

李颖,赵国强,陈怀亮,等. 基于冬小麦农业气候分区的 WOFOST 模型参数标定. 应用气象学报,2021,32(1):38-51.
DOI: 10.11898/1001-7313.20210104

基于冬小麦农业气候分区的 WOFOST 模型参数标定

李 颖¹⁾²⁾ 赵国强¹⁾³⁾* 陈怀亮¹⁾⁴⁾ 余卫东¹⁾²⁾ 苏 伟⁵⁾ 程耀达⁶⁾

¹⁾(中国气象局·河南省农业气象保障与应用技术重点实验室,郑州 450003)

²⁾(河南省气象科学研究所,郑州 450003)

³⁾(河南省气象局,郑州 450003) ⁴⁾(哈尔滨市气象局,哈尔滨 150028)

⁵⁾(中国农业大学土地科学与技术学院,北京 100083) ⁶⁾(郑州大学生态与环境学院,郑州 450001)

摘 要

以 1981—2010 年河南省 113 个气象观测站影响冬小麦生长及产量形成的主要气象因素为区划指标,利用 K 均值聚类算法,将河南省划分为 5 个农业气候生态区。根据 2013—2017 年地面农业气象观测数据,利用 Sobol 全局敏感性分析方法,各分区选择总敏感指数大于 0.01 的作物参数,得到 9 种敏感参数。以产量与叶面积指数为代价函数,采用差分进化马尔科夫链蒙特卡洛方法对敏感参数进行分区标定,并使用 2018—2019 年观测数据进行验证。结果表明:分区进行参数标定时,叶面积指数动态模拟精度和产量模拟精度均显著优于使用默认参数或整个研究区使用同一套优化参数时的精度,其中,使用分区调参后验平均值模拟关键生育期叶面积指数的总均方根误差为 0.655,其模拟产量的均方根误差为 672.016 kg·hm⁻²。该方法将农业气候学知识与差分进化马尔科夫链蒙特卡洛优化算法相结合,通过合理、高效地分区域标定作物模型参数,可为作物模型区域应用和模型参数调整优化提供科学依据。

关键词: K 均值聚类; 农业气候区划; 全局敏感性分析; 参数标定

引 言

作物生长模型起步于 20 世纪 60 年代,其实质是用数学方法表达作物生长过程^[1]。伴随作物生长理论的完善、系统科学和计算机技术的发展,作物生长模型经历了从理论走向实用的发展历程^[2]。在各国对粮食生产的持续关注下,以应用为目的的作物生长模型迅速发展,既可用于现代化和科学化的作物种植管理,也可辅助政策的分析制定^[3],已成为农业研究最有力的工具之一^[4]。随着作物生长模型向应用多元化方向发展,其区域化大范围应用的要求也日益迫切^[5],其中典型应用包括区域作物生长监测和产量预测,准确进行大面积作物生长监测和产量预测对加强作物生长调控和保障粮食安全具

有重要意义^[6-7]。

现有的作物模型,如 WOFOST, CERES, APSIM 和 WheatSM 等,都具有较强的机理性和实用性,已在小麦、水稻、玉米等作物的长势监测和产量评估中得到广泛应用^[8-9]。通过对作物模型进行参数的严格标定,在单点作物生育期的模拟误差普遍小于 3 d,产量的估算误差一般不超过 15%^[10-13]。然而,随着作物模型在区域上的升尺度应用,其区域模拟效果受到来自模型品种参数、外部输入数据等因素不确定性增大的影响,模型的区域模拟精度较难满足实际应用需求^[14]。

为提高作物模型在区域应用的精度,必须针对特定区域对其进行本地化,即完成模型参数的本地化校准^[15]。当区域面积较大时,气候条件差异较大,作物品种不同,需对区域进行合理分区,结合科

2020-09-30 收到,2020-11-02 收到再改稿。

资助项目: 国家自然科学基金项目(41805090),中国气象局·河南省农业气象保障与应用技术重点开放实验室开放基金项目(AMF201802, AMF201807),河南省气象局气象科学技术研究项目重点项目(KZ201803)

* 通信作者, 邮箱: z13756188998@163.com

学高效的优化算法对模型参数分区标定。本文以 WOFOST 模型用于河南省大范围冬小麦生长模拟与产量预测为目标,通过划分不同农业气候生态区,各区分别对模型参数进行标定,将农业气候学知识与科学高效的优化算法相结合,通过合理、高效地对研究区域分区标定,提高作物模型在区域应用的精度,为作物模型区域应用和模型参数调整优化提供科学依据。河南全省农业气候生态区划分采用 K 均值聚类算法,模型参数的敏感性分析选择 Sobol 全局敏感性分析方法,敏感参数的标定采用差分进化马尔科夫链蒙特卡洛方法^[16]。

1 研究区和研究方法

1.1 研究区概况

本文以河南省为研究区。河南省位于中国中部,黄河中下游地区,西起太行山和豫西山地东麓,南至大别山北麓,东西长约 580 km,南北跨约 550 km,全省土地面积为 $1.67 \times 10^5 \text{ km}^2$,空间范围在 $110^\circ 21' \sim 116^\circ 39' \text{E}$, $31^\circ 23' \sim 36^\circ 22' \text{N}$ 。河南省位于中国冬小麦核心主产区,平原面积广阔,土壤肥沃,其小麦种植面积占全国小麦种植面积的 1/4 以上,年产量约占全国的 26%,为我国小麦产量第一大省^[17]。

1.2 WOFOST 模型

WOFOST 模型由荷兰瓦赫宁根大学与世界粮食研究中心共同研发^[18]。该模型是一个动态解释性模型,以逐日为步长对作物生长发育进行模拟,可描述光合作用、呼吸作用、蒸腾作用、干物质分配等作物基本生理过程,并可描述这些过程受环境影响情况,可模拟水分限制条件下的产量、营养限制条件下的产量和潜在产量^[19]。

WOFOST 模型以逐日气象数据为驱动,基于积温理论建立作物生长发育模块,并考虑作物对光长的敏感性和叶片的生长及衰老过程,对叶面积生长分两个阶段进行描述。模型中干物质的积累和分配则是经过呼吸作用的消耗后被分配到作物的各个器官。作物在每个时刻获得的总干物质重量,被按照一定的比例系数分配给地上部分和地下部分,地上部分再被分配到贮存器官和茎、叶等器官。现有研究表明,WOFOST 模型在河南省及其相邻省份适用良好^[19-20]。

1.3 数据来源与处理

本研究中冬小麦农业气候分区数据来源于 1981—2010 年河南全省 113 个气象观测站冬小麦全生育期逐日气象数据,WOFOST 模型运行所需气象数据来源于研究区内 20 个气象观测站 2013—2019 年的逐日观测,包括日最低气温、日最高气温、日平均水汽压、日平均瞬时风速和日降水量。由于 WOFOST 模型的气象驱动数据包括太阳辐射量,利用 Hargreaves 辐射公式^[21]由大气温差计算得到(式(1)),将 6 种气象要素按照 WOFOST 作物模型需要的格式建立对应数据库文件。

$$R_s = K_{Rs} \sqrt{(T_{\max} - T_{\min})} R_a, \quad (1)$$

式(1)中, R_a 为天顶辐射(单位: $\text{MJ} \cdot \text{m}^{-2} \cdot \text{d}^{-1}$), T_{\max} 为最高气温(单位: $^\circ\text{C}$), T_{\min} 为最低气温(单位: $^\circ\text{C}$), K_{Rs} 为调整系数。在内陆地区,陆地占主体,气团不会受到广阔水体的强烈影响, K_{Rs} 约为 0.16^[22],河南属于内陆地区,取 0.16。

