

doi:10.6041/j.issn.1000-1298.2020.06.021

基于叶面积指数的河北中部平原夏玉米单产预测研究

李 俐^{1,2} 许连香^{1,2} 王鹏新^{2,3} 齐 璇^{2,3} 王 蕾^{2,3}

(1. 中国农业大学土地科学与技术学院, 北京 100083; 2. 农业农村部农业灾害遥感重点实验室, 北京 100083;
3. 中国农业大学信息与电气工程学院, 北京 100083)

摘要: 为解决玉米单产预测的时效性和业务化问题,以河北中部平原为研究区域,选取与籽粒产量密切相关的叶面积指数(LAI)作为遥感特征参数,对研究区 2016—2018 年夏玉米单产进行预测研究。基于求和自回归移动平均 (ARIMA) 模型及径向基神经网络(RBFNN)分别逐像素预测研究区域的 LAI,结果表明,基于 ARIMA 模型的 LAI 预测精度比 RBF 神经网络的预测精度高,1 步、2 步 LAI 预测结果的 RMSE 较 RBF 神经网络分别降低了 0.18、0.14 m²/m²,更适合于河北中部平原的夏玉米单产预测。基于 LAI 监测数据和加权 LAI 与夏玉米单产的相关性研究成果,并结合基于 ARIMA 模型的 LAI 预测数据,得到 2016—2018 年夏玉米监测单产和向前 1 旬、2 旬和 3 旬的单产预测结果。结果表明,无论是县域尺度还是像素尺度,向前 1、2、3 旬夏玉米的单产预测精度均较高,2016—2018 年县域尺度预测单产与监测单产间最大相对误差仅为 3.73%。

关键词: 夏玉米; 单产预测; 叶面积指数; 求和自回归移动平均模型; 径向基神经网络

中图分类号: S127 文献标识码: A 文章编号: 1000-1298 (2020)06-0198-11 **OSID:** 

Summer Maize Yield Forecasting Based on Leaf Area Index

LI Li^{1,2} XU Lianxiang^{1,2} WANG Pengxin^{2,3} QI Xuan^{2,3} WANG Lei^{2,3}

(1. College of Land Science and Technology, China Agricultural University, Beijing 100083, China
2. Key Laboratory of Remote Sensing for Agri-Hazards, Ministry of Agriculture and Rural Affairs, Beijing 100083, China
3. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China)

Abstract: The large-scale crop yield forecasting is of great significance to grasp the state of national grain production timely and accurately and carry out effective grain macro-control. To improve the timeliness of maize yield forecasting, taking central plain of Hebei Province as the study area, yield forecasting was carried out for period during 2016 to 2018. The Savitzky – Golay filtered leaf area index, closely related to maize growth and yield, was selected as the characteristic parameter. The LAI data from early July 2010 to late August 2018 were used as the modeling data, and the LAI data from early September to late September of each year from 2016 to 2018 were used as the test data. The LAI data were extracted pixel by pixel to form a one-dimensional time series as the input data of the model. Based on the autoregressive integrated moving average (ARIMA) model and radial basis function (RBF) neural network, LAI data of the study area were forecasted pixel by pixel. And the average absolute error and root mean square error were used to evaluate the prediction accuracy of the two models. The results showed that the accuracy of LAI forecasting based on the ARIMA model was better than that of RBF neural network. The RMSE of step-1 and step-2 LAI forecasting results was 0.18 m²/m² and 0.14 m²/m² respectively, which was lower than that of RBF neural network, indicating that the ARIMA model was more suitable for forecasting summer maize yield per unit area in the central plain of Hebei Province. Based on the research correlation of weighted LAI and summer maize yield, and ARIMA LAI forecasting results, the summer maize yield forecasting models were developed at intervals of 1-ten day, 2-ten day, 3-ten day before the harvest. The results showed that the forecasting accuracy of maize yield per unit area at 1-ten day, 2-ten day, 3-ten day intervals was high in both the county scale and pixel scale, and the maximum relative error between the forecasting and monitoring of yield per unit area in the county (district) scale from 2016 to 2018 was only 3.73%. The method can be used to forecast summer maize yield at 30 days before the harvest.

Key words: summer maize; yield forecast; leaf area index; autoregressive integrated moving average model; radial basis function neural network

收稿日期: 2019-10-16 修回日期: 2019-11-19
基金项目: 国家重点研发计划项目 (2016YFD0300603-3)
作者简介: 李俐 (1976—), 女, 副教授, 博士, 主要从事微波遥感及其在农业中的应用研究, E-mail: lilixh@cau.edu.cn
通信作者: 王鹏新 (1965—), 男, 教授, 博士生导师, 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

0 引言

近年来,随着经济全球化的深入,国际国内粮食市场竞争日益激烈,作物产量信息受到社会各界越来越多的关注。玉米是我国主要粮食作物之一^[1],对玉米进行大区域长势监测和单产预测,能够帮助农业经营者和相关部门制定合理的生产计划,并及时优化调整玉米种植结构,同时对保障粮食安全、促进粮食贸易也具有重要意义。

目前,作物产量预测方法主要基于作物生长机理模型和经验回归模型^[2]。作物生长机理模型通过输入气象、土壤、田间管理和作物品种遗传特性等参数,可以在单点尺度上动态模拟作物生长发育的变化及产量的形成过程,适用于点尺度的作物产量预测^[3];基于数据同化方法耦合遥感观测数据和作物生长机理模型,可实现区域尺度的作物产量预测^[4],但由于过程复杂、输入参数众多,在一定程度上限制了其在大面积作物产量预测方面的广泛应用。经验回归模型通常选取与产量相关的特征参数,与作物单产建立统计相关模型,预测作物产量,是一种简便且符合业务化运行要求的大范围作物产量预测方法^[5]。

经验回归模型常采用的特征参数包括归一化植被指数 (Normalized difference vegetation index, NDVI)、比值植被指数 (Ratio vegetation index, RVI)、差值植被指数 (Difference vegetation index, DVI) 和绿度植被指数 (Greenness vegetation index, GVI) 等^[6],这些植被指数可间接反映作物的长势和生物量信息。随着定量遥感技术的发展,作物叶面积指数 (Leaf area index, LAI) 和生物量等越来越多的陆地生物物理量可通过遥感技术获得,这类参数是作物产量形成的物质基础,可决定产量可能达到的最高上限^[7]。有研究表明,叶面积指数与光合作用、蒸腾作用以及净初级生产力等密切相关,因此常作为监测作物长势与预测作物产量的特征参数^[8]。目前,基于作物叶面积指数进行作物产量预测常基于单个生育期或某个时段的叶面积指数,可实现作物收获前一段时间的预测单产^[9]。然而,作物单产是作物在整个生长阶段长势的综合反映^[10],不同阶段的作物长势对作物单产的影响不同,仅依靠单个生育期或某个时段 LAI 反映的长势信息进行单产预测会影响单产预测精度。王鹏新等^[11]综合考虑作物不同生育时期 LAI 对产量的影响程度,利用随机森林回归法赋予不同生育时期 LAI 不同权重,得到加权 LAI,以反映作物的综合长势,进一步提高单产预测的准确性,但获取作物整个生长阶段 (出苗期—成熟期) LAI 所需时间周期较长。因此,可尝试利用预

测模型获取夏玉米收获前生育时期的 LAI 数据,并在此基础上进行夏玉米的早期单产预测研究。

本文以河北中部平原为研究区,以夏玉米为研究对象,选取与籽粒产量密切相关的 LAI 作为遥感特征参数,利用求和自回归移动平均模型 (Autoregressive integrated moving average model, ARIMA) 和径向基函数 (Radial basis function, RBF) 神经网络对该地区夏玉米的 LAI 进行预测,并基于 LAI 监测数据和预测数据,结合河北中部平原加权 LAI 与产量的相关性研究成果,对研究区 2016—2018 年夏玉米单产进行预测,以期提高夏玉米单产预测的准确性和时效性。

