

基于田间环境及气象数据的甘蔗产量预测方法

李修华^{1,2} 李婉¹ 张木清² 温标堂³ 叶志鹏³ 张云皓¹

(1. 广西大学电气工程学院, 南宁 530004; 2. 广西甘蔗生物学重点实验室, 南宁 530004;
3. 广西捷佳润科技股份有限公司, 南宁 530004)

摘要: 对2008—2017年间广西某甘蔗地装设的田间物联网获取的田间环境及气象数据进行了分析,研究通过这些数据预测甘蔗产量的可行性。对田间的环境数据(土壤含水率和土壤温度)和气象数据(降水量和空气温度)进行预处理,分别运用BP神经网络及遗传算法优化初始阈值的BP神经网络(GA-BP)对所选地块的甘蔗产量进行预测。结果表明,GA-BP神经网络模型的预测精度明显高于BP神经网络模型, R^2 达到0.989 4,平均相对误差仅为0.64%。说明用GA-BP神经网络预测甘蔗产量是有效、可行的。
关键词: 甘蔗; 产量预测; 环境数据; 气象数据; BP神经网络; 遗传算法
中图分类号: S566.1; S24 **文献标识码:** A **文章编号:** 1000-1298(2019)S0-0233-04

Sugarcane Yield Prediction Method Based on Field Environmental and Meteorological Data

LI Xiuhua^{1,2} LI Wan¹ ZHANG Muqing² WEN Biaotang³ YE Zhipeng³ ZHANG Yunhao¹

(1. School of Electrical Engineering, Guangxi University, Nanning 530004, China
2. Guangxi Key Laboratory for Sugarcane Biology, Nanning 530004, China
3. Guangxi JJR Technology Co., Ltd., Nanning 530004, China)

Abstract: The agricultural internet of things has been gradually popularized in agricultural planting management, but most of the application scenarios are mainly focusing on guiding growers to fertilize and irrigate crops from its acquired field data. Environmental and meteorological data of the crop's entire growth season can also reflect the microscopic and macroscopic environment of crop growth, and hence influencing the final yield. Therefore, the environmental data (soil moisture content and soil temperature) and meteorological data (precipitation and air temperature) of a sugarcane field collected from a field IoTs system were used to build up yield predict models. The collected data was ranged from the year of 2008 to 2017, and the original BP neural network and an improved BP neural network based on genetic algorithm (GA-BP) were adopted to predict the yield. Comparing the prediction results of these two models, it was found that the prediction accuracy of the GA-BP neural network model was significantly higher than that of the BP neural network model, with its R^2 reaching 0.989 4, and its average relative prediction error was only 0.64%, which was also obviously lower than that of BP neural network model (5.66%). The result indicated that the GA-BP neural network was effective and feasible in predicting sugarcane production.
Key words: sugarcane; yield prediction; environmental data; meteorological data; BP neural network; genetic algorithm

0 引言

广西是我国最大的甘蔗产区,也是我国蔗糖供应的重点区域^[1]。甘蔗的生长与气温、湿度、光照

强度、降雨量等因子密切相关^[2-4]。甘蔗产量预测是制糖工业一项重要的技术管理工作,对优化甘蔗生产作业及提高统筹管理水平具有指导作用^[5-7]。近年来,国内外对于农作物产量预测研究成果

收稿日期: 2019-04-23 修回日期: 2019-05-27
基金项目: 国家自然科学基金项目(31760342)和广西科技重大专项(桂科 AA18118037)
作者简介: 李修华(1983—),女,副教授,博士,主要从事作物检测和农业物联网研究,E-mail: lixh@gxu.edu.cn

