

陈嘉铎, 吴烨荣, 甘锦洪, 等. 改进的 BP 神经网络在水稻产量预测中的应用[J]. 智能计算机与应用, 2025, 15(11): 170–177.
DOI: 10.20169/j.issn.2095-2163.251128

改进的 BP 神经网络在水稻产量预测中的应用

陈嘉铎¹, 吴烨荣², 甘锦洪², 林桢菀², 金亚文², 李承璋²

(1 南宁师范大学 物理与电子学院, 广西 南宁 530299; 2 广州华商学院 人工智能学院, 广州 511300)

摘要: 针对传统 BP 神经网络在产量预测中存在的精度差、准确度低等问题, 提出了改进的 BP 神经网络水稻产量预测模型 (CGA-BP)。首先, 建立水稻产量数据、气象环境数据 (年降雨量、年日照时数、平均气温) 和田间情况数据 (农作物受灾面积、水稻播种面积); 其次, 在传统的遗传算法中对选择、交叉和变异这些算子的概率进行改进, 通过改进的遗传算法对 BP 神经网络模型进行优化; 最后, 应用 BP 神经网络模型、遗传算法优化的 BP 神经网络模型 (GA-BP) 以及改进的 BP 神经网络模型 (CGA-BP) 对湖南省水稻产量进行预测分析。实验结果表明: CGA-BP 神经网络模型的水稻产量预测精准度明显高于 GA-BP 神经网络模型和 BP 神经网络模型, MAE、MAPE 以及 RMSE 分别是 0.045 3, 0.017 3 和 0.052 6, 与 BP 预测模型相比分别提高了 5.99%, 2.29% 和 6.84%, 与 GA-BP 模型相比分别提高了 4.23%, 1.60% 和 4.46%。CGA-BP 神经网络模型下的平均相对误差仅为 2.27%, 较 BP 神经网络降低了 1.75%, 较 GA-BP 神经网络降低了 1.06%。CGA-BP 神经网络模型能够更加科学、合理地在水稻产量进行预测, 可行且有效。

关键词: BP 神经网络; 遗传算法; 水稻产量预测

中图分类号: S11+7

文献标志码: A

文章编号: 2095-2163(2025)11-0170-08

Application of improved BP neural network in rice yield prediction

CHEN Jiahua¹, WU Yerong², GAN Jinhong², LIN Chengchun², JIN Yawen², LI Chengzhang²

(1 School of Physics and Electronics, Nanning Normal University, Nanning 530299, China;

2 School of Artificial Intelligence, Guangzhou Huashang College, Guangzhou 511300, China)

Abstract: Aiming at the problems such as poor accuracy and low precision existing in the traditional BP neural network in yield prediction, an improved BP neural network rice yield prediction model (CGA-BP) is proposed. Firstly, establish data on rice yield, meteorological environment (annual rainfall, annual sunshine duration, average temperature), and field conditions (affected area of crops, sown area of rice); Then, in the traditional genetic algorithm, the probabilities of operators such as selection, crossover and mutation were improved, and the BP neural network model was optimized through the improved genetic algorithm. Finally, the BP neural network model, the BP neural network model optimized by genetic algorithm (GA-BP), and the improved BP neural network model (CGA-BP) were applied to conduct the prediction and analysis of rice yield in Hunan Province. The improved BP neural network model (CGA-BP) was used to predict and analyze the rice yield in Hunan Province. The experimental results show that: Under the CGA-BP neural network model, the prediction accuracy of rice yield was significantly higher than that of the GA-BP neural network model and the BP neural network model. The MAE, MAPE and RMSE were 0.045 3, 0.017 3 and 0.052 6 respectively, which increased by 5.99%, 2.29% and 6.84% respectively compared with the BP prediction model. Compared with the GA-BP model, they have increased by 4.23%, 1.60% and 4.46% respectively. The average relative error under the CGA-BP neural network model is only 2.27%, which is 1.75% lower than that of the BP neural network and 1.06% lower than that of the GA-BP neural network. It has been verified that the CGA-BP neural network model can predict the rice yield more scientifically and reasonably, indicating that it is feasible and effective to optimize the BP neural network using the improved genetic algorithm in the prediction of rice yield.

