doi:10.6041/j.issn.1000-1298.2026.17.006

区域冬小麦单产统计数据降尺度制图方法

李　贺[1]　杜若琪[1,2]　黄　翀[1]　刘庆生[1]　张俊艳[1]　张佐君[3]　兰苛意[4]　陶琨健[1]

(1. 中国科学院地理科学与资源研究所地理信息科学与技术全国重点实验室, 北京 100101;
2. 中国地质大学(武汉)地理与信息工程学院, 武汉 430078; 3. 辽宁师范大学地理科学学院, 大连 116029;
4. 重庆师范大学地理与旅游学院, 重庆 401331)

摘要: 区域冬小麦单产空间分布是作物精细化管理的重要基础,但受限于实测样本稀缺,精细制图仍面临挑战。农业统计数据连续稳定且权威,但其向像元尺度单产的降尺度应用不足,如何在保持统计一致性的同时刻画空间异质性,是亟待解决的问题。以中国北方冬小麦主产区为研究对象,提出一种区域冬小麦单产统计数据降尺度制图方法。首先,综合植被指数、气象、土壤及地形数据,利用多元线性回归(Multiple linear regression, MLR)、随机森林(Random forest, RF)和轻量级梯度提升机(Light gradient boosting machine, LightGBM)3 类回归模型,构建县域尺度单产模型;其次,在保持统计一致性的前提下,引入相对单产权重,将县级平均单产推至像元尺度,生成冬小麦单产精细化空间分布数据;最后,利用野外实测样点对降尺度结果进行精度验证,并基于 SHAP(SHapley additive exPlanations)可解释性框架解析单产空间变化的驱动因子。结果表明:LightGBM 模型在县域尺度估产表现最优(R^2 为 0.79,RMSE 为 581.51 kg/hm^2,MAE 为 401.48 kg/hm^2, MRE 为 8.20%, NRMSE 为 9.72%),降尺度像元单产与实测数据具有较好一致性(R^2 为 0.65, RMSE 为 623.32 kg/hm^2, MAE 为 540.65 kg/hm^2, MRE 为 8.45%, NRMSE 为 9.76%);研究区冬小麦降尺度像元单产介于 943.96 ~ 9 453.06 kg/hm^{2}之间,以中高产水平为主导的双峰分布特征,呈现东高西低且平原高、山地低的格局;SHAP 可解释性分析发现,海拔、土壤粉粒含量、全磷含量、坡度和水分胁迫是研究区冬小麦单产高低变化的主导驱动因子,低产区受地形与土壤结构性约束显著,高产区则以养分因子为主要限制。本研究所提冬小麦单产统计数据降尺度方法可为单产统计数据降尺度提供方法参考,相关结果可为作物精细化管理、冬小麦精准种植决策提供数据支撑。

关键词: 冬小麦; 单产制图; 统计数据; 降尺度; 环境信息; 遥感

中图分类号: S17; TP79　　**文献标识码**: A　　**文章编号**: 1000-1298(2026)17-0065-11　　**OSID**:

Downscaling Mapping and Driving Factor Analysis of Regional Winter Wheat Yield Statistical Data

Li He[1]　Du Ruoqi[1,2]　Huang Chong[1]　Liu Qingsheng[1]　Zhang Junyan[1]　Zhang Zuojun[3]
Lan Keyi[4]　Tao Kunjian[1]

(1. *State Key Laboratory of Resources and Environmental Information System*, *Institute of Geographic Sciences and Natural Resources Research*, *Chinese Academy of Sciences*, *Beijing* 100101, *China*
2. *School of Geography and Information Engineering*, *China University of Geosciences* (*Wuhan*), *Wuhan* 430078, *China*
3. *School of Geographical Sciences*, *Liaoning Normal University*, *Dalian* 116029, *China*
4. *School of Geography and Tourism*, *Chongqing Normal University*, *Chongqing* 401331, *China*)

Abstract: Spatial distribution of winter wheat yield is fundamental for precision crop management. However, fine-scale mapping remains challenging due to the scarcity of field-observed yield samples. Agricultural statistical data are continuous, stable and authoritative, but their application in pixel-scale yield downscaling is still limited. A key challenge is how to characterize spatial heterogeneity while maintaining statistical consistency. The main winter wheat producing area in northern China was taken as the research object, and a downscaling mapping method for regional winter wheat yield statistics was proposed. Firstly, based on the data of vegetation index, meteorology, soil and terrain, a county-scale

收稿日期: 2026-05-19　修回日期: 2026-06-01

基金项目: 国家重点研发计划项目(2023YFD1900300、2023YFD1900100)

作者简介: 李贺(1987—),男,副研究员,博士,主要从事农业地理智能研究,E-mail: lih@lreis.ac.cn

通信作者: 黄翀(1975—),男,研究员,博士,主要从事资源环境遥感研究,E-mail: huangch@lreis.ac.cn

yield model was constructed by using multiple linear regression (MLR), random forest (RF), and light gradient boosting machine (LightGBM). Secondly, on the premise of maintaining statistical consistency, the relative yield weight was introduced to push the average yield at the county level to the pixel scale, and the refined spatial distribution data of winter wheat yield per unit area were generated. Finally, the accuracy of the downscaling results was verified by field measured samples, and the driving factors of spatial variation of yield per unit area were analyzed based on SHapley additive exPlanations (SHAP) framework. The results showed that the LightGBM model performed best at the county scale (R^2 = 0.79, RMSE = 581.51 kg/hm^2, MAE = 401.48 kg/hm^2, MRE = 8.20%, NRMSE = 9.72%), and the downscaled pixel yield was in good agreement with the measured data (R^2 = 0.65, RMSE = 623.32 kg/hm^2, MAE = 540.65 kg/hm^2, MRE = 8.45%, NRMSE = 9.76%). The yield of winter wheat ranged from 943.96 kg/hm^2 to 9 453.06 kg/hm^2, which was dominated by medium-high yield level and showed a bimodal distribution, showing a spatial distribution of high in the east and low in the west, high in the plain and low in the mountain. SHAP-based interpretability analysis further indicated that elevation, soil silt content, total phosphorus content, slope and water stress were the main driving factors affecting winter wheat yield variation. Low-yield areas were significantly constrained by topography and soil structure, while high-yield areas were mainly limited by nutrient factors. The proposed downscaling method can provide a useful reference for statistical yield downscaling and can support precision crop management and winter wheat cultivation decision-making.

Key words: winter wheat; yield mapping; statistical data; downscaling; environmental information; remote sensing

0 引言

粮食作物单产是关乎国计民生的基础性指标，对保障国家粮食安全具有重要意义。然而，受粮食需求持续增长、耕地资源约束加剧以及极端气候频发等多重因素影响，全球粮食安全正面临严峻挑战[1-2]。因此，明确粮食单产状况、挖掘单产提升潜力、优化区域生产结构成为当前农业发展的核心方向。获取粮食单产空间分布信息，不仅可以为刻画农业生产格局与区域差异提供重要基础，还对保障粮食安全与推进精准农业管理具有重要意义，是农业研究领域核心任务之一。