较多。文献[8]分别运用边界线分析法、线性回归法以及神经网络算法对水稻产量进行预测,研究表明,神经网络模型的预测精度高于线性回归模型、边界分析模型的预测精度。文献[9]分别以每个星期的降雨量、最高和最低气温、早晨和黄昏的相对湿度作为神经网络模型的输入,预测小麦产量,研究了结构不同的神经网络对小麦产量的预测效果。文献[10]以降雨量和 LAI 参数为输入量,分别用多元线性回归算法和神经网络算法对土豆产量进行预测,并且对比分析了这两种模型的预测效果。文献[11]提出了一种改进的遗传 BP 算法,通过多元逐步回归法简化 BP 网络的输入因子,运用改进的遗传 BP 算法,对甘蔗产量实例数据进行了验证和分析,研究表明,改进的遗传 BP 算法总体预测效果最好。文献[12]通过灰色关联度分析确定影响粮食产量的主要指标,利用人工神经网络将这些指标应用到产量预测中,研究表明,人工神经网络预测的平均误差为 0.63%,预测精度较高。文献[13]采用去噪声修正后的冬小麦抽穗期 NDVI 峰值与单产进行回归分析,建立冬小麦产量预测模型,结果表明,模型的平均估产误差约为 7.49%,基于冬小麦抽穗期 NDVI 峰值的产量预测方法在中国冬小麦主产区(河北、河南和山东)具有一定的应用潜力。文献[14]针对晋城市的粮食产量,对线性回归模型、灰色 GM(1,1)模型和灰色多元线性回归组合模型 3 种粮食产量预测方法进行研究,结果表明,灰色多元线性回归组合模型最适宜于晋城市粮食产量的预测。

目前,国内外在农作物产量预测方面多以气象数据为研究手段进行大面积(全国、全省或全县)的宏观预测,很少将气象数据与农作物田间的环境数据相结合来微观预测某一地块的农作物产量。农业物联网在农业种植管理中已逐渐普及,但大部分应用场景还处于指导种植户对农作物给水施肥方面。实际上,作物整个生长期的环境与气象信息能反映其生长的微观及宏观条件,从而间接反映作物产量。因此,本文分析 2008—2017 年甘蔗田间环境数据(土壤含水率和土壤温度)与气象数据(降水量和空气温度),研究基于 BP 神经网络及遗传算法优化的 BP 神经网络的甘蔗产量预测方法。

1 材料与方法

1.1 数据来源

实验数据采用广西捷佳润科技股份有限公司提供的安吉实验灌溉站的环境数据和气象数据,以及安吉实验灌溉站提供的甘蔗产量信息。数据覆盖的

时间范围为 2008—2017 年,共 10 年。其中环境数据包括土壤含水率和土壤温度,气象数据包括降水量和空气温度,均为日平均数据。

1.2 方法

主要采用遗传算法^[15-17]将 BP 神经网络^[18-19]的预测输出和期望输出间的误差绝对值的和作为个体适应度,进行选择、交叉和变异操作,使个体适应度尽可能地小,从而训练出最优的 BP 神经网络初始权值和阈值参数。基于遗传算法的 BP 神经网络(GA-BP)算法的流程如图 1 所示。

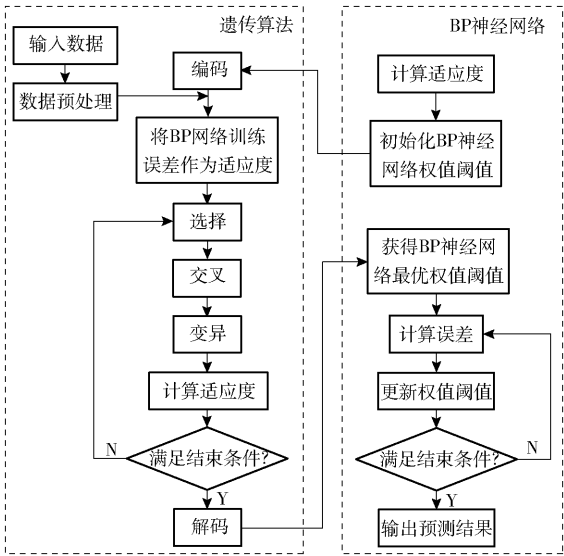


图 1 基于遗传算法的 BP 神经网络算法流程图
Fig. 1 Flow chart of BP neural network algorithm based on genetic algorithm

2 结果分析

2.1 模型建立

考虑到安吉实验灌溉站种植的甘蔗属于春植甘蔗,故只选取了甘蔗生长周期内,即每年的 3—12 月的环境数据和气象数据进行处理。由于数据覆盖 3—12 月的甘蔗整个生长期,数据量大,不利于建模。因此将每类数据进一步按月进行归类压缩,选取月土壤含水率均值、月土壤温度最小值、月土壤温度最大值、月土壤温度均值、月降水量、月平均气温、月最高气温和月最低气温这 8 个因子作为环境与气象因子,最终得到 80 个环境与气象数据。待预测的变量为甘蔗产量。为了得到可靠稳定的甘蔗产量预测模型,建模时采用了 3 折交叉验证的方法。

