据获取方式成本高、效率低,

收稿日期: 2024-04-08 修订日期: 2024-07-12

基金项目: 国家重点研发计划项目 (2022YFB3903504); 中国科学院战略性先导科技专项 (XDA0440405); 国家自然科学基金项目 (72221002)

作者简介: 王永昊, 研究方向为遥感信息提取技术。

Email: wyh990312@foxmail.com

*通信作者: 史文娇, 博士, 研究员, 研究方向为全球变化与农业系统。

Email: shiwj@reis.ac.cn

与利用遥感手段进行低成本、高效率的农作物估产目标相违背,严重制约了基于遥感技术的VPM等模型在地块尺度上快速、准确实现估产工作的能力^[15]。

为解决大范围估产参数确定难、成本高的问题,本研究在传统VPM估产模型基础上,提出了VPM与转产系数相结合的估产模型;并以吉林省玉米为研究对象,基于Sentinel-2遥感数据完成了10m分辨率的玉米估产研究。

1 研究方法与数据来源

1.1 研究区概况

本研究选取中国吉林省内玉米年播种总面积超过 $4 \times 10^4 \text{ hm}^2$ 的30个区县作为研究区(图1)。吉林省($40^\circ 52' \sim 46^\circ 18' \text{ N}$, $121^\circ 38' \sim 131^\circ 19' \text{ E}$)是中国重要的粮食生产基地,2021年全省玉米产量达 $3.20 \times 10^7 \text{ t}$,位列全国第二。吉林省地形表现为东南高、西北低的趋势,并以东北-西南向纵贯全省的大黑山脉为界,分为东部长白山地和西部松辽平原两大区域。吉林省属于温带大陆性气候,冬季长而严寒,夏季短而炎热,年平均降水量 $500 \sim 600 \text{ mm}$,年均温在 $-2 \sim 6 \text{ }^\circ \text{C}$ 之间。研究区内粮食作物主要以玉米为主,据吉林省统计局统计,2016—2021年间研究区的年平均玉米播种面积和产量分别为 $3.88 \times 10^6 \text{ hm}^2$ 和 $2.81 \times 10^7 \text{ t}$,分别占研究区粮食作物总播种面积和总产量的77.32%和80.01%。

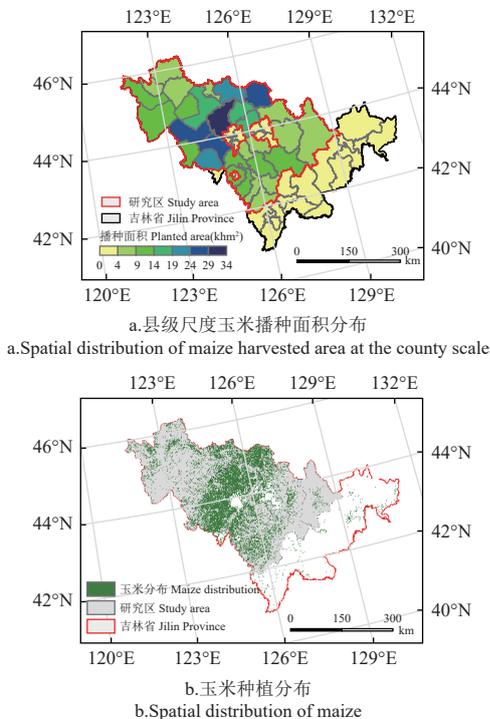


图1 研究区概况
Fig.1 Study area

1.2 数据来源

本研究所使用的数据包括遥感数据、气象数据、玉米种植分布数据以及农业统计数据。遥感数据包括欧洲航天局的Sentinel-2数据和美国宇航局的MOD09A1数

据,其中Sentinel-2数据空间分辨率为10m,时间分辨率为5d;MOD09A1数据空间分辨率为500m,时间分辨率为8d。气象数据主要包括光合有效辐射(photosynthetically active radiation, PAR)和气温,其中PAR数据来源于全球陆表特征参量数据产品数据集(global land surface satellite, GLASS),空间分辨率为5km,时间分辨率为1d;气温数据采用国家青藏高原数据中心(https://data.tpdc.ac.cn/)发布的“中国1km分辨率逐月平均气温数据集”^[16]。玉米种植分布数据来源于国家生态科学数据中心(http://www.nesdc.org.cn/)发布的“2016—2021年中国玉米种植分布数据集(crop distribution, CD)”,空间分辨率为30m,时间分辨率为1a^[17]。此外,为探讨玉米种植分布数据的空间分辨率对研究结果的影响,本研究选择该平台发布的“全国三大粮食作物1km种植分布数据集”中玉米种植分布数据用于对比研究^[18]。2016—2021年研究区内30个区县的玉米总产量和播种面积来源于吉林省统计局^[19](http://tjj.jl.gov.cn/)。最后,本研究将不同空间分辨率的栅格类型数据重采样为10m空间分辨率数据,进而展开后续估产工作。