Key words: BP neural network; genetic algorithm; rice yield prediction

基金项目: 国家级大学生创新创业训练项目 (202312621004)。

作者简介: 陈嘉铎 (2002—), 男, 硕士研究生, 主要研究方向: 人工智能; 吴烨荣 (2002—), 男, 学士, 主要研究方向: 人工智能, 深度学习, 软件开发; 甘锦洪 (2002—), 男, 硕士研究生, 主要研究方向: 人工智能; 林桢菀 (2003—), 女, 学士, 主要研究方向: 人工智能; 李承璋 (2003—), 男, 学士, 主要研究方向: 人工智能, 计算机视觉, 自然语言处理。

通信作者: 金亚文 (1993—), 女, 学士, 主要研究方向: 人工智能。Email: 461783643@qq.com。

收稿日期: 2024-03-05

0 引言

中国是农业大国,水稻在中国粮食作物中占据着很重要的地位^[1]。近年来,随着大数据和人工智能技术的应用,水稻产量预测研究也取得了不少进展,国内外学者已经提出了多种水稻产量预测模型,如基于时间序列分析的统计模型、机器学习模型、基于统计模型和机器学习模型结合起来的混合模型。基于统计自回归积分滑动平均模型(ARIMA 模型)建立在时间序列数据的基础上,考虑时间的序列性、自回归和滑动平均等因素,得出未来一段时间内的产量预测结果。谢琦等^[2]以云南省红河州水稻产量为研究对象,运用 ARIMA 模型对该地水稻产量进行预测分析,结果表明 ARIMA 模型可以有效地预测水稻产量,其预测结果的平均误差率为 5.8% 左右。机器学习模型,如随机森林(RF)、支持向量机(SVM)、神经网络(NN)等,这些模型基于数据驱动,通过对大量历史的数据学习和对水稻产量变化规律及其特征关系的挖掘,预测未来产量;Sukartini N 等^[3]使用基于随机森林的方法来预测泰国的水稻产量,研究的数据包括历史天气数据、土地利用数据、社会经济数据等,显著提高水稻产量的预测精度,为农业生产提供了一种新的预测方法,有望帮助农民和政府更好地进行决策;Mo T 等^[4]使用支持向量机模型预测缅甸的水稻产量,取得较为准确的预测结果,为农业生产提供了一种新的预测方法,可用于优化农业生产;Zhang X 等^[5]采用基于神经网络的方法来预测中国的水稻产量,利用空间和时间上的多种数据来建立预测模型,为农业生产提供了一种新的预测方法,同时也为研究中国水稻生产的规律和变化趋势提供了新的数据支持。统计模型和机器学习模型结合的混合模型是将统计模型和机器学习模型结合起来,考虑使用历史数据的线性趋势和周期性以及生长期不同环境因素的影响,通过时间序列分析和概率模型来构建一种更可靠的预测模型。Zhang X 等^[6]提出了一种基于时间序列和人工神经网络(ANN)的混合模型,用于预测水稻产量,首先采用时间序列分析方法确定产量的趋势和周期性信息,其次将气象数据、土壤环境数据、农业管理数据等与产量进行关联分析,构建 ANN 模型,采用遗传算法和误差反向传播算法对模型参数进行优化,通过将时间序列和 ANN 模型的预测结果进行集成,提高了预测的准确性和稳定性;Rabbi M F 等^[7]提出了一种基于时间序列分析和模糊聚类递归

(Adaptive Resonance Theory, ART)算法的混合模型,用于孟加拉国水稻产量的预测,将历史数据中的水稻产量与环境因素等多种影响因素建立关联,利用时间序列分析得到产量的趋势、季节性变化和残差等信息,利用模糊聚类递归算法将这些影响因素分成几个聚类。每个聚类中的数据通过适当的权重被分配到神经网络中,用于对孟加拉国水稻产量进行预测。本文提出了一种改进的 BP 神经网络水稻产量预测模型,对遗传算法中的选择、交叉和变异算子进行了改进,利用改进的遗传算法对 BP 神经网络模型进行优化。

1 基础理论

1.1 BP 神经网络

BP(Back Propagation)神经网络是一种典型的多层前馈神经网络,通过误差逆向传播算法进行学习训练^[8]。BP 神经网络是目前使用最为广泛、结构最直观、工作原理最容易理解的神经网络,其结构简单、可塑性强且拥有较强的数据拟合能力^[9]。BP 神经网络主要包含输入层、隐含层和输出层,在训练过程中,神经网络不断调整输入层与隐含层以及隐含层和输出层之间的权值和阈值,当神经网络输出值与目标值一致或者达到迭代次数时训练停止,这种神经网络拥有较强的泛化能力^[10]。BP 神经网络神经元的激活函数使用 Sigmoid 函数,表达式如下:

$$f_s(u) = \frac{1}{1 + e^{-u}} \quad (1)$$

神经元的输入数据经过加权求和后减去阈值,表达式如下:

$$u_j = \sum_{i=1}^n w_{ij}x_i - o_j \quad (2)$$

其中, x_i 表示输入数据; w_{ij} 表示权值; o_j 表示阈值。

局部输出用 p 表示,联立(1)式和(2)式,可以得到神经元的局部输出值,表达式如下:

$$p = f_s\left(\sum_{j=1}^n w_j o_j + \theta\right) \quad (3)$$

传统的 BP 神经网络通过正向传播将数据输入到网络中,根据输出与期望输出之间的误差,运用反向传播算法对网络权值和阈值进行优化调整,这一过程通过梯度下降方法实现,不断减小误差直至达到预期值。权值与阈值的更新过程如下式:

$$w(k+1) = w(k) - \gamma \frac{\partial F}{\partial w} \quad (4)$$

$$\theta(k+1) = \theta(k) - \gamma \frac{\partial F}{\partial \theta} \quad (5)$$

其中, F 表示神经网络输出值与目标值之间的误差; w 表示更新后的权值; γ 表示学习率; θ 表示更新后的阈值。

对于处理数据, BP 神经网络虽具有较强的自适应、泛化以及非线性映射能力, 但其在网络训练中容易陷入局部极小值、误差梯度变化非常小、收敛速度特别慢^[11], 而遗传算法具有较强的全局搜索最优值特性^[11-12]。因此, 本文考虑利用遗传算法来优化 BP 神经网络。

1.2 遗传算法

1.2.1 遗传算法原理

Holland^[13] 于 20 世纪 70 年代提出遗传算法 (Genetic Algorithm, GA)。遗传算法是一种模拟自然选择和遗传机制的优化算法, 用于解决搜索和优化问题, 该算法是根据大自然中生物体进化规律而提出的^[14]。通过模拟生物进化过程, 逐步改进生成的个体以获得更好的解决方案。开始时, 随机生成一组个体, 称为种群。每个个体都代表问题的一个潜在解决方案, 接着对种群中的每个个体, 通过一个适应度函数来评估其解决问题的能力, 适应度函数会为每个个体分配一个适应度分数, 用于衡量其在解决问题上的优劣, 接着根据适应度分数, 选择一些个体作为“父代”, 通常适应度高的个体被选中的概率更高, 即模拟了生物进化中的“适者生存”原则。选中的父代个体按照某种规则进行交叉操作, 生成新的个体, 称为“子代”, 交叉操作模拟了生物进化中的基因重组; 紧接着对一些子代进行变异操作, 以引入新的遗传信息, 变异操作是随机的, 模拟了生物进化中的基因突变; 最后, 将新生成的子代个体与旧的个体种群合并, 并根据一定的策略选择哪些个体保留, 以更新种群, 通常选择适应度高的个体保留下来, 而适应度低的个体被替换掉。算法会不断重复上述步骤, 直到满足某个终止条件, 如达到最大迭代次数、找到满足特定条件的解等。一旦算法终止, 最终的种群中的个体就代表了问题的解决方案, 通常从种群中选择适应度最高的个体作为最终结果。通过重复上述步骤, 遗传算法不断演进, 逐步优化种群中的个体, 从而在搜索空间中寻找问题的最优解或接近最优解。

1.2.2 适应度函数的选择

适应度函数用于评估每个个体 (染色体) 解决

特定问题的能力或优劣程度。适应度函数将问题的解转化为一个数值, 该数值越大表示个体越适应问题, 即越接近问题的最优解。本文选择误差平方的倒数作为适应度函数, 具体表示如下式:

$$F = \frac{1}{2} \sum_{j=1}^n (P_j^{k-1}(\mathbf{w}, \mathbf{x}) - y_j)^2 \quad (6)$$

