

基于BP神经网络的黄山市茶叶产量预测研究

钱锦 王佳佳* 李佳振 黄可云 李娇娇 陈尧庭

黄山学院数学与统计学院, 安徽省黄山市, 245021

摘要: 茶叶产量的准确预测对于农业生产规划与管理具有重要意义。本研究基于2003—2023年的黄山市茶叶产量数据, 系统分析了影响茶叶产量的主要因子, 包括茶园采摘面积、年平均气温、年平均降水量及累计日照时数等。以此为输入特征, 构建了BP神经网络预测模型, 用于量化茶叶产量变化规律。试验结果表明, 模型预测值与实际产量相关系数超过0.7, 整体平均误差百分比仅为0.65%, 显示出较高的预测精度与稳定性。该模型能够有效反映茶叶产量的年度波动趋势, 为黄山市茶叶种植规划、生产管理和政策制定提供了科学依据。研究验证了神经网络方法在非线性和农业产量预测中的实用性与可靠性。

关键词: 茶叶产量; BP神经网络; 预测模型; 气象因子; 非线性建模

1. 引言

茶产业作为中国传统优势产业, 在促进区域经济发展、推动乡村振兴和传承茶文化等方面发挥着重要作用。黄山市地处皖南山区, 山高谷深, 生态环境优越, 气候条件独特, 是我国重要的产茶区之一。黄山毛峰、太平猴魁、祁门红茶等享誉中外的名茶, 奠定了黄山市在全国茶叶产业中的重要地位。近年来, 黄山市坚持把茶产业作为富民增收的支柱产业和生态产业, 大力推动茶园扩建、产业结构优化和品牌建设。据统计, 黄山市茶产业在复杂的国内外经济形势下依然保持较快增长, 2023年, 黄山市茶产业持续保持稳健发展态势, 茶园面积稳定在80万亩左右, 茶叶总产量约4.7万吨, 较2020年进一步提升。同时, 茶产业综合产值达到260亿元, 产业规模与质量同步提升。与此同时, 茶叶进出口额增长超过60%, 进一步凸显了黄山茶产业在国内国际“双循环”格局下的市场活力。

在国家层面, 茶产业的发展被赋予了更高的战略价值。自“十四五”规划实施以来, 中央高度重视“三农”工作, 强调推进脱贫攻坚与乡村振兴的有效衔接。习近平总书记在多次考察中指出, 要统筹推进茶文化、茶产业和茶科技, 将茶产业作为乡村振兴的重要支柱。相关政策文件提出要提高产业发展补助资金在乡村振兴中的比重, 为包括黄山市在内的茶叶主产区提供了政策支持与发展机遇。茶产业不仅承担着农民增收和农村绿色发展的重要任务, 还在“一带一路”倡议推动下不断拓展国际市场。目前黄山市茶叶出口市场主要集中在摩洛哥、阿尔及利亚、乌兹别克斯坦等“一带一路”沿线国家, 数亿人口规模为茶叶消费提供了广阔空间, 也为区域经济高质量发展注入了新动力。

然而, 茶叶产量受自然环境和措施的多重影响, 具有显著的时空差异性。茶树属于喜温、喜湿、耐阴的多年生木本植物, 其生长发育与气温、降水、光照等气象条件密切相关。适宜的温度和充足的水分有助于茶树旺盛生长, 而过高或过低的气温、极端干旱或洪涝灾害则会导致减产甚至绝收。日照时数对茶树光合作用、养分积累和产量品质同样具有重要影响, 但过强的光照又可能造成叶片灼伤, 降低茶叶品质。因此, 在茶叶主产区, 气象条件波动及灾害性天气的频发, 常常给茶叶生产带来不确定性和挑战。近年来, 黄山市部分地区出现洪水、干旱等极端天气事件, 对茶叶产量造成了较大影响。与此同时, 黄山茶产业还面临夏秋茶资源利用不足、单位面积产量偏低、深加

工水平有限等问题, 这些制约因素进一步凸显了加强茶叶产量预测与调控研究的重要性。

目前, 国内外对茶叶产量预测的研究相对有限。已有的研究多采用模型构建的方法, 利用历史产量数据和气象要素数据来提高预测精度。在国内, 研究者们广泛尝试了基于神经网络的预测方法。例如, 有学者提出自适应增强BP神经网络模型, 有效降低了产量预测误差^[1]; 也有学者利用统计年鉴数据, 通过Matlab建立BP神经网络, 实现了对未来若干年的茶叶产量预测^[2]; 此外, 梯度提升决策树等集成学习方法也被引入茶叶产量预测, 结果显示其决定系数可达0.9, 预测精度较高^[3]。国外研究同样关注气象、土壤、植被指数等因子对茶叶产量的影响。如有学者将AquaCrop模拟模型与机器学习结合, 对巴基斯坦茶园产量进行预测, 平均绝对误差仅为0.123^[4]; 也有研究通过提取NDVI指数表征茶树生长状况, 并结合温度、降水等因子建立模型, 获得了较好的拟合效果(R^2 达0.79)。此外, 集成优化算法的混合模型在孟加拉国茶叶产量预测中显著提升了精度, 最

