

# 基于遥感多参数和 VMD - GRU 的冬小麦单产估测

郭丰玮<sup>1,2</sup> 王鹏新<sup>1,2</sup> 刘峻明<sup>3</sup> 李红梅<sup>4</sup>

(1. 中国农业大学信息与电气工程学院, 北京 100083;

2. 农业农村部农机作业监测与大数据应用重点实验室, 北京 100083;

3. 中国农业大学土地科学与技术学院, 北京 100193; 4. 陕西省气象局, 西安 710014)

**摘要:** 为充分挖掘时间序列遥感参数的时序信息和趋势信息,并进一步提升冬小麦估产精度,以陕西省关中平原为研究区域,选取与冬小麦长势密切相关的生育时期尺度的条件植被温度指数(VTCI)、叶面积指数(LAI)和光合有效辐射吸收比率(FPAR)作为遥感参数,构建耦合变分模态分解(VMD)与门控循环单元(GRU)神经网络的估产模型。应用VMD算法将各个时间序列遥感参数分解为多组平稳的本征模态函数(IMF)分量,选取与原始时间序列遥感参数高度相关的IMF分量进行特征重构,并将重构特征作为GRU网络的输入,以构建冬小麦组合估产模型。结果表明,VMD-GRU组合估产模型决定系数为0.63,均方根误差为448.80 kg/hm<sup>2</sup>,平均相对误差为8.14%,相关性达到极显著水平( $P < 0.01$ ),其精度优于单一估产模型精度,表明该组合估产模型能够提取非平稳时间序列数据的多尺度、多层次特征,并充分挖掘冬小麦各生育时期遥感参数间的内在联系,获得准确单产估测结果的同时提升了估产模型的可解释性。

**关键词:** 冬小麦; 产量估测; 变分模态分解; 门控循环单元; 遥感参数

中图分类号: TP79

文献标识码: A

文章编号: 1000-1298(2024)01-0164-11

OSID:



## Yield Estimation of Winter Wheat Based on Multiple Remotely Sensed Parameters and VMD - GRU

GUO Fengwei<sup>1,2</sup> WANG Pengxin<sup>1,2</sup> LIU Junming<sup>3</sup> LI Hongmei<sup>4</sup>

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

2. Key Laboratory of Agricultural Machinery Monitoring and Big Data Applications,

Ministry of Agriculture and Rural Affairs, Beijing 100083, China

3. College of Land Science and Technology, China Agricultural University, Beijing 100193, China

4. Shaanxi Provincial Meteorological Bureau, Xi'an 710014, China)

**Abstract:** In order to fully exploit the time-series information and trend information of time-series remotely sensed parameters and further improve the yield estimation accuracy of winter wheat, vegetation temperature condition index (VTCI), leaf area index (LAI) and fraction of photosynthetically active radiation (FPAR), which were closely related to the growth and development of winter wheat, were selected as remotely sensed parameters, and a neural network was constructed based on variational mode decomposition (VMD) and gated recurrent unit (GRU). The VMD algorithm was applied to decompose each remotely sensed parameter series into multiple sets of intrinsic mode function (IMF) components, and the IMF components that were highly correlated with the original remotely sensed parameter series were selected for feature reconstruction, and the reconstructed features were used as the input of the GRU network to develop a combined model for yield estimation of winter wheat. The results showed that the VMD-GRU model for yield estimation had a coefficient of determination of 0.63, root mean squared error of 448.80 kg/hm<sup>2</sup>, and mean relative error of 8.14%, with a highly significant correlation level ( $P < 0.01$ ), and its accuracy was better than that of the single model for yield estimation, indicating that the combined model for yield estimation can extract multi-scale and multi-level features of non-stationary

收稿日期: 2023-06-22 修回日期: 2023-07-16

基金项目: 国家自然科学基金项目(42171332)

作者简介: 郭丰玮(1999—),男,博士生,主要从事定量遥感及其在农业中的应用研究,E-mail: fengwei\_guo@163.com

通信作者: 王鹏新(1965—),男,教授,博士,主要从事定量遥感及其在农业中的应用研究,E-mail: wangpx@cau.edu.cn

time series and fully explore the internal linkage between remotely sensed parameters in each growth stage of winter wheat to obtain accurate yield estimation results and improve interpretability of model for yield estimation.

**Key words:** winter wheat; yield estimation; variational mode decomposition; gated recurrent unit; remotely sensed parameter

## 0 引言

小麦作为主要口粮作物之一,是中国粮食产业的重要组成部分,对维护民生稳定和保障国家粮食安全有重要作用<sup>[1]</sup>。因此,在小麦收获前及时且准确地预测产量,对粮食政策制定、粮食市场调节和国民经济发展具有重大意义<sup>[2]</sup>。传统的作物产量估测方法主要依赖田间实地调查,过程繁琐且规模有限<sup>[3]</sup>。近年来,遥感技术凭借快速、简便、宏观、无损以及客观的优势广泛应用于农业生产领域<sup>[4]</sup>。目前遥感估产方法以数据同化和统计回归 2 种方法最为有效<sup>[5]</sup>。统计估产模型基于表征作物长势的遥感参数与产量间的回归关系建立经验模型,实现作物单产估测。相较于数据同化方法,传统统计回归方法更易基于遥感参数和单产数据建立估产模型,但在捕捉两者间的非线性关系方面仍存在挑战。

干旱会引起作物水分亏缺和 CO<sub>2</sub> 摄入量减少,影响作物生长发育和有机质合成,进而造成作物减产和长势变差<sup>[6]</sup>。王鹏新等<sup>[7]</sup>在归一化植被指数(Normalized difference vegetation index, NDVI)和地表温度(Land surface temperature, LST)的散点图呈三角形区域分布的基础上,提出了基于条件植被温度指数(Vegetation temperature condition index, VTCI)的近实时干旱监测方法,用以表征作物生长过程中的水分胁迫状况,已在关中平原地区取得广泛应用。除水分胁迫状况外,作物生长状态也与产量密切相关,其对产量的影响可以通过能反映作物生长过程的遥感参数进行表征。其中,叶面积指数(Leaf area index, LAI)和光合有效辐射吸收比率(Fraction of photosynthetically active radiation, FPAR)具有表征作物干物质积累量和光合作用的能力,是作物估产的重要指标。基于此,选用 VTCI、LAI 和 FPAR 作为冬小麦产量估测模型的输入参数,通过综合多个遥感参数表征的信息可实现更准确的作物单产估测。

作物产量的形成是其生长过程的综合表达,受到气候资源、土壤养分和田间管理等多方面的影响,通常为非线性过程<sup>[8]</sup>。随着机器学习和深度学习的快速发展,基于机器学习和深度学习的估产模型能够捕捉遥感参数与产量间的非线性关系,在作物

估产领域引起广泛关注<sup>[9]</sup>。作物单产估测研究通常使用作物主要生育期的时间序列遥感参数和产量数据,而长短期记忆(Long short-term memory, LSTM)神经网络在处理长时间序列数据问题时具有强大的记忆能力,收敛速度快<sup>[10]</sup>。ZHANG 等<sup>[11]</sup>结合植被指数和气候数据提出了基于 LSTM 的田间玉米产量预测方法。结果表明,该估产模型能够表征环境因素对产量的累积影响,获得了准确且稳定的估产结果。门控循环单元(Gated recurrent unit, GRU)在保证较高输出精度的基础上,简化了 LSTM 的结构,提高了模型计算效率<sup>[12]</sup>。WANG 等<sup>[13]</sup>利用 GRU 的时间序列记忆优势并结合卷积神经网络(Convolutional neural network, CNN)进行旬时间尺度的县级冬小麦产量估测,揭示了遥感参数和产量间的累积效应,提高了基于深度学习方法的估产模型的可解释性。然而,各原始时间序列遥感参数仍存在非线性、非平稳特性,若仅使用神经网络进行产量预测,存在原始特征信息利用率低、遥感参数深层信息挖掘不足的问题。变分模态分解(Variational mode decomposition, VMD)可以将原始序列分解为一组具有较强规律性的本征模态函数(Intrinsic mode function, IMF)分量,以充分且有效地提取原始序列的特征信息,并降低数据非平稳性<sup>[14]</sup>。文献[15]基于 VMD 方法分析了内蒙古近 34 年 NDVI 周期特征及其与气候的关系,以提取 NDVI 和气候因子的特征分量。结果表明,NDVI 与降水具有较好的周期一致性,且去除周期变量前后的 NDVI 均与气温、降水呈显著正相关。文献[16]提出将 VMD 与包络线去除(Continuum removal, CR)和欧氏距离(Euclidean distance, ED)相结合,构建农作物光谱信息监测模型,以实现相似光谱间不同程度污染的弱差信息的有效提取。结果表明,VMD 能够有效获取光谱信息,且 VMD-CR-ED 模型能够有效区分玉米叶片受污染的不同胁迫程度。上述研究表明,VMD 分解得到的 IMF 分量能够揭示原始序列在不同尺度上的变化规律,并蕴含详细的趋势和动态变化信息。深度学习模型强大的特征提取能力能够挖掘 VMD 分解序列的深层信息,因此将 VMD 与深度学习模型相结合开展研究成为当前热点之一。LU 等<sup>[17]</sup>以新疆 11 个站点 1993—2016 年的作物蒸散