其中, F 是误差函数; P 表示整体输出; \mathbf{w} 是权矢量; 而 \mathbf{x} 则表示输入矢量。

$$E(j) = \frac{1}{F(j)} \quad (7)$$

其中, E 代表适应度; 而 j 表示迭代次数。

1.3 遗传算法的改进

在遗传算法中, 如果选择、交叉和变异的概率被设置为固定值, 可能导致算法在某些情况下收敛得很快, 而在其他情况下可能会很慢, 即算法对于不同类型的问题表现不一致, 难以在不同问题上取得良好的性能。传统遗传算法的选择概率、交叉概率和变异概率使用的是固定的常数, 这样在训练的前期和后期容易产生局部最优解, 使得算法缺乏对问题动态性的适应能力^[15]。如果交叉和变异的概率太低, 算法可能难以引入足够的变化; 如果概率太高, 可能导致搜索空间中的多样性不足。为使算法能够更好地适应问题的不同阶段和特性, 避免算法过早收敛到局部最优解, 提高搜索效率, 本文对选择、交叉和变异这些算子的概率进行了优化; 引入了自适应机制, 使得这些概率值能够根据种群的动态变化进行调整, 这样的自适应机制旨在提高算法的性能和鲁棒性, 使其更适应不同类型的问题。适应度值的不断调整, 预测模型的选择、交叉和变异概率也得到了相应调整, 确保模型能够有效地调整各个参数, 以达到最佳状态。

1.3.1 对选择算子的改进

选择算子用于从当前种群中选择个体, 以便进行后续的交叉和变异操作。选择算子的目的是根据每个个体的适应度值, 将更适应解的个体更有可能地被选中, 从而促使优秀的基因传递给下一代。传统的遗传算法通常采用“竞标赛选择”, 效果受到竞标赛选择规模的影响, 过小的规模可能导致选择的不稳定性, 过大的规模可能减缓算法的收敛速度, 并且弱个体在竞标赛中被选中的概率较低, 可能导致了一部分潜在优秀的基因丧失。本文首先采用排序算法对种群个体进行排序, 然后根据其排名分配选择概率。适应度较高的个体获得较高的选择概率, 不同于轮盘赌选择, 基于排名而不是适应度值来分

配概率,这样可以减少适应度值差异较大的个体之间的选择概率差距。使用这种方法对选择算子进行改进,个体被选中的概率如下式:

$$p = \frac{p_0}{1 - (1 - p_0)^\alpha} (1 - p_0)^{\alpha-1} \quad (8)$$

其中, α 表示种群的数量; p_0 表示最优个体被选到的概率; b 表示通过排序算法后得到的第 n 个个体在整个排序序列的位置; p 为个体被选中的概率。

1.3.2 对交叉算子的改进

交叉操作模拟了生物学中的基因重组过程,通过将两个或多个个体的基因片段进行组合,生成新的个体,从而促使优秀的特征传递给下一代。传统的遗传算法的工作中,通常会将交叉概率设定为一个合适的常数值,这个值通常在 0.4~0.9。交叉概率过高时,个体之间的遗传信息传递过于频繁,容易导致种群过早收敛到局部最优解,陷入早熟状态,还可能导致种群中的个体趋向于相似,减少了种群的多样性,算法的全局搜索能力降低,还会导致子代个体的遗传信息稀释,使得优秀特征难以传递给后代。交叉概率过低时,种群中的优秀特征传递速度较慢,可能导致算法的收敛速度变慢,还可能导致算法在搜索过程中过于保守,难以跳出局部最优解,还会导致种群中的个体变异和交叉的可能性减小,使得适应度值的提升变得困难。因此,本文将动态地调整交叉概率,使其根据种群中的适应度变化而做出相对应的改变,由此改进交叉算子,改进的交叉概率如下:

$$P_j = \begin{cases} P_{jmax}, & E < E_{mean} \\ P_{jmax} - \frac{P_{jmax} - P_{jmin}}{n_{max}} \times n, & E \geq E_{mean} \end{cases} \quad (9)$$

其中, E 表示交叉操作的两个个体在种群中的适应度最大值; E_{mean} 表示平均适应度; n 表示迭代次数; n_{max} 表示最大迭代次数。

在交叉概率 P 的设置过程中,将最小交叉概率 P_{jmin} 的值设置为 0.4,将最大交叉概率 P_{jmax} 的值设置为 0.9。

1.3.3 对变异算子的改进

变异算子用于在个体的染色体中引入一些随机性,以增加种群的多样性,避免陷入局部最优解。变异操作模拟了生物学中的基因突变过程,通过对个体的某些基因进行随机改变,引入新的基因组合,产生新的个体。传统的遗传算法一般会将变异概率设定为一个较小的常数值,通常在 0.001~0.01。变异