作物单产研究的尺度可分为统计单元尺度和地块/像元尺度两类。统计单元尺度常以县/市级作物单产统计数据作为输入数据，刻画行政单元的平均单产水平；而地块尺度多采用实地采样或地块级观测数据[3]。相较于后者，统计资料更易获取和收集，故统计单元尺度的研究较为广泛、方法相对成熟。在早期主要依赖统计调查和基于作物生理参数的经验模型[4]，但受气象、土壤和管理等多因素综合影响，单产形成过程具有时间特定性和非线性，造成单产估测成本高、外推性差，难以满足大范围单产估测需求[5]。遥感数据因其丰富的光谱信息、高时空分辨率，以及较强的空间异质性表现能力，在单产估测中应用广泛[6-7]。单产遥感估测的发展大致经历 4 个阶段：基于回归分析的经验模型，即将各类遥感植被指数与单产建立统计关系，该方法虽简单，但机理性不足、时空可迁移性弱[8]；基于光能利用率的参数模型，通过估算作物对光合有效辐射的利用情况来反演生物量，该方法能反映辐射对单产的影响，但对生长过程刻画有限[9]；利用数据同化方法将遥感信息融入 WOFOST、DSSAT 等作物生长机理模型中，该方法机理性强但参数需求多、计算效率低[10]；机器学习驱动的单产遥感估测，方法能够从多源时空数据中自动挖掘非线性关系，可显著提升计算效率，但也存在模型黑箱等问题[11]。

尽管目前研究在统计单元尺度取得了较高精度，但其结果通常仅反映行政单元的平均单产水平，难以揭示县域内部地块单产的空间异质性，而精细化空间分异信息是精准农业开展变量管理与差异化调控的重要基础[12-13]。当前研究中，模型输出尺度通常与训练数据尺度一致，即基于县域统计数据训练的模型，最终获得县域尺度的单产结果；若要获取像元/地块级的精细单产数据，理论上需匹配同尺度的实测样本进行模型训练与验证。然而，大范围内像元/地块级的单产实测数据获取成本高、覆盖范围有限，难以支撑区域尺度精细化单产制图。为此，部分研究尝试将县域尺度模型直接应用于像元尺度进行单产精细制图[14]。但该方法忽略了不同空间尺度下环境因子与单产响应关系的尺度效应，即县域尺度建立的统计关系难以直接刻画像元尺度的非线性过程，易引入系统偏差，且缺乏与统计数据一致性的约束，影响结果的可靠性。

将县域或其他统计单元平均单产向像元尺度单

产进行转换，本质上是在统计约束条件下实现空间信息细化表达的过程，即在保持县域统计单产一致的前提下，合理刻画县域内部单产的相对差异，实现单产的空间再分配。针对尺度效应难题和实测数据难以获取的问题，部分学者围绕作物单产空间降尺度开展研究，证实降尺度方法与统计数据约束在产量精细尺度转换中具有较好的适用性，可有效兼顾统计一致性与空间分布信息[15-16]。在这一过程中，农业统计数据具有不可替代的作用，能够提供连续、权威的区域单产信息，是当前单产研究的重要数据来源[17]。

然而，现有研究多将农业统计单产数据用于精度验证或统计尺度单产分析[18-19]，对其在精细尺度单产空间构建中的潜力挖掘不足，缺乏对融合农业统计数据与遥感数据以实现单产合理降尺度的探索。同时，现有统计单产降尺度方法仍存在两方面局限：多依赖单一植被指数进行权重构建，难以刻画地形、土壤、气象等多环境因子的综合作用；在统计一致性与空间异质性之间难以取得平衡，直接模型下推方法易造成偏差，难以保证结果与统计数据的一致。

鉴于此，本研究针对冬小麦开展以下两方面研究：融合冬小麦遥感长势信息和土壤、地形、气象等环境驱动因素，构建县域尺度冬小麦单产模型，通过多模型对比筛选最优框架，并结合农业统计数据约束，实现县域尺度单产平均数据到像元尺度单产数据的降尺度制图；基于单产统计数据降尺度分布结果和 SHAP 可解释性框架，定量解析冬小麦单产空间变化的驱动因子。旨在为缺乏大范围连续实测样本条件下县域尺度单产统计向精细化农业管理尺度的过渡，提供一种可解释、可推广的研究方法。

1　研究区与数据

1.1　研究区概况

以黄淮海平原、黄土高原等中国北方农业区为研究区域。区域涵盖高原、山地、丘陵、平原等地形，海拔从 0 m 到 5 238 m，地势起伏大；范围广泛，涉及山东、河北、河南、江苏、安徽、陕西、甘肃等多个省份；涵盖多种自然环境和社会经济条件。该区域是我国重要的粮食主产区，冬小麦为主要粮食作物之一，10 月上旬播种，次年 3 月上旬返青，4 月下旬进入抽穗期，6 月上中旬收获。结合冬小麦种植分布数据及单产统计数据，选取种植面积大于1 000 hm^2 及单产大于 1 t/hm^2 的 493 个县区，作为研究县域（图 1）。

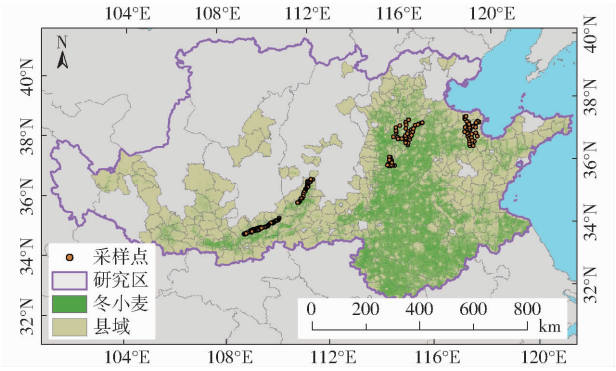


图 1　研究区位置

Fig. 1　Location of study area

1.2　数据来源与处理

研究使用 Sentinel－2 卫星数据来表征冬小麦的生长状态，数据空间分辨率为 10 m。使用美国国家航空航天局发布的 MODIS 陆表产品来表征冬小麦生长期间的地表温度状况与水分消耗过程。MODIS 地表温度（Land surface temperature，LST）数据选用 MOD11A2 产品，空间分辨率为 1 000 m。MODIS 蒸散发（Evapotranspiration，ET）与潜在蒸散发（Potential evapotranspiration，PET）数据选用 MOD16A2 产品，该数据能够较好反映陆地生态系统的实际蒸散发过程及其潜在水分需求，空间分辨率为 500 m。基于谷歌地球引擎平台（Google earth engine，GEE）获取 2023 年 Sentinel－2 的 L2A 数据和 MODIS 产品，筛选云量低于 20% 的影像，并进行去云、镶嵌、裁剪、投影等预处理，重采样至 30 m。

土壤属性数据来自国家地球系统科学数据中心发布的中国高分辨率国家土壤信息格网基本属性 90 m 数据集（https://www.geodata.cn/）[20]，包含全国范围的土壤物理和化学等基本属性。DEM 数据来自 NASA EarthData 网站下载的 ASTER GDEM 第 3 版数据集（https://www.earthdata.nasa.gov），空间分辨率为 30 m。

冬小麦 2023 年种植分布数据来自国家生态科学数据中心（https://www.nesdc.org.cn）。该数据集是基于时间权重的动态时间规整方法生成的 2001—2024 年全国 11 个省 30 m 空间分辨率冬小麦种植分布图[21]，总体精度达 91.17%。各县市 2023 年冬小麦单产统计数据由各地级市统计年鉴、区域统计年鉴、国民经济和社会发展统计公报、各县市政务信息公开中收集整理得到，对于记录缺失的少部分县市，使用该区域临近 3 年的单产均值代替。

2023 年 6 月，于山东省、河北省、山西省和陕西省 4 个省份 23 个县区开展实地测产，共收集 179 个样点的冬小麦单产数据（图 1）。在每个样点所在地块按照对角线取 3 点，设置 1 m^2 样方并确定穗数；依

据冬小麦生长状况随机选取足够数量的麦穗于实验室进行脱粒,记录穗粒数,并任意数出 1 000 粒种子称量;最后依据单产为单位面积成穗数、穗粒数、千粒质量、测产系数的乘积(其中测产系数常用 0.85),计算样点冬小麦单产[22]。