量时间序列为研究对象,基于 VMD 和反向传播神经网络 (Back propagation neural network, BPNN) 建立作物蒸散量预测模型,以实现有效的短期干旱监测。结果表明,VMD 有效降低了日作物蒸散量时间序列的非平稳性,且基于 VMD-BPNN 组合模型得到的预测结果具有较高的准确性。褚江东等<sup>[18]</sup>基于分解-集成的思想构建 VMD-LSTM 模型,进行陆地水储量 (Terrestrial water storage, TWS) 的有效监测并探究了其时空变化分布及驱动因素,结果表明,地下水储量与 TWS 的变化趋势具有较好的空间相关性,且对西北干旱半干旱地区和华北地区的 TWS 变化占据主导作用,而湿润半湿润地区的 TWS 变化与土壤水储量和冠层水储量的相关性更强。目前 VMD 已广泛应用于机械故障诊断、电力预测和气象监测等领域<sup>[15-19]</sup>,但将 VMD 与深度学习和遥感技术相结合用以作物长势监测和估产的研究较少。此外,过往的研究大多直接应用时间序列遥感参数进行作物估产,对蕴含其中的趋势信息和深层信息探索较少。将序列分解技术与多变量时间序列预测模型相结合,以特征性强的重构序列作为输入,将有助于提高预测模型的可靠性和可解释性。

本文以关中平原为研究区域,以生育时期尺度的时间序列 VTCI、LAI 和 FPAR 遥感参数为研究对象,基于 VMD 方法进行序列分解,以揭示遥感参数在不同尺度上的变化规律,并降低原始序列的非平

稳性。此外,结合 GRU 网络处理时间序列的记忆优势,建立遥感参数序列深层信息与产量间的关系,探讨 VMD-GRU 组合模型估产表现,以期进一步提升冬小麦估产结果的精度,为农作物长势监测和单产估测提供新的思路和方法。

## 1 材料与方法

### 1.1 研究区概况

关中平原覆盖范围为  $106^{\circ}22' \sim 110^{\circ}24' E$ ,  $33^{\circ}57' \sim 35^{\circ}39' N$ ,包括西安市、宝鸡市、咸阳市、渭南市、铜川市和杨凌国家农业高新技术产业示范区 (图 1),总面积约为  $3.4 \times 10^4 \text{ km}^2$ ,平均海拔约 500 m。区域内年平均气温在  $6 \sim 13^{\circ}C$  之间,年平均降水量为 550 ~ 700 mm,属于典型的温带大陆性季风气候,由降水时空分布不均导致的干旱灾害时常发生<sup>[20]</sup>。该区域地势平坦,土壤肥沃,光照充足,是陕西省主要的粮食生产基地,其中灌溉区种植结构为冬小麦与夏玉米轮作,旱作区则以冬小麦种植为主。

冬小麦一般在 10 月上旬或中旬播种,次年 6 月上旬收获。结合关中平原作物实际生长情况,将越冬后冬小麦的主要生长阶段划分为返青期 (3 月上旬至 3 月中旬)、拔节期 (3 月下旬至 4 月中旬)、抽穗-灌浆期 (4 月下旬至 5 月上旬) 和乳熟期 (5 月中旬至 5 月下旬)<sup>[21]</sup>。

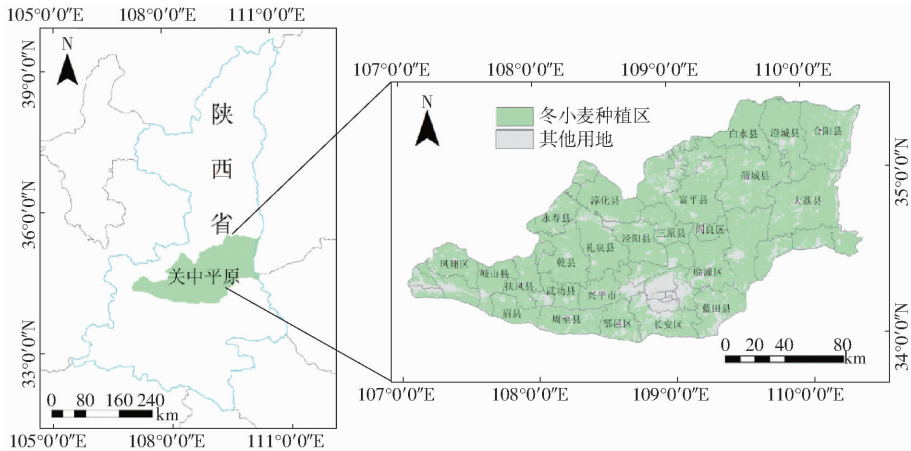


图 1 研究区域地理位置及作物种植区

Fig. 1 Geographical location of study area and crop planting areas

### 1.2 数据来源与预处理

#### 1.2.1 数据来源

采用的遥感数据为 MODIS 地表温度产品 (MYD11A1)、地表反射率产品 (MYD09GA)、叶面积指数及光合有效辐射吸收比率产品 (MCD15A3H) 和地表覆盖类型产品 (MCD12Q1)。所采用的 2011—2021 年各县冬小麦单产数据来源于其所在地级市统计局发布的统计年鉴。本文基于关中平原

2011—2021 年 24 个县的生育时期尺度的 VTCI、LAI 和 FPAR 及冬小麦单产数据开展研究。

#### 1.2.2 冬小麦种植区提取

采用 Terra 和 Aqua 卫星上搭载的 MODIS 传感器获取的地表覆盖类型产品 (MCD12Q1) 进行冬小麦种植区的提取。利用 MRT (MODIS reprojection tools) 工具对该产品进行预处理,并叠加研究区行政边界矢量图,选择国际地圈-生物圈计划

(International Geosphere - Biosphere Programme, IGBP) 分类方案中代码 14 对应的农田/自然植被嵌入土地覆盖类型的提取结果, 得到关中平原各县(区)冬小麦种植区(图 1)。

### 1.2.3 生育时期 VTCI 计算

选取 2011—2021 年关中平原每年 3—5 月的 Aqua - MODIS 日地表温度产品 (MYD11A1) 和地表反射率产品 (MYD09GA), 经 MRT 进行投影转换、重采样、影像拼接和裁剪等操作, 获取空间分辨率为 1 km、时间分辨率为 1 d 的 LST 和 NDVI 产品数据。以旬为时间单位, 应用最大值合成技术分别生成旬时间尺度的 LST 和 NDVI 最大值合成产品; 基于多年某一旬的 LST 和 NDVI 最大值合成产品应用最大值合成技术分别生成多年旬尺度 LST 和 NDVI 最大值合成产品。基于多年某一旬的 LST 最大值合成产品应用最小值合成技术生成多年旬尺度 LST 最大-最小值合成产品; 基于多年旬尺度 LST 和 NDVI 最大值合成产品确定研究区域 VTCI 的热边界, 基于多年旬尺度 LST 最大-最小值合成产品和多年旬尺度 NDVI 最大值合成产品确定研究区域 VTCI 的冷边界, 并以此计算旬尺度 VTCI<sup>[7,22]</sup>

$$VTCI = \frac{L_{\max}(N_i) - L(N_i)}{L_{\max}(N_i) - L_{\min}(N_i)} \quad (1)$$

$$\text{其中} \quad L_{\max}(N_i) = a + bN_i \quad (2)$$

$$L_{\min}(N_i) = a' + b'N_i \quad (3)$$

式中 VTCI——条件植被温度指数

$N$ ——归一化植被指数

$L$ ——地表温度

$L_{\max}(N_i)$ 、 $L_{\min}(N_i)$ ——在研究区域内, 当 NDVI 等于  $N_i$  时的所有像素 LST 的最大值和最小值

$L(N_i)$ ——某一像素的 NDVI 值为  $N_i$  时的 LST

$a$ 、 $b$ 、 $a'$ 、 $b'$ ——待定系数, 由研究区域 NDVI 和 LST 散点图近似获得

根据冬小麦生育时期的划分, 逐像素取某一生育时期内包含的多个旬尺度 VTCI 的平均值作为该生育时期的 VTCI 值; 基于关中平原各县(区)冬小麦种植区分布图, 对各县(区)所包含像素的 VTCI 值取平均值作为该县(区)该年该生育时期的 VTCI 值。