1.3 研究方法

1.3.1 VPM模型基本原理

VPM模型作为光能利用率模型^[20]的典型代表之一,被广泛应用于全球植被生产力的模拟^[21]。VPM模型中将植被冠层吸收光合有效辐射的比例分为叶绿素吸收部分和非光合植被吸收部分。由于估产所需变量为NPP,而传统VPM模型模拟对象为GPP,因此结合GPP与NPP的转换系数进行调整。主要算法如下:

$$V_{\text{NPP}} = V_{\text{GPP}} \cdot V_{\text{CUE}} \quad (1)$$

$$V_{\text{GPP}} = \varepsilon_g \cdot V_{\text{FPAR}_{\text{chl}}} \cdot \text{PAR} \quad (2)$$

$$\varepsilon_g = \varepsilon_0 \cdot T_{\text{scalar}} \cdot P_{\text{scalar}} \cdot W_{\text{scalar}} \quad (3)$$

式中 V_{NPP} 为净初级生产力; V_{GPP} 为总初级生产力; V_{CUE} 表示碳利用率,即GPP转换成NPP的转化系数; ε_g 表示光能利用率, g/MJ ; $V_{\text{FPAR}_{\text{chl}}}$ 为叶绿素吸收部分的植被冠层吸收光合有效辐射比例; PAR 表示光合有效辐射, MJ/m^2 ; ε_0 表示最大光能利用率, g/MJ ; T_{scalar} 、 P_{scalar} 和 W_{scalar} 分别表示温度、水分和叶龄在冠层尺度上对最大光能利用率的胁迫系数^[22]。

进一步,NPP转化为单产的计算公式如下^[23]:

$$Y = \frac{V_{\text{NPP}} \cdot (1-b) \cdot V_{\text{HI}}}{c \cdot (1-\omega)} \quad (4)$$

式中 Y 为农作物单产, t/hm^2 ; b 为作物地下部分生物量较之整株植物所占的比例; V_{HI} 为收获指数; c 为植物碳素含量; ω 为作物收获后储藏期的含水率系数。依据前人研究^[10], b 、 V_{HI} 、 c 、 ω 分别取值为0.1、0.49、0.45、14%。

1.3.2 转产系数的构建与计算

根据式(1)~(4),传统VPM模型中农作物单产的计算公式可以综合表达为

$$Y = \frac{T_{scalar} \cdot P_{scalar} \cdot W_{scalar} \cdot V_{FPAR_{chl}} \cdot PAR \cdot V_{CUE} \cdot (1-b) \cdot V_{HI} \cdot \varepsilon_0}{c \cdot (1-\omega)} \quad (5)$$

式中 ε_0 、 V_{HI} 、 V_{CUE} 、 c 、 b 和 ω 等关键参数通常为固定值，忽略了不同作物长势、环境、气候条件、作物品种和农业管理条件下参数的差异性，进而导致模拟结果存在较大偏差； PAR 、 $V_{FPAR_{chl}}$ 、 T_{scalar} 、 P_{scalar} 和 W_{scalar} 为动态观测值。为解决关键参数值获取难、成本高等问题，本研究提出 VPM 动态观测指数 $APAR_{e_g}$ 和转产系数 a 以代替多个区域化固定参数和动态观测变量，以对原模型参数进行整合。

本文所定义的 VPM 动态观测指数 $APAR_{e_g}$ 计算如下：

$$APAR_{e_g} = T_{scalar} \cdot P_{scalar} \cdot W_{scalar} \cdot V_{FPAR_{chl}} \cdot PAR \quad (6)$$

本文依照 VPM 原理提出将众多相对固定的关键参数综合成为一个固定的转产系数 a 。结合以后可以具备降低估产成本，提高估产精度优势，如果具备更为精细和丰富产量数据则可以得到更小范围的转产系数，以代替原模型中的相对固定的众多关键参数。本研究定义的转产系数 a 计算公式如下：

$$a = \frac{V_{CUE} \cdot (1-b) \cdot V_{HI} \cdot \varepsilon_0}{c \cdot (1-\omega)} \quad (7)$$