2 研究方法

为获取精细的冬小麦单产空间分布数据,研究利用农业统计、遥感、土壤等多源数据,构建区域冬小麦单产统计数据降尺度制图方法,将县域统计数据合理下推至像元尺度,实现冬小麦单产的空间精细化表达,并分析单产空间变化的驱动因子(图 2)。具体地说,首先,综合植被指数、气象、地形和土壤 4 类因素,以县域为单位,采用多元线性回归(Multiple linear regression,MLR)、随机森林(Random forest,RF)和轻量级梯度提升机(Light gradient boosting machine,LightGBM)3 种回归方法,构建基于生长状况和环境信息的县域尺度冬小麦单产模型;然后,结合训练好的最优模型与统计数据,构建基于统计数据约束的冬小麦单产降尺度方法,得到研究区 30 m 像元尺度冬小麦单产空间分布图;最后,通过野外实测数据验证冬小麦降尺度制图精度,并采用 SHAP 框架分析冬小麦单产空间分布的驱动因素。

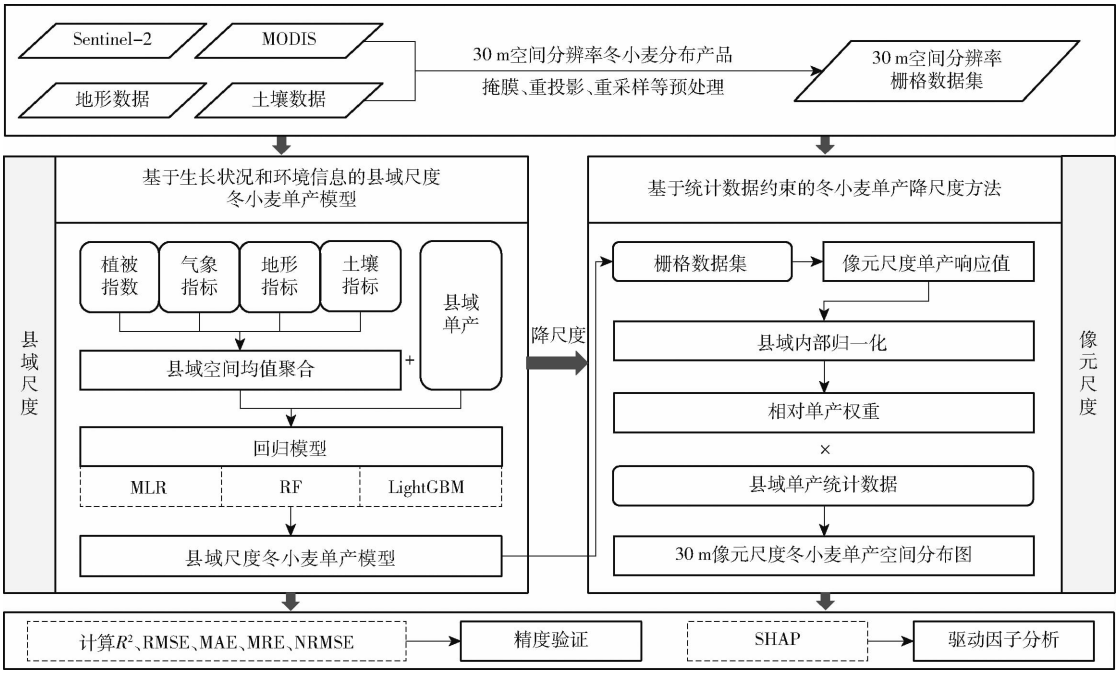


图 2 研究技术路线图

Fig. 2 Technical framework of study

2.1 基于生长状况和环境信息的县域尺度冬小麦单产模型

2.1.1 特征构建

研究选取冬小麦生长信息、气象、土壤、地形 4 类特征变量(表 1)。

植被指数可以有效表征作物的生长状况[3],研究选取冬小麦单产制图中常用的 13 个植被指数(表 1)[23]。结合研究区冬小麦生长物候历,选择 5 月 1—31 日(抽穗开花—成熟期),计算各植被指数的均值与最大值,共得到 26 个指标。其中,均值反映冬小麦成熟期整体长势水平,最大值表征冬小麦生长高峰期最优冠层状态。

气象条件是单产模型构建的关键环境因素之一[24]。研究显示,在冬小麦单产制图中,相较于地面观测温度数据,地表温度 LST 数据的效果更好[25]。因此,研究使用 LST 数据作为温度指标。考虑到除了霜冻等极端情况下,夜间地表温度对植物生长的影响较小,故研究仅提取白天地表温度(LST_Day_1km)波段,来表征冬小麦生长期间的地表温度状况[25]。含水率是影响冬小麦生长的另一个重要因子,研究采用蒸散发 ET,以及 ET 与潜在蒸散发 PET 的比值来表征冬小麦水分胁迫程度。与植被指数类似,采用研究区冬小麦主要生长季期间(3 月 1 日—5 月 31 日)的时序统计特征,包括平均 LST、最大 LST、平均 ET,以及 ET 与 PET 比值的平均值。

土壤状况和地形作为影响冬小麦生长的重要背景因子,对冬小麦单产具有长期稳定的控制作用。研究选取土壤砂粒、土壤粉粒、土壤黏粒、有机碳、氮、磷、钾含量,以及容重、厚度、阳离子交换量、pH 值 11 个土壤属性指标和海拔、坡度 2 个地形指标。

以上所有变量均基于冬小麦空间分布数据进行

表 1　冬小麦单产模型输入指标

Tab. 1　Input variables for winter wheat yield model

指标类型	指标名称	指标缩写
植被指数	归一化差异植被指数(NDVI)	ndvi_mean、ndvi_max
	修正归一化差异植被指数(MNDVI)	mndvi_mean、mndvi_max
	绿色归一化差异植被指数(GNDVI)	gndvi_mean、gndvi_max
	蓝色归一化差异植被指数(BNDVI)	bndvi_mean、bndvi_max
	增强型植被指数(EVI)	evi_mean、evi_max
	双波段增强型植被指数(EVI2)	evi2_mean、evi2_max
	比值植被指数(RVI)	rvi_mean、rvi_max
	差分植被指数(DVI)	dvi_mean、dvi_max
	叶绿素指数(绿光)(CIgreen)	cigreen_mean、cigreen_max
	归一化差异红边指数(NDREI)	ndrei_mean、ndrei_max
	土壤调节植被指数(SAVI)	savi_mean、savi_max
	优化土壤调节植被指数(OSAVI)	osavi_mean、osavi_max
	归一化差异水分指数(NDMI)	ndmi_mean、ndmi_max
气象	地表温度(LST)	lst_mean、lst_max
	蒸散发(ET)	et
	水分胁迫状况(以 ET/PET 代表，其中 PET 为潜在蒸散发)	et_pet
土壤	土壤容重、pH 值、有机碳含量、土体厚度、全磷含量、全钾含量、全氮含量、砂粒含量、粉粒含量、黏粒含量、土壤阳离子交换量	
地形	海拔、坡度	

掩膜、重采样和重投影操作，以避免非目标地物的光谱干扰，并与冬小麦分布数据的分辨率和投影坐标系保持一致。在县域尺度冬小麦单产模型构建阶段，利用县域行政区矢量数据对 43 个变量统计特征进行空间均值聚合，所得到的县域尺度变量作为模型的输入特征。像元级变量则应用于模型应用和单产降尺度阶段。