### 1.2.4 生育时期 LAI 和 FPAR 计算

选取 2011—2021 年关中平原每年 3—5 月的 Aqua - MODIS 叶面积指数及光合有效辐射吸收比率产品 (MCD15A3H), 获取空间分辨率为 500 m、时

间分辨率为 4 d 的 LAI 和 FPAR 产品数据, 并基于上包络线 Savitzky - Golay (S - G) 滤波对原始 LAI 和 FPAR 数据进行平滑处理, 以消除因云和大气产生的噪声, 经 S - G 滤波平滑处理后的 LAI 和 FPAR 数据更符合冬小麦的生长情况<sup>[23]</sup>。

VTCI 和 FPAR 的取值范围均为  $[0, 1]$ , 为使 LAI 与两者具有相同的取值范围, 对经 S - G 滤波处理后的 LAI 进行归一化处理。应用最大值合成技术, 逐像素取每旬包含的多时相 LAI 和 FPAR 的最大值作为该旬 LAI 值和 FPAR 值, 逐像素取各生育时期所包含的多旬 LAI 和 FPAR 的最大值作为该生育时期 LAI 值和 FPAR 值。根据关中平原各县(区)冬小麦种植分布图, 对各县(区)包含像素的 LAI 和 FPAR 取平均值作为该县(区)该年该生育时期的 LAI 和 FPAR 值。

VTCI 产品的空间分辨率为 1 km, 而 LAI 和 FPAR 产品的空间分辨率为 500 m, 为使三者空间分辨率一致, 基于最近邻法对 VTCI 进行重采样, 以获取空间分辨率为 500 m 的 VTCI 产品。此外, 冬小麦产量取值范围为  $[2\ 000, 7\ 000]$  kg/hm<sup>2</sup>, 为避免因特征与标签量纲不同而给模型训练带来的干扰, 对冬小麦单产数据进行归一化处理, 将所有数据取值范围统一为  $[0, 1]$ 。

### 1.2.5 各时间序列遥感参数及单产生成

以县为单位, 按照年份顺序构建各县 2011—2021 年冬小麦 4 个生育时期时间序列遥感参数及单产。根据各县在关中平原的空间分布情况, 按照自西向东的方位依次串联各县 11 年各生育时期时间序列遥感参数及单产, 将各县时间序列遥感参数及单产纳入关中平原整体的时间进程之中, 最终形成 24 县 11 年 4 个生育时期时间序列 VTCI、LAI、FPAR 和冬小麦单产, 长度为 264 a, 时间间隔为 1 a, 共计 13 个时间序列。

## 1.3 研究方法

### 1.3.1 变分模态分解

VMD 是一种自适应、非递归的信号分解方法, 其自适应性体现在可以根据输入信号的特性确定所给序列的模态分解个数  $K$ , 并匹配每种模态的最优中心频率的有限带宽, 以便平衡处理序列各部分存在的噪声<sup>[19]</sup>。VMD 的实质是变分问题, 通过变分约束条件, 自适应地将输入信号  $x(t)$  同时分解为  $K$  个有限带宽的 IMF 分量  $u(t)$ , 同时要求最小化所有 IMF 分量的带宽之和。具体过程可分为变分问题的构造过程与求解过程两部分<sup>[24]</sup>。

#### (1) 构造过程

对于每个模态函数  $u_k(t)$ , 通过希尔伯特

(Hillbert)变换得到对应的解析函数,用以获得其单边频谱,通过混合解析函数与所对应的中心频率指数项,将 IMF 分量的频谱移动到基频带,模态函数  $u_k(t)$  带约束的变分问题模型为

$$\begin{cases} \min_{|u_k|, |\omega_k|} \left\{ \sum_{k=1}^K \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] * e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s. t. } \sum_{k=1}^K u_k(t) = x(t) \end{cases} \quad (4)$$

式中  $\partial_t(\cdot)$ ——时间  $t$  的偏导数  
 $\delta(t)$ ——单位脉冲函数  
 $*$ ——卷积运算符号  
 $\omega_k$ ——频率

(2) 求解过程

为方便求取最优解,引入拉格朗日乘子和二次惩罚因子对式(4)的变分问题模型进行重构。其中,拉格朗日乘子保证约束条件的严格性,二次惩罚因子有助于算法收敛和提升重构精度。重构后无约束变分问题模型为

$$L_{u_k, \omega_k, \lambda} = \alpha \sum_{k=1}^K \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] * e^{-j\omega_k t} \right\|_2^2 + \left\| x(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 + \left\langle \lambda(t), x(t) - \sum_{k=1}^K u_k(t) \right\rangle \quad (5)$$

式中  $\alpha$ ——惩罚因子  
 $\lambda$ ——拉格朗日算子

基于交替方向乘子法对 IMF 分量、中心频率以及拉格朗日算子不断迭代更新,求取无约束变分问题的最优解,并根据式(6)~(8)更新 IMF 分量、中心频率和拉格朗日算子。

$$\hat{u}_k^{n+1}(\omega) = \frac{x(\omega) - \sum_{k=1}^K \hat{u}_k^{n+1}(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \quad (6)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \quad (7)$$

$$\lambda^{(n+1)}(\omega) = \lambda^{(n)}(\omega) + \tau \left( x(\omega) - \sum_{k=1}^K \hat{u}_k^{n+1}(\omega) \right) \quad (8)$$

式中  $\hat{u}_k^{n+1}(\omega)$ —— $u_k(t)$  在傅里叶域第  $n$  次的迭代值  
 $\omega_k^{n+1}$ —— $\omega_k$  第  $n+1$  次的迭代值  
 $x(\omega)$ —— $x(t)$  的傅里叶变换  
 $\lambda^{(n+1)}(\omega)$ —— $\lambda(t)$  在傅里叶域第  $n$  次的迭代值

$\tau$ ——噪声容忍度  $\omega$ ——频率

当满足

$$\frac{\sum_{k=1}^K \left\| (\hat{u}_k^{n+1}(\omega) - \hat{u}_k^n(\omega)) \right\|_2^2}{\left\| \hat{u}_k^n(\omega) \right\|_2^2} < \varepsilon \quad (9)$$

式中  $\varepsilon$ ——收敛精度,取  $10^{-7}$   
 算法停止。

VMD 算法使用的参数有:模态分解数  $K$ 、惩罚因子  $\alpha$ 、收敛精度  $\varepsilon$ 、噪声容忍度  $\tau$  和中心频率初始值  $init$ 。基于经验对除模态分解数  $K$  外的 4 个参数取值,其中  $\alpha$  为 2 000,  $\varepsilon$  为  $10^{-7}$ ,  $\tau$  为 0,  $init$  为 1。模态分解数  $K$  为 VMD 算法的重要参数,且要在算法开始前确定。当  $K$  过小时,原始序列中的部分重要成分被过滤,有效信息未被充分挖掘;当  $K$  过大时,相邻的 IMF 分量的中心频率重叠,会产生多余的噪声。为优化 VMD 序列分解结果,本文基于中心频率法<sup>[25]</sup>寻找最优模态分解数  $K$ 。其核心思想为:令  $K$  在参数范围内从小到大取值,分析各个 IMF 分量的中心频率是否处于稳定状态。在  $K$  增大的过程中,当最后一个 IMF 分量的中心频率保持相对稳定时,此时的  $K$  即可认为是最优分解模态数。

1.3.2 GRU 神经网络

作为 LSTM 的一种变体,GRU 凭借其门控机制(重置门和更新门)及递归结构,在处理时间序列数据的长期依赖问题时具有显著优势。重置门用于决定历史信息的遗忘程度,数值越小,被遗忘的历史信息越多。更新门决定信息输入和历史状态信息的保留,数值越大,被传递的历史信息越多<sup>[26]</sup>。GRU 通过当前时刻的输入信息  $x_t$  和上一时刻的隐藏状态信息  $h_{t-1}$ ,得出当前时刻的隐藏状态信息  $h_t$ (图 2)。GRU 的门控机制和递归结构能够依据冬小麦先前的生长信息和当前的输入信息进行自我更新,以充分挖掘各生育时期遥感参数与产量间的内在联系,从而有效表征作物生长过程中的依赖关系。GRU 模型计算式为<sup>[12,26]</sup>

$$z_t = \sigma(w_z \cdot [h_{t-1}, x_t] + b_z) \quad (10)$$

$$r_t = \sigma(w_r \cdot [h_{t-1}, x_t] + b_r) \quad (11)$$

$$\tilde{h}_t = \tanh(w_h \cdot [r_t \cdot h_{t-1}, x_t] + b_h) \quad (12)$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \quad (13)$$