小相对误差降至11%^[5]。这些研究成果为本研究提供了参考, 也显示出将机器学习方法与多源数据结合进行茶叶产量预测的可行性和价值。

在此背景下, 本研究以黄山市作为研究区域, 基于2003—2023年的茶叶产量及地面气象观测数据, 重点选取茶园采摘面积、年平均气温、年平均降水量和累计日照时数等关键影响因子, 构建基于误差反向传播算法的多层BP神经网络预测模型。通过对模型进行训练与验证, 探讨其在黄山市茶叶产量预测中的适用性和精度表现。本研究的创新点在于结合长期尺度的茶叶产量数据与多维气象因子, 利用神经网络模型对黄山市及区县层面的茶叶产量进行模拟与预测。研究结果不仅为区域茶叶生产的宏观调控和科学管理提供决策支持, 也对推动农业智能化和气候适应性研究具有一定的示范意义。

2 数据与研究方法

2.1 数据来源与基本情况

本研究选取了黄山市辖区内屯溪区、黄山区、徽州、歙县、休宁县、黟县和祁门县七个县区2003—2023年的茶叶产量及采摘面积数据, 数据主要来自《黄山市统计年鉴》。产量数据以吨为单位, 采摘面积以公顷为单位。统计结果显示, 近二十年来黄山市茶叶产业保持稳定发展, 全市年均茶叶产量为25411.19吨,

年均采摘面积为46048.38公顷。不同县区间差异明显，歙县因地形适宜、茶园规模较大，产量和面积均位居全市首位，而屯溪区因城区面积限制，产量和面积相对较小。总体而言，山区县茶叶生产规模更大，贡献更为突出。数据在整理过程中对缺失值进行了补充，利用线性插值和空间邻近站点数据进行校正，确保时序的完整性与一致性。同时，为避免极端年份异常值对分析结果的干扰，采用 3σ 原则剔除了个别异常数据点，从而为后续模型构建提供可靠的数据基础。

2.2 气象数据处理与指标选取

在气象数据方面，本研究收集了七个县区2003—2023年的年均气温、年均日降水量和年累计日照时数。主要数据来源为黄山市统计年鉴，部分缺失区县数据（如徽州区）通过ArcGIS空间插值方法获取。为保证各气象指标在模型中的可比性，研究中对所有气象数据进行了统一量纲转换与标准化处理，消除了数值区间差异带来的影响。指标选取上，结合茶树的生态学特性与前人研究成果，最终确定了四类因子：①茶园采摘面积，直接反映生产规模和茶园管理水平；②年平均气温，衡量热量条件对茶树萌芽和生长的影响；③年平均日降水量，代表水分供给对茶树生长发育的重要性；④年累计日照时数，体现光照资源对光合作用与养分积累的支撑作用。这些因子的共同作用决定了茶叶的产量与品质。从生理学角度来看，茶树喜温喜湿而又依赖充足光照，因此产量对降水与日照条件的响应更为敏感，而年均气温则在一定范围内对产量影响有限，这一点也将在后续的相关性分析结果中得到验证。

2.3 研究方法与分析框架

在研究方法上，首先利用SPSS软件对茶叶产量及采摘面积与气象因子之间的相关性进行了统计检验，分析其显著性和作用方向，为确定关键驱动因子提供依据。相关性分析表明，茶叶产量及采摘面积与年均降水量和累计日照时数存在较强相关性，而与年均气温的相关性不显著。基于此结果，本研究构建了以采摘面积、年均气温、年均日降水量和年累计日照时数为输入因子的多层BP神经网络预测模型，茶叶产量作为输出变量。模型采用误差反向传播算法进行训练，设置1—2个隐含层，激活函数选用Sigmoid型函数，并通过交叉验证确定学习率和迭代次数，以提升模型的泛化能力。通过模型训练与测试，结果显示预测产量与实际观测值之间的相关系数达到0.7以上，平均误差仅为0.65%，表明该模型在区域尺度的茶叶产量预测中具有较高精度和实用性。该研究框架不仅揭示了气象因子对茶叶产量的影响机理，也为茶产业宏观调控、市场供需预测及应对气候变化风险提供了科学支撑。