2.1.2　模型训练

研究以县域冬小麦单产统计数据作为因变量，以各类环境与遥感指标在县域范围内的空间均值为自变量，构建基于生长状况和环境信息的县域尺度冬小麦单产模型。

考虑到冬小麦单产形成过程受多种环境因子的影响，响应关系通常表现出显著的非线性特征、变量交互效应及空间异质性，不同的模型在处理上述问题时具有差异。因此，研究选取 MLR、RF 和 LightGBM 3 类回归模型进行对比分析。

2.2　基于统计数据约束的冬小麦单产降尺度

在构建县域尺度冬小麦单产模型的基础上，充分利用单产统计数据，将其下推至像元尺度，得到研究区 30 m 冬小麦单产空间分布图，实现冬小麦单产的空间精细化表达。

2.2.1　相对单产权重构建

相对单产权重是利用训练好的县域尺度冬小麦单产模型在像元尺度上对冬小麦单产的响应，描述不同像元之间单产的相对差异，而非直接得到绝对单产值。相对单产权重综合反映植被生长状况以及气象条件、土壤属性、地形因素等多种环境因子对冬小麦单产形成的综合影响，表示像元尺度单产的空间分布差异，为后续引入统计数据约束提供合理的空间分配基础。

首先，基于前文构建并验证的最优县域尺度冬小麦单产模型，将像元尺度遥感与环境指标输入模型，在保持模型结构和参数不变的前提下，对研究区内所有冬小麦像元进行逐像元回归，输出像元尺度单产响应值

$$\hat{y}_p = f(\boldsymbol{x}_p) \tag{1}$$

式中　$\hat{y}_p$——像元 p 的单产响应值，kg/hm^2

$f(\cdot)$——最优县域尺度单产模型

$\boldsymbol{x}_p$——像元 p 对应的多源环境与遥感特征向量

此时模型输出的不是像元绝对单产，而是反映在给定环境条件下不同像元之间的单产相对差异。

为消除模型得到的响应值在不同县域内部可能存在的系统偏差，突出像元间的相对差异，须对该响应值进行归一化处理，构建像元相对单产权重。对于县域内的任一像元 p，其相对单产权重 w_p 定义为

$$w_p = \frac{\hat{y}_p}{\bar{\hat{y}}_i} \tag{2}$$

式中　$\bar{\hat{y}}_i$——县域 i 内所有冬小麦像元响应值的平均值，kg/hm^2

通过归一化保证权重具有无量纲特性，以反映县域内部的相对空间差异。

2.2.2　统计数据约束下的像元尺度单产一致化

像元相对单产权重可以有效反映某县域内冬小麦单产的空间格局差异，但单产绝对数值仍受模型系统误差和尺度差异的影响，难以直接与统计数据保持一致。为此，研究引入统计数据约束下的像元尺度单产一致化方法。

在保持县域内部像元相对单产差异不变的前提下，通过对像元相对单产权重进行线性缩放，使县域内所有冬小麦像元的平均单产等于对应的统计单产值。像元尺度最终单产 Y_p 的计算公式为

$$Y_p = w_p y_i \tag{3}$$

式中　y_i——县域 i 的单产统计数据，kg/hm^2

由于相对单产权重在县域尺度内已满足均值为 1 的条件，式(3)能够确保县域尺度像元单产的空间平均值严格等于统计单产。

基于统计数据约束的冬小麦单产降尺度方法可

以在保持统计数据一致的同时,充分利用遥感与环境因子刻画冬小麦单产的空间异质性,有效避免回归模型直接应用在像元尺度上产生偏差的问题,为像元尺度冬小麦单产制图提供新的研究思路。

2.3 精度验证

为评估冬小麦单产统计数据降尺度制图结果的可靠性与精度,研究从县域尺度模型性能及像元尺度制图精度 2 方面评估。

在县域尺度上,采用空间留一交叉方法对冬小麦单产模型进行验证。该方法以县域行政单元作为最小空间验证单元,在每一次验证过程中,将研究区内的一个县域样本整体剔除作为独立测试集,其余所有县域样本作为训练集用于模型构建[27]。模型训练完成后,对留出的县域计算单产,并将结果与对应的统计单产观测值进行对比分析。该过程在研究区内所有冬小麦种植县域上依次重复,使每个县域均在不同迭代中作为独立的空间测试样本参与模型评估。

在像元尺度上,采用野外实测数据对降尺度后的冬小麦单产空间制图结果进行验证,考虑到实测样点数据代表该地区样方的单产均值,将其与 3×3 窗口有效像元均值进行对比,并构建实测值与估测值散点图,检验二者的一致性。

以上验证采用决定系数(Coefficient of determination, R^2)、均方根误差(Root mean square error, RMSE)、平均绝对误差(Mean absolute error, MAE)、平均相对误差(Mean relative error, MRE)和归一化均方根误差(Normalized root mean square error, NRMSE)作为评价指标。

2.4 单产驱动因子分析

为了揭示冬小麦单产统计数据降尺度结果的空间分布差异的主要驱动因素,研究引入 SHAP(SHapley additive exPlanations)可解释性框架对最优单产模型进行解析。

SHAP 不仅能够量化每个特征对冬小麦单产空间分布结果的全局重要性,还能揭示特征值变化对单产结果高低的正负影响及非线性响应[28]。其核心思想源于合作博弈论中的 Shapley 值,通过计算某一特征加入模型前后估算值的边际变化,将模型的总估算值分解为各个特征贡献值的和。

在计算输入变量 SHAP 值的基础上,首先通过所有样本中各特征平均绝对 SHAP 值识别影响冬小麦单产空间分布的关键驱动因子;然后,结合 Beeswarm 分布图分析不同特征取值对单产变化方向与作用强度的影响;最后,依据冬小麦单产统计数据降尺度结果,将像元单产前 25% 定义为高产组,后 25% 定义为低产组,分别统计两组中各特征的平均绝对 SHAP 值差异,对比不同单产水平下的主导限制因子。

3 结果与分析

3.1 县域尺度冬小麦单产模型结果与精度

由县域尺度单产模型得到各县域冬小麦单产空间分布(图 3a~3c)。冬小麦单产大于 7 000 kg/hm^2 的县域多集中于黄淮海平原中部,在 MLR 模型中山东半岛东部县域冬小麦单产高于另外 2 个模型。单产小于 4 000 kg/hm^2 的县域集中在黄土高原地区,但在 LightGBM 模型中,河北省东部沿渤海地区单产相较于另外 2 个模型较低。

对于 MLR、RF 和 LightGBM 3 类模型在县域尺度的单产拟合性能(图 3d~3f),图中 *** 表示极显著($p<0.001$)。从精度指标来看,3 种模型均能较好地刻画各类指标与冬小麦单产的关系,R^2 均高于 0.7,拟合效果较好。其中,LightGBM 模型表现最优(R^2 为 0.79),RF 次之(R^2 为 0.76),MLR 相对较低(R^2 为 0.74)。从绝对误差角度来看,LightGBM 模型的 RMSE(581.51 kg/hm^2)最低,相较于 RF(620.16 kg/hm^2)和 MLR(642.74 kg/hm^2)分别降低约 6.2% 和 9.5%。同样地,LightGBM 模型的 MAE(401.48 kg/hm^2)也最低,比 RF(437.77 kg/hm^2)和 MLR(466.83 kg/hm^2)分别降低约 8.3% 和 14.0%。从相对误差角度来看,LightGBM 模型的 MRE(8.20%)和 NRMSE(9.72%)均为三者中最低,说明该模型在不同单产水平下具有更稳定的预测能力。从散点分布特征来看,LightGBM 模型的估算值与实测值分布最为集中,且更接近 1∶1参考线,系统性偏差较小;而 RF 模型的高产低估、低产高估情况较为明显;MLR 模型散点分布较为分散。