2.1.1 BP 神经网络模型

由以上数据描述可知,BP 神经网络的输入层节点数和输出层节点数分别为 80 和 1。隐含层节点数根据公式法和试凑法,按照预测误差最小的原则最终确定为 21。设置网络学习的收敛误差为 0.000 65,

学习速度为 0.05,最大训练次数为50 000。当收敛误差达到要求后,训练结束,得到 BP 神经网络对甘蔗产量的预测结果如表 1 所示。

表 1 BP 神经网络模型预测结果

Tab.1 Prediction results of BP neural network model

年份	预测产量/ (t·hm ⁻²)	实际产量/ (t·hm ⁻²)	预测误差/ (t·hm ⁻²)	相对 误差/%	平均相对 误差/%
2008	106.22	97.82	8.40	8.59	5.66
2009	102.00	100.40	1.60	1.59	
2010	102.63	101.92	0.71	0.70	
2011	97.82	102.00	-4.18	-4.10	
2012	117.28	102.63	14.65	14.27	
2013	121.20	121.20	0	0	
2014	80.65	80.65	0	0	
2015	101.92	122.88	-20.96	-17.06	
2016	100.40	106.22	-5.82	-5.48	
2017	122.88	117.28	5.60	4.77	

由表 1 可知,BP 神经网络产量预测模型验证集的平均相对误差为 5.66%,说明 BP 神经网络模型可以应用在甘蔗产量预测方面,但个别预测结果误差较大,仍需要改进。

2.1.2 基于遗传算法的 BP 神经网络模型

由于 BP 神经网络模型预测结果并不理想,且 BP 神经网络中的初始权值阈值是随机得到的,而遗传算法可以优化 BP 神经网络的初始权值和阈值^[20],使 BP 神经网络预测结果更优,故考虑用 GA-BP 算法提高预测精度。为了对比 GA-BP 神经网络与 BP 神经网络,在训练 GA-BP 神经网络时,参数选取与 BP 神经网络一致。

由于实数编码方式不需要进行数值转换,可以直接在解的表现型上进行遗传算法的操作,因此本文中的个体编码方式采用的是实数法,由于输入层的节点数为 80,隐含层的节点数为 21,输出层的节点数为 1,则共有 1 701 个权值和 22 个阈值,遗传算法需要优化的参数为 1 723 个,即染色体长度为 1 723。将遗传算法中的参数设置为:进化代数 100,种群规模 40,交叉概率 0.7,变异概率 0.01。以训练好的 GA-BP 神经网络对甘蔗产量的预测结果如表 2 所示。

由表 2 可知,GA-BP 神经网络模型对甘蔗产量的预测精度明显高于 BP 神经网络模型。

2.2 模型对比

分别用 BP 神经网络算法和 GA-BP 神经网络算法对甘蔗产量进行预测,将两种模型得到的预测值与实际值进行相关性分析,结果如图 2 所示。

通过对比发现,GA-BP 神经网络模型的验证集平均相对误差(0.64%)明显低于BP神经网络模

表 2 基于遗传算法的 BP 神经网络模型预测结果
Tab.2 Prediction results of BP neural network model with genetic algorithm

年份	预测产量/ (t·hm ⁻²)	实际产量/ (t·hm ⁻²)	预测误差/ (t·hm ⁻²)	相对 误差/%	平均相对 误差/%
2008	97.80	97.82	-0.02	-0.02	0.64
2009	100.41	100.40	0.01	0.01	
2010	101.92	101.92	0	0	
2011	99.96	102.00	-2.04	-2.00	
2012	102.79	102.63	0.16	0.16	
2013	118.73	121.20	-2.47	-2.04	
2014	80.63	80.65	-0.02	-0.02	
2015	122.89	122.88	0.01	0.01	
2016	106.41	106.22	0.19	0.18	
2017	119.56	117.28	2.28	1.94	