式中  $\sigma$ ——Sigmoid 函数

$z_t$ —— $t$  时刻更新门状态

$r_t$ —— $t$  时刻重置门状态

$h_{t-1}$ —— $t-1$  时刻的隐藏状态信息

$h_t$ —— $t$  时刻的隐藏状态信息

$\tilde{h}_t$ —— $t$  时刻的候选隐藏状态信息

$b_z, b_r, b_h$ ——各函数偏移  
 $x_t$ —— $t$ 时刻模型输入信息  
 $w_z$ ——更新门中与  $h_{t-1}$  和  $x_t$  连接的权重矩阵  
 $w_r$ ——重置门中与  $h_{t-1}$  和  $x_t$  连接的权重矩阵  
 $w_h$ ——候选隐藏层中与  $h_{t-1}$  和  $x_t$  连接的权重矩阵

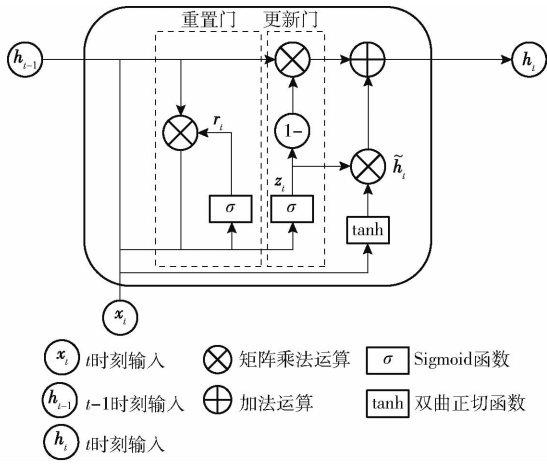


图 2 门控循环单元的结构

Fig. 2 Structure of gated recurrent unit

### 1.3.3 相关性检验

相关性检验是验证变量间相关程度或共有特征的有效手段。为量化各 IMF 分量与原始时间序列遥感参数的相关程度,本文选择皮尔逊相关系数进行各 IMF 分量与原始时间序列遥感参数间的相关性检验。当相关系数不小于 0.30 时,此 IMF 分量被认为与原始时间序列遥感参数间存在一定的相关性。基于此,可确定用于序列重构的有效 IMF 分量<sup>[27]</sup>。

### 1.4 VMD - GRU 估产模型

基于 VMD 提取非平稳时间序列中蕴含的多尺度、多层次特征,并利用 GRU 捕获时间序列中的长期依赖关系能力构建 VMD - GRU 冬小麦单产估测模型。本文基于中心频率法和经验取值确定 VMD 分解模态数  $K$  和其余参数,并对 4 个生育时期的时间序列 VTCI、LAI 和 FPAR 遥感参数进行 VMD 分解,获得多组平稳 IMF 分量。基于各 IMF 分量和原始时间序列遥感参数间的皮尔逊相关系数进行相关性检验,确定有效 IMF 分量,并进行序列重构。具体公式为

$$f = \sum_{i=1}^M IMF_i \quad (14)$$

式中  $f$ ——重构序列  $M$ ——有效模态数

$IMF_i$ ——第  $i$  个模态分量

此外,为减少重构特征序列的数据分布对估产模型的影响,采用 Min - Max 法对重构序列进行归一化,以提高估产模型训练速度和预测精度。基于此,以 4 个生育时期的 VTCI、LAI 和 FPAR 的归一化

重构特征序列作为模型的输入,以关中平原各县(区)的单产数据作为标签,进行 VMD - GRU 估产模型构建。

基于 Keras 框架搭建 GRU 模型,GRU 模型包含 1 层输入层、2 层 GRU 层、1 层全连接层和 1 层输出层(图 3)。为了防止过拟合现象对预测结果的影响,在第 1 层 GRU 层和全连接层前加入 Dropout 层,丢弃率分别设置为 0.30 和 0.25,用以降低模型训练过程的复杂性。重构后的归一化特征以时间序列  $\{x_1, x_2, \dots, x_t\}$  的形式输入到 GRU 模型中,其中  $x_t$  为冬小麦 4 个生育时期(返青期、拔节期、抽穗-灌浆期和乳熟期)的遥感参数(VTCI、LAI 和 FPAR)组成的向量。将 2011—2021 年 VTCI、LAI、FPAR 遥感参数重构特征序列及对应的单产数据按照比例 8:2 随机划分训练集和测试集。经超参数寻优和网络结构调整后,确定最大迭代次数为 500,GRU 神经元个数为 64,激活函数为 ReLU,学习率为 0.001,模型优化方法为 Adam,损失函数为均方误差。基于以上最优参数训练估产模型并得到冬小麦单产估测结果。

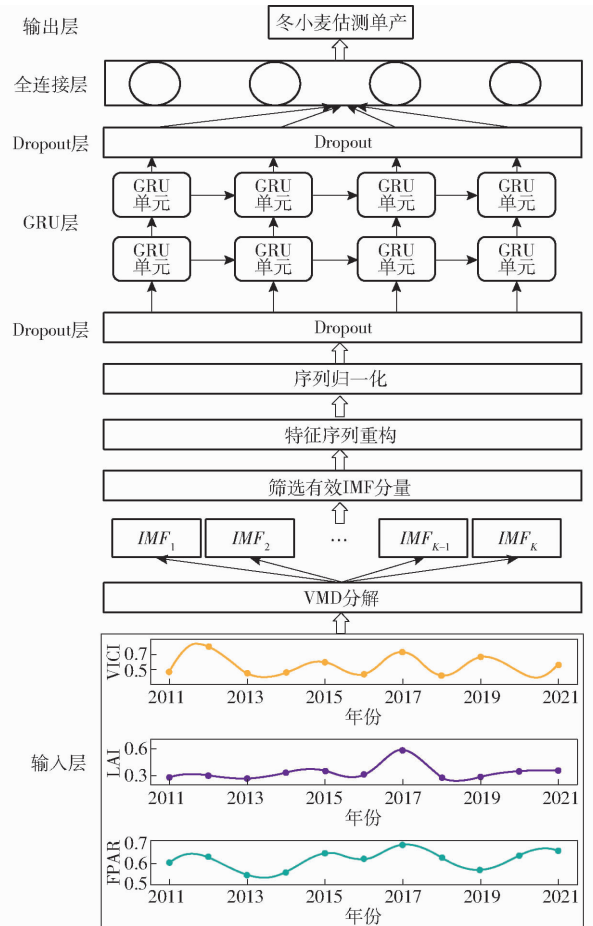


图 3 VMD - GRU 估产模型结构图

Fig. 3 Structure of VMD - GRU model for yield estimation

### 1.5 冬小麦单产估测模型性能评估

选用决定系数 ( $R^2$ )、均方根误差 (RMSE) 和平

均相对误差 (MRE) 作为冬小麦估产模型的精度评价指标。通过计算并对比单一 GRU 估产模型、单一 VMD 估产模型和 VMD-GRU 组合模型输出的估产结果与实际单产间的  $R^2$ 、RMSE 和 MRE 在同一测试集上的精度, 得出最优估产模型, 并对陕西省关中平原 2016—2022 年的冬小麦单产进行逐像素模拟, 以分析估产结果的空间分布特点。

## 2 结果与分析

### 2.1 VMD 序列分解

#### 2.1.1 模态分解数寻优

将模态分解数  $K$  的范围设置为  $[1, 8]$ , 以抽穗-灌浆期时间序列 VTCI 为例, 基于不同  $K$  设置对该时间序列进行 VMD 分解。表 1 是对应  $K$  下的各 IMF 分量的中心频率, 可以看出, 当  $K$  取 1~5 时, 最后一个 IMF 分量中心频率分别为  $1.35 \times 10^{-5}$ 、0.272、0.273、0.364、0.454 Hz, 有明显上升趋势, 说明最后一个 IMF 分量的中心频率没有达到稳定状态, 原始序列处于欠分解状态。而当  $K$  取 5~8 时, 最后一个 IMF 分量的中心频率均为 0.454 Hz, 表明当  $K \geq 5$  时, 最后一个 IMF 分量的中心频率相对稳定。此时, 若继续增大  $K$ , 分解得到的 IMF 分量增多, 会出现模态重叠, 造成序列过分解。因此, 应用中心频率法确定抽穗-灌浆期时间序列 VTCI 的最优模态分解数  $K$  为 5。