综上,LightGBM 模型在绝对误差、相对误差及散点分布方面均优于 MLR、RF 模型,因此选取 LightGBM 作为后续冬小麦单产降尺度制图与 SHAP 驱动机制解析的基础模型。

3.2 冬小麦单产统计数据降尺度结果与精度

利用最优县域尺度冬小麦单产模型(LightGBM 模型),结合基于统计数据约束的单产降尺度方法,得到研究区冬小麦降尺度单产空间分布(图 4)。研究区冬小麦像元单产介于 943.96~9 453.06 kg/hm^2 之间,空间异质性显著,呈现出东高西低且平原高、山地低的空间分布格局。像元统计显示,研究区平均单产为 6 383 kg/hm^2,中位数为 6 601 kg/hm^2,中位数高于均值,单产整体分布略向高产端偏斜(图 5)。

从空间格局上看,高产区主要集中在黄淮海平

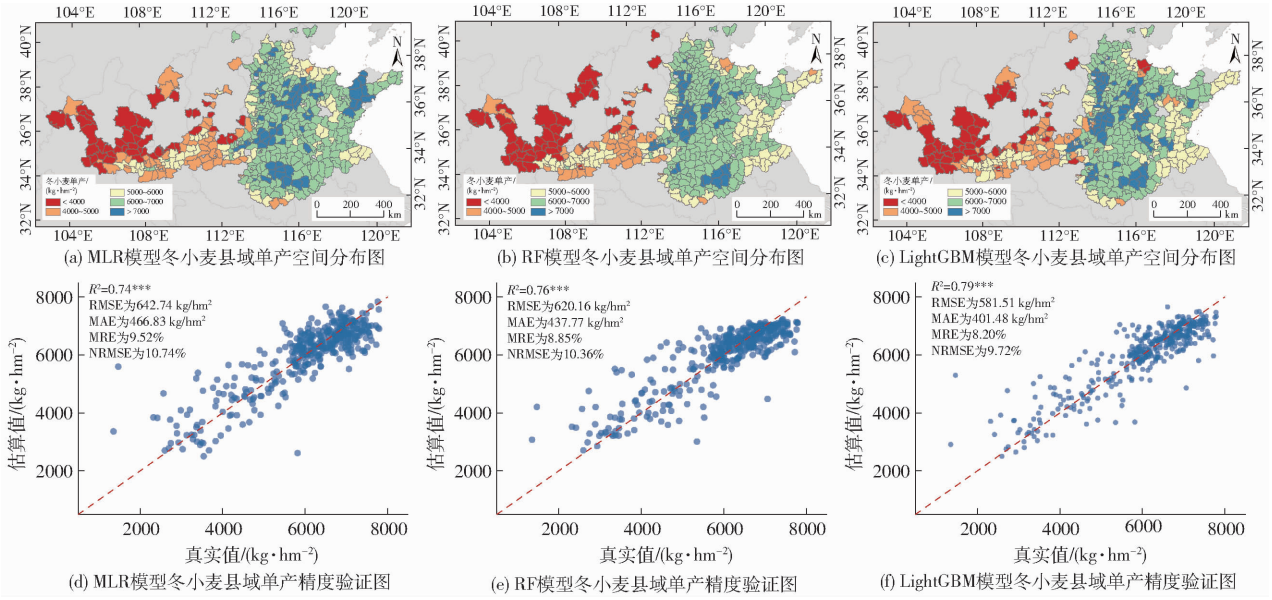


图 3　MLR、RF 和 LightGBM 模型冬小麦县域单产空间分布及精度对比

Fig. 3　Spatial distribution and accuracy comparison of county-level winter wheat yield estimation using MLR, RF, and LightGBM models

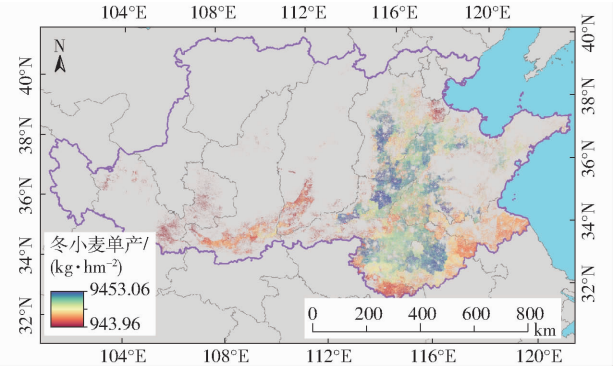


图 4　中国北方 30 m 分辨率冬小麦单产空间分布制图(2023 年)

Fig. 4　Mapping of 30 m resolution winter wheat yield spatial distribution in northern China (2023)

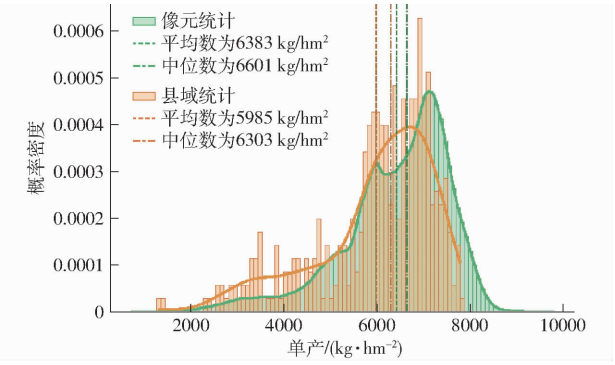


图 5　中国北方冬小麦单产空间分布制图结果的频率分布直方图(2023 年)

Fig. 5　Frequency distribution histogram of winter wheat yield downscaling results in northern China (2023)

原,包括河北省南部、河南省中部及北部、山东省西部和安徽省西北部,该区域地势平坦、灌溉条件良好、耕作制度稳定,单产普遍在 7 000 kg/hm² 以上,局部核心区超过 8 500 kg/hm²,形成连片高值斑块。中等产区多分布于平原向丘陵过渡地带,单产集中在 5 500 ~ 7 000 kg/hm² 之间;而低产区主要位于西部黄土丘陵区和秦岭北麓山区,受地形破碎、水资源条件和土壤肥力限制,冬小麦单产多低于 5 000 kg/hm²,且呈现明显的破碎化特征。另外沿海盐渍化区域由于土壤盐分和地下水位的影响,单产也相对较低。

像元单产频率分布呈现双峰形态,表明研究区存在两种不同单产水平。第 1 主峰位于 6 000 kg/hm² 附近,对应生态条件脆弱的雨养农业区或山地丘陵地区;第 2 主峰则在 7 100 kg/hm² 附近,对应于资源条件优越的灌溉农业主产区或平原等地势平坦地区。同时,直方图双峰均偏右,说明研究区以中高产水平为主导。4 000 kg/hm² 以下的低产像元占比较小,为 4.77%;8 000 kg/hm² 以上的高产像元比例约为 3.01%,6 000 ~ 7 000 kg/hm² 以及 7 000 ~ 8 000 kg/hm² 的占比最多,分别达到 33.85%、31.89%。与县域冬小麦单产统计数据分布情况相同,两者的单产中位数均大于平均数,整体分布均偏右,且都呈现双峰特点。另外,像元单产核密度曲线在 2 000 ~ 9 000 kg/hm² 区间内平滑波动,未出现明显的异常值或断层,且均值与中位数的差异较小,单产降尺度结果在整体上具有较好的稳定性。

基于野外实测样本对该降尺度结果进行验证,结果表明估算值与实测值具有较好的一致性(图 6),图中 *** 表示极显著($p < 0.001$)。定量评价结果显示,R^2 为 0.65,RMSE 为 623.32 kg/hm²,MAE 为 540.65 kg/hm²,相对误差结果中,MRE 为 8.45%,NRMSE 为 9.76%;从散点分布特征来看,