1 材料与方法

1.1 研究区概况

选择河北省中部平原的保定市、石家庄市、廊坊市、沧州市和衡水市的部分或全部地区为研究区域,其覆盖范围为 114°32′ ~ 117°36′E, 36°57′ ~ 39°50′N, 面积约为 $5.3 \times 10^4 \text{ km}^2$, 如图 1 所示。该地区属半干旱半湿润大陆性季风气候,雨热同期,年均降水量范围为 350 ~ 700 mm,且时空分布不均,降水主要集中在夏季,占全年的 65% ~ 70%,降水量由南向北逐渐减少。该地区是我国重要的玉米生产基地^[12],冬小麦-夏玉米轮作是其典型的种植模式。

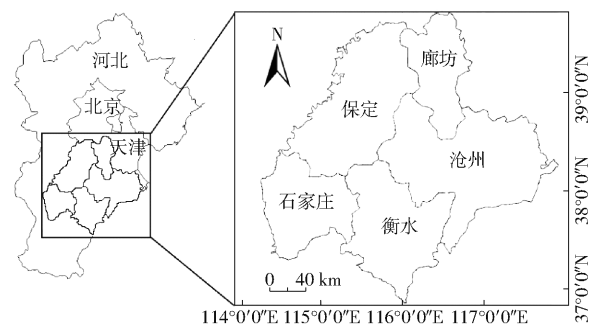


图1 研究区域位置

Fig.1 Location of study area

1.2 数据获取与处理

1.2.1 LAI 数据预处理

选取 2010—2018 年夏玉米生育期 (7—9 月), 基于 Terra 和 Aqua 卫星上的 MODIS 传感器获得的 MODIS LAI 产品 MCD15A3H。与 MOD15A2 和 MYD15A2 产品相比, MCD15A3H 产品具有较高的时空分辨率,其时间分辨率为 4 d,空间分辨率为 500 m,有利于作物长势和物候的监测。利用 MRT 对原始 LAI 影像进行镶嵌、重采样、投影转换和裁剪等预处理,并且输出投影统一为 Lambert 投影。

1.2.2 Savitzky - Golay 滤波平滑处理

为消除由于云、大气等因素的噪声影响引起的

LAI 数据骤降现象,本文应用上包络线 Savitzky - Golay (S - G) 滤波对预处理后 LAI 进行平滑处理^[13]。首先对 LAI 时序数据的长期变化趋势进行拟合。再通过局部循环迭代 S - G 滤波使得拟合结果更接近于 LAI 时序数据的上包络曲线。以饶阳县某像素为例,图2 为其经过 S - G 滤波(平滑多项式次数为 2,窗口尺寸为 5)处理后的 MODIS LAI 时间序列曲线和原始 LAI 时间序列曲线。可以看出,原始 LAI 时间序列波谷到波峰的曲线呈锯齿状的不规则波动,而经 S - G 滤波处理后的 LAI 曲线在保持原有曲线基本形状的基础上保证了曲线的平稳变化,更符合作物生长的物候特征,有利于进行时间序列 LAI 的变化趋势分析及作物长势信息提取。最后逐像素取每旬所包含的多时相 LAI 最大值作为该旬的 LAI 值,通过叠加研究区域夏玉米种植区图(图 3),得到研究区域夏玉米种植区 2010—2018 年每年 7—9 月以旬为单位的 LAI 时间序列数据。

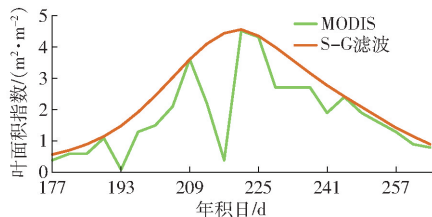


图2 S - G 滤波后 LAI 时间序列与原始 LAI 时间序列对比

Fig. 2 Comparison of S - G filtered LAI and original LAI

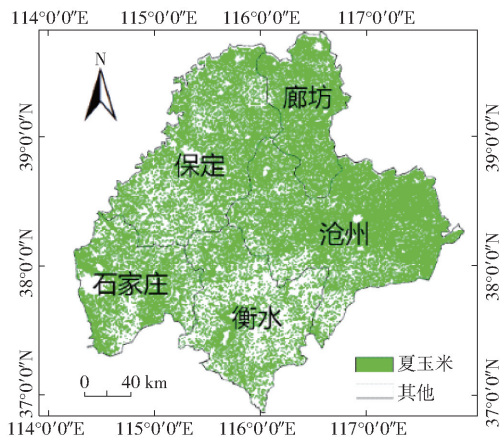


图3 河北中部平原夏玉米种植区(2017 年)

Fig. 3 Summer maize planting area in central plain of Hebei Province (2017)

1.3 夏玉米种植区的获取

研究表明,地表植被覆盖类型不同,其 LAI 预测精度也不同^[14]。本文主要研究夏玉米的 LAI 预测,为消除其他地物类型的影响,对研究区域地物类型进行了划分,并在进行 LAI 时间序列预测研究时利用分类信息对夏玉米种植区进行掩膜处理,而棉花、林地和草地等其他地物类型排除在外。具体农作物分类及夏玉米种植区域提取通过文献[15]提出的

应用一阶差分法和重构 LAI 的傅里叶变换的谐波特征识别方法实现,研究区域作物识别的总体精度为 82.51%,夏玉米的识别精度为 88.5%。

2 研究方法

2.1 LAI 预测模型

2.1.1 ARIMA 预测模型

ARIMA(p, d, q) 模型^[16-17]可看作 ARMA(p, q) 模型的推广,参数 p 为自回归阶数, q 为移动平均阶数, d 为差分阶数,可应用于非平稳时间序列预测。利用 ARIMA 模型进行预测的基本步骤^[18]为:

(1)数据平稳化处理:首先利用单位根检验 (Augmented dickey-fuller, ADF) 判断 LAI 时间序列的平稳性。若为非平稳序列,则采用 d 阶差分处理将其转换为平稳序列,对平稳后的 LAI 时间序列拟合 ARMA(p, q) 模型

$$\nabla^d X_t = \frac{\theta(B)}{\phi(B)} \varepsilon_t \quad (1)$$

其中

$$\begin{cases} \phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p \\ \theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q \end{cases} \quad (2)$$

式中 X_t ——LAI 样本时间序列数据

∇^d ——序列经过 d 阶逐期差分

θ_i, ϕ_i ——模型参数

B ——后移算子

ε_t ——白噪声序列

(2)模型定阶:ARMA(p, q) 模型的参数 p, q 可通过平稳时间序列的自相关函数 (Autocorrelation function, ACF) 和偏自相关函数 (Partial autocorrelation function, PACF) 确定。若序列的自相关函数拖尾,而偏自相关函数在 p 步截尾,则可建立 AR(p) 模型;若序列的偏自相关函数拖尾,而自相关函数在 q 步截尾,则可建立 MA(q) 模型;若序列的自相关函数和偏自相关函数均拖尾,则可建立 ARMA(p, q) 模型。然后利用最小信息准则 (Akaike information criterion, AIC) 进行模型优选,逐像素确定自回归阶数 p 和移动平均阶数 q 。AIC 准则利用似然函数估计值最大的原则来确定适用的模型, AIC 值最小的模型即为最优模型。

(3)参数估计:模型定阶后,对选定模型中的参数 θ_i, ϕ_i 进行估计。常用的估计方法有矩估计、最小二乘估计、极大似然估计等。极大似然估计可充分利用序列值的信息,精度较高,故本文利用极大似然估计法对参数 θ_i, ϕ_i 进行估计。