转产系数 a 受研究区环境、作物类型、作物品种和管理条件等影响，且需依托大量实测数据，获取不易。本研究通过作物统计单产数据和 $APAR_{e_g}$ 之间的线性关系计算得到。计算如下：

$$Y = a \cdot APAR_{e_g} \quad (8)$$

式中 $APAR_{e_g}$ 为 VPM 动态观测指数。最后结合各地级市不同区县的 $APAR_{e_g}$ 与对应区县的统计单产建立截距为 0 的线性回归模型，进而模拟得到各地级市的转产系数 a 。

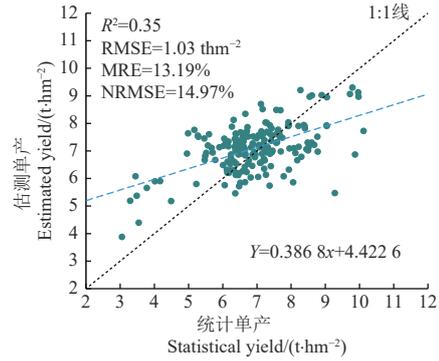
1.4 模型验证

本研究将 2016—2021 年间研究区各区县 $APAR_{e_g}$ 与其对应的统计单产数据集以 7:3 的比例随机分为训练集和测试集，首先通过训练集得到地级市的转产系数 a ，计算出研究区各区县的估产结果后，进而在验证集上获取最终的模型估产精度结果。本研究所采用的精度评价指标包括决定系数^[24]（coefficient of determination, R^2 ）、均方根误差（root mean square error, RMSE）、平均相对误差（mean relative error, MRE）和归一化均方根误差^[25]（normalized root mean squared error, NRMSE）。

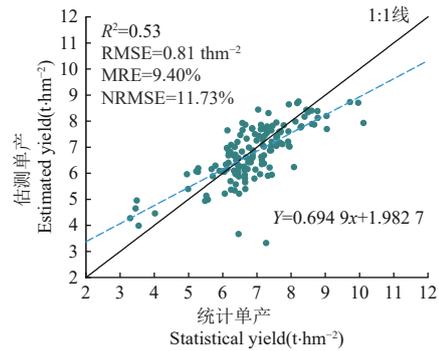
2 结果与分析

2.1 精度验证

相比传统 VPM 模型，本研究方法估产精度相对更高。对比结果显示（图 2a、2b），本研究模型估测单产和统计单产拟合程度更高， R^2 （0.53）高于传统 VPM 模型（0.35）；估产结果误差（RMSE=0.81 t/hm²、MRE=9.40%、NRMSE=11.73%）也较传统 VPM 估产模型（RMSE=1.03 t/hm²、MRE=13.19%、NRMSE=14.97%）要更小。



a. 传统 VPM
a. Traditional VPM (vegetation photosynthesis model)



b. VPM combining yield conversion coefficient a

注： R^2 代表决定系数，RMSE 代表均方根误差，MRE 代表平均相对误差，NRMSE 代表归一化均方根误差。
Note: R^2 stands for coefficient of determination, RMSE stands for root mean square error, MRE stands for mean relative error, NRMSE stands for normalized root mean square error.

图 2 模型精度对比

Fig.2 Comparison of model accuracy

在空间上（图 3a、3b），本研究区内有 12 个县区（40%）的 RMSE 有所下降。主要原因是由于本研究模型可以得到更为本地化的转产系数，并且转产系数比原固定的关键参数更加接近于真实值，使估产结果因地区品种、气候等因素的差异引起的误差有所减小。

2.2 吉林省玉米单产估测

基于 VPM 模型原理与转产系数，本研究估测了 2016—2021 年吉林省 30 个区县玉米单产（图 4）。研究区内各地级市转产系数如表 1 所示，其中四平市转产系数最高，通化和白城市的转产系数则较低。估产结果表明，研究区的高产区主要集中在吉林省中部，并呈现由中部向西北和东南地区递减的趋势。这主要受到地形、气候因素的影响，如吉林省东南地区多为山地，玉米单产偏低；而吉林省西部和中部多为平原地区，并且河流分布广泛，有利于作物生长，玉米单产偏高。

从时间上看（图 5），研究区玉米单产高值主要出现在 2016 年和 2017 年，平均单产分别达到 8.60 和 9.02 t/hm²；2018—2021 年的平均单产则为 6.62 t/hm²。与统计结果对比显示，估测单产结果与统计单产在数值分布以及年际变化趋势上基本保持一致，但个别年份略有偏差。如统计单产最低值（6.40 t/hm²）出现在 2018 年，而估测单产最低值出现在 2020 年（6.15 t/hm²）。出现该问题的原因可能是未考虑农业技术进步所带来的