农业气象观测数据包括作物生长数据、土壤数据和田间管理信息,其中作物生长数据主要包括冬小麦播种期、关键生育期、产量等;土壤数据包括土壤重量含水率、土壤相对湿度、降水量与灌溉量、地下水位深度、土壤水分总贮存量、土壤有效水分贮存量、田间持水量、土壤容重、凋萎湿度等;田间管理信息主要为灌溉时间和灌溉量、施肥时间和施肥量。本研究所需农业气象观测数据、土壤数据和田间管理信息来源于研究区内 20 个农业气象试验站和农业气象观测站 2013—2019 年观测数据以及 2019 年田间调查的叶面积指数,剔除异常值,整理为标准数据格式。

1.4 研究方法

针对 WOFOST 模型用于大范围冬小麦生长模拟与产量预测目的,首先根据 30 年气象观测数据,利用 K 均值聚类算法将研究区河南省划分为不同的农业气候生态区,在每一个区间利用 Sobol 全局敏感性分析方法筛选出 WOFOST 模型的敏感参数,采用差分进化马尔科夫链蒙特卡洛方法对模型的敏感参数进行自动标定,并利用实测数据对模型分区标定结果进行验证。

1.4.1 K 均值聚类算法

K 均值聚类算法属于聚类分析方法,具有快速收敛的优势,其基本思想是以距离为相似度进行分

类。给定聚类数目 K 后,首先随机把数据集划分为 K 个簇,计算每个簇平均值得到 K 个初始聚类中心,计算每簇中每个样本与每个聚类中心的距离,将其重新分配到距离最近的聚类中心所在的簇,再计算每个新簇的中心。不断重复该过程,直至聚类中心不再发生变化,则聚类结束^[23]。

1.4.2 Sobol 方法

Sobol 方法是一种典型的全局敏感性分析方法^[24-25],经验证是最为有效的模型参数敏感性分析方法之一^[26]。Sobol 方法基于方差分解的思想,将参数化模型^[16]表示为

$$y = f(X, \theta)。$$

式(2)中, y 为目标函数; X 为驱动数据; θ 为参数向量。

将模型总方差分解为单个参数的影响和参数之间组合的影响:

$$D(y) = \sum_{i=1}^k D_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^k D_{ij} + \cdots + D_{1,\dots,k}。(3)$$

式(3)中, D_i 为参数 θ_i 对目标函数 y 的一阶偏方差, D_{ij} 为参数 θ_i 和 θ_j 之间相互作用的二阶偏方差, k 为参数的维数。

对式(3)进行归一化,通过偏方差和总方差之比计算参数敏感性指标。

一阶敏感性指标

$$S_i = \frac{D_i}{D}。$$

二阶敏感性指标

$$S_{ij} = \frac{D_{ij}}{D}。$$

总敏感性指标

$$S_{O_i} = \sum S_{(i)}。$$

其中, S_{O_i} 表示包含所有参数 θ_i 的敏感性,即参数 θ_i 单独对目标函数 y 的影响程度及与其他参数相互作用对 y 影响程度的总和^[27]。

1.4.3 差分进化马尔科夫链蒙特卡洛方法

马尔科夫链蒙特卡洛方法是一种贝叶斯后验采样方法,可以提供趋近于真实后验分布情况的样本序列^[28]。该方法的基本原理见文献^[29],在求解随机变量的期望或随机事件的发生概率时,为了克服蒙特卡洛随机抽样所需样本量随维数增加而指数增大,导致计算速度过慢的问题,通过构造马尔科夫链的极限平稳分布情况模拟计算积分,即利用先验知

识从建议概率分布抽样产生候选样本值,建立状态转移规则并判断状态是否转移,从而产生马尔科夫链,马尔科夫链的各个状态为随机变量的样本值,在进行足够多次的迭代后,马尔科夫链的转移概率将趋于平稳分布,得到目标的后验分布^[30]。由于马尔科夫链蒙特卡洛方法收敛效率较低,本研究使用 Braak^[31] 提出的差分进化马尔科夫链蒙特卡洛方法对待优化参数的后验分布进行估算。该方法解决了传统马尔科夫链蒙特卡洛方法中使用单一马尔科夫链,易发生局部收敛的问题,可以同时构建多条马尔科夫链估算参数的后验分布,并在不同链之间进行相互学习,大大提高了计算效率^[16]。

2 结果与分析

2.1 河南省冬小麦农业气候区划

本研究采用 K 均值聚类算法,根据河南省气候特点和冬小麦生产品种布局将全省划分为不同的农业气候生态区^[32-33]。依据气候条件对冬小麦全生育期影响情况,利用专家知识法,选取 1981—2010 年全省 113 个气象观测站冬小麦全生育期降水(单位: mm)、全生育期大于等于 0℃ 积温(单位:℃·d)、全生育期日照时数(单位:h)、冬前积温(单位:℃·d)、3—4 月日照时数(单位:h)、5 月降水量(单位: mm)、3—4 月雨日(单位:d)、5 月平均气温日较差(单位:℃)、3—4 月降水量(单位:mm)、5 月日照时数(单位:h)和 5 月雨日(单位:d)为区划指标进行 K 均值聚类,结合冬小麦生产品种的布局情况,完成河南省冬小麦农业气候分区(图 1)。

如图 1 所示,河南省冬小麦生产分为 5 个农业气候生态区,基本呈纬向分布。其中 I 区分布在河南省北部,包括安阳、濮阳等地区,该区日照充足,但大于等于 0℃ 积温较少,冬季温度较低,常有冻害发生。春旱机率较高,多数年份小麦生长受到一定水分胁迫。II 区主要分布在河南省中北部,包括新乡、焦作、郑州、开封、商丘、洛阳等地区,这些地区春季降水变率大,自然降水偏少,春季低温霜冻对小麦有一定影响。III 区以河南省中南部为主,包括漯河、平顶山、周口、驻马店中北部、南阳中东部等地区。该区正常年份降水量可以满足小麦需要,但春季连阴雨天气较多,光照不足。IV 区主要包括河南省南部

的信阳、驻马店南部和南阳西部等地区,该区小麦全生育期大于等于 0℃ 积温在 2500℃ · d 以上,冬前气温高,春季多雨,灌浆期间高温、多雨、日较差较小。Ⅴ区主要分布在河南省西部的三门峡等地,该区干旱少雨,小麦生育期内降水量不足 300 mm,春季日照充足,小麦生育后期常出现干热风。该分区结果一方面可体现纬度对冬小麦种植品种和生长发育的重要影响,另一方面也反映出河南省复杂地形影响气候、耕作方式等,进而对冬小麦种植品种和生

长发育造成影响。河南省北部、西部、南部由太行山、伏牛山、桐柏山和大别山沿省界环绕,西南部为南阳盆地、中东部为黄淮海平原,图 1 中的分区良好地吻合该地形特点,显示本研究对河南省冬小麦农业气候分区结果的合理性。

作物模型中的积温参数可通过气象观测数据准确计算,在完成河南省冬小麦农业气候分区后,利用 30 年气象资料计算各区小麦播种至出苗、出苗至开花和开花至成熟各生育阶段的积温(表 1)。