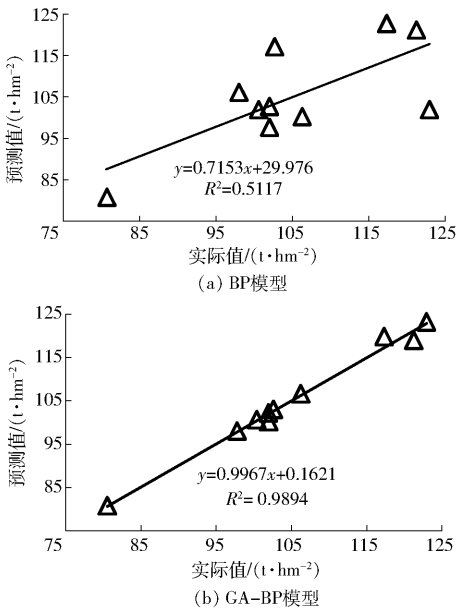


图 2 两种模型预测值与实测值间的相关性分析
Fig.2 Correlation analysis between predicted and measured values of two models

型(5.66%),而预测值与实际值之间的决定系数 R^2 从 0.511 7 增加到了 0.989 4,说明 GA-BP 神经网络算法对甘蔗产量的预测能力明显优于 BP 神经网络算法。

目前建立的预测模型是基于全生长期数据的事后预测模型,没有考虑到当季作物生长期内的动态预测需求,下一步还可尝试基于任一生长期或时间节点,建立对当季作物产量的预报模型,以便为作物管理提供更多决策支持。