影响,例如,新的种植品种、育种技术或农业机械的引入可能会在某些年份提高产量,而在其他年份可能没有同样的效果;偶然因素,如在农业生产中,偶然因素(如意外事件、天灾等)可能会导致个别年份的产量偏差,这些因素可能无法完全通过统计或估测方法考虑到。

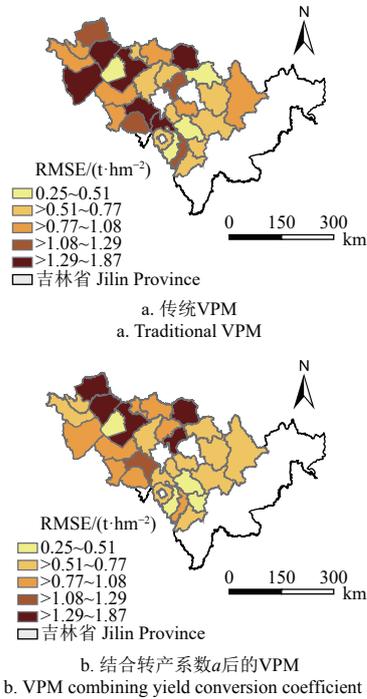


图3 模型均方根误差 RMSE 空间分布对比
Fig.3 Comparison of model RMSE (root mean square error)

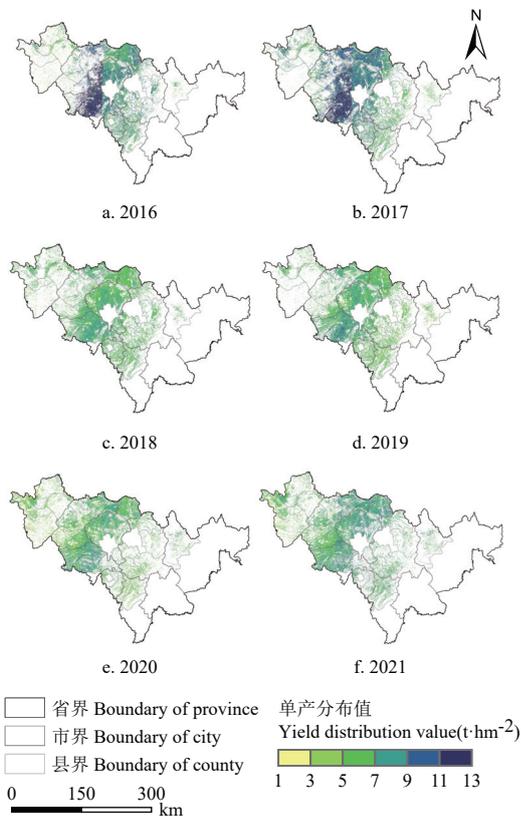


图4 2016—2021 年研究区玉米估测单产空间分布
Fig.4 Spatial distribution of estimated maize yield from 2016 to 2021 in the study area

表1 吉林省各地级市转产系数 α
Table 1 Yield conversion coefficient α of each city in Jilin

Province			
城市 City	转产系数 Yield conversion coefficient/($\times 10^{-2}$)	城市 City	转产系数 Yield conversion coefficient/($\times 10^{-2}$)
长春	1.21	四平	1.33
通化	1.08	吉林	1.10
松原	1.19	白城	1.05

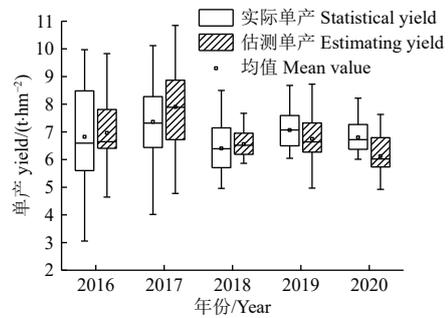


图5 2016-2021 年研究区玉米估测单产与统计单产对比
Fig.5 Comparison of estimated and statistical maize yield from 2016 to 2021 in the study area

2.3 敏感性分析

本研究通过对比调整转产系数的精细程度(市级与省级转产系数)、玉米种植分布的空间分辨率(30 m 与 1 km)以及遥感数据的空间分辨率(10 与 500 m),对本研究中估产模型进行敏感性分析(表2)。