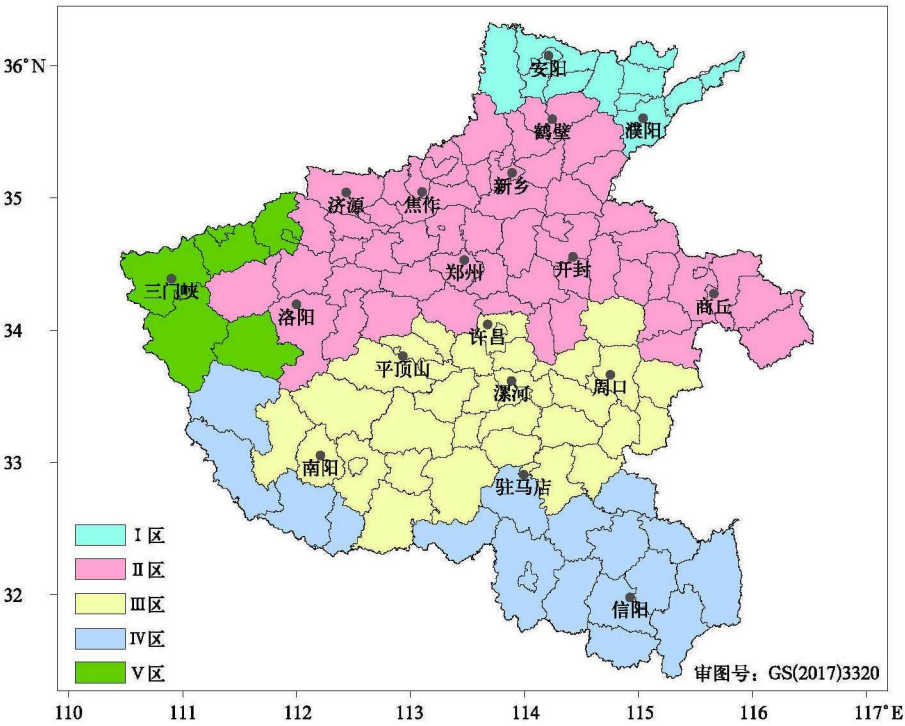


图 1 河南省冬小麦农业气候分区

Fig. 1 Agro-climatic division of winter wheat in Henan Province

表 1 不同区域冬小麦积温(单位:℃ · d)

Table 1 Accumulated temperature parameters of winter wheat in different zones(unit:℃ · d)					
生育阶段	I 区	II 区	III 区	IV 区	V 区
播种至出苗	1170	1180	1100	1150	1250
出苗至开花	600	620	610	600	610
开花至成熟	110	120	120	130	120

2.2 敏感参数分析

应用 Python 中的 SALib 灵敏度分析库,排除不适合自动标定的参数,对 WOFOST 模型的 21 种小麦生长发育参数进行敏感性分析,分析过程包括待分析参数输入、参数样本生成、模型模拟过程以及

分析结果输出。本研究通过设置待分析参数围绕默认值变化的比例确定其取值范围(表 2)。

使用 Saltelli 采样器对表 2 中的参数进行随机采样,设置采样数为 10000,产生 (21+2) × 10000 = 230000 组样本输入进 WOFOST 模型运转,输出冬

小麦模拟产量,从而建立从输入参数空间到输出结果空间的映射关系。利用 Sobol 方法筛选出对冬小麦模拟产量变化影响较大的参数。参数的实际采样值由如下公式计算:

$$a_e = a_0 \times f.$$
 (7)

式(7)中, a_e 表示参数实际采样值, f 表示本次采样比率, a_0 表示该参数默认值。

对于某些默认值为 0 的参数,采用如下公式计算参数实际采样值:

$$a_e = \frac{a_{\max} - a_{\min}}{2} \times (f - 1).$$
 (8)

式(8)中, a_{\max} 表示该参数最大值, a_{\min} 表示该参数最小值。

在河南省 5 个农业气候生态区分别进行 21 种作物参数的敏感性分析,每个分区选择 5 个观测站

点,使用 2013—2017 年的气象数据驱动模型。经对比,各分区总敏感性指数与一阶敏感性指数和二阶敏感性指数的筛选结果基本一致。以总敏感性指数为代表,河南 5 个农业气候生态区的参数敏感性分析结果见图 2。

由图 2 可知,尽管各区敏感性分析结果中待分析参数的总敏感性指数不同,参数的敏感程度排序也存在差异,但在 5 个农业生态区,达到 0.05 的显著性水平,且总敏感性指数大于 0.01 的参数筛选结果一致,均为 FL,FR,FO,SP,SL05,AM00,AM10,AM13 和 SL00。具体地,总敏感性指数大于 0.05 的参数相同,共 4 个:FL,FR,FO 和 SL05。总敏感性指数小于 0.05 且大于 0.01 的参数相同,共 5 个(不同区顺序不同):AM00,AM10,AM13,SL00 和 SP。

表 2 待分析参数及其围绕默认值的变化比例

Table 2 The proportion of changes of parameters to be analyzed around the default value

参数	定义	最小比例	最大比例
AM00	发育期为 0 时 CO ₂ 最大同化速率	0.7	1.5
AM10	发育期为 1 时 CO ₂ 最大同化速率	0.7	1.5
AM13	发育期为 1.3 时 CO ₂ 最大同化速率	0.7	1.5
AM20	发育期为 2 时 CO ₂ 最大同化速率	0.7	1.5
SL00	发育期为 0 时比叶面积	0.7	1.5
SL05	发育期为 0.5 时比叶面积	0.7	1.5
SL20	发育期为 2 时比叶面积	0.7	1.5
FL	总物质分配到叶片的比例	0.5	2.0
FO	总物质分配到储存器官的比例	0.8	1.2
FR	总物质分配到根的比例	0.5	2.0
SP	35℃ 时叶片的生命周期	0.7	1.5
TB	出苗最低温度	0.5	1.5
TD	初始总干物质重量	0.9	1.3
TE	出苗最高有效温度	0.9	1.3
TM00	日平均温度为 0℃ 时 CO ₂ 最大同化速率减小因子	0.7	1.5
TM10	日平均温度为 10℃ 时 CO ₂ 最大同化速率减小因子	0.7	1.3
TM15	日平均温度为 15℃ 时 CO ₂ 最大同化速率减小因子	0.7	1.5
TM25	日平均温度为 25℃ 时 CO ₂ 最大同化速率减小因子	0.7	1.5
TM35	日平均温度为 35℃ 时 CO ₂ 最大同化速率减小因子	0.7	1.5
RD	根的相对死亡速率	0.9	1.0
RG	叶面积指数最大日增量	0.9	1.3

2.3 参数标定与精度验证

对敏感性分析筛选出的 9 种敏感参数进行本地化标定,使用差分进化马尔科夫链蒙特卡洛方法对其后验分布进行估算。差分进化马尔科夫链蒙特卡洛种群(并行链)数目设置为 5 条,每次采样都进行一次链更新,并通过方差法计算各个参数当前的收

敛指标 R ,通过 R 判断是否达到收敛。当存在某个参数的 $R > 1.1$ 时,认为链未收敛;当所有参数的 $R < 1.1$ 时,认为链达到收敛。收敛后再次进行 1000 次采样,由此得到参数的后验样本和分布。差分进化马尔科夫链蒙特卡洛方法中使用的似然函数为

$$\lg L = \lg L_{\text{lai}} + \lg L_{\text{yield}}.$$
 (9)