表 1 抽穗-灌浆期时间序列 VTCI 各 IMF 分量的中心频率

Tab. 1 Center frequency of each IMF of VTCI series at heading-filling stage Hz

K	模态分量							
	IMF <sub>1</sub>	IMF <sub>2</sub>	IMF <sub>3</sub>	IMF <sub>4</sub>	IMF <sub>5</sub>	IMF <sub>6</sub>	IMF <sub>7</sub>	IMF <sub>8</sub>
1	$1.35 \times 10^{-5}$							
2	$1.34 \times 10^{-5}$	0.272						
3	$1.32 \times 10^{-5}$	0.183	0.273					
4	$1.29 \times 10^{-5}$	0.091	0.272	0.364				
5	$1.26 \times 10^{-5}$	0.091	0.183	0.272	0.454			
6	$1.26 \times 10^{-5}$	0.091	0.183	0.272	0.363	0.454		
7	$1.25 \times 10^{-5}$	0.091	0.183	0.250	0.273	0.363	0.454	
8	$1.23 \times 10^{-5}$	0.090	0.114	0.183	0.272	0.322	0.364	0.454

基于上述过程, 得到 4 个生育时期的时间序列 VTCI、LAI 和 FPAR 的最优模态分解数  $K$  (表 2), 可以看出, 同一遥感参数的取值大致相同, 且基于中心频率法确定的 4 个生育时期的 LAI 和 FPAR 序列的最优模态分解数  $K$  一致, 这可能是因为 LAI 和 FPAR 来自同一遥感产品, 反演过程大致相同, 具有高度相似性和相关性。

表 2 各时间序列遥感参数的最优  $K$  值

Tab. 2 Optimal  $K$  values of each remotely sensed parameter series

遥感参数	返青期	拔节期	抽穗-灌浆期	乳熟期
VTCI	5	6	5	5
LAI	6	6	6	6
FPAR	6	6	6	6

#### 2.1.2 序列分解结果与分析

基于最优参数设置下的 VMD 算法, 对 4 个生育时期的时间序列 VTCI、LAI 和 FPAR 进行分解, 得到各组 IMF 分量的时域图 (图 4)。结果表明, 各遥感参数序列经 VMD 分解后得到的 IMF 分量频率依次增大, 较原始序列非平稳性明显降低。各生育时期时间序列 LAI 和 FPAR 的同一 IMF 分量具有相似性, 特别是 IMF<sub>3</sub>、IMF<sub>5</sub> 和 IMF<sub>6</sub> 分量, 序列重合率较高。而各生育时期时间序列 VTCI 受模态分解数  $K$  不同和波动程度的影响, 其分解得到的 IMF 分量表现出独立的无规则性, 这一点在高频分量中表现尤为突出。

每个 IMF 分量能够独立表征原始序列中蕴含的非线性趋势变化信息。由图 4 可以看出, IMF<sub>1</sub> 和 IMF<sub>2</sub> 能够反映原始时间序列的均值或趋势信息, 其余分量蕴含对应时刻的周期变化信息或噪声信息, 表明 VMD 可以有效地提取原始序列的特征信息, 并降低特征序列的非平稳性, 经分解得到 IMF 分量能够揭示原始序列在不同特征尺度上固有的变化规律, 并蕴含了详细的趋势和动态变化信息。

### 2.2 有效 IMF 分量选取

VMD 分解后的 IMF 分量蕴含了原始序列的均值信息、趋势信息、局部特征和深层信息, 相较于原始时间序列遥感参数, 其特征性更强。然而, 由于参数设置和序列本身的复杂性, 分解得到的 IMF 分量中不可避免地存在无关噪声。为进一步降低无关噪声对模型训练过程和预测精度的影响<sup>[25]</sup>, 采用皮尔逊相关系数量化各 IMF 分量与原始序列间的相关程度 (表 3), 以 0.3 为相关系数阈值, 确定用于重构的有效 IMF 分量。结果显示, 时间序列 LAI 和 FPAR 的低频分量与原始时间序列遥感参数的相关性高, 其中, IMF<sub>1</sub> 分量相关性最大, IMF<sub>2</sub> 分量次之。4 个生育时期的时间序列 VTCI 则表现出与 LAI 和 FPAR 不同的特性, 特别是返青期时间序列 VTCI, 其与各 IMF 分量间的相关系数逐渐增大, IMF<sub>5</sub> 分量与原始序列具有最强的相关性, 这可归因为时间序列 VTCI 本身能够表征作物水分胁迫情况, 对环境变化的敏感性强, 易受降水年际变化的显著影响, 序列波动性强, 导致表征均值信息和趋势信息的低频分量

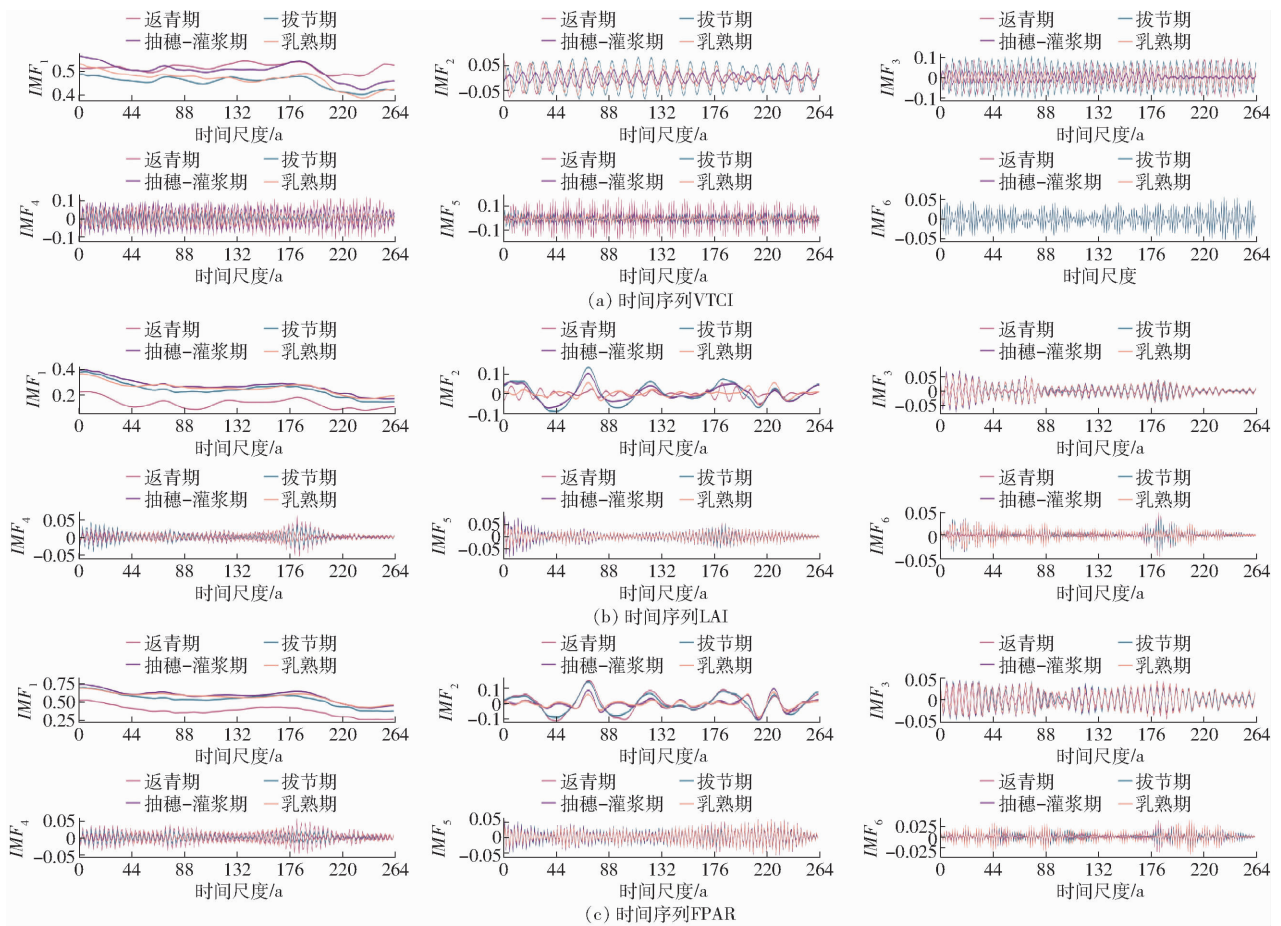


图 4 各时间序列遥感参数分解结果

Fig. 4 Decomposition result of each remotely sensed parameter series

表 3 各时间序列遥感参数与 IMF 分量的相关系数

Tab. 3 Correlation coefficients of each remotely sensed parameter series with IMF components