表2 敏感性分析表
Table 2 Sensitivity analysis table $t \cdot \text{hm}^{-2}$

项目 Item	转产系数 α		1 km CD	MOD09A1	
	市级 City-level	省级 Province-level			
城市 City	白城	1.29	1.36	1.24	2.13
	四平	0.94	1.30	1.34	1.74
	长春	1.07	1.13	1.08	1.11
	松原	0.93	0.96	1.03	0.99
	通化	0.71	0.84	1.09	0.79
年份 Year	吉林	0.59	0.71	0.96	0.59
	2016	1.38	1.39	1.61	1.84
	2017	0.91	0.98	0.97	1.38
	2018	0.80	1.04	1.13	1.27
	2019	0.67	0.86	1.01	1.22
	2020	0.87	0.94	0.98	0.94
2021	0.77	0.94	0.98	1.31	

注: 从左到右 4 列数据分别是使用不同尺度(市级、省级)的转产系数、不同空间分辨率的玉米分布数据(1 km)和遥感数据(500 m)估产结果的 RMSE 值。

Note: Four columns from left to right were the RMSE of estimated maize yield by using different scale conversion coefficients(city level, provincial level), maize distribution data (1 km), and remote sensing data (500 m) with different spatial resolutions, respectively.

具体来说,在时间上和空间上与使用地级市级转产系数、省级转产系数、1 km 分辨率玉米种植分布数据和 500 m 分辨率 MOD09A1 遥感数据的估产结果进行精度对比。结果显示,结合地级市级转产系数、10 m 空间分辨率的遥感数据和 30 m 空间分辨率作物种植分布数据,模拟效果更好。具体而言,在市级层面上,使用 MOD09A1 遥感数据的估产结果中,个别地级市的 RMSE 出现高值,并且在逐年的对比中也总是表现出较大的 RMSE;使用省级转产系数和使用 1 km 玉米种植分布数据的估产结果

中, RMSE 在时间和空间上基本相同, 但从整体上看, 前者的 RMSE 略低。综合来看, 各数据对估产结果影响从大到小依次是遥感数据空间分辨率、作物分布数据空间分辨率和所使用的转产系数精细程度。

3 讨论

为解决传统 VPM 模型中区域化相对固定的参数多且较难获取的问题, 本研究基于传统 VPM 模型估产原理, 分别提出转产系数 a 和 VPM 动态观测指数 $APAR_{e_s}$ 以代替多个区域化固定参数以及多个动态观测变量, 简化了获取区域性的关键参数过程。本研究模型通过 $APAR_{e_s}$ 和单产的线性关系得到各地市的转产系数, 以替代传统的一个省一个参数的问题。研究发现, 各地级市的转产系数存在较大差异 (表 1), 并且能够有效减少省级尺度转产系数带来的估产误差 (图 5)。本研究模型与传统 VPM 模型模拟结果对比发现, 本研究模型模拟结果中 R^2 提升 0.23, RMSE、MRE 和 NRMSE 分别下降了 0.23 t/hm²、3.66% 和 3.53%, 模拟精度更高。此外, 与其他研究方法相比, 本研究模型也具有一定的优势。如王鹏新等^[26] 使条件植被温度指数和叶面积指数结合机器学习进行华北地区玉米估产时, R^2 为 0.30; 周西嘉等^[27] 使用 Sentinel-2 数据进行 LAI 重建进而得到冬小麦的估产结果, 其 R^2 为 0.52。除模拟精度外, 本研究基于 Sentinel-2 遥感数据实现了 10 m 空间分辨率的玉米估产, 在空间分辨率上具有一定的优势^[28]。本研究所使用数据的开源性相较于部分使用无人机和非公开数据集具有成本优势和容易获取的优势^[29]; 最后, 本文模型的估产机理性相较于深度学习等人工智能模型相比, 具有较强的可解释性等特点^[30]。

本研究也存在一定的局限性和不确定性。首先, 从数据角度来看, 1) 光学遥感数据、气象数据作为该模型的重要输入变量, 受时间分辨率、多数据空间分辨率不一致和云等因素的影响较大, 可能导致模拟精度降低; 2) 本模型使用的作物种植分布范围数据存在误差因素, 限制了其应用范围; 3) 地块级作物实际单产和统计数据缺乏导致本研究未能实现获取更为精细尺度的县级转产系数; 4) 本文模型采用气象数据和遥感数据对研究区的玉米进行估产, 然而未考虑土壤、地形、管理措施等多种因素影响。其次, 从方法和试验过程来看, 由于数据和 GEE 平台计算限制, 未采取更高时空分辨率数据和时空融合方法估测值。此外, 本研究模型对吉林省 30 个县区玉米单产估产效果较好, 但对其他地区或多作物估产研究的适用性并不清楚。因此, 进一步的改进方案可以在提升输入数据分辨率、降低作物分布数据误差、增补地块级实际单产数据和增加更多的估产因子等方面进行设计, 以此来提高估产模型的数据丰富度和作物适用性, 提高估产结果的精度。