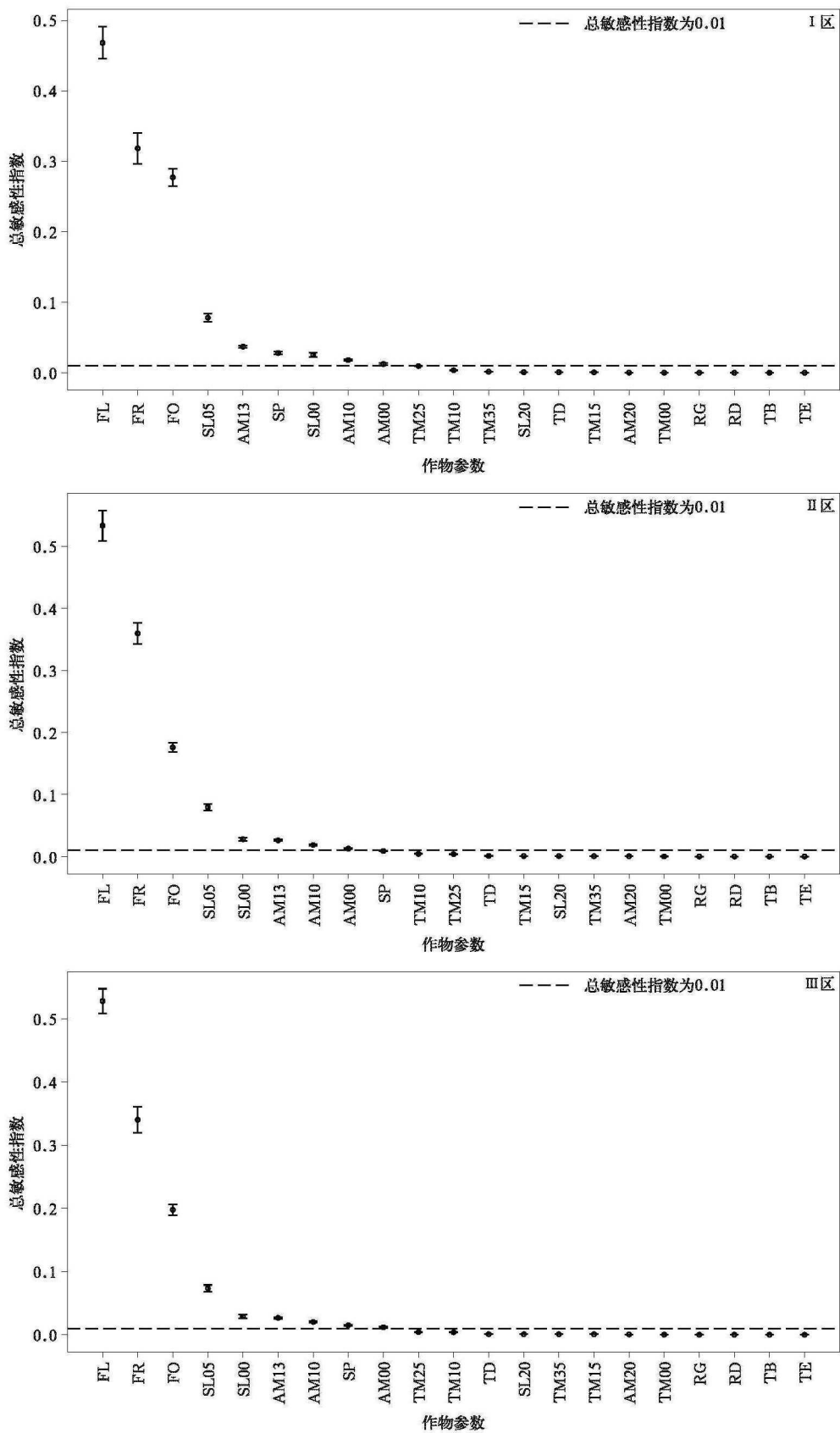
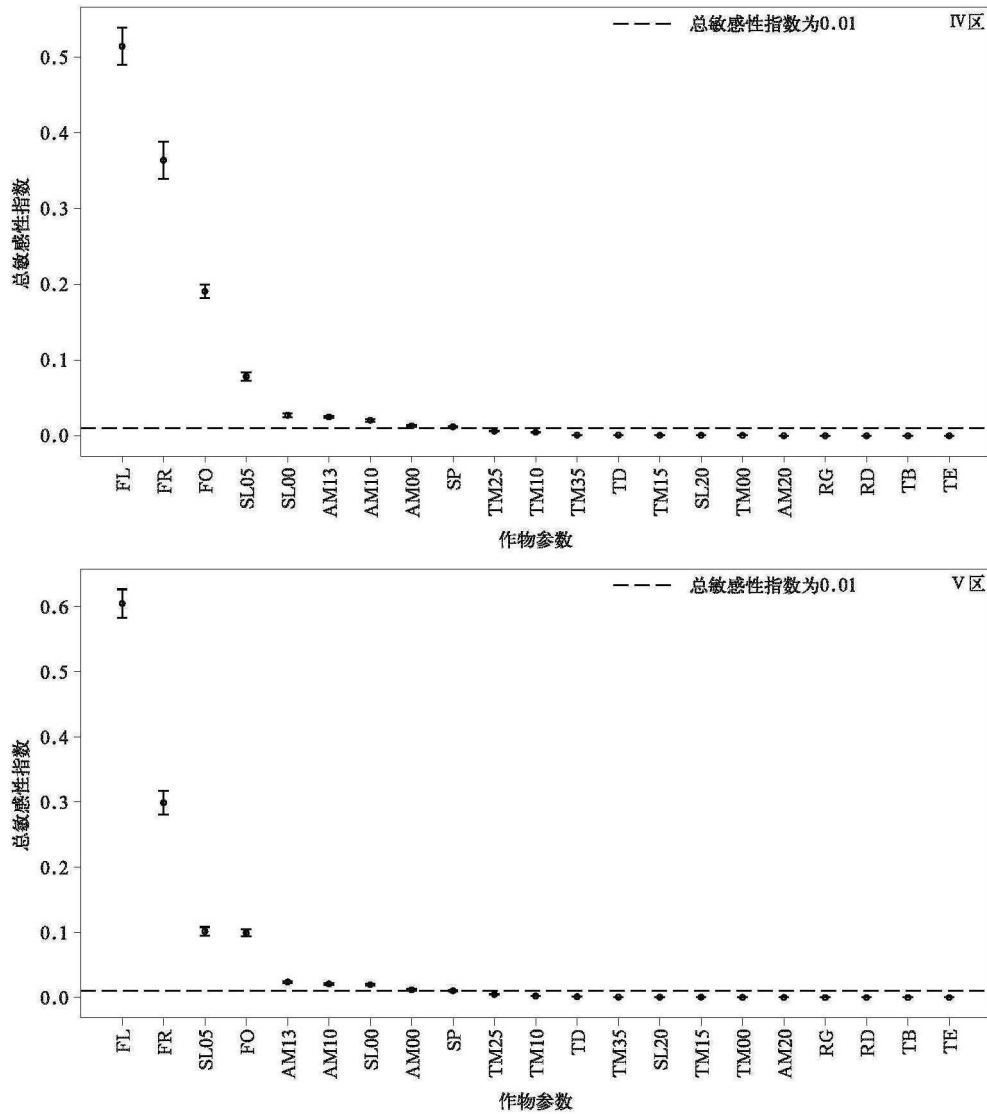


图 2 5 个农业气候生态区参数的总敏感性指数

Fig. 2 Total sensitivity index of parameters in five agro-climatic ecological zones



续图 2

式(9)中, L 表示总似然函数。 L_{lai} 表示叶面积指数的似然函数, L_{yield} 表示产量的似然函数,具体如下:

$$\lg L_{lai} = -0.5(p - p_{obs})^T \cdot V^{-1}(p - p_{obs}) - 0.5K \cdot \lg(2\pi) - \lg(\det V). \quad (10)$$

式(10)中, p 和 p_{obs} 分别表示不同生育期叶面积指数的模型模拟值和观测值, V 表示叶面积指数观测值的协方差矩阵, K 表示空间维数,即叶面积指数观测值数量。

$$\lg L_{yield} = -0.5\left(\frac{z - z_{obs}}{\sigma}\right)^2 - 0.5\lg(2\pi) - \lg\sigma. \quad (11)$$