(4)模型检验:对已建立模型进行显著性检验,若模型残差序列是白噪声,认为模型拟合显著有效;否则说明残差序列中仍有信息未被提取,需要重新

拟合模型。

2.1.2 RBF神经网络模型

径向基神经网络^[19-20]是一种3层前向神经网络,网络结构较为简单、训练简洁并且收敛速度快,能够逼近任意非线性函数。其基本思想是用径向基函数作为隐含层节点的“基”构成隐含层空间,利用隐含层对输入矢量进行变换,将低维的模式输入数据映射到高维空间,然后通过隐含层节点的输出加权求和得到输出,其结构如图4所示。

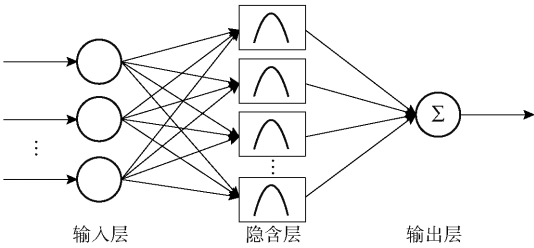


图4 径向基(RBF)神经网络结构示意图
Fig.4 RBF neural network structure diagram

径向基神经网络中常用的径向基函数是高斯基函数,设神经网络的径向基向量为 $\mathbf{H}=[h_1, h_2, \dots, h_k]^T$, h_k 为基函数,其表达式为

$$h_i = \exp \left(\frac{\| \mathbf{x}_s - \mathbf{c}_i \|^2}{2\sigma_i^2} \right) \quad (i=1,2,\dots,k; s=1,2,\dots,S) \tag{3}$$

式中 $\mathbf{x}_s$ ——隐含层第 s 个神经元的输入向量
 $\mathbf{c}_i$ ——隐含层第 i 个神经元高斯基函数的中心
 σ_i ——隐含层第 i 个神经元高斯基函数的方差
 $\| \cdot \|$ ——欧式范数,表示输入向量与中心向量的距离
 k ——隐含层神经元个数
 S ——输入向量总数

假设 RBF 神经网络隐含层和输出层的连接权值为 $\mathbf{W}=[w_1, w_2, \dots, w_k]^T$,则输出层的输出为

$$O = \mathbf{W}^T \mathbf{H} = w_1 h_1 + w_2 h_2 + \dots + w_k h_k \tag{4}$$

将式(3)代入式(4),可得输出表达式为

$$O = \sum_{i=1}^k w_i \exp \left(\frac{1}{2\sigma_i^2} \| \mathbf{x}_s - \mathbf{c}_i \|^2 \right) \tag{5}$$

基于自组织选取中心的 RBF 神经网络学习算法进行参数的求解,该方法由两个阶段组成,一是自组织学习阶段:求解隐含层基函数的中心与方差;二是监督学习阶段:求解隐含层到输出层权值的阶段。具体步骤为:

(1)基于 K-均值聚类方法求基函数中心^[21]
初始化聚类中心。随机选取 k 个训练样本作为聚类中心 $\mathbf{c}_i (i=1,2,\dots,k)$ 。将输入的训练样本集合按最近邻规则分组。即按照 $\mathbf{x}_s$ 与中心 $\mathbf{c}_i$ 之间的欧氏距离将 $\mathbf{x}_s$ 分配给中心为 $\mathbf{c}_i$ 的输入样本集合 ϑ_s 。

重新调整聚类中心,计算各个聚类集合中训练样本平均值,即新的聚类中心 $\mathbf{c}_i$,直到新的聚类中心不再发生变化时,所得到的 $\mathbf{c}_i$ 为 RBF 神经网络最终的基函数中心。

(2)求解方差 σ_i

本文 RBF 神经网络的基函数为高斯基函数,其方差 σ_i 计算式为

$$\sigma_i = \frac{c_{\max}}{\sqrt{2k}} \quad (i=1,2,\dots,k) \tag{6}$$

式中 $c_{\max}$ ——所选取中心间的最大距离

(3)计算隐含层和输出层间的权值

隐含层至输出层之间神经元的连接权值 $\mathbf{W}$,可利用最小二乘法直接计算,即

$$\mathbf{W} = \exp \left(\frac{k}{c_{\max}^2} \| \mathbf{x}_s - \mathbf{c}_i \|^2 \right) \tag{7}$$

RBF 神经网络训练过程中,径向基函数神经网络可自动增加径向基神经元,直到达到所设立的误差为止。另外,径向基函数的扩展速度对网络的收敛性影响很大,一般不宜过大。经过反复试验和调试,设定均方误差为 0.01,径向基函数扩展速度为 0.5。

将 2010 年 7 月上旬—2018 年 8 月下旬的 LAI 数据作为分析建模数据,2016—2018 年每年 9 月上旬—下旬的 LAI 数据作为检验数据。逐像素提取多旬 LAI 建模数据组成一维时间序列,分别作为模型的输入数据,运用 ARIMA 模型对河北中部平原所有像素未来的 LAI 变化状况进行多步预测^[22]。RBF 神经网络利用相同的输入数据,设置输入层节点个数为 3,输出层节点个数为 1^[23],基于滚动预测^[24]的方法对 LAI 时间序列进行多步预测。在预测过程中,将得到的预测值作为下一步预测的输入来计算出下一步的预测值,如此迭代,即完成 LAI 向前 1、2、3 步的预测。

2.2 夏玉米单产预测方法

根据该地区夏玉米主要生育期划分,一般 7 月上旬至 7 月中旬为出苗—拔节期、7 月下旬至 8 月上旬为拔节—抽雄期、8 月中旬至 9 月上旬为抽雄—乳熟期、9 月中旬至 9 月下旬为乳熟—成熟期,分别取各生育时期内所包含的多旬 LAI 的平均值作为该生育时期的 LAI 值。基于文献[25]各生育期 LAI 的权重系数:出苗—拔节期为 0.148 3,拔节—抽雄期为 0.360 8,抽雄—乳熟期为 0.274 5,乳熟—成熟期为 0.216 4。另外,由于文献[25]中为对比分析不同特征参数与夏玉米单产的线性关系,对各生育期权重系数对应的多旬 LAI 值均按最大值为 7,最小值为 0 进行了归一化处理,值域范围为[0,1]。

因此,在使用上述权重系数计算各县(区)加权 LAI 时需对旬 LAI 进行同样的归一化处理。

对河北中部平原 53 个县(区)2010—2015 年夏玉米主要生育期的加权 LAI 与玉米单产进行回归分析^[25],得到夏玉米单产的回归模型为: $Y = 5\,651X + 4\,493$, $P < 0.001$ 。其中, X 表示加权 LAI, Y 表示夏玉米单产(kg/hm^2)。基于 2010—2015 年的单产监测结果与夏玉米实际单产结果,应用线性回归分析的方法分析两者之间的相关性(图 5),可以发现,在样本数 $n = 318$ 情况下,监测单产与实际单产的相关系数(r)为 0.487,达到显著性水平($P < 0.001$),决定系数(R^2)为 0.237,说明基于粒子群优化投影寻踪法的估产模型的精度较高。

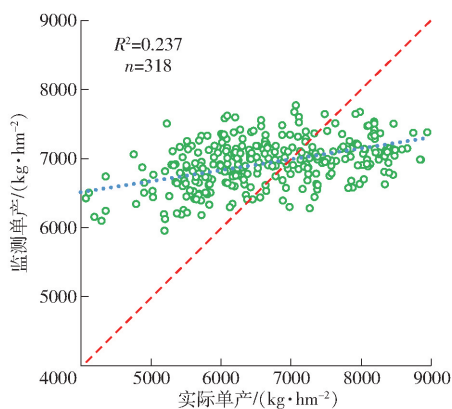


图5 基于粒子群优化投影寻踪法的监测单产与实际单产散点图

Fig.5 Scattered plot of actual yields and estimated ones based on projection pursuit with particle swarm optimization