遥感参数	生育时期	IMF 分量						有效 IMF 分量
		IMF <sub>1</sub>	IMF <sub>2</sub>	IMF <sub>3</sub>	IMF <sub>4</sub>	IMF <sub>5</sub>	IMF <sub>6</sub>	
VTCI	返青期	0.166	0.222	0.368	0.453	0.734		IMF <sub>3</sub> , IMF <sub>4</sub> , IMF <sub>5</sub>
	拔节期	0.273	0.505	0.613	0.279	0.351	0.284	IMF <sub>2</sub> , IMF <sub>3</sub> , IMF <sub>5</sub>
	抽穗-灌浆期	0.444	0.258	0.228	0.650	0.432		IMF <sub>1</sub> , IMF <sub>4</sub> , IMF <sub>5</sub>
	乳熟期	0.483	0.472	0.378	0.449	0.376		IMF <sub>1</sub> , IMF <sub>2</sub> , IMF <sub>3</sub> , IMF <sub>4</sub> , IMF <sub>5</sub>
LAI	返青期	0.765	0.397	0.258	0.287	0.278	0.205	IMF <sub>1</sub> , IMF <sub>2</sub>
	拔节期	0.694	0.651	0.264	0.172	0.258	0.159	IMF <sub>1</sub> , IMF <sub>2</sub>
	抽穗-灌浆期	0.715	0.598	0.295	0.149	0.285	0.161	IMF <sub>1</sub> , IMF <sub>2</sub>
	乳熟期	0.736	0.421	0.356	0.186	0.299	0.227	IMF <sub>1</sub> , IMF <sub>2</sub> , IMF <sub>3</sub>
FPAR	返青期	0.703	0.679	0.201	0.199	0.190	0.119	IMF <sub>1</sub> , IMF <sub>2</sub>
	拔节期	0.759	0.629	0.207	0.129	0.225	0.116	IMF <sub>1</sub> , IMF <sub>2</sub>
	抽穗-灌浆期	0.819	0.467	0.219	0.130	0.260	0.151	IMF <sub>1</sub> , IMF <sub>2</sub>
	乳熟期	0.795	0.416	0.254	0.162	0.284	0.206	IMF <sub>1</sub> , IMF <sub>2</sub>

难以描述原始序列中的主要信息,而高频分量的波形和走势更符合原始时间序列 VTCI 的变化趋势。

将有效 IMF 分量进行叠加重构得到的特征序列,既能降低原始序列的非平稳性,又保留了蕴含冬小麦生长的多尺度变化规律的特征信息,以此构建冬小麦估产模型,有助于提升估产模型的可靠性和可解释性。

### 2.3 冬小麦估产模型构建和精度评价

为进一步挖掘冬小麦各个生长阶段遥感参数与单产间的内在联系,进而有效表征作物生长过程中的依赖关系,将归一化的重构特征序列和单产数据作为 GRU 模型的输入,以构建 VMD-GRU 冬小麦估产模型。为进一步直观描述估产模型的性能,将 VMD-GRU 估产模型与单一 GRU 估产模型和单一



VMD 估产模型在同一测试集下的估产结果与实际单产进行回归分析(图5)。结果显示,VMD-GRU 组合估产模型的  $R^2$ 、RMSE 和 MRE 均优于单一 VMD 估产模型和单一 GRU 估产模型,分别为 0.63、448.80 kg/hm<sup>2</sup> 和 8.14%。相较于 GRU 估产模型,组合估产模型的 RMSE 降低 53.30 kg/hm<sup>2</sup>,MRE 降

低 0.18 个百分点。结果表明,相较于波动性较强的原始时间序列遥感参数,经 VMD 分解和特征筛选、重构后的序列平稳性更强,时序特征更易提取。因此,VMD-GRU 估产模型既能提取非平稳时间序列遥感参数的多尺度、多层次特征,又能捕获蕴含其中的长期依赖关系,在县域尺度估产方面可靠性更强。

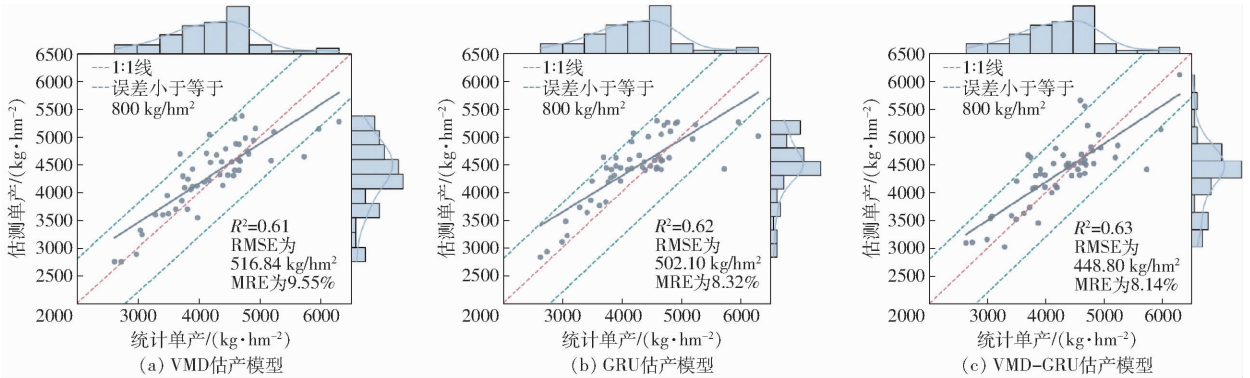


图5 实际单产和估测单产回归分析图

Fig.5 Regression analysis of actual yields and estimated yields

从散点图可看出,3 个模型的估测单产与实际单产基本拟合 1:1 线,但 VMD-GRU 估产模型的斜率更接近 1。原因为组合估产模型在低产县和高产县有更准确的估产结果,学习到更多蕴含在低产样本和高产样本中的特征,但当产量大于 5 000 kg/hm<sup>2</sup> 时,估产模型精度有所下降,仍存在高产低估现象。这应归因于本文使用的数据中高于 5 000 kg/hm<sup>2</sup> 的样本较少,数据样本不平衡导致在估产模型训练时,高产样本的估产结果趋向于中等产量。这也反映出深度学习模型的有效训练需要大量样本,样本不足或不均衡可能会降低模型精度。散点图上方和右侧分别代表实际单产和估测单产的分布直方图和概率密度曲线,可以看出,VMD-GRU 估产模型的估产范围在 3 000 ~ 6 000 kg/hm<sup>2</sup> 之间,相较于单一 VMD 估产模型和单一 GRU 估产模型,更符合测试集中实际单产的分布情况。

综上所述,经 VMD 分解、筛选和重构后得到的特征序列既能降低原始时间序列遥感参数的非平稳特性,又保留了其中与作物长势相关的趋势信息和深层信息。基于蕴含冬小麦生长的多尺度变化规律的重构特征序列构建的 VMD-GRU 冬小麦估产模型实现了最优的产量估测性能,并进一步提升了估产模型的可解释性。

#### 2.4 冬小麦估测单产空间分析

为进一步验证估产模型在不同年份的泛化能力,基于生育时期尺度的多参数构建的 VMD-GRU 组合估产模型对关中平原 2016—2022 年的单产进行逐像素模拟(图6)。结果表明,关中平原的估产结果呈西高东低的空间分布,即西部地区产量高,中

部地区产量居中,东部地区产量较低。西部地区的凤翔区、岐山县、扶风县、眉县以及中部地区的郿邑区、阎良区产量较高,大多高于 5 000 kg/hm<sup>2</sup>,东部地区的白水、澄城、合阳产量较低。关中平原西部和中部地区为冬小麦主要种植区,其估产结果空间分布与关中平原实际单产的空间分布特点一致,表明 VMD-GRU 组合估产模型在关中平原冬小麦单产估测方面表现出良好的适用性。

### 3 讨论

基于 VMD 方法分解各时间序列遥感参数,将经过模态筛选和重构后的特征序列作为 GRU 神经网络模型的输入,从而有效学习与作物生长密切相关的遥感参数序列中的长期依赖关系,实现了较为准确的冬小麦单产估测。然而,不同遥感参数对于产量估测都有独特的信息,深度学习能够解开共线性变量的影响。因此,未来研究可以整合土壤属性和气象等数据,实现更好的估产性能。此外,未来可开展多源数据协同反演和多源数据融合研究,实现多源、多平台数据融合,获取更高质量的遥感数据,以提高估产模型精度。

对时间序列 VTCI、LAI 和 FPAR 进行 VMD 分解,结果表明,VMD 分解得到的 IMF 分量的平稳性较强,为时序特征的提取提供了有效途径。本文基于中心频率观察法确定  $K$ ,存在一定程度的主观性和局限性。未来可通过灰狼优化、麻雀算法等参数寻优方法确定模态分解数  $K$ ,以提高 VMD 参数确定过程的可解释性和准确性。