3 模型构建与实证分析

3.1 BP神经网络预测模型构建

BP神经网络是一种典型的多层前馈神经网络，由输入层、隐藏层和输出层组成，其学习机制主要包括信号的正向传播与误差的反向传播两个过程。在本研究中，以黄山市2003—2023年茶叶产量及相关气象因子数据为基础，构建了基于误差反向传播算法的多层BP神经网络模型。在模型运行过程中，输入层接收采摘面积、年均气温、年均降水量和累计日照时数等关键因子，并通过隐藏层逐步提取非线性特征；输

出层则输出预测的茶叶产量值。通过误差反向传播机制，模型不断修正连接权值，使预测结果逐渐逼近真实值，直至误差达到设定阈值或完成最大迭代次数。

在数据集划分方面，本研究随机抽取71.5%的样本作为训练集，剩余28.5%的样本作为测试集。为消除不同指标量纲差异的影响，对输入特征数据进行了标准化处理。

3.2 预测结果与精度分析

模型预测结果如图1所示，整体上预测值与实际值高度吻合，大多数样本点分布在1:1线附近。整体相关系数超过0.7，表明该神经网络模型具有良好的拟合效果。进一步分析显示，休宁县和黟县的预测效果最佳，相关系数达到0.8以上；黄山区和歙县也表现出较高的预测精度。总体而言，该模型能够有效反映各区县茶叶产量的时空差异。

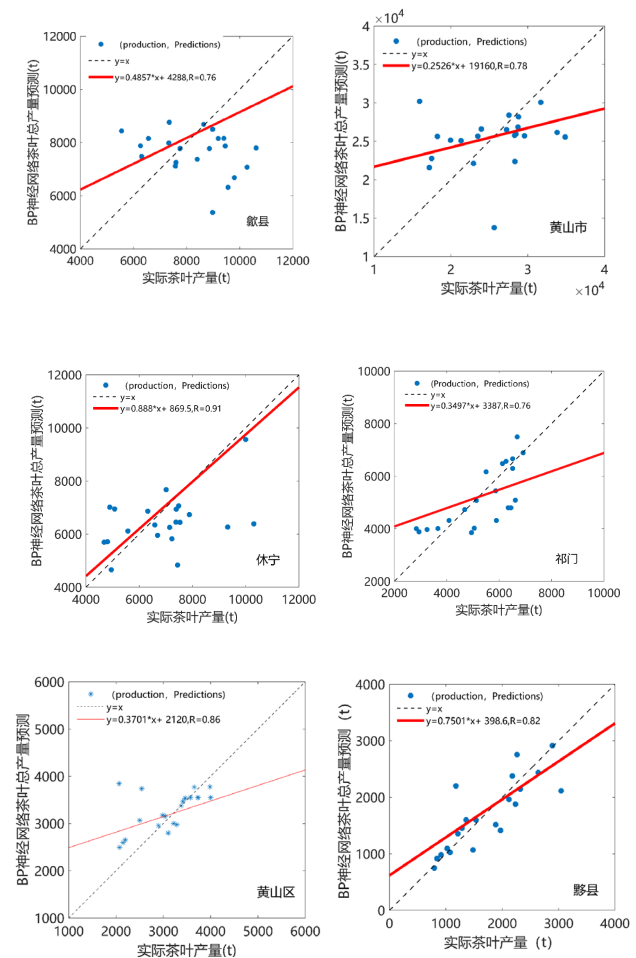


图1 茶叶产量预测结果图

在误差分析方面，如表1所示，黟县的平均误差为1.90%，显示预测值与实际值高度一致；休宁县和黄山区的平均误差分别为6.85%和5.52%，处于可接受范围；歙县误差为7.95%，说明仍有进一步优化空间。而祁门县的误差较大，达到25.29%，提示该地区茶叶产量可能受独特的地理环境、管理方式或气象波动影响，从而削弱了模型的适用性。整体来看，黄山市总体平均误差仅为0.65%，验证了该模型在区域尺度上

的可靠性与适用性。

表1 黄山市各区县茶叶产量预测结果及误差分析

地区	平均误差百分比	实际产量(吨)	预测产量(吨)
黟县	1.90%	36254	35563.96
休宁	6.85%	145695	35713.7
黄山区	5.52%	64976	68563.4
祁门	25.29%	145695	108848.3
歙县	7.95%	174446	160580.1
黄山市	0.65%	533635	530177.4

3.3 结果讨论

研究表明,基于BP神经网络的多因子产量预测模型能够较为准确地揭示黄山市茶叶产量与气象因子之间的关系。特别是在产量较高的休宁县和产量相对较低的黟县,预测效果均表现优异,进一步证明了气象条件对茶叶生长的显著影响。然而,对于祁门县等误差偏大的地区,还需结合更精细化的气象数据与茶园管理因素,以提升模型的适应性与预测精度。整体而言,本研究所提出的模型不仅在统计层面上实现了对茶叶产量的较高精度预测,也为区域茶叶生产的宏观管理和政策制定提供了量化参考。