利用加权 LAI 与夏玉米单产间的线性关系,结合 LAI 预测数据,可在玉米收获前 1—3 旬进行夏玉米单产预测。其中,向前 1 旬的单产预测过程为:利用 7 月上旬至 9 月中旬的 LAI 监测数据和 9 月下旬的 LAI 预测数据,得到加权 LAI,再根据上述产量回归模型得到夏玉米单产 1 旬预测结果;向前 2 旬的单产预测过程为:利用 7 月上旬至 9 月上旬的 LAI 监测数据和 9 月中旬、9 月下旬的 LAI 预测数据,得到加权 LAI,再根据产量回归模型得到夏玉米单产 2 旬预测结果;向前 3 旬的单产预测过程为:利用 7 月上旬至 8 月下旬的 LAI 监测数据和 9 月上旬至 9 月下旬的预测数据,得到加权 LAI,再根据产量回归模型得到夏玉米单产 3 旬预测结果。

应用相对误差(Relative error, RE)、绝对误差(Absolute error, AE)与均方根误差(Root mean square error, RMSE)作为河北中部平原夏玉米单产预测精度的评价指标。首先,分别统计各县(区)预测单产及监测单产的平均产量,并计算各县(区)平均产量的相对误差,以对县域尺度上单产预测结果

进行精度评价。

逐像素计算预测单产与监测单产间的相对误差和绝对误差,得到预测单产相对误差的空间分布图并统计每旬预测单产结果所有像素绝对误差的最大值、最小值、平均值和均方根误差,对像素尺度上单产预测结果进行精度评价。

3 结果与分析

3.1 叶面积指数预测结果及精度评价

基于 ARIMA 模型和 RBF 神经网络分别逐像素对河北中部平原 LAI 进行预测,图 6b、6c 从上至下依次为基于两模型 2017 年 8 月下旬分别向前预测 1、2、3 步得到的 2017 年 9 月上旬至下旬 LAI 预测结果。ARIMA 模型和 RBF 神经网络 1 步、2 步和 3 步预测结果都基本反映了叶面积指数监测结果的时空分布特征。从时间上看,9 月上旬至下旬属于夏玉米的花粒期阶段(包括抽雄—乳熟期和乳熟—成熟期)中后期,此时夏玉米从以营养生长为主转向以生殖生长为主,生殖器官生长速度加快,而营养器官(叶片等)生长速率减缓,植株下部叶龄较大叶片开始枯黄,LAI 开始缓慢下降。两模型 LAI 预测结果均准确反映了夏玉米 LAI 在 9 月上旬至 9 月下旬间变化特点,叶面积指数随着时间的增长而降低,并且受降水等环境因子的影响,各像素叶面积指数的下降速率不同,降水较丰沛的地区 LAI 下降较慢,而降水偏少的地区 LAI 下降较快,主要是因为降水充足可延缓叶片衰老,这与文献[26]的研究结果一致。以石家庄市藁城区为例,其 2017 年 9 月降水较常年偏多,两模型 LAI 预测结果均反映了藁城区各像素 LAI 较其他同期降水偏少的县区(新乐市、高邑县等)下降速率慢,表明两模型对 LAI 变化趋势预测较为准确。从空间上看,预测结果显示河北中部平原叶面积指数具有较明显的区域特征,沧州市西部及衡水市南部地区 LAI 偏高,廊坊市及沧州市东部地区 LAI 偏低,整体特征与监测结果较吻合。

对比分析两个模型的 LAI 预测精度,计算得到两模型 2016—2018 年 1 步、2 步和 3 步预测与监测结果的绝对误差(预测值与监测值的差)和绝对误差频数分布图(共 $3 \times 22\,985$ 个像素,图 7)。结果表明,1 步预测结果的绝对误差分布较为集中,两模型峰值十分接近,ARIMA 模型为 $0.04 \text{ m}^2/\text{m}^2$,RBF 神经网络模型为 $0.07 \text{ m}^2/\text{m}^2$ 。不同的是 LAI 预测结果的误差范围,对比频数大于 100 时绝对误差的分布范围,ARIMA 模型主要分布在 $[-1.86 \text{ m}^2/\text{m}^2, 1.43 \text{ m}^2/\text{m}^2]$,而 RBF 神经网络主要分布在