4 结论

本研究基于 VPM 模型估产原理, 整合原模型各

动态变量形成动态观测指数 $APAR_{e_s}$, 将原模型中的固定关键参数整合为转产系数, 提出了一种新的思路来解决传统模型中固定参数多且难以获取的问题, 对大规模、低成本精确作物估产工作具有一定的现实意义。

1) 相较于传统综合植被光合作用模型 (vegetation photosynthesis model, VPM), 本研究提出的结合转产系数的估产模型估产精度更高 (决定系数分别为 0.35 和 0.53, 均方根误差分别为 1.03 和 0.81 t/hm²、平均相对误差分别为 13.19% 和 9.40%, 归一化均方根误差分别为 14.97% 和 11.73%)。改进后的模型, 既保持了良好的机理性, 又可以在缺少实测参数的情况下, 基于县级统计数据完成低成本、大面积的地块尺度作物估产。

2) 本研究结合 Sentinel-2 遥感数据, 得到 2016—2021 年吉林省 10 m 分辨率玉米单产分布图, 显示中部高产 (7~13 t/hm²), 东西部较低 (5~9 t/hm²)。年际变化显示, 2016—2017 年单产较高, 而 2018—2021 年偏低。

3) 使用精细转产系数、高分辨率遥感及玉米种植分布数据能提升估产精度。模型精度的敏感性程度依次为: 遥感数据空间分辨率、作物分布数据空间分辨率、转产系数精细度。

[参 考 文 献]

- [1] 贺鹏, 王婧姝, 曹晨斌, 等. 基于多源遥感数据融合的运城盆地夏玉米估产研究[J]. 中国农业资源与区划, 2023, 44(3): 1-10.
HE Peng, WANG Jingshu, CAO Chenbin, et al. Yield estimation of summer maize in Yuncheng Basin based on fusion of multi-source remote sensing data[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2023, 44(3): 1-10. (in Chinese with English abstract)
- [2] 王耀民, 陈皓锐, 陈俊英, 等. 光谱指数筛选方法与统计回归算法结合的水稻估产模型对比[J]. 农业工程学报, 2021, 37(21): 208-216.
WANG Yaomin, CHEN Haorui, CHEN Junying, et al. Comparison of rice yield estimation model combining spectral index screening method and statistical regression algorithm[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CASE), 2021, 37(21): 208-216. (in Chinese with English abstract)
- [3] KERN A, BARCZA Z, MARJANOVIĆ H, et al. Statistical modelling of crop yield in Central Europe using climate data and remote sensing vegetation indices[J]. Agricultural and forest meteorology, 2018, 260: 300-320.
- [4] WIT A, BOOGAARD H, FUMAGALLI D, et al. 25 years of the WOFOST cropping systems model[J]. Agricultural Systems, 2019, 168: 154-167.
- [5] ZHANG L, ZHANG Z, LUO Y, et al. Integrating satellite-derived climatic and vegetation indices to predict smallholder maize yield using deep learning[J]. Agricultural and Forest Meteorology, 2021, 311: 108666.
- [6] WU B, ZHANG M, ZENG H, et al. Challenges and opportunities in remote sensing-based crop monitoring: A