散点整体沿 1∶1 线分布，未表现出明显的系统性偏离，但存在高产低估情况。总体来看，在县域单产统计数据约束下构建的降尺度方法，其效果能满足研究区像元尺度冬小麦单产空间分布的需求。

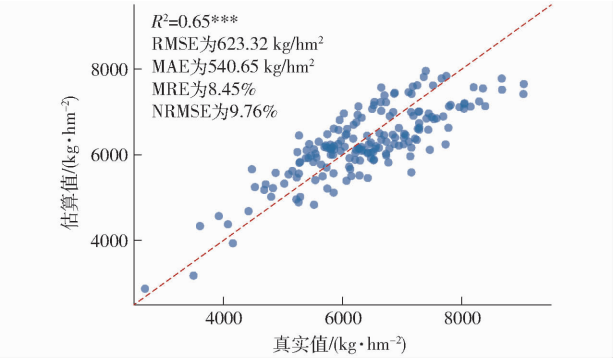


图 6 中国北方冬小麦单产降尺度结果与实测结果对比(2023 年)

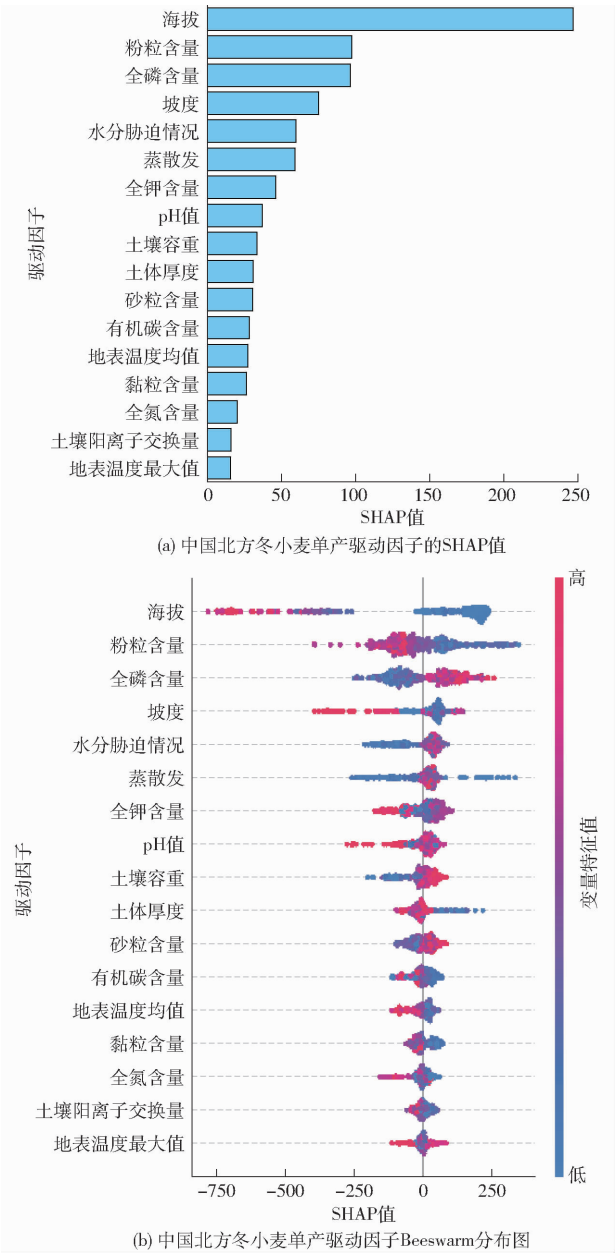
Fig. 6 Comparison between downscaled winter wheat yield results and field-measured data in northern China (2023)

3.3 冬小麦单产空间变化驱动因子分析

为揭示冬小麦单产统计数据降尺度结果空间分布差异的驱动因子，研究基于 SHAP 可解释性框架进行分析。由于植被指数仅指示冬小麦生长状况，而非形成单产空间差异的驱动因子，因此对气象、地形和土壤因素共 17 个指标的贡献度与作用方向进行定量解析。

基于 SHAP 值的平均绝对重要性排序(图 7a)，海拔是对冬小麦单产解释力最强的变量，其次为土壤粉粒含量、全磷含量、坡度和水分胁迫状况。上述变量累计贡献超过总解释力的 60%，表明地形条件、土壤颗粒组成、养分供应及水分胁迫是调控研究区冬小麦单产的主导因素；而全氮含量、阳离子交换量及地表温度最大值等变量的边际贡献相对较弱。

通过 SHAP 依赖图解析的单因子非线性调控效应显示(图 7b)，在地形方面，高海拔对应负向 SHAP 值，主要通过降低积温、增加低温胁迫抑制单产，而低海拔则凭借更优的热量条件正向促进单产[29]；陡坡因加剧水土流失、降低土壤保肥能力呈现负向贡献，缓坡则因水土条件稳定对单产具有正向调控作用[30]。在土壤颗粒组成方面，较高的砂粒含量对应正向 SHAP 值，这可能与砂粒为主的土壤透气性更佳、利于根系发育有关；而高粉粒与高黏粒含量则对应负向 SHAP 值，高粉粒易引发土壤板结、高黏粒会限制水分渗透，二者均不利于冬小麦根系伸展与水肥吸收[31]。在土壤养分方面，充足的氮磷钾含量可增强冬小麦光合效率与养分吸收[32]，对应正向 SHAP 值；过高的 pH 值会使土壤盐渍化严重，降低



(a) 中国北方冬小麦单产驱动因子的SHAP值

(b) 中国北方冬小麦单产驱动因子Beeswarm分布图

图 7 中国北方冬小麦单产驱动因子的 SHAP 值及 Beeswarm 分布图(2023 年)

Fig. 7 SHAP values and Beeswarm distribution of winter wheat yield driving factors in northern China (2023)

养分有效性，不利于单产形成[33]。在水分与热量条件方面，冬小麦水分满足程度不足会抑制冬小麦蒸腾与光合速率，对应负向 SHAP 值；而充足的蒸散反映冬小麦旺盛的生理活动，与单产正相关；地表温度过高会加速冬小麦呼吸消耗、引发热胁迫，从而降低单产。

高低产区的 SHAP 值对比进一步揭示驱动机制的异质性(图 8)。在低产区中，海拔、坡度的相对贡献显著高于其他变量，在低产区驱动因素中占据绝对优势地位，同时粉粒含量、土层厚度及蒸散发的重要性也较强，说明低产部分主要受不利地形和土壤条件的结构性约束。相反，在高产区内，虽然海拔仍

是最主要的驱动因素，但相对贡献占比相较于低产区明显下降，而氮、磷、钾、有机质等土壤养分状况的贡献占比上升。这表明，低产区应以减缓胁迫和改善立地条件为核心，高产区则应以培肥与养分精准调控为重点。这可为分区治理和差异化提升策略提供量化依据。

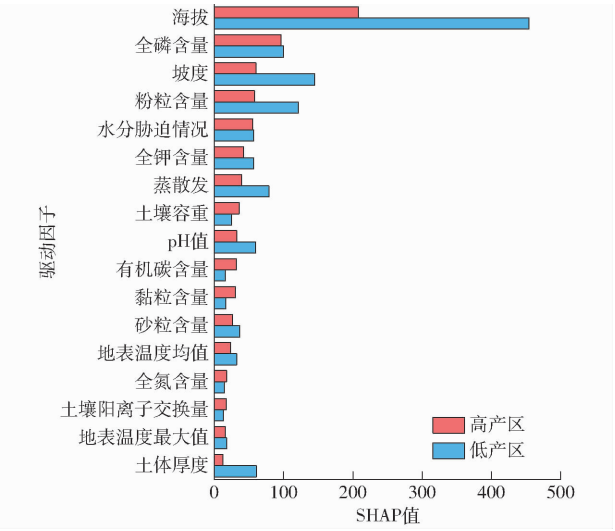


图 8　中国北方不同冬小麦单产水平下单产驱动因子的 SHAP 值对比（2023 年）

Fig. 8　Comparison of SHAP values of yield driving factors under different winter wheat yield levels in northern China (2023)