式(11)中, z 和 z_{obs} 分别表示成熟期产量的模型模拟值和观测值, σ 表示产量观测值的标准差。

对各区使用 2013—2017 年多站点实测数据联

合标定,本研究中参数标定所用建模数据和验证数据见表 3。5 个农业生态区待优化参数的后验分布见表 4~表 8,可以看到,多种作物参数在各区之间差异明显,可能体现不同区域适宜种植品种的不同,这也表明在大区域进行分区调参的必要性。将河南省作为一个整体区域使用 2013—2017 年多站点实测数据联合标定,整个研究区使用同一套参数的后验分布见表 9。

9 种参数后验分布的平均值、中值以及最大似然值优化结果(代价函数取最优解时的一组参数值)见表 4~表 9,分别代入 WOFOST 模型进行模拟,使用 2018—2019 年观测数据进行验证,模拟及验证结果见图 3。由图 3 可知,分区标定后,在各子区模拟的叶面积指数与实际观测值均吻合较好,综合 5

表 3 参数标定所用建模数据和验证数据

Table 3 Modeling data and verification data used for parameter calibration

研究区	优化数据		验证数据	
	年份	站点	年份	站点
I 区	2013	林州	2018—2019	汤阴
	2014	汤阴		
	2015	安阳		
	2016	濮阳		
	2017	范县		
II 区	2013	郑州	2018—2019	郑州
	2014	商丘		
	2015	伊川		
	2016	济源		
	2017	郑州		
III 区	2013	许昌	2018—2019	黄泛区
	2014	许昌		
	2015	黄泛区		
	2016	驻马店		
	2017	南阳		
IV 区	2013	信阳	2018—2019	潢川
	2014	正阳		
	2015	新野		
	2016	潢川		
	2017	固始		
V 区	2013	卢氏	2018—2019	卢氏
	2014	三门峡		
	2015	三门峡		
	2016	卢氏		
	2017	卢氏		

表 4 河南省 I 区敏感参数的后验分布

Table 4 The posteriori distribution of sensitive parameters in Zone I of Henan Province

参数	平均值	中值	最大似然值	均方根误差	95%置信区间
AM00	1.369	1.414	1.470	0.00396	[1.361, 1.377]
AM10	1.351	1.374	1.426	0.00301	[1.345, 1.357]
AM13	0.729	0.722	0.704	0.00070	[0.728, 0.731]
SL00	1.443	1.455	1.500	0.00130	[1.440, 1.445]
SL05	0.774	0.770	0.707	0.00127	[0.771, 0.776]
FL	1.904	1.932	1.994	0.00220	[1.900, 1.909]
FO	0.857	0.856	0.860	0.00023	[0.856, 0.857]
FR	1.423	1.429	1.487	0.00131	[1.421, 1.426]
SP	0.831	0.829	0.826	0.00026	[0.831, 0.832]

表 5 河南省 II 区敏感参数的后验分布

Table 5 The posteriori distribution of sensitive parameters in Zone II of Henan Province

参数	平均值	中值	最大似然值	均方根误差	95%置信区间
AM00	0.761	0.750	0.704	0.00176	[0.757, 0.764]
AM10	1.397	1.410	1.412	0.00274	[1.392, 1.403]
AM13	0.793	0.785	0.760	0.00235	[0.788, 0.798]
SL00	1.350	1.361	1.458	0.00385	[1.342, 1.357]
SL05	0.772	0.759	0.737	0.00195	[0.768, 0.775]
FL	1.823	1.848	1.973	0.00496	[1.813, 1.833]
FO	0.921	0.920	0.901	0.00055	[0.919, 0.922]
FR	0.946	0.964	1.112	0.00573	[0.934, 0.957]
SP	0.822	0.824	0.824	0.00034	[0.822, 0.823]

表 6 河南省Ⅲ区敏感参数的后验分布

Table 6 The posteriori distribution of sensitive parameters in Zone Ⅲ of Henan Province

参数	平均值	中值	最大似然值	均方根误差	95%置信区间
AM00	0.715	0.711	0.702	0.00040	[0.714, 0.716]
AM10	0.891	0.872	0.869	0.00319	[0.885, 0.897]
AM13	0.738	0.730	0.700	0.00089	[0.737, 0.740]
SL00	1.352	1.376	1.498	0.00296	[1.346, 1.358]
SL05	1.082	1.067	1.144	0.00383	[1.075, 1.090]
FL	1.326	1.334	1.215	0.00426	[1.317, 1.334]
FO	0.809	0.807	0.802	0.00023	[0.809, 0.810]
FR	0.790	0.787	0.794	0.00397	[0.782, 0.798]
SP	0.932	0.938	0.940	0.00042	[0.932, 0.933]

表 7 河南省Ⅳ区敏感参数的后验分布

Table 7 The posteriori distribution of sensitive parameters in Zone Ⅳ of Henan Province

参数	平均值	中值	最大似然值	均方根误差	95%置信区间
AM00	1.446	1.460	1.496	0.00094	[1.444, 1.447]
AM10	1.118	1.127	0.744	0.00472	[1.109, 1.128]
AM13	1.418	1.441	1.497	0.00151	[1.415, 1.421]
SL00	1.174	1.188	0.918	0.00304	[1.168, 1.180]
SL05	0.799	0.777	0.711	0.00164	[0.796, 0.802]
FL	1.090	1.076	1.297	0.00250	[1.085, 1.095]
FO	1.033	1.021	1.124	0.00076	[1.031, 1.034]
FR	0.625	0.595	0.602	0.00198	[0.621, 0.629]
SP	0.834	0.829	0.828	0.00027	[0.833, 0.834]

表 8 河南省Ⅴ区敏感参数的后验分布

Table 8 The posteriori distribution of sensitive parameters in Zone Ⅴ of Henan Province

参数	平均值	中值	最大似然值	均方根误差	95%置信区间
AM00	1.082	1.070	1.333	0.00562	[1.071, 1.093]
AM10	0.732	0.724	0.705	0.00068	[0.730, 0.732]
AM13	0.728	0.721	0.708	0.00057	[0.727, 0.729]
SL00	1.481	1.484	1.492	0.00036	[1.480, 1.481]
SL05	0.718	0.715	0.708	0.00026	[0.717, 0.718]
FL	1.880	1.885	1.792	0.00203	[1.876, 1.884]
FO	0.937	0.935	0.915	0.00046	[0.936, 0.938]
FR	0.532	0.527	0.502	0.00059	[0.531, 0.533]
SP	0.738	0.741	0.748	0.00030	[0.738, 0.739]

表 9 河南省同一套敏感参数的后验分布

Table 9 The posteriori distribution of sensitive parameters for the whole Henan Province

参数	平均值	中值	最大似然值	均方根误差	95%置信区间
AM00	0.758	0.747	0.756	0.00170	[0.755, 0.761]
AM10	0.724	0.717	0.709	0.00072	[0.722, 0.725]
AM13	0.714	0.711	0.702	0.00044	[0.714, 0.715]
SL00	1.483	1.485	1.499	0.00048	[1.482, 1.484]
SL05	0.712	0.709	0.701	0.00032	[0.711, 0.712]
FL	1.855	1.853	1.856	0.00178	[1.851, 1.858]
FO	0.901	0.899	0.892	0.00046	[0.900, 0.902]
FR	0.553	0.544	0.538	0.00149	[0.550, 0.556]
SP	0.896	0.896	0.906	0.00015	[0.895, 0.896]