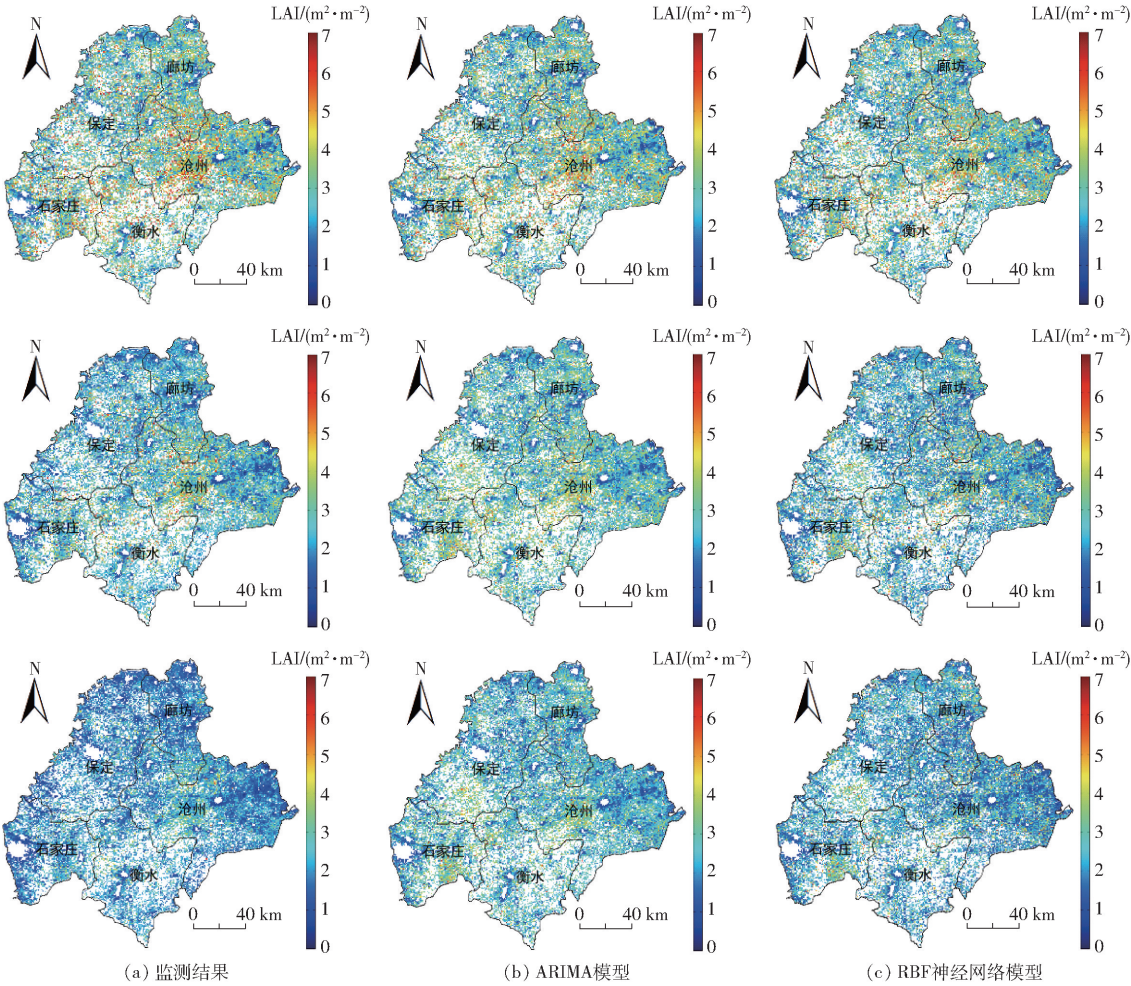


图6 2017年9月预测结果及实际监测结果

Fig. 6 Forecasting and monitoring results of September 2017

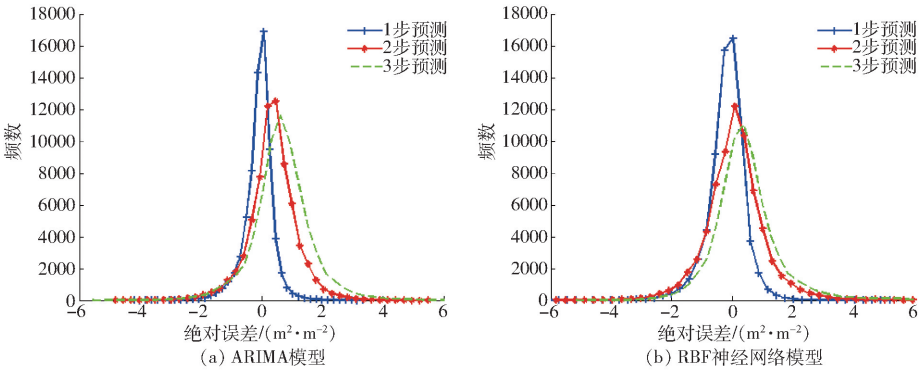


Fig. 7 Frequency distributions of absolute errors of forecasting results of two models in September 2017

$[-2.56\text{ m}^2/\text{m}^2, 1.77\text{ m}^2/\text{m}^2]$,较 ARIMA 模型误差分布更为分散。随预测步长增加,两模型误差范围均呈增大趋势,3 步预测结果误差分布较 2 步预测结果分散。另外,逐像素计算得到两模型 2016—2018 年 1 步、2 步和 3 步预测结果的平均绝对误差和均方根误差(表 1),结果表明,基于 ARIMA 模型 1、2 步预测结果的 MAE 和 RMSE 均低于基于 RBF 神经网络的误差,MAE 分别降低了 $0.12、0.05\text{ m}^2/\text{m}^2$, RMSE 分别降低了 $0.18、0.14\text{ m}^2/\text{m}^2$,3 步预测结果

的 MAE 虽较 RBF 神经网络的 MAE 高 $0.10\text{ m}^2/\text{m}^2$,但两者 RMSE 相等。整体来看,ARIMA 模型预测结果准确性和稳定性更好,预测结果反映的 LAI 变化与实际情况更为吻合,更适合河北中部平原的夏玉米 LAI 预测。

3.2 玉米单产预测结果及精度评价

利用 LAI 遥感监测结果结合基于 ARIMA 模型的 LAI 预测结果,采用粒子群优化投影寻踪法确定的夏玉米单产回归模型得到 2016—2018 年的河北

表1 ARIMA 模型和 RBF 神经网络模型预测误差的统计分析

Tab.1 Statistical results of forecasting errors of ARIMA model and RBF neural network model m^2/m^2

预测步长	模型	MAE	RMSE
1 步	ARIMA	0.33	0.48
	RBF 神经网络	0.45	0.66
2 步	ARIMA	0.67	0.87
	RBF 神经网络	0.72	1.01
3 步	ARIMA	0.93	1.16
	RBF 神经网络	0.83	1.16

中部平原夏玉米单产监测结果和 1—3 旬单产预测结果(图 8)。由图 8 单产监测结果可得,河北中部平原西部玉米单产最高,南部和北部次之,东部单产最低。2016 年,河北中部平原 53 县(区)平均单产为 6 912 kg/hm²,西部大部分县(区)高于 7 000 kg/hm²,少部分县(区)接近 8 000 kg/hm²;2017 年,大部分地区单产约为 6 800 kg/hm²,东部地区单产较低,约为 6 300 kg/hm²;2018 年单产略高于 2017 年,西部大部分县(区)单产在 7 000 kg/hm²左右,东北部地区单产相对偏低,在 6 300 kg/hm²左右。从预测结果来看,向前 1 旬、2 旬和 3 旬预测单产均与监测单产较吻合,预测单产时空变化规律与监测产量一致,整体预测精度较高。

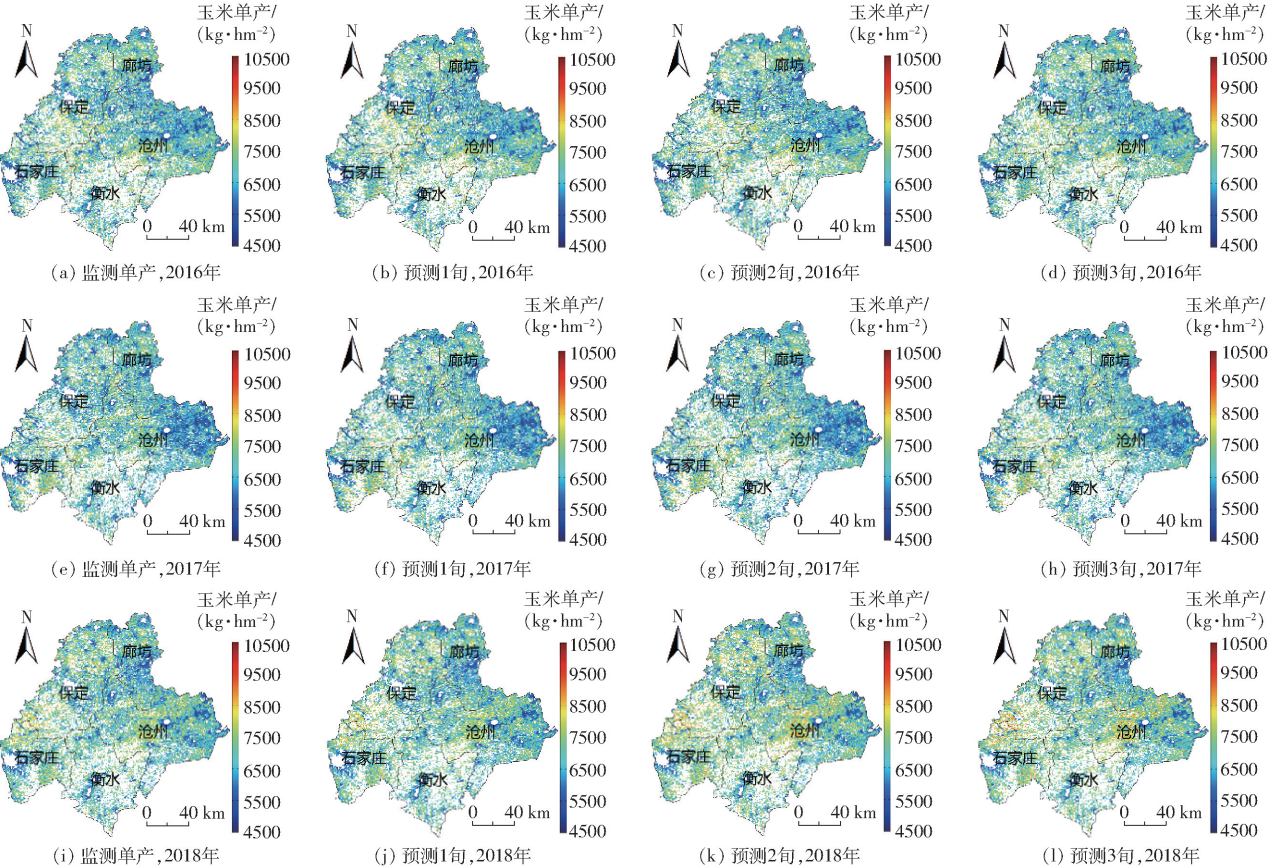


图 8 河北中部平原夏玉米产量监测和预测结果

Fig. 8 Monitoring and forecasting yields of summer maize in central plain of Hebei