4. 结论与建议

茶产业是黄山市农业经济的重要组成部分,对区域经济增长、农村经济发展以及农民增收具有重要推动作用。因此,科学预测茶叶产量对于制定产业发展规划、优化资源配置和提升市场调控能力具有重要参考价值。

本研究基于2003—2023年黄山市七个区县的茶叶产量数据及地面气象要素数据,选取了茶园采摘面积、年平均气温、年平均降水量和累计日照时数等关键因子,构建了黄山市茶叶产量预测模型。实验结果显示,大部分样本紧密分布在1:1线附近,相关系数均达到0.7以上,表明模型在区域茶叶产量预测中具有较高的可靠性。进一步分析发现,产量较高的休宁县以及产量相对较低的黄山区和黟县预测效果更为显著,相关系数超过0.8,这表明气象因子与茶叶产量之间存在显著相关性。整体来看,黄山市茶叶产量预测的平均误差仅为0.65%,显示出模型在全市范围内具有良好的适用性。

尽管整体预测精度较高,但部分地区仍存在较大误差,如祁门县的预测误差较高,这可能与该地区的特殊地理环境、管理方式差异、社会经济因素以及数据采集精度等因素有关。此外,黄山市各区县茶叶产量数据存在明显差异,导致训练数据不均衡,从而影响了模型的稳定性和泛化能力。在此基础上,建议在未来的茶叶生产与管理中,应进一步加强各区县长期产量及气象数据的积累和更新,以提升模型训练的稳定性与预测精度;同时,可结合茶园实际管理情况合理调整种植布局、采摘面积和管理措施,以提高高产区的产量潜力并改善低产区的生产条件;此外,应综

合考虑当地政策扶持、劳动力流动、农药及肥料使用等社会经济因素,以增强预测结果的实用性和科学性;进一步引入土壤质量、病虫害监测和遥感数据,构建更全面的指标体系,将有助于优化预测模型。通过持续改进和完善预测方法,黄山市茶叶产量预测模型不仅可以为全市及各区县提供可靠的产量预测,也能够为茶产业的科学规划、政策制定和管理优化提供重要参考。

参考文献

- [1] 陈冬梅,韩文炎,周贤锋,等.基于自适应增强的BP模型的浙江省茶叶产量预测[J].茶叶科学,2021,41(4):13.
 - [2] 王郑霞,肖来胜.基于神经网络的茶叶产量建模与预测方法研究[J].电脑知识与技术,2011,7(22):5427-5429.
 - [3] 王美洁,姜同强.基于BP神经网络的茶叶产量预测研究[J].福建茶叶,2019,41(2):27.
 - [4] BATOOL D, SHAHBAZ M, SHAHZAD ASIF H, et al. A Hybrid Approach to Tea Crop Yield Prediction Using Simulation Models and Machine Learning [J]. Plants, 2022, 11(15): 1925.
 - [5] Phan P, Chen N, Xu L, et al. Using multi-temporal MODIS NDVI data to monitor tea status and forecast yield: a case study at Tanuyen, Laichau, Vietnam[J]. Remote Sensing, 2020, 12(11): 1814.
 - [6] 胡克满,胡海燕.基于灰色神经网络的茶叶产量预测算法[J].浙江农业科学,2019,60(04):577-579+583.
 - [7] 袁泉,胡玉才,孙永厚.基于BP神经网络的预测方法应用研究[J].农业系统科学与综合研究,1998,(04):261-263.
 - [8] 丁鹏,徐爱俊,周素茵.基于梯度提升决策树多特征结合的茶叶产量预测[J].西南农业学报,2021,(007):034.
 - [9] 徐涛,闫彪,何曹尹宣,等.基于图像处理和BP神经网络对武夷岩茶杂质识别方法[J].武夷学院学报,2025,44(06):41-47.
 - [10] 李聪,郭军,杨鹏至,等.基于BP神经网络的Se元素生物有效性预测研究——以黄沙街镇茶叶Se为例[J].四川地质学报,2022,42(04):673-676.
- 基金项目:安徽省教育厅重点项目《环境规制对安徽省经济高质量发展的影响研究》(2023AH051355),省级大学生创新创业计划训练项目《空间视角下环境规制的经济高质量发展效应研究——以安徽省为例》(S202410375094),国家级大学生创新创业计划训练项目《基于误差反向传播算法的多层BP神经网络模型的黄山市茶叶产量预测研究》(202410375047)。
- 课题项目:大学生科技创新项目《大数据驱动下的应用统计模型对碳源汇的精准反演》(XSKY202410);省级大学生创新创业计划训练项目《区域数字经济发展评价及组态视角下驱动路径研究》(s202510375134)