3 结论

(1)运用 BP 神经网络模型对甘蔗产量预测效果较差,预测值与实测值的 R^2 仅为 0.511 7。

(2)GA-BP 神经网络模型的预测精度明显高于 BP 神经网络模型, R^2 达到 0.989 4,说明用 GA-BP 神经网络预测甘蔗产量是有效和可行的。

参考文献

- [1] 李炳杨. 广西甘蔗种植现状、问题及对策[J]. 热带农业科学, 2018, 38(4): 119–127.
LI Bingyang. The present situation, problems and countermeasures of sugarcane cultivation in Guangxi[J]. Chinese Journal of Tropical Agriculture, 2018, 38(4): 119–127. (in Chinese)
- [2] 肖凌, 梁玉莲, 杨宇红, 等. 南宁市甘蔗气候产量预测模型[J]. 广东农业科学, 2013, 40(4): 174–176, 238.
XIAO Ling, LIANG Yulian, YANG Yuhong, et al. Sugarcane climatic yield predicting model of Nanning[J]. Guangdong Agricultural Sciences, 2013, 40(4): 174–176, 238. (in Chinese)
- [3] 何燕, 谭宗琨, 丁美花, 等. 制约广西甘蔗产量及蔗糖分含量的关键气象条件研究[J]. 安徽农业科学, 2008, 36(8): 3181–3184.
- [4] 付建涛, 孙东磊, 陈立君, 等. 温度和降水变化与甘蔗产量的关系分析[J]. 热带农业科学, 2018, 38(6): 13–17.
FU Jiantao, SUN Donglei, CHEN Lijun, et al. Relationship analysis of the temperature and precipitation with the yield of sugarcane[J]. Chinese Journal of Tropical Agriculture, 2018, 38(6): 13–17. (in Chinese)
- [5] 李奇伟. 现代甘蔗改良技术[M]. 广州: 华南理工大学出版社, 2000.
- [6] 阮红燕. 气候变化对广西甘蔗生产潜力影响的模拟研究[D]. 南宁: 广西大学, 2018.
RUAN Hongyan. Climate change projects impacts on sugarcane productivity in Guangxi[D]. Nanning: Guangxi University, 2018. (in Chinese)
- [7] 欧钊荣, 谭宗琨, 何燕, 等. BP神经网络模型在广西原料蔗产量预报中的应用[J]. 中国农业气象, 2008, 29(2): 213–216.
OU Zhaorong, TAN Zongkun, HE Yan, et al. Application of BP neural network in yield predication of sugarcane (*Saccharum officinarum* L.) in Guangxi Province[J]. Chinese Journal of Agrometeorology, 2008, 29(2): 213–216. (in Chinese)
- [8] EZRIN M H, AMIN M S M, ANUAR A R, et al. Relationship between rice yield and apparent electrical conductivity of paddy soils[J]. American Journal of Applied Sciences, 2010, 7(1): 63.
- [9] MEENA P K, HARDAHA M K, KHARE D, et al. Neural method for site-specific yield prediction[C]//Proceedings of the Third International Conference on Soft Computing for Problem Solving. Springer India, 2014: 237–246.
- [10] FORTIN J G, ANCTIL F, PARENT L É, et al. Site-specific early season potato yield forecast by neural network in eastern Canada[J]. Precision Agriculture, 2011, 12(6): 905–923.
- [11] 徐永春, 张森文. 改进遗传神经网络在甘蔗产量预测中的应用[J]. 华南农业大学学报, 2010, 31(3): 102–104.
XU Yongchun, ZHANG Senwen. The Application on the sugar cane forecast with the neural networks using improved genetic algorithms[J]. Journal of South China Agricultural University, 2010, 31(3): 102–104. (in Chinese)
- [12] 姜新. 河南省粮食产量影响因素和预测方法研究[J]. 中国农学通报, 2019, 35(1): 154–158.
JIANG Xin. Influencing factors and forecasting methods of grain yield in Henan Province[J]. Chinese Agricultural Science Bulletin, 2019, 35(1): 154–158. (in Chinese)
- [13] 黄健熙, 罗倩, 刘晓暄, 等. 基于时间序列 MODIS NDVI 的冬小麦产量预测方法[J]. 农业机械学报, 2016, 47(2): 295–301.
HUANG Jianxi, LUO Qian, LIU Xiaoxuan, et al. Winter wheat yield forecasting based on time series of MODIS NDVI[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(2): 295–301. (in Chinese)
- [14] 和文超, 师学义, 邓青云, 等. 土地利用规划修编中粮食产量预测方法比较[J]. 农业工程学报, 2011, 27(12): 348–352.
HE Wenchao, SHI Xueyi, DENG Qingyun, et al. Comparison of grain yield prediction methods in land use planning[J]. Transactions of the CSAE, 2011, 27(12): 348–352. (in Chinese)
- [15] 马晓敏, 王新. 基于遗传算法的 BP 神经网络改进[J]. 云南大学学报(自然科学版), 2013, 35(增刊2): 34–38.
MA Xiaomin, WANG Xin. An improved BP neural network algorithm based on genetic algorithm[J]. Journal of Yunnan University(Natural Sciences Edition), 2013, 35(Supp. 2): 34–38. (in Chinese)
- [16] 李海青, 赵又群, 闫茜, 等. 基于遗传神经网络的车轮匹配整车逆动力学[J]. 华中科技大学学报(自然科学版), 2019, 47(5): 27–32.
LI Haiqing, ZHAO Youqun, YAN Xi, et al. Inverse dynamics of vehicle matching with mechanical elastic wheels based on genetic neural network[J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2019, 47(5): 27–32. (in Chinese)
- [17] 陈英义, 程倩倩, 成艳君, 等. 基于 GA-BP 神经网络的池塘养殖水温短期预测系统[J]. 农业机械学报, 2017, 48(8): 172–178.
CHEN Yingyi, CHENG Qianqian, CHENG Yanjun, et al. Short-term prediction system of water temperature in pond aquaculture based on GA-BP neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(8): 172–178. (in Chinese)
- [18] 张漫, 李婷, 季宇寒, 等. 基于 BP 神经网络算法的温室番茄 CO₂ 增施策略优化[J]. 农业机械学报, 2015, 46(8): 239–245.
ZHANG Man, LI Ting, JI Yuhan, et al. Optimization of CO₂ enrichment strategy based on BPNN for tomato plants in greenhouse[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(8): 239–245. (in Chinese)
- [19] 刘雪, 李亚妹, 刘娇, 等. 基于 BP 神经网络的鲜鸡蛋货架期预测模型[J]. 农业机械学报, 2015, 46(10): 328–334.
LIU Xue, LI Yamei, LIU Jiao, et al. BP neural network based prediction model for fresh egg's shelf life[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(10): 328–334. (in Chinese)
- [20] 郁磊, 史峰, 王辉, 等. MATLAB 智能算法 30 个案例分析[M]. 北京: 北京航空航天大学出版社, 2011.