3.2.1 县域尺度单产预测结果精度评价

计算得到河北中部平原各县(区)夏玉米 2016—2018 年的监测单产与向前 1—3 旬预测单产之间的相对误差及其统计直方图见图 9,对县域尺度夏玉米单产预测结果精度进行评价。结果表明,向前 1 旬的单产预测精度较高,2016—2018 年各县(区)向前 1 旬的相对误差十分接近,均分布在 1% 以内,最大值为 2016 年定兴县的 0.79%。向前 2 旬较向前 1 旬的相对误差有所增加,误差分布相对分散,最大值为 2018 年献县的 3.53%。向前 3 旬的相对误差分布与向前 2 旬相似,并且误差分布不均匀,误差最小为 2017 年海兴县的 1.21%,而最大为 2018 年献县的 3.73%。从年际变化来看,各县(区)向前 1—3 旬的单产预测结果在不同年份的相对误差均十分接近,误差波动最大的为向前 3 旬中献县,2018 年较 2017 年增大了 3.64 个百分点,其余相对误差的波动值均小于该值。总体来说,向前 1 旬、2 旬和 3 旬的玉米单产预测结果均与监测结果十分接近,预测精度虽然会随预测步长增加而降低,但整体预测精度较高,各县(区)单产预测结果相对误差均在 4% 以内,说明基于 ARIMA 模型的 LAI 预测数据可以较好地反映玉米生育后期的长势变化及干物质向籽粒转移的能力。

为进一步验证县域尺度夏玉米单产预测的精度,基于2016—2018年的单产预测结果与夏玉米监测结果,应用线性回归分析的方法分析两者之间的相关性,见图10,可以发现,在样本数 $n=159$ 的情

况下,向前1旬、2旬和3旬预测单产与监测单产均呈显著的正相关($P<0.001$),决定系数(R^2)分别为0.998、0.960和0.947,说明基于该方法的县域尺度预测单产精度较高。

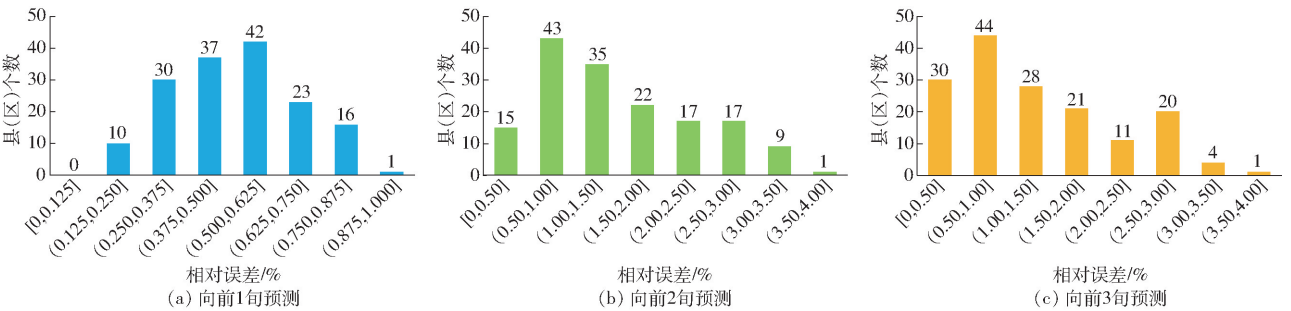


图9 河北中部平原各县(区)玉米单产预测的相对误差

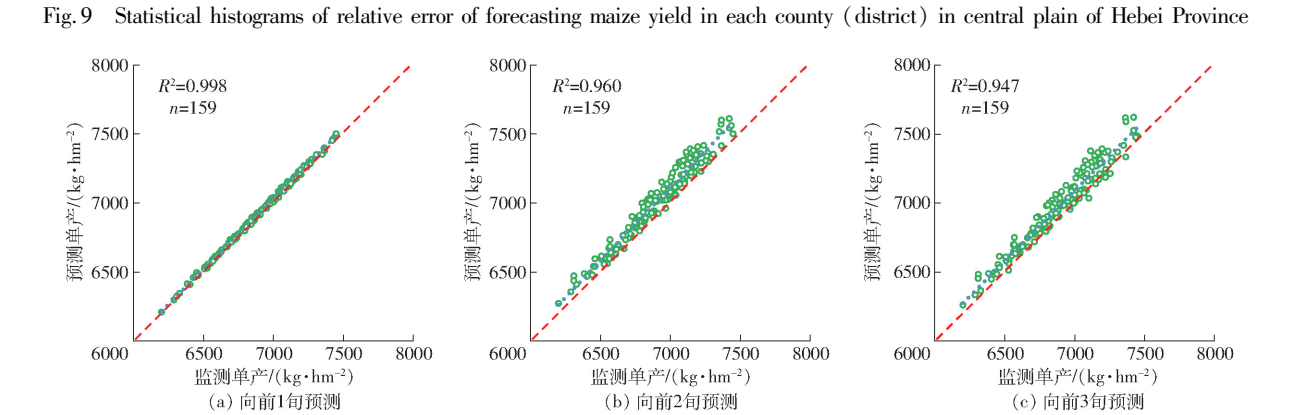


图10 河北中部平原预测单产与监测单产散点图

Fig. 10 Scattered plots of forecasting yields and monitoring ones in central plain of Hebei Province

3.2.2 像素尺度单产预测结果精度评价

逐像素计算河北中部平原预测单产与监测单产的相对误差,得到2016—2018年3种预测单产相对误差的空间分布图(图11),对像素尺度夏玉米单产预测结果精度进行评价。结果表明,2016—2018年向前1旬预测单产相对误差分布相似,分别有89.8%、92.0%和93.2%像素相对误差小于1%,表明向前1旬预测单产精度较好;随预测时间的增加,2016—2018年向前2旬和向前3旬单产预测结果相对误差均呈增加趋势,且2016年和2017年向前2旬和3旬单产预测结果相对误差空间分布相似,2018年向前2旬和3旬单产预测结果相对误差较2016年和2017年偏大,其向前3旬单产预测结果具有较大不确定性,但仍有90.3%像素相对误差均小于5%,表明不同年份间夏玉米向前1旬、2旬和3旬单产预测结果精度存在一定差异,但整体来看不同年份间单产预测精度均较高。

逐像素计算河北中部平原2016—2018年预测单产与监测单产的绝对误差,并统计其绝对误差的最大值、最小值、均值和均方根误差(表2)。结果表明,2016—2018年1旬预测单产RMSE分布于

27.47~32.17 kg/hm²,且各像素绝对误差均不大于288 kg/hm²,表明向前1旬的单产预测精度较高且十分稳定。误差随预测时间的增加而增大,3旬预测单产的各项误差指标较前两旬都有明显的增加,个别像素绝对误差最大达到1285 kg/hm²,但考虑到统计像素较多(每旬22985个),且绝大多数像素绝对误差较低,因此可认为整体预测精度较高,表明基于该方法可实现对研究区域内夏玉米收获前3旬进行准确的单产预测。

4 讨论

针对以往单产预测模型采用单一生育期或单一时段的特征参数进行单产预测,本研究综合考虑夏玉米主要生育期的特征参数,并且选择精度较高的基于ARIMA模型的LAI预测数据代替夏玉米生育后期的LAI监测数据,实现了夏玉米收获前1—3旬进行单产预测,整体预测精度较高,各县(区)单产预测结果相对误差均在4%以内。LAI预测精度的关键是预测模型的选择,在今后的研究中可尝试其他预测模型,如遗传算法优化的RBF神经网络、ARIMA-RBF组合预测模型等,以期进一步提高