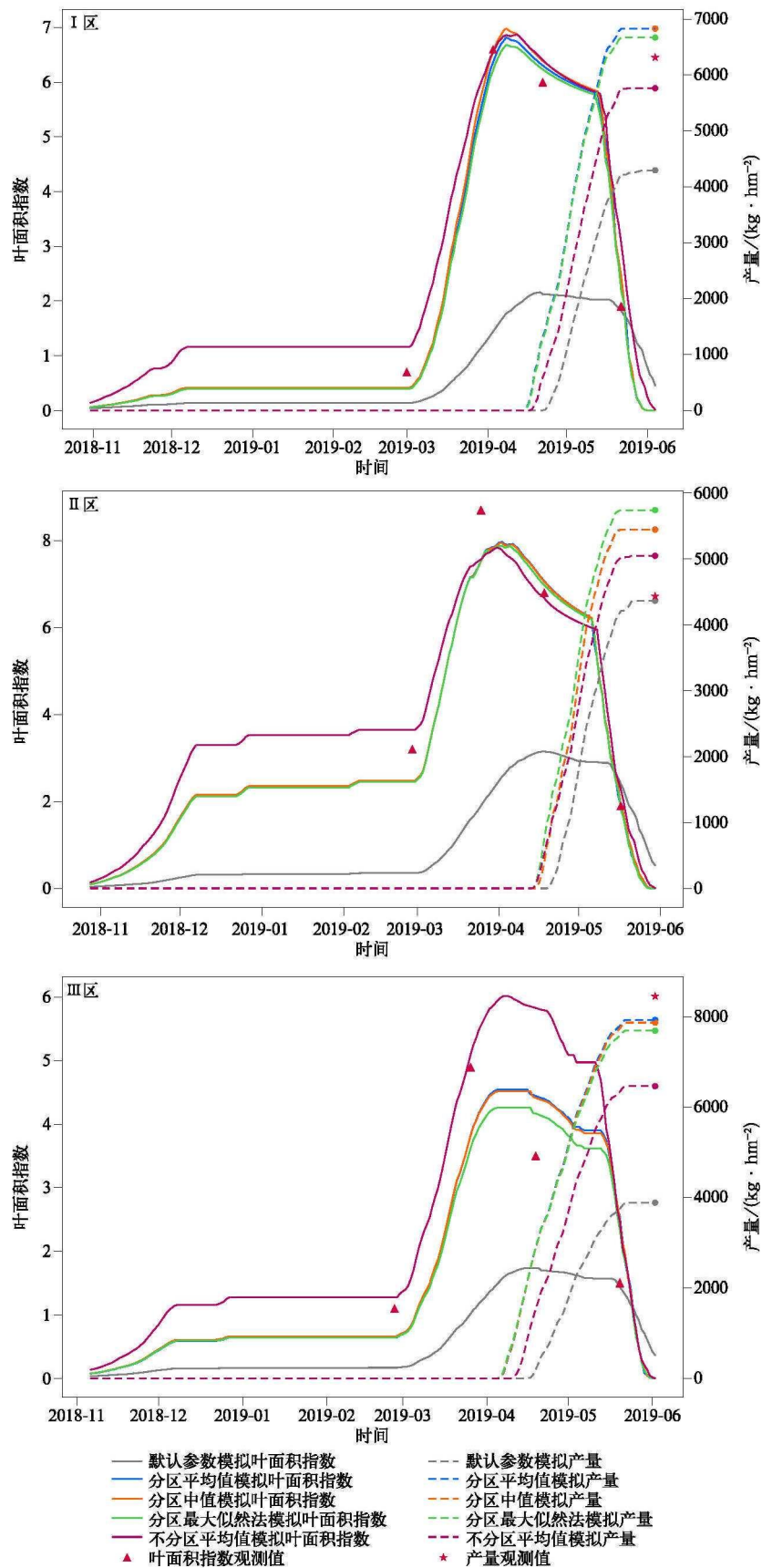
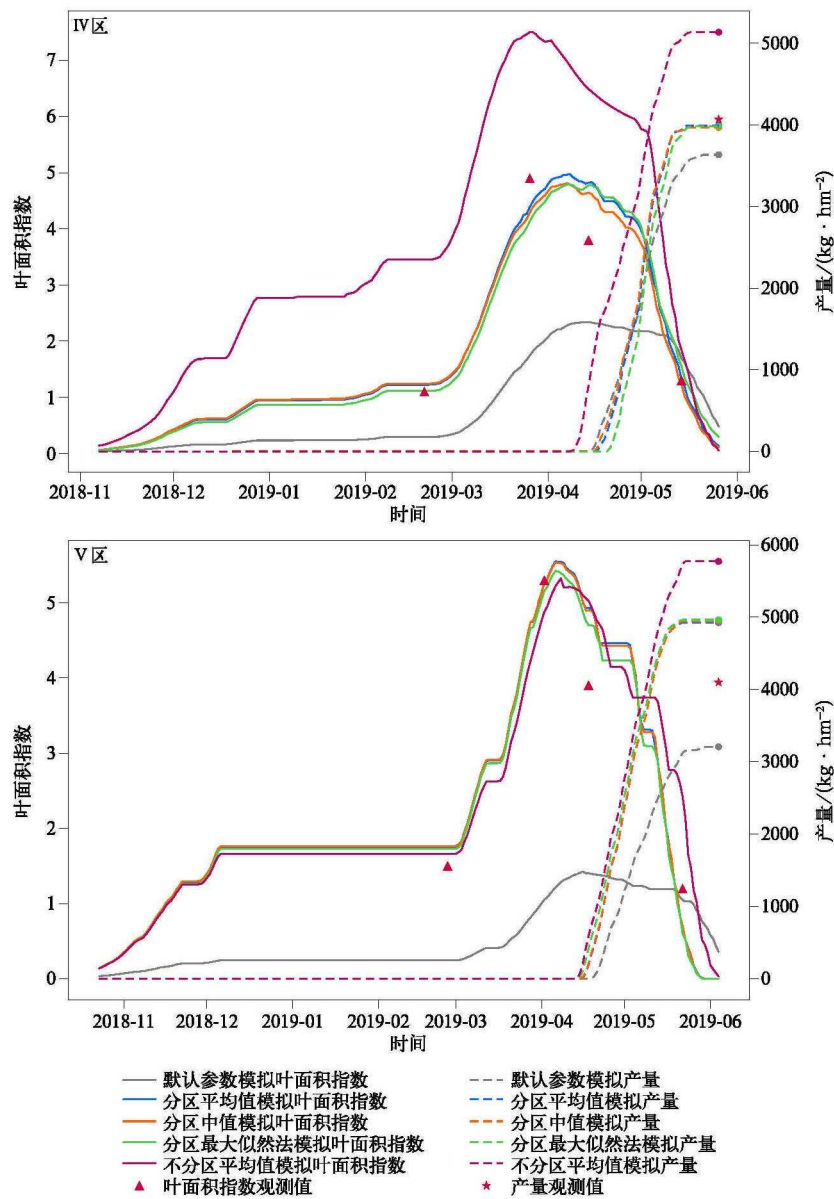


图 3 2019 年产量与叶面积指数验证结果

Fig. 3 Verification results of yield and leaf area index in 2019



续图 3

个分区的模拟结果,使用分区调参后验平均值模拟关键生育期叶面积指数的总均方根误差为 0.655,其中,模拟返青期叶面积指数的均方根误差为 0.428,模拟拔节期叶面积指数的均方根误差为 0.753,模拟抽穗期叶面积指数的均方根误差为 0.796,模拟乳熟期叶面积指数的均方根误差为 0.578。使用默认参数模拟关键生育期叶面积指数的总均方根误差为 2.897,整个研究区使用同一套优化参数(后验平均值)模拟关键生育期叶面积指数的总均方根误差为 1.277,对比可知,经分区调参,模型动态模拟冬小麦多个关键生育期叶面积指数的精度均明显提高,表明基于农业气候分区标定 WO-