- review[J]. *National Science Review*, 2023, 10(4): nwac290.
- [7] POTTER C, RANDERSON J, FIELD C, et al. Terrestrial ecosystem production: A process model based on global satellite and surface data[J]. *Global Biogeochemical Cycles*, 1993, 7(4): 811-841.
- [8] RUNNING W, NEMANI R, HEINSCH A, et al. A continuous satellite-derived measure of global terrestrial primary production[J]. *Bioscience*, 2004, 54(6): 547-560.
- [9] XIAO X, HOLLINGER D, ABER J, et al. Satellite-based modeling of gross primary production in an evergreen needleleaf forest[J]. *Remote Sensing of Environment*, 2004, 89(4): 519-534.
- [10] YAN H, DU W, ZHOU Y, et al. Satellite-based evidences to improve cropland productivity on the high-standard farmland project regions in Henan Province, China[J]. *Remote Sensing*, 2022, 14(7): 1724.
- [11] YAN H, FU Y, XIAO X, et al. Modeling gross primary productivity for winter wheat-maize double cropping system using MODIS time series and CO₂ eddy flux tower data[J]. *Agriculture, Ecosystems & Environment*, 2009, 129(4): 391-400.
- [12] HUANG X, MA M, WANG X, et al. The uncertainty analysis of the MODIS GPP product in global maize croplands[J]. *Frontiers of Earth Science*, 2018, 12: 739-749.
- [13] 黄晓娟. 基于全球 FLUXNET 站点的 MODIS GPP 产品算法改进 [D]. 重庆: 西南大学, 2019.
HUANG, Xiaojuan. Improved MODIS GPP Product Algorithm Based on Global FLUXNET Sites[D]. Chongqing: Southwest University, 2019. (in Chinese with English abstract)
- [14] 谢光辉, 韩东倩, 王晓玉, 等. 中国禾谷类大田作物收获指数和秸秆系数[J]. *中国农业大学学报*, 2011, 16(1): 1-8.
XIE Guanghui, HAN Dongqian, WANG Xiaoyu, et al. Harvest index and residue factor of cereal crops in China[J]. *Journal of China Agricultural University*, 2011, 16(1): 1-8. (in Chinese with English abstract)
- [15] 袁文平, 蔡文文, 刘丹, 等. 陆地生态系统植被生产力遥感模型研究进展[J]. *地球科学进展*, 2014, 29(5): 541-550.
YUAN Wenping, CAI Wenwen, LIU Dan, et al. Satellite-based vegetation production models of terrestrial ecosystem: an overview[J]. *Advances In Earth Science*, 2014, 29(5): 541-550. (in Chinese with English abstract)
- [16] PENG S, DING Y, WEN Z, et al. Spatiotemporal change and trend analysis of potential evapotranspiration over the Loess Plateau of China during 2011-2100[J]. *Agricultural and forest meteorology*, 2017, 233: 183-194.
- [17] SHEN R, DONG J, YUAN W, et al. A 30 m Resolution distribution map of maize for China based on landsat and sentinel images. *Journal of Remote Sensing*[J], 2022, 2022: 12.
- [18] LUO Y, ZHANG Z, CHEN Y, et al. ChinaCropPhen1km: a high-resolution crop phenological dataset for three staple crops in China during 2000-2015 based on leaf area index (LAI) products[J]. *Earth System Science Data*, 2020, 12(1): 197-214.
- [19] 吉林省统计局. 吉林统计年鉴 [M] 北京; 中国统计出版社. 2017-2022
Statistics Bureau of Jilin. *Jilin Statistical Yearbook*[M] Beijing; China Statistics Press. 2017-2022 (in Chinese with English abstract)
- [20] MONTEITH J, Moss C. Climate and the efficiency of crop production in Britain[J]. *Philosophical Transactions of the Royal Society of London B, Biological Sciences*, 1977, 281(980): 277-294;
- [21] XIAO X, ZHANG Q, BRASWELL B, et al. Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data[J]. *Remote sensing of environment*, 2004, 91(2): 256-270.
- [22] XIAO X, BOIES S, LIU J, et al. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images[J]. *Remote Sensing of Environment*, 2005, 95(4): 480-492.
- [23] LIU G, WANG X, BAIOCCHI G, et al. On the accuracy of official Chinese crop production data: Evidence from biophysical indexes of net primary production[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2020, 117(41): 25434-25444.
- [24] 赫晓慧, 罗浩田, 乔梦佳. 基于 CNN-RNN 网络的中国冬小麦估产[J]. *农业工程学报*, 2021, 37(17): 124-132.
HE Xiaohui, LUO Haotian, QIAO Mengjia, et al. Yield estimation of winter wheat in China based on CNN-RNN network[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CASE)*, 2021, 37(17): 124-132. (in Chinese with English abstract)
- [25] 张悦, 王鹏新, 张树誉, 等. 基于双参数和 Morlet 多时间尺度特性的冬小麦单产估测[J]. *农业机械学报*, 2021, 52(10): 243-254.
ZHANG Yue, WANG Pengxin, ZHANG Shuyu, et al. Yield estimation of winter wheat based on two parameters and morlet multi-scale characteristics[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(10): 243-254. (in Chinese with English abstract)
- [26] 王鹏新, 齐璇, 李俐, 等. 基于随机森林回归的玉米单产估测[J]. *农业机械学报*, 2019, 50(7): 237-245.
WANG Pengxin, QI Xuan, LI Li, et al. Estimation of maize yield based on random forest regression[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(7): 237-245. (in Chinese with English abstract)
- [27] 周西嘉, 张悦, 王鹏新, 等. 基于 Sentinel 的时间序列田块尺度 LAI 重建与冬小麦估产[J]. *农业机械学报*, 2022, 53(8): 173-185.
ZHOU Xijia, ZHANG Yue, WANG Pengxin, et al. Reconstruction of time series LAI and winter wheat yield estimation at field scales based on sentinel satellites[J]. *Transactions of the chinese society for agricultural machinery*, 2022, 53(8): 173-185. (in Chinese with English abstract)
- [28] 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 Qinbo, 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 with English abstract)
- [29] 王汇涵, 张泽, 康孝岩, 等. 基于 Sentinel-2A 的棉花种植面积提取及产量预测[J]. *农业工程学报*, 2022, 38(9): 205-214.
- WANG Huihan, ZHANG Ze, KANG Xiaoyan, et al. Cotton planting area extraction and yield prediction based on Sentinel-2A[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CASE)*, 2022, 38(9): 205-214.
- (in Chinese with English abstract)
- [30] 孙少杰, 吴门新, 庄立伟, 等. 基于 CNN 卷积神经网络和 BP 神经网络的冬小麦县级产量预测[J]. *农业工程学报*, 2022, 38(11): 151-160.
- SUN Shaojie, WU Men Xin, ZHUANG Liwei, et al. Forecasting winter wheat yield at county level using CNN and BP neural networks[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CASE)*, 2022, 38(11): 151-160. (in Chinese with English abstract)