4　讨论

4.1　区域冬小麦单产统计数据降尺度制图方法优势

本研究构建基于生长状况和环境信息的县域尺度冬小麦单产模型，并在统计数据约束下开展单产降尺度，实现县域尺度平均单产向像元尺度单产的空间细化。不同于直接对像元单产进行回归预测的思路，研究将像元单产视为统计单产在空间上的再分配结果，通过构建模型得到环境驱动的相对单产权重，重构县域内部单产的空间格局，从而有效表达县域内部单产的空间异质性。野外实测样点验证结果表明，即使存在局部小幅偏差，该方法仍能解释 65% 的实测单产空间变化，可满足区域像元尺度单产空间制图的精度需求。

该方法的优势主要体现在 3 方面：统计一致性，以统计单产作为刚性约束，保证降尺度结果在县域尺度上与统计数据严格一致，避免纯数据驱动模型易出现的系统偏差；空间显式与可解释性，模型以冬小麦生长状况和环境因子为驱动，单产空间差异具有明确的物理与生态学意义，可揭示地形、土壤与气象条件对单产空间分布的差异化影响；能够刻画县域内部冬小麦单产的空间格局及异质性，在缺乏大规模实测样本的条件下，仍可进行像元尺度单产制图，为开展冬小麦差异化管理提供数据基础。

4.2　区域冬小麦单产统计数据降尺度制图方法应用价值

高分辨率、可解释的冬小麦单产空间信息是实施精准农业决策的重要前提。研究生成的 30 m 分辨率冬小麦单产空间分布，以及对应的驱动因子分析结果，可为从县域宏观调控到田块精准管理的决策提供科学支撑。

降尺度结果可清晰揭示县域内部单产空间分布异质性，同一县域内单产最大差异达 8 500 kg/hm^2 以上，弥补了传统县域平均数据掩盖田块尺度差异的短板，可直接作为精准管理分区、变量施肥灌溉的基础单元，有效解决传统“一刀切”管理导致的高产田养分不足、低产田水肥浪费的双重低效问题。结合 SHAP 驱动因子解析结果，可将研究区划分为 3 类区域并匹配相应的提升措施[34]：黄土丘陵与秦岭北麓低产胁迫区，重点改良地形土壤、缓解水分胁迫；平原-丘陵过渡带中产潜力区，强化灌溉配套与测土配方施肥；黄淮海平原高产优化区，开展养分精准调控与水肥耦合管理，规避通用管理方案的盲目性，最大化挖掘各区域冬小麦增产潜力。

此外，研究成果可满足不同农业经营主体的分层级决策需求：为农业管理部门的高标准农田布局、粮食生产功能区精细化划定提供数据支撑；为规模化种植经营主体优化田间管理方案、提升投入产出比提供依据；为基层农技推广部门与小农户精准识别低产制约因素、开展针对性田间改良提供靶向指导，助力小农户与现代农业有机衔接。

4.3　环境因子对冬小麦单产空间分异的影响

研究区跨越黄淮海平原、黄土高原等多种地貌单元，冬小麦单产空间变化受地形、土壤、气象的协同调控，最终呈现出东高西低、平原高而山地低的空间格局。这种空间分布格局，本质上是以地形为基础约束、土壤为肥力约束、气象为生境调控协同作用的结果，且在不同区域表现出显著差异。

在地形层面，海拔与坡度是决定冬小麦的主导因子。黄淮海平原地势低平，为水肥运移与机械化作业提供良好条件，单产主要受土壤与水热条件控制[35-38]。而黄土丘陵区与秦岭北麓山区，海拔较高、坡度较大，高海拔带来的积温不足、低温胁迫直接限制冬小麦光合积累与生育进程[29]，陡坡引发的水土流失则进一步加剧土壤保肥、保水能力下降[30]，此时地形因子通过影响气温、土壤养分等因素，成为单产的主导因子，其 SHAP 贡献占比超 20%。

在土壤层面,颗粒组成与养分含量决定单产水平。平原区土壤结构良好、养分充足,有利于形成高产[31];而西部低产区,粉粒比例偏高、养分流失严重,土壤结构与养分的双重劣势,进一步约束单产状况[30]。

气象因子中,水分胁迫(ET/PET)起关键调节作用,其强度与灌溉条件密切相关。在平原灌溉区,人工灌溉有效缓解了自然降水不足带来的水分胁迫,水分限制较弱;而雨养农业区中冬小麦生长完全依赖自然降水,干旱频发,成为制约单产提升的主要瓶颈。此外,沿海盐渍化区域虽水分条件充足,但土壤 pH 值偏高引发的盐渍化问题,降低了土壤养分有效性,形成特殊的低产类型。

4.4 研究局限

本研究充分利用农业统计数据,构建考虑环境因素的区域冬小麦单产统计数据降尺度制图方法,揭示冬小麦单产的空间异质性,并分析其驱动因子,但仍存在一定局限。研究仅基于 2023 年单年数据开展分析,冬小麦单产对气候年际变化的响应未充分体现,未来可增加多年份单产与环境数据,进一步提升模型的泛化能力与驱动因子分析的稳定性;同时可纳入耕作制度、农业投入等人为管理因子,完善单产空间变化驱动因子分析的完整性,为更精准的区域粮食安全保障提供技术支撑。

5 结论

(1)以冬小麦分布广泛的中国北方农业区为研究区域,综合 Sentinel-2、MODIS、土壤与地形等多源数据,建立基于统计数据约束的单产降尺度方法,实现县域尺度单产平均数据到像元尺度单产数据的降尺度制图,得到研究区 2023 年 30 m 分辨率冬小麦单产分布结果。

(2)基于生长状况和环境信息的最优县域尺度冬小麦单产模型为 LightGBM 模型,R^2 达到 0.79;由基于统计数据约束的单产降尺度方法得到的像元尺度单产数据 R^2 为 0.65;研究区冬小麦单产介于 943.96 ~ 9 453.06 kg/hm^2 之间,平均单产为 6 383 kg/hm^2,空间异质性显著,东高西低,平原高、山地低。

(3)海拔是对冬小麦单产统计数据降尺度结果空间分布差异影响最大的环境因子,其次为土壤粉粒含量、全磷含量、坡度和水分胁迫状况,累计贡献超过总解释力的 60%;低产地区主要受地形和土壤条件的结构性约束,高产地区驱动因素向氮、磷、钾、有机质等土壤养分状况转变。

参 考 文 献

[1] Blesh J, Hoey L, Jones A D, et al. Development pathways toward "zero hunger"[J]. World Development, 2019, 118: 1-14.

[2] Von Braun J, Afsana K, Fresco L O, et al. Food systems: seven priorities to end hunger and protect the planet[J]. Nature (London), 2021, 597(7874): 28-30.

[3] Joshi A, Pradhan B, Gite S, et al. Remote-sensing data and deep-learning techniques in crop mapping and yield prediction: a systematic review[J]. Remote Sensing, 2023, 15(8): 2014.