FOST 模型可准确模拟作物生长过程。在估产精度方面,综合 5 个分区模拟结果,使用默认参数模拟产量的均方根误差为 2282.192 kg·hm⁻²,整个研究区使用同一套优化参数(后验平均值)模拟产量的均方根误差为 1311.303 kg·hm⁻²,使用分区调参后验平均值模拟产量的均方根误差为 672.016 kg·hm⁻²,使用分区调参后验中值模拟产量的均方根误差为 684.622 kg·hm⁻²,使用分区最大似然值参数优化结果模拟产量的均方根误差为 796.436 kg·hm⁻²。对比可知,使用本研究提出的基于冬小麦农业气候分区的差分进化马尔科夫链蒙特卡罗参数标定方法,显著优于默认参数以及不分区调参的产量

模拟精度,其中以使用分区调参后验平均值模拟产量的精度最优,较使用默认参数时产量模拟误差减小 70.55%,较全区使用同一套优化参数(平均值)时产量模拟误差减小 48.75%。

3 结论与讨论

本研究采用 K 均值聚类算法根据 1981—2010 年气象数据将河南省划分为 5 个农业气候生态区,在各区分别计算作物模型中的积温参数,再根据 2013—2017 年的观测数据利用 Sobol 全局敏感性分析方法选择其他敏感性参数,随后采用差分进化马尔科夫链蒙特卡洛方法对敏感参数进行标定,并用 2018—2019 年的观测数据进行验证,得到结论如下:

1) 河南省 5 个农业气候生态区,整体呈纬向分布,良好吻合研究区的地形特点。冬小麦农业气候分区结果可为作物模型分区标定提供科学的分区依据。

2) 对 5 个农业气候生态区分区进行 WOFOST 模型参数的敏感性分析,结果表明:不同冬小麦品种筛选出的敏感参数具有较高一致性。

3) 在农业气候分区的基础上,使用差分进化马尔科夫链蒙特卡洛方法优化 WOFOST 模型参数,模型模拟冬小麦多个关键生育期叶面积指数的精度和产量预测精度均显著优于使用默认参数或整个研究区使用同一套优化参数时的模拟精度。

本研究以河南省为研究区,将农业气候学知识与科学高效的优化算法相结合,有效实现作物模型区域应用时模型参数的优化调整,但存在以下不足:一方面,本研究局限在省域范围内,若区域面积进一步扩大,冬小麦种植品种有更大变化时,尚需进一步研究模型的敏感参数是否发生变化以及模型输出的不确定性将如何变化等问题;另一方面,本研究使用差分进化马尔科夫链蒙特卡洛方法通过最优化代价函数自动为模型敏感参数赋值,不同参数在不同区间差异较大,有待根据不同区间冬小麦实际种植品种进行测量分析,进一步验证采用差分进化马尔科夫链蒙特卡洛方法为作物参数自动赋值的准确性。下一步用于整个黄淮海地区可进一步优化本研究提出的基于农业气候分区的参数标定方法,实现 WOFOST 模型在大区域应用的本地化工作。

此外,本研究以近 30 年的气象观测数据为依据

对研究区进行分区,但各区内模型模拟效果仍受到来自模型品种参数、外部输入数据等因素的不确定性影响。为进一步提高作物模型的区域化应用精度,可采用网格单元运转模型,并在该过程中将遥感信息与作物模型耦合解决模型空间扩展过程中格点参数获取困难的问题。

参 考 文 献

- [1] Edwards D, Hamson M. Guide to Mathematical Modeling. Boca Raton Florida, US: CRC Press, Inc., 1990.
- [2] 谢云, James R K. 国外作物生长模型发展综述. 作物学报, 2002, 28(2): 190-195.
Xie Y, James R K. A review on the development of crop modeling and its application. *Acta Agronomica Sinica*, 2002, 28(2): 190-195.
- [3] Boote K J, Jones J W, Pickering N B. Potential uses and limitations of crop models. *Agronomy Journal*, 1996, 88: 704-716.
- [4] 林忠辉, 莫兴国, 项月琴. 作物生长模型研究综述. 作物学报, 2003, 29(5): 750-758.
Lin Z H, Mo X G, Xiang Y Q. Research advances on crop growth models. *Acta Agronomica Sinica*, 2003, 29(5): 750-758.
- [5] 高永刚, 顾红, 姬菊枝, 等. 近 43 年来黑龙江气候变化对农作物产量影响的模拟研究. 应用气象学报, 2007, 18(4): 532-538.
Gao Y G, Gu H, Ji J Z, et al. Simulation study of climate change impact on crop yield in Heilongjiang Province from 1961 to 2003. *J Appl Meteor Sci*, 2007, 18(4): 532-538.
- [6] 张蕾, 侯英雨, 郑昌玲, 等. 作物长势评估指数设计与应用. 应用气象学报, 2019, 30(5): 543-554.
Zhang L, Hou Y Y, Zheng C L, et al. The construction and application of assessing index to crop growing condition. *J Appl Meteor Sci*, 2019, 30(5): 543-554.
- [7] 陈怀亮, 李颖, 田宏伟, 等. 利用亚像元尺度信息改进区域冬小麦生长的模拟. 生态学杂志, 2018, 37(7): 2221-2228.
Chen H L, Li Y, Tian H W, et al. Improvement of regional-scale winter wheat growth modeling with sub-pixel information. *Chinese Journal of Ecology*, 2018, 37(7): 2221-2228.
- [8] 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. *Agric For Meteorol*, 2011, 151(12): 1843-1855.
- [9] 李颖, 陈怀亮, 田宏伟, 等. 同化遥感信息与 WheatSM 模型的冬小麦估产. 生态学杂志, 2019, 38(7): 2258-2264.
Li Y, Chen H L, Tian H W, et al. Estimation of winter wheat yield based on coupling remote sensing information and WheatSM model. *Chinese Journal of Ecology*, 2019, 38(7): 2258-2264.
- [10] 孙琳丽, 马玉平, 景元书, 等. 基于约束性分析的数据与作物模