Estimating maize yield in Jilin Province of China using VPM model combined with conversion coefficient

WANG Yonghao¹, WANG Minglei^{1,2}, YAN Huimin^{1,2}, YANG Jianyu³, SHI Wenjiao^{1,2*}

(1. Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China; 2. College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, 100049, China; 3. College of Land Science and Technology, China Agricultural University, Beijing 100193, China)

Abstract: The objective of this study is to estimate the crop yields in the key corn-producing counties of Jilin Province, China. An accurate, efficient, scalable, and cost-effective model was developed using Sentinel-2 remote sensing data. High spatial and temporal resolution was offered along with the comprehensive meteorological data. A robust framework was built to estimate the maize yield. The limitations of traditional estimation were examined using ground surveys or lower-resolution satellite imagery. These were time-consuming, resource-intensive, and prone to errors, due to sampling biases or limited coverage. Sentinel-2 data was incorporated to provide a continuous and consistent view of crop growth patterns over a large area. The vegetation productivity model (VPM) was integrated to calibrate the yield conversion coefficient. VPM approach was used to estimate the crop biomass, according to the vegetation indices from remote sensing data. The biomass was converted directly into the yield. A yield conversion coefficient was also required to consider the agronomic conditions and crop varieties in the study area. The accuracy and relevance of the model were then enhanced to fine-tune the coefficient with the local yield data. The dynamic variables were integrated into the dynamic observation index in the VPM model. The relatively stable parameters were integrated into the conversion coefficient. The accuracy of yield estimation of the improved model ($R^2=0.53$, RMSE=0.81, MRE=9.40%, NRMSE=11.73%) was superior to the traditional models ($R^2=0.35$, RMSE=1.03 t/hm², MRE=13.19%, NRMSE=14.97%). The obtained model was then applied to estimate the corn yields in the target counties of Jilin Province, where the yield range of maize per unit area was found to be 7-13 t/hm². There was a distinct spatial pattern, where the higher yields were concentrated in the central regions and then gradually decreased towards the peripheries. This pattern was aligned with the geographical features, including soil fertility, irrigation availability, and climatic conditions. The high-resolution Sentinel-2 data was used to better capture these subtle variations in the yield patterns. Sensitivity analysis was conducted to further validate the robustness of the model. A systematic investigation was implemented to explore the impact of various factors, including the spatial resolution of remote sensing data, the vegetation indices, and the calibrated conversion coefficient. The precision of yield estimation was enhanced to employ the precise yield conversion coefficients, high-resolution remote sensing data, and maize planting distribution data. The rest factors were the spatial resolution of remote sensing data, the spatial resolution of crop distribution data, and the fineness of yield conversion coefficients. Furthermore, the future research direction was compared to further improve the accuracy of the model. The significant implications were achieved in agricultural modernization and food security. The timely and accurate information was obtained on crop yields, in order to optimize the planting strategies and decision-making on resource use. The findings can provide valuable insights into the influencing factors on agricultural productivity, and sustainability, particularly for food security at the regional and global levels.

Keywords: models; crops; yield; Sentinel-2; Jilin Province; maximum light use efficiency