[4] 卢必慧,邱琳,李卫国,等. 基于遥感的作物产量和产量差估算研究进展[J]. 江苏农业学报,2023,39(3):881-894.
Lu Bihui, Qiu Lin, Li Weiguo, et al. Research progress on crop yield and yield gap estimation based on remote sensing[J]. Jiangsu Journal of Agricultural Sciences, 2023,39(3):881-894. (in Chinese)

[5] 吴炳方. 全国农情监测与估产的运行化遥感方法[J]. 地理学报, 2000, 55(1): 25-35.
Wu Bingfang. Operational remote sensing methods for agricultural statistics[J]. Acta Geographica Sinica, 2000, 55(1):25-35. (in Chinese)

[6] Wang S, Azzari G, Lobell D B. Crop type mapping without field-level labels: random forest transfer and unsupervised clustering techniques[J]. Remote Sensing of Environment, 2019, 222: 303-317.

[7] 王鹏新,杜江莉,张悦,等. 基于遥感多参数和 CNN-Transformer 的冬小麦单产估测[J]. 农业机械学报, 2024, 55(3): 173-182.
Wang Pengxin, Du Jiangli, Zhang Yue, et al. Yield estimation of winter wheat based on multiple remotely sensed parameters and CNN-Transformer[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(3): 173-182. (in Chinese)

[8] Wang J, Zhang J, Bai Y, et al. Integrating remote sensing-based process model with environmental zonation scheme to estimate rice yield gap in Northeast China[J]. Field Crops Research, 2020, 246: 107682.

[9] Lobell D B. The use of satellite data for crop yield gap analysis[J]. Field Crops Research, 2013, 143: 56-64.

[10] Curnel Y, De Wit A J W, Duveiller G, et al. Potential performances of remotely sensed LAI assimilation in WOFOST model based on an OSS experiment[J]. Agricultural and Forest Meteorology, 2011, 151(12): 1843-1855.

[11] Reichstein M, Camps-Valls G, Stevens B, et al. Deep learning and process understanding for data-driven earth system science[J]. Nature, 2019, 566(7743): 195-204.

[12] Becker-Reshef I, Justice C, Sullivan M, et al. Monitoring global croplands with coarse resolution earth observations: the Global Agriculture Monitoring (GLAM) project[J]. Remote Sensing, 2010, 2(6): 1589-1609.

[13] Shirsath P B, Sehgal V K, Aggarwal P K. Downscaling regional crop yields to local scale using remote sensing[J]. Agriculture, 2020, 10(3): 58.

[14] Wang J, Wang P, Tian H, et al. A deep learning framework combining CNN and GRU for improving wheat yield estimates using time series remotely sensed multi-variables[J]. Computers and Electronics in Agriculture, 2023, 206: 107705.
[15] Impollonia G, Croci M, Amaducci S. Upscaling and downscaling approaches for early season rice yield prediction using Sentinel-2 and machine learning for precision nitrogen fertilisation[J]. Computers and Electronics in Agriculture, 2024, 227: 109603.
[16] Hu Q, Yin H, Friedl M A, et al. Integrating coarse-resolution images and agricultural statistics to generate sub-pixel crop type maps and reconciled area estimates[J]. Remote Sensing of Environment, 2021, 258: 112365.
[17] Pittman K, Hansen M C, Becker-Reshef I, et al. Estimating global cropland extent with multi-year MODIS data[J]. Remote Sensing, 2010, 2(7): 1844-1863.
[18] Chang J, Hansen M C, Pittman K, et al. Corn and soybean mapping in the United States using MODIS time-series data sets [J]. Agronomy Journal, 2007, 99(6): 1654-1664.
[19] Xiao X, Boles 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.
[20] Liu F, Wu H, Zhao Y, et al. Mapping high resolution national soil information grids of China[J]. Science Bulletin, 2022, 67(3): 328-340.
[21] Dong J, Fu Y, Wang J, et al. Early season mapping of winter wheat in China based on Landsat and Sentinel images[J]. Earth System Science Data, 2020, 12(4): 3081-3095.
[22] 中华人民共和国农业部办公厅. 农业部办公厅关于印发《全国粮食高产创建测产验收办法(试行)》的通知[R]. 中华人民共和国农业部公报,2008(6):37-40.
[23] Muruganantham P, Wibowo S, Grandhi S, et al. A systematic literature review on crop yield prediction with deep learning and remote sensing[J]. Remote Sensing, 2022, 14(9): 1990.
[24] Cao B, Yu L, Zhong L, et al. Physics-guided deep learning for crop yield estimation[J]. European Journal of Agronomy, 2026, 172: 127850.
[25] 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.
[26] 王鹏新,韩弘炜,李明启,等. 基于 LSTM-AM 的冬小麦单产估测模型构建及其可解释性分析[J]. 农业机械学报, 2026,57(6): 206-214.
Wang Pengxin, Han Hongwei, Li Mingqi, et al. Construction and interpretability analysis of winter wheat yield estimation model based on LSTM-AM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2026,57(6): 206-214. (in Chinese)
[27] Tiwari P, Poudel K P, Yang J, et al. Marginal agricultural land identification in the Lower Mississippi Alluvial Valley based on remote sensing and machine learning model[J]. International Journal of Applied Earth Observation and Geoinformation, 2023, 125: 103568.
[28] Nohara Y, Matsumoto K, Soejima H, et al. Explanation of machine learning models using improved Shapley additive explanation[C]// Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, 2019.
[29] Prajapat A L, Saxena R, Sharma M, et al. Growing degree days requirement and yield of wheat cultivars as influenced by irrigation scheduling and time of sowing[J]. International Journal of Plant & Soil Science, 2022, 34(3): 28-35.
[30] Wang L, Li Y, Wu J, et al. Effects of the rainfall intensity and slope gradient on soil erosion and nitrogen loss on the sloping fields of Miyun reservoir[J]. Plants, 2023, 12(3): 423.
[31] Bhatt R, Kunal K, Garg A K, et al. Soil texture: unlocking fertility and productivity for better crops[J]. Agriculture Association of Textile Chemical and Critical Reviews, 2025, 13(3): 396-408.
[32] Guo J, Jia Y, Chen H, et al. Growth, photosynthesis, and nutrient uptake in wheat are affected by differences in nitrogen levels and forms and potassium supply[J]. Scientific Reports, 2019, 9(1): 1248.
[33] Hartemink A E, Barrow N J. Soil pH-nutrient relationships: the diagram[J]. Plant and Soil, 2023, 486(1): 209-215.
[34] Chen W, Kong X, Liao Y. Exploring differentiated improvement strategies of cultivated land quality in China[J]. Engineering Agriculture, 2024, 11(2): 282-291.
[35] Han J J, Xia X P, Liu H, et al. Assessment of the winter wheat yield gap for smallholder farmers in the Loess Plateau via boundary line analysis[J]. Plants, 2025, 14(21): 3375.
[36] 吴冰洁,王靖,唐建昭,等. 华北平原冬小麦产量变异的气象影响因子分析[J]. 中国农业气象, 2018, 39(10): 623-635.
Wu Bingjie, Wang Jing, Tang Jianzhao, et al. Meteorological influencing factors on variation in winter wheat yield in the North China Plain[J]. Chinese Journal of Agrometeorology, 2018, 39(10): 623-635. (in Chinese)
[37] 任贺靖,路凯超,蔡甲冰,等. 基于水热耦合的冬小麦-夏玉米产量响应与变化预测[J]. 农业机械学报, 2025,56(2): 429-443, 484.
Ren Hejing, Lu Kaichao, Cai Jiabing, et al. Yield responses and predictions of future change for winter wheat-summer maize based on water-heat coupling[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025,56(2): 429-443, 484. (in Chinese)
[38] Lu C, Fan L. Winter wheat yield potentials and yield gaps in the North China Plain[J]. Field Crops Research, 2013, 143: 98-105.