- 型同化方法. 应用气象学报, 2013, 24(3): 287-296.
- Sun L L, Ma Y P, Jing Y S, et al. Assimilation of observations with crop growth model based on the constrained analysis of parameters. *J Appl Meteor Sci*, 2013, 24(3): 287-296.
- [11] 黄健熙, 黄海, 马鸿元, 等. 基于 MCMC 方法的 WOFOST 模型参数标定与不确定性分析. 农业工程学报, 2018, 34(16): 113-119.
- Huang J X, Huang H, Ma H Y, et al. Markov Chain Monte Carlo based WOFOST model parameters calibration and uncertainty analysis. *Transactions of the Chinese Society of Agricultural Engineering*, 2018, 34(16): 113-119.
- [12] 马玉平, 王石立, 王馥棠. 作物模拟模型在农业气象业务应用中的研究初探. 应用气象学报, 2005, 16(3): 293-303.
- Ma Y P, Wang S L, Wang F T. A preliminary study on the application of crop simulation models in agrometeorological services. *J Appl Meteor Sci*, 2005, 16(3): 293-303.
- [13] 刘布春, 王石立, 庄立伟, 等. 基于东北玉米区域动力模型的低温冷害预报应用研究. 应用气象学报, 2003, 14(5): 616-625.
- Liu B C, Wang S L, Zhuang L W, et al. Study of low temperature damage prediction applications in EN, China based on a scaling up maize dynamic model. *J Appl Meteor Sci*, 2003, 14(5): 616-625.
- [14] 秦鹏程, 刘敏, 万素琴, 等. 不完整气象资料下基于作物模型的产量预报方法. 应用气象学报, 2016, 27(4): 407-416.
- Qin P C, Liu M, Wan S Q, et al. Methods for yield forecast based on crop model with incomplete weather observations. *J Appl Meteor Sci*, 2016, 27(4): 407-416.
- [15] 许伟, 秦其明, 张添源, 等. SCE 标定结合 EnKF 同化遥感和 WOFOST 模型模拟冬小麦时序 LAI. 农业工程学报, 2019, 35(14): 166-173.
- Xu W, Qin Q M, Zhang T Y, et al. Time-series LAI simulation of winter wheat based on WOFOST model calibrated by SCE and assimilated by EnKF. *Transactions of the Chinese Society of Agricultural Engineering*, 2019, 35(14): 166-173.
- [16] 张琨. 遥感蒸散发模型参数敏感性分析与优化方法研究. 兰州: 兰州大学, 2018.
- Zhang K. Parameter Sensitivity Analysis and Optimization for Remote Sensing Based Evapotranspiration Model. Lanzhou: Lanzhou University, 2018.
- [17] 段国辉, 田文仲, 温红霞, 等. 近 13 a 河南省高产冬小麦产量构成及亲本利用演变分析. 山西农业科学, 2020, 48(2): 148-153.
- Duan G H, Tian W Z, Wen H X, et al. Analysis on yield components and parent utilization evolution of high yield winter wheat in Henan province in the past 13 years. *Journal of Shanxi Agricultural Sciences*, 2020, 48(2): 148-153.
- [18] Hijmans R J, Guiking-Lens I M, Van Diepen C A. WOFOST 6.0: User's Guide for the WOFOST 6.0 Crop Growth Simulation Model. Wageningen: DLO Winand Staring Centre, 1994.
- [19] 兴安, 卓志清, 赵云泽, 等. 基于 EFAST 的不同生产水平下 WOFOST 模型参数敏感性分析. 农业机械学报, 2020, 51(2): 161-171.
- Xing A, Zhuo Z Q, Zhao Y Z, et al. Sensitivity analysis of WOFOST model crop parameters under different production levels based on EFAST method. *Transactions of the Chinese Society for Agricultural Machinery*, 2020, 51(2): 161-171.
- [20] 张弘, 刘伟昌, 李树岩. WOFOST 模型对河南省冬小麦模拟的适用性分析. 气象与环境科学, 2019, 42(1): 34-40.
- Zhang H, Liu W C, Li S Y. Applicability analysis of WOFOST model for winter wheat in Henan. *Meteorological and Environmental Sciences*, 2019, 42(1): 34-40.
- [21] 邱美娟, 宋迎波, 王建林, 等. 山东省冬小麦产量动态集成预报方法. 应用气象学报, 2016, 27(2): 191-200.
- Qiu M J, Song Y B, Wang J L, et al. Integrated technology of yield dynamic prediction of winter wheat in Shandong province. *J Appl Meteor Sci*, 2016, 27(2): 191-200.
- [22] Allen R G, Pereira L S, Raes D, et al. Crop Evapotranspiration: FAO Irrigation and Drainage Paper No. 56. FAO, Rome, Italy, 1998.
- [23] 张文君, 顾行发, 陈良富, 等. 基于均值-标准差的 K 均值初始聚类中心选取算法. 遥感学报, 2006, 10(5): 715-721.
- Zhang W J, Gu X F, Chen L F, et al. An algorithm for initialization of K-Means clustering based on Mean-standard deviation. *Journal of Remote Sensing*, 2006, 10(5): 715-721.
- [24] Sobol I M. Sensitivity estimates for nonlinear mathematical models. *Math Model Comput Exp*, 1993, 1(4): 407-414.
- [25] Sobol I M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, 2001, 55(1-3): 271-280.
- [26] Tang Y, Reed P, Wagener T, et al. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Hydrology and Earth System Sciences*, 2007, 3(6): 793-817.
- [27] Nossent J, Elsen P, Bauwens W. Sobol sensitivity analysis of a complex environmental model. *Environmental Modelling & Software*, 2011, 26: 1515-1525.
- [28] 符天凡. 基于聚类的随机梯度马尔科夫链蒙特卡洛算法. 上海: 上海交通大学, 2018.
- Fu T F. Clustering-based Stochastic Gradient Markov Chain Monte Carlo. Shanghai: Shanghai Jiao Tong University, 2018.
- [29] Hasting W. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika*, 1970, 57: 97-109.
- [30] 孙玫. MCMC 算法及其应用. 应用数学进展, 2018, 7(12): 1626-1637.
- Sun M. MCMC algorithm and its application. *Advances in Applied Mathematics*, 2018, 7(12): 1626-1637.
- [31] Braak C J. A Markov Chain Monte Carlo version of the genetic algorithm differential evolution. *Stats and Computing*, 2006, 16: 239-249.
- [32] 周广胜, 何奇瑾, 汲玉河. 适应气候变化的国际行动和农业措施研究进展. 应用气象学报, 2016, 27(5): 527-533.

Zhou G S, He Q J, Ji Y H. Advances in the international action and agricultural measurements of adaptation to climate change. *J Appl Meteor Sci*, 2016, 27(5): 527-533.

[33] 郭建平. 气候变化对中国农业生产的影响研究进展. 应用气象

学报, 2015, 26(1): 1-11.

Guo J P. Advances in impacts of climate change on agricultural production in China. *J Appl Meteor Sci*, 2015, 26(1): 1-11.

WOFOST Model Parameter Calibration Based on Agro-climatic Division of Winter Wheat

Li Ying¹⁾²⁾ Zhao Guoqiang¹⁾³⁾ Chen Huailiang¹⁾⁴⁾ Yu Weidong¹⁾²⁾ Su Wei⁵⁾ Cheng Yaoda⁶⁾

¹⁾ (CMA • Henan Key Laboratory of Agrometeorological Support and Applied Technique, Zhengzhou 450003)

²⁾ (Henan Institute of Meteorological Sciences, Zhengzhou 450003)

³⁾ (Henan Meteorological Service, Zhengzhou 450003) ⁴⁾ (Harbin Meteorological Bureau, Harbin 150028)

⁵⁾ (College of Land Science and Technology, China Agricultural University, Beijing 100083)

⁶⁾ (School of Ecology and Environment, Zhengzhou University, Zhengzhou 450001)

Abstract

Crop model parameter calibration is an important work of extending point-scale crop model to regional application. Using K-means method with the main meteorological factors affecting the growth and yield formation of winter wheat obtained from 113 meteorological stations from 1981 to 2010 as zoning indicators, Henan Province is divided into five different agro-climatic ecological zones and the cumulative temperature parameters are calculated for each zone. Based on the observations during 2013—2017, nine sensitive parameters are obtained by using Sobol global sensitivity analysis method to analyze and select crop parameters with total sensitivity index greater than 0.01. The sensitive parameters selected from different agro-climatic ecological zones of different winter wheat varieties are highly consistent. A cost function is constructed with yield and leaf area index(LAI), and each partition is calibrated for sensitive parameters using Differential Evolution Markov Chain method. The results show that the simulated leaf area index in the different agro-climatic ecological zones are in good agreement with the observed values, the root mean square error (RMSE) using the posterior mean value of regional parameters adjustment to simulate the LAI of key growth periods is 0.655, which is obviously higher than that of using default parameters or using the same set of optimized parameters in the whole study area. Results show that the WOFOST model based on agro-climatic division can accurately simulate the growth process of crops. In terms of yield estimation accuracy, the yield simulation accuracy of regional parameter adjustment is also significantly improved. The best accuracy of simulated yield is achieved by using the posterior mean of regional parameters and RMSE is $672.016 \text{ kg} \cdot \text{hm}^{-2}$, 70.55% reduction than the yield simulation error when using the default parameters, or 48.75% reduction than the yield simulation error when the same set of optimized parameters (posterior mean) are used for the entire area. The method takes advantage of the knowledge of agro-climatology with the scientific and efficient Differential Evolution Markov Chain optimization algorithm to provide a scientific and theoretical basis for the application of crop models and optimization of regional model parameters through rational and efficient zonal calibration of the study area.

Key words: K-means analysis; agro-climatic zoning; global sensitivity analysis; parameter calibration