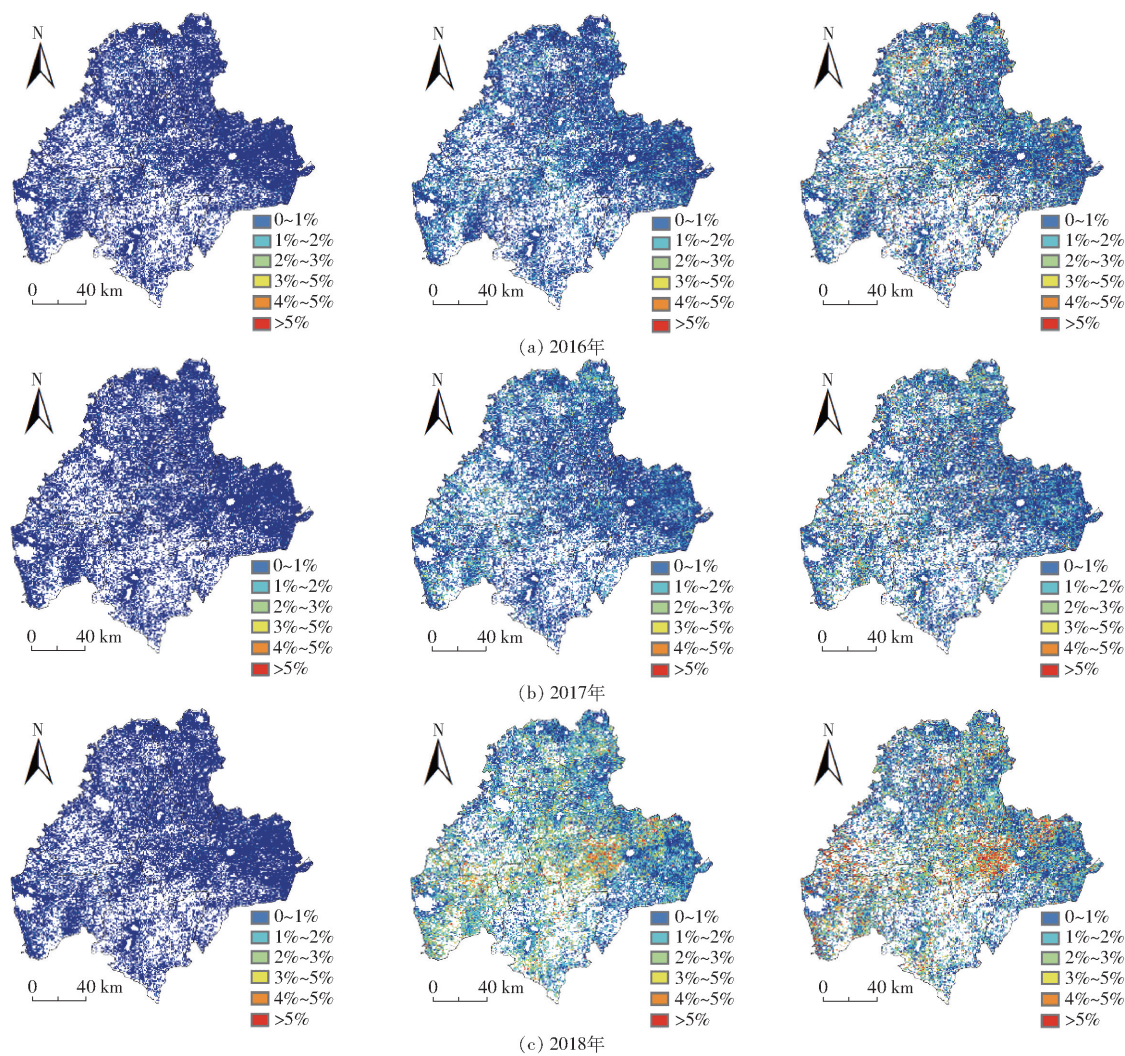


图 11 河北中部平原夏玉米单产预测相对误差空间分布

Fig. 11 Relative error spatial distributions of forecasting summer maize yield in central plain of Hebei Province

表 2 河北中部平原夏玉米单产绝对误差分布统计

Tab.2 Absolute errors of monitoring and forecasting yields of summer maize in central plain of Hebei													kg/hm ²
年份	向前 1 旬预测				向前 2 旬预测				向前 3 旬预测				
	最大值	最小值	均值	RMSE	最大值	最小值	均值	RMSE	最大值	最小值	均值	RMSE	
2016	288	-177	33.62	32.17	643	-643	45.50	73.33	1 285	-1263	52.82	123.51	
2017	244	-177	20.68	29.50	665	-643	80.99	82.03	1 130	-975	56.76	113.52	
2018	244	-177	16.55	27.47	753	-510	169.94	141.75	1 042	-931	159.36	162.64	

LAI 模型预测精度。除此之外,夏玉米研究区识别的精度也会在一定程度上对 LAI 预测精度产生影响,本文采用的夏玉米种植区提取方法在一定程度上保证了算法精度,在今后的研究中可通过提高夏玉米识别精度进一步减少混合地物对 LAI 预测精度的影响。

夏玉米单产的预测精度除了受 LAI 预测精度影响外,还受各生育期 LAI 的权重以及加权 LAI 与夏玉米单产间的回归方程影响,不同的赋权方法可能导致夏玉米单产预测精度也不同,合理的赋权方法对保证基于该方法的夏玉米单产预测精度十分重

要。另外,影响夏玉米单产的因素有很多,除了受到作物自身的生理因素影响外,还受到作物生长的生态环境条件的综合影响,如水分、养分、温度、光照等,它们可能成为最终产量形成的胁迫因子,使得产量发生增减的波动。今后的研究中,需要综合考虑多种因素对作物最终单产的影响,提高作物单产预测模型的普适性和准确性。

5 结论

(1)基于 ARIMA 模型的 LAI 1、2 步预测结果的 MAE 和 RMSE 均低于基于 RBF 神经网络的误差,

MAE 分别降低了 0.12、0.05 m²/m², RMSE 分别降低了 0.18、0.14 m²/m², 3 步预测结果的 MAE 虽较 RBF 神经网络高 0.10 m²/m², 但两者 RMSE 相等。整体来看, ARIMA 模型预测结果的准确性和稳定性更好, 更适合于河北中部平原的夏玉米 LAI 预测。

(2) 基于粒子群优化投影寻踪法确定的单产回归模型及 ARIMA 模型, 对 2016—2018 年河北中部

平原进行向前 1 旬、2 旬和 3 旬夏玉米的单产预测, 结果表明, 无论是县域尺度还是像素尺度, 各旬单产预测结果均与监测结果十分接近, 并且随预测时间的增加预测结果的不确定性增大, 但整体预测精度较高, 2016—2018 年县域尺度预测单产与监测单产间最大相对误差仅为 3.73%。

参 考 文 献

[1] 陈亮. 农作物生长动态监测与定量评价[M]. 北京:气象出版社,2016.

[2] 刘峻明,和晓彤,王鹏新,等. 长时间序列气象数据结合随机森林法早期预测冬小麦产量[J]. 农业工程学报,2019,35(6):158-166.
LIU Junming, HE Xiaotong, WANG Pengxin, et al. Early prediction of winter wheat yield with long time series meteorological data and random forest method[J]. Transactions of the CSAE, 2019, 35(6):158-166. (in Chinese)

[3] 王鹏新,孙辉涛,解毅,等. 基于 LAI 和 VTCI 及粒子滤波同化算法的冬小麦单产估测[J/OL]. 农业机械学报,2016,47(4):248-256.
WANG Pengxin, SUN Huitao, XIE Yi, et al. Winter wheat yield estimation based on particle filter assimilation algorithm and remotely sensed LAI and VTCI[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(4):248-256. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160433&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2016.04.033. (in Chinese)

[4] 黄健熙,黄海,马鸿元,等. 遥感与作物生长模型数据同化应用综述[J]. 农业工程学报,2018,34(21):144-156.
HUANG Jianxi, HUANG Hai, MA Hongyuan, et al. Review on data assimilation of remote sensing and crop growth models[J]. Transactions of the CSAE, 2018, 34(21):144-156. (in Chinese)

[5] 朱再春,陈联裙,张锦水,等. 基于信息扩散和关键期遥感数据的冬小麦估产模型[J]. 农业工程学报,2011,27(2):187-193.
ZHU Zaichun, CHEN Lianqun, ZHANG Jinshui, et al. Winter wheat yield estimation model based on information diffusion and remote sensing data at major growth stages[J]. Transactions of the CSAE, 2011, 27(2):187-193. (in Chinese)

[6] 赵英时. 遥感应用分析原理与方法[M]. 北京:科学出版社,2003.