使用 GRU 神经网络模型提取重构序列的时序

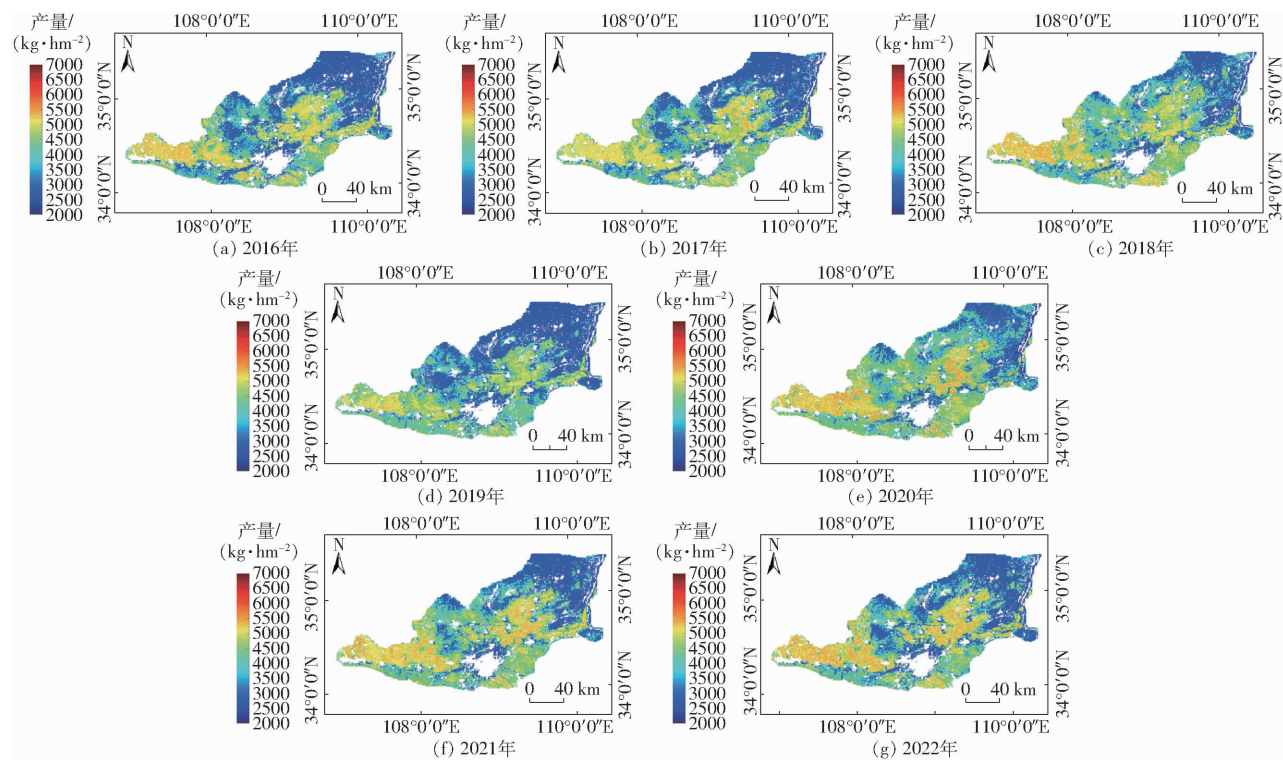


图 6 关中平原 2016—2022 年冬小麦区域单产估测结果

Fig. 6 Regional yield estimation of winter wheat in Guanzhong Plain from 2016 to 2022

特征,总体上模型性能稳定,但在高产样本的估测上存在估产结果偏低现象。这可归因为用于估产模型训练的样本不平衡,对于高产样本的特征学习不够充分。而深度学习模型强大的预测能力依赖于样本规模,因此,可以从样本角度提升算法精度。通过生成对抗网络、变分自编码器等样本增强方法对数据集进行有效扩充,将有助于提升模型稳定性。

## 4 结论

(1)将序列分解技术、相关性检验与时间序列预测模型相结合,按照分解-筛选-重构-预测的思想,构建了基于 VMD-GRU 的冬小麦单产组合估测模型。

(2)基于中心频率法确定最优模态分解数  $K$  并对生育时期尺度的时间序列遥感参数进行 VMD 分

解,得到的各组 IMF 分量规律性较强,能够表征原始时间序列遥感参数中蕴含的非线性趋势变化信息。基于相关性检验结果确定有效 IMF 分量进行特征重构,提升了特征序列的平稳性和可解释性,并保留原始序列中与作物长势相关的趋势信息和深层信息。

(3)基于蕴含冬小麦生长的多尺度变化规律的重构特征序列构建的 VMD-GRU 组合估产模型实现了最优的产量估测性能并提升了估产模型的可解释性,在县域尺度单产估测中具有良好的适用性。其得到的冬小麦单产估测结果与实际产量的  $R^2$ 、RMSE 和 MRE 分别为 0.63、448.80 kg/hm<sup>2</sup> 和 8.14%,且显著性检验结果  $P < 0.01$ ,达到极显著水平。基于最优估产模型逐像素进行关中平原冬小麦单产估测,结果表明,冬小麦产量空间分布呈西高东低的特点,与实际单产分布特点一致。

## 参 考 文 献

- [1] 代瑞熙,许世卫. 中国小麦绿色全要素生产率时空特征及影响因素[J]. 农业工程学报, 2022, 38(8): 304-314. DAI Ruixi, XU Shiwei. Spatiotemporal characteristics and influencing factors of the green total factor productivity of wheat in China[J]. Transactions of the CSAE, 2022, 38(8): 304-314. (in Chinese)
- [2] 甘甜,李雷,李红叶,等. 基于多源遥感数据和机器学习算法的冬小麦产量预测研究[J]. 麦类作物学报, 2022, 42(11): 1419-1428. GAN Tian, LI Lei, LI Hongye, et al. Winter wheat yield prediction based on multi-source remote sensing data and machine learning algorithms[J]. Journal of Triticeae Crops, 2022, 42(11): 1419-1428. (in Chinese)
- [3] TIAN H R, WANG P X, TANSEY K, et al. A deep learning framework under attention mechanism for wheat yield estimation using remotely sensed indices in the Guanzhong Plain, PR China[J]. International Journal of Applied Earth Observation and Geoinformation, 2021, 102: 102375.