[7] 徐新刚,吴炳方,蒙继华,等. 农作物单产遥感估算模型研究进展[J]. 农业工程学报,2008,24(2):290-298.
XU Xingang, WU Bingfang, MENG Jihua, et al. Research advances in crop yield estimation models based on remote sensing[J]. Transactions of the CSAE, 2008, 24(2):290-298. (in Chinese)

[8] 林卉,梁亮,张连蓬,等. 基于支持向量机回归算法的小麦叶面积指数高光谱遥感反演[J]. 农业工程学报,2013,29(11):139-146.
LIN Hui, LIANG Liang, ZHANG Lianpeng, et al. Wheat leaf area index inversion with hyperspectral remote sensing based on support vector regression algorithm[J]. Transactions of the CSAE, 2013, 29(11):139-146. (in Chinese)

[9] 任建强,陈仲新,周清波,等. 基于叶面积指数反演的区域冬小麦单产遥感估测[J]. 应用生态学报,2010,21(11):2883-2888.
REN Jianqiang, CHEN Zhongxin, ZHOU Qingbo, et al. LAI-based regional winter wheat yield estimation by remote sensing[J]. Chinese Journal of Applied Ecology, 2010, 21(11):2883-2888. (in Chinese)

[10] 王蕾,王鹏新,李俐,等. 河北省中部平原玉米长势遥感综合监测[J]. 资源科学,2018,40(10):2099-2109.
WANG Lei, WANG Pengxin, LI Li, et al. Integrated maize growth monitoring based on gray correlation analysis and remote sense data in the central plain of Hebei Province [J]. Resources Science, 2018, 40(10):2099-2109. (in Chinese)

[11] 王鹏新,齐璇,李俐,等. 基于随机森林回归的玉米单产估测[J/OL]. 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/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(7):237-245. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190726&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.07.026. (in Chinese)

[12] 薛昌颖,刘荣花,马志红. 黄淮海地区夏玉米干旱等级划分[J]. 农业工程学报,2014,30(16):147-156.
XUE Changying, LIU Ronghua, MA Zhihong. Drought grade classification of summer maize in Huang-Huai-Hai area [J]. Transactions of the CSAE, 2014, 30(16):147-156. (in Chinese)

[13] BIAN Jinhu, LI Ainong, SONG Mengqiang, et al. Reconstruction of NDVI time-series datasets of MODIS based on Savitzky-Golay filter[J]. Journal of Remote Sensing, 2010, 14(4):725-741.

[14] 柴琳娜,屈永华,张立新,等. 基于自回归神经网络的时间序列叶面积指数估算[J]. 地球科学进展,2009,24(7):756-768.
CHAI Linna, QU Yonghua, ZHANG Lixin, et al. Estimating time series leaf area index based on recurrent neural networks[J]. Advances in Earth Science, 2009, 24(7):756-768. (in Chinese)

[15] 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)

- [16] BOX G E P, JENKINS G M. Time series analysis forecasting and control[M]. San Francisco: Holden Day, 1976: 525.
- [17] SHARAFI M, GHAEM H, TABATABAEE H R, et al. Forecasting the number of zoonotic cutaneous leishmaniasis cases in south of Fars Province, Iran using seasonal ARIMA time series method[J]. Asian Pacific Journal of Tropical Medicine, 2017, 10(1): 79 – 86.
- [18] 左志宇, 毛罕平, 张晓东, 等. 基于时序分析法的温室温度预测模型[J]. 农业机械学报, 2010, 41(11): 173 – 177, 182.
ZUO Zhiyu, MAO Hanping, ZHANG Xiaodong, et al. Forecast model of greenhouse temperature based on time series method[J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(11): 173 – 177, 182. (in Chinese)
- [19] 李志刚, 刘伯颖, 李玲玲, 等. 基于小波包变换及 RBF 神经网络的继电器寿命预测[J]. 电工技术学报, 2015, 30(14): 233 – 240.
LI Zhigang, LIU Boying, LI Lingling, et al. Life prediction of relay based on wavelet packet transform and RBF neural network[J]. Transactions of China Electrotechnical Society, 2015, 30(14): 233 – 240. (in Chinese)
- [20] 田苗, 王鹏新, 侯姗姗, 等. 基于相空间重构与 RBF 神经网络模型的面上干旱预测研究[J]. 干旱地区农业研究, 2013, 31(6): 164 – 168, 180.
TIAN Miao, WANG Pengxin, HOU Shanshan, et al. Research of regional drought forecasting based on phase space reconstruction and RBF neural network model[J]. Agricultural Research in the Arid Areas, 2013, 31(6): 164 – 168, 180. (in Chinese)
- [21] 刘志杰, 季令, 叶玉玲, 等. 基于径向基神经网络的铁路货运量预测[J]. 铁道学报, 2006, 28(5): 1 – 5.
LIU Zhijie, JI Ling, YE Yuling, et al. Study on prediction of railway freight volumes based on RBF neural network[J]. Journal of the China Railway Society, 2006, 28(5): 1 – 5. (in Chinese)
- [22] 田苗, 王鹏新, 韩萍, 等. 基于 SARIMA 模型和条件植被温度指数的干旱预测[J/OL]. 农业机械学报, 2013, 44(2): 109 – 116.
TIAN Miao, WANG Pengxin, HAN Ping, et al. Drought forecasts based on SARIMA models and vegetation temperature condition index[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(2): 109 – 116. http://www.jcsam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20130221&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2013.02.021. (in Chinese)
- [23] 范伟, 林瑜阳, 李钟慎. 遗传算法优化的 BP 神经网络压电陶瓷蠕变预测[J]. 电机与控制学报, 2018, 22(7): 91 – 96.
FAN Wei, LIN Yuyang, LI Zhongshen. Prediction model of the creep of piezoceramic based on BP neural network optimized by genetic algorithm[J]. Electric Machines and Control, 2018, 22(7): 91 – 96. (in Chinese)
- [24] 姜春雷, 张树清, 张策, 等. 基于 SARIMA – BP 神经网络组合方法的 MODIS 叶面积指数时间序列建模与预测[J]. 光谱学与光谱分析, 2017, 37(1): 189 – 193.
JIANG Chunlei, ZHANG Shuqing, ZHANG Ce, et al. Modeling and predicting of MODIS leaf area index time series based on a hybrid SARIMA and BP neural network method[J]. Spectroscopy and Spectral Analysis, 2017, 37(1): 189 – 193. (in Chinese)
- [25] 王鹏新, 齐璇, 李俐, 等. 基于粒子群优化投影寻踪的玉米单产估测[J]. 农业工程学报, 2019, 35(13): 145 – 153.
WANG Pengxin, QI Xuan, LI Li, et al. Estimation of maize yield based on projection pursuit with particle swarm optimization[J]. Transactions of the CSAE, 2019, 35(13): 145 – 153. (in Chinese)
- [26] 王宁珍, 邓振镛, 张谋草, 等. 陇东黄土高原气候变化对玉米叶面积生长的影响研究[J]. 干旱地区农业研究, 2006, 24(2): 189 – 194.
WANG Ningzhen, DENG Zhenyong, ZHANG Moucao, et al. Discussion on water shortage and countermeasures in northwest regions of China[J]. Agricultural Research in the Arid Areas, 2006, 24(2): 189 – 194. (in Chinese)