- [4] 陈仲新,任建强,唐华俊,等. 农业遥感研究应用进展与展望[J]. 遥感学报, 2016, 20(5): 748-767.  
CHEN Zhongxin, REN Jianqiang, TANG Huajun, et al. Progress and perspectives on agricultural remote sensing research and applications in China[J]. Journal of Remote Sensing, 2016, 20(5): 748-767. (in Chinese)
- [5] 刘峻明,周舟,和晓彤,等. 基于NDWI和卷积神经网络的冬小麦产量估测方法[J]. 农业机械学报, 2021, 52(12): 273-280.  
LIU Junming, ZHOU Zhou, HE Xiaotong, et al. Winter wheat yield estimation method based on NDWI and convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(12): 273-280. (in Chinese)
- [6] 刘宪锋,傅伯杰. 干旱对作物产量影响研究进展与展望[J]. 地理学报, 2021, 76(11): 2632-2646.  
LIU Xianfeng, FU Bojie. Drought impacts on crop yield: progress, challenges and prospect[J]. Acta Geographica Sinica, 2021, 76(11): 2632-2646. (in Chinese)
- [7] 王鹏新,龚健雅,李小文. 条件植被温度指数及其在干旱监测中的应用[J]. 武汉大学学报(信息科学版), 2001, 26(5): 412-418.  
WANG Pengxin, GONG Jianya, LI Xiaowen. Vegetation temperature condition index and its application for drought monitoring[J]. Geomatics and Information Science of Wuhan University, 2001, 26(5): 412-418. (in Chinese)
- [8] 赵龙才,李粉玲,常庆瑞. 农作物遥感识别与单产估算研究综述[J]. 农业机械学报, 2023, 54(2): 1-19.  
ZHAO Longcai, LI Fenling, CHANG Qingrui. Review on crop type identification and yield forecasting using remote sensing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(2): 1-19. (in Chinese)
- [9] FORTIN J G, ANCTIL F, PARENT L, et al. Site-specific early season potato yield forecast by neural network in Eastern Canada[J]. Precision Agriculture, 2011, 12(6): 905-923.
- [10] WU Jun, HU Kui, CHENG Yiwei, et al. Data-driven remaining useful life prediction via multiple sensor signals and deep long short-term memory neural network[J]. ISA Transactions, 2020, 97: 241-250.
- [11] ZHANG Liangliang, ZHANG Zhao, LUO Yuchuan, 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.
- [12] CHO K, MERRIENBOER B V, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[J]. arXiv: 1406. 1078, 2014.
- [13] WANG J, WANG P X, TIAN H R, 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.
- [14] DRAGOMIRETSKIY K, ZOSSO D. Variational mode decomposition[J]. IEEE Transactions on Signal Processing, 2014, 62(3): 531-544.
- [15] WANG H F, KANG C Y, TIAN Z X, et al. Vegetation periodic changes and relationships with climate in Inner Mongolia based on the VMD method[J]. Ecological Indicators, 2023, 146: 109764.
- [16] 杨可明,孙彤彤,张伟,等. 玉米光谱变分模态分解及其铜铅污染的VMD-CR-ED测度模型[J]. 红外与毫米波学报, 2018, 37(1): 80-86.  
YANG Keming, SUN Tongtong, ZHANG Wei, et al. Spectral variational mode decomposition and VMD-CR-ED model on measuring copper and lead pollution of corn[J]. Journal of Infrared and Millimeter Waves, 2018, 37(1): 80-86. (in Chinese)
- [17] LU Y J, LI T, HU H, et al. Short-term prediction of reference crop evapotranspiration based on machine learning with different decomposition methods in arid areas of China[J]. Agricultural Water Management, 2023, 279: 108175.
- [18] 褚江东,粟晓玲,吴海江,等. 2002—2021年中国陆地水储量及其组分变化分析[J]. 水资源保护, 2023, 39(3): 170-178.  
CHU Jiangdong, SU Xiaoling, WU Haijiang, et al. Analysis of terrestrial water storage and its component changes in China from 2002 to 2021[J]. Water Resources Protection, 2023, 39(3): 170-178. (in Chinese)
- [19] 王维高,魏云冰,滕旭东. 基于VMD-SSA-LSSVM的短期风电预测[J]. 太阳能学报, 2023, 44(3): 204-211.  
WANG Weigao, WEI Yunbing, TENG Xudong. Short-term wind power forecasting based on VMD-SSA-LSSVM[J]. Acta Energetica Sinica, 2023, 44(3): 204-211. (in Chinese)
- [20] 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)
- [21] 李艳,王鹏新,刘峻明,等. 基于条件植被温度指数的冬小麦主要生育时期干旱监测效果评价——II. 改进的层次分析法和变异系数法组合赋权[J]. 干旱地区农业研究, 2014, 32(1): 236-239.  
LI Yan, WANG Pengxin, LIU Junming, et al. Evaluation of drought monitoring effects in the main growing stages of winter wheat by using the vegetation temperature condition index—II. Improved analytic hierarchy process and variation coefficient method[J]. Agricultural Research in the Arid Areas, 2014, 32(1): 236-239. (in Chinese)

- Sundarban delta of India using frequency ratio (FR), fuzzy logic (FL), logistic regression (LR) and random forest (RF) models[J]. *Regional Studies in Marine Science*, 2021, 42: 101624.
- [41] GHOLIZADEH M, JAMEI M, AHMADIANFAR I, et al. Prediction of nanofluids viscosity using random forest (RF) approach[J]. *Chemometrics and Intelligent Laboratory Systems*, 2020, 201: 104010.
- [42] PRASAD R, DEO R C, LI Y, et al. Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition[J]. *Geoderma*, 2018, 330: 136 - 161.
- [43] MA Z, SHI Z, ZHOU Y, et al. A spatial data mining algorithm for downscaling TMPA 3B43 V7 data over the Qinghai-Tibet plateau with the effects of systematic anomalies removed[J]. *Remote Sensing of Environment*, 2017, 200: 378 - 395.
- [44] HENDERSON B L, BUI E N, MORAN C J, et al. Australia-wide predictions of soil properties using decision trees[J]. *Geoderma*, 2005, 124(3): 383 - 398.
- [45] BUI M, LIU Z. Simple allele-discriminating PCR for cost-effective and rapid genotyping and mapping[J]. *Plant Methods*, 2009, 5(1): 1 - 8.
- [46] RAJ A, CHAKRABORTY S, DUDA B M, et al. Soil mapping via diffuse reflectance spectroscopy based on variable indicators: an ordered predictor selection approach[J]. *Geoderma*, 2018, 314: 146 - 159.
- [47] 陈俊英, 王新涛, 张智韬, 等. 基于无人机-卫星遥感升尺度的土壤盐渍化监测方法[J]. *农业机械学报*, 2019, 50(12): 161 - 169.  
CHEN Junying, WANG Xintao, ZHANG Zhitao, et al. Soil salinization monitoring method based on UAV - satellite remote sensing scale-up[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(12): 161 - 169. (in Chinese)
- [48] 张智韬, 陈钦达, 黄小鱼, 等. 基于加权算法的空-天遥感升尺度土壤含盐量监测模型[J]. *农业机械学报*, 2022, 53(9): 226 - 238, 251.  
ZHANG Zhitao, CHEN Qinda, HUANG Xiaoyu, et al. UAV - satellite remote sensing scale-up monitoring model of soil salinity based on dominant class variability-weighted method [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(9): 226 - 238, 251. (in Chinese)
- [49] ZHANG J, ZHANG Z, CHEN J, et al. Estimating soil salinity with different fractional vegetation cover using remote sensing [J]. *Land Degradation & Development*, 2021, 32(2): 597 - 612.
- [50] PENG J, BISWAS A, JIANG Q, et al. Estimating soil salinity from remote sensing and terrain data in southern Xinjiang, China[J]. *Geoderma*, 2019, 337: 1309 - 1319.
- [51] 张智韬, 杜瑞麒, 杨帅, 等. 水盐交互作用对河套灌区土壤光谱特征的影响[J]. *农业工程学报*, 2020, 36(18): 153 - 164.  
ZHANG Zhitao, DU Ruiqi, YANG Shuai, et al. Effects of water-salt interaction on soil spectral characteristics in Hetao Irrigation Areas of Inner Mongolia, China[J]. *Transactions of the CSAE*, 2020, 36(18): 153 - 164. (in Chinese)
- [52] HE Y, ZHANG Z, XIANG R, et al. Monitoring salinity in bare soil based on Sentinel-1/2 image fusion and machine learning [J]. *Infrared Physics & Technology*, 2023, 131: 104656.

~~~~~

(上接第 174 页)

- [22] 孙威, 王鹏新, 韩丽娟, 等. 条件植被温度指数干旱监测方法的完善[J]. *农业工程学报*, 2006, 22(2): 22 - 26.  
SUN Wei, WANG Pengxin, HAN Lijuan, et al. Further improvement of the approach to monitoring drought using vegetation and temperature condition indexes from multi-years' remotely sensed data[J]. *Transactions of the CSAE*, 2006, 22(2): 22 - 26. (in Chinese)
- [23] 王鹏新, 荀兰, 李俐, 等. 基于时间序列叶面积指数傅里叶变换的作物种植区域提取[J]. *农业工程学报*, 2017, 33(21): 207 - 215.  
WANG Pengxin, XUN Lan, LI Li, et al. Extraction of planting areas of main crops based on Fourier transformed characteristics of time series leaf area index products[J]. *Transactions of the CSAE*, 2017, 33(21): 207 - 215. (in Chinese)
- [24] 贺志将, 李前, 王彦超, 等. 基于 VMD - LSTM 的奶牛动态称量算法[J]. *农业机械学报*, 2022, 53(增刊 2): 234 - 240.  
HE Zhijiang, LI Qian, WANG Yanchao, et al. Cow dynamic weighing algorithm based on VMD - LSTM[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(Supp. 2): 234 - 240. (in Chinese)
- [25] 王俊, 李霞, 周昔东, 等. 基于 VMD 和 LSTM 的超短期风速预测[J]. *电力系统保护与控制*, 2020, 48(11): 45 - 52.  
WANG Jun, LI Xia, ZHOU Xidong, et al. Ultra-short-term wind speed prediction based on VMD - LSTM[J]. *Power System Protection and Control*, 2020, 48(11): 45 - 52. (in Chinese)
- [26] 王鹏新, 王婕, 田惠仁, 等. 基于遥感多参数和门控循环单元网络的冬小麦单产估测[J]. *农业机械学报*, 2022, 53(9): 207 - 216.  
WANG Pengxin, WANG Jie, TIAN Huiren, et al. Yield estimation of winter wheat based on multiple remotely sensed parameters and gated recurrent unit neural network[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(9): 207 - 216. (in Chinese)
- [27] ASUERO A G, SAYAGO A, GONZÁLEZ A G. The correlation coefficient: an overview[J]. *Critical Reviews in Analytical Chemistry*, 2007, 36(1): 41 - 59.