概率过高会导致个体的基因频繁发生变化,使得种群失去了稳定性,降低了种群的多样性,还会导致优秀特征被破坏,使算法的收敛速度变慢,还会导致算法陷入随机搜索的状态,难以有效地收敛到优秀解等问题。变异概率过低时,种群的变异可能会不足以跳出局部最优解,使得算法陷入局部最优,导致优秀特征的传递速度变慢,种群中优秀的基因难以快速传递给后代。在遗传算法中种群存在两种适应度,即个体的适应度和相对平均适应度。初始阶段,种群中的个体适应度会低于相对平均适应度,因此,为了保留那些具有优秀基因的个体,会将变异概率设定为一个较小的数值。在中期,由于个体的适应度不断提高,后期种群中个体的适应度将会高于相对平均适应度。此时为了提升算法的局部搜索能力,又会将变异概率设定为一个较大的数值,这种方式需要进行多次的人工调整,而人工调整往往受到个体经验和主观判断的影响,可能会导致对参数的理解和调整存在偏差,还要耗费大量的时间和精力,特别是在参数空间较大或复杂的情况下,可能需要进行大量的试验和评估。为了克服人工调整的缺点,本文通过在遗传算法的运算过程中,自动修改变异概率数值,来对变异算子进行改进,改进后的变异概率如下式:

$$P_b = \begin{cases} P_{bmin}, & E < E_{mean} \\ P_{bmax} - \frac{P_{bmax} - P_{bmin}}{n_{max}} \times n, & E \geq E_{mean} \end{cases} \quad (10)$$

其中, E 表示变异操作的两个个体在种群中的适应度最大值; E_{mean} 表示平均适应度; n 是迭代次数; n_{max} 表示迭代次数的上限。

在初始化遗传算法时,变异概率的最小值 P_{bmin} 设置为 0.001,变异概率的最大值 P_{bmax} 设置为 0.01。

2 改进遗传算法优化 BP 神经网络预测模型

2.1 预测模型的框架设计

BP 网络对初始权重的选择很敏感,不同的初始权重可能导致不同的训练结果,选择合适的网络结构包括隐藏层的神经元数量,对于 BP 网络的性能影响很大,但很难事先确定最佳结构,并且在 BP 神经网络优化过程中容易陷入局部最优解,难以找到全局最优解。因此,本文引入遗传算法,更有效避免陷入局部最优解的特性,提高神经网络的性能,优化 BP 神经网络;帮助减少适应度值差异较大的个体之间的选择概率差距,对遗传算法中的选择算子进行

改进;动态地调整交叉概率,根据种群中的适应度变化而做出相对应的改变,改进交叉算子;最后,根据适应度会在遗传算法的运算过程中不断变化的特点,通过变异概率数值的自动修改,对变异算子进行改进,从而获得改进的遗传算法,再将这种改进的遗传算法应用到 BP 神经网络预测模型中,得到 CGA-BP 神经网络模型。

改进遗传算法优化 BP 神经网络预测模型的具体操作流程如下:首先,在国家统计局中查找到 1949-2021 年的湖南省水稻产量和 5 个影响水稻产量的数据共 73 组;其次,对所获取的数据进行预处理,采集到的数据可能含有异常值,通过对数据进行清洗,得到正确的湖南省水稻产量数据;将数据按 8 : 2 划分为训练集数据 58 组,测试集数据 15 组;最终,将测试集数据导入训练好的模型中进行预测和评估。改进遗传算法优化 BP 神经网络预测模型的设计流程图如图 1 所示。

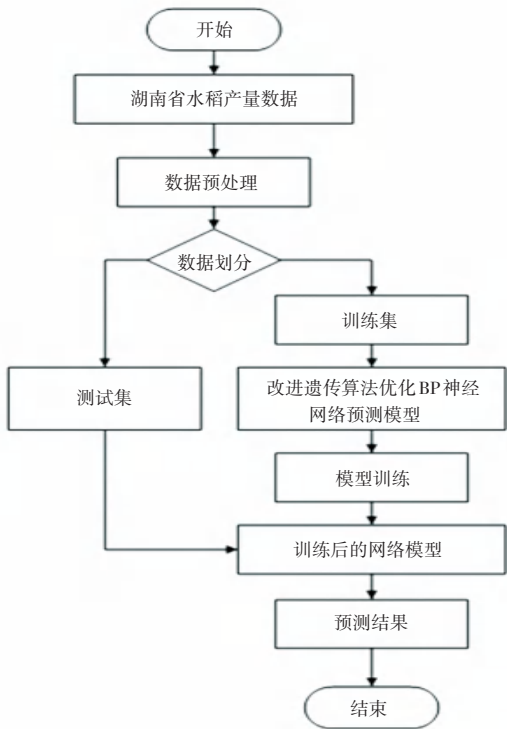


图 1 改进遗传算法优化 BP 神经网络预测模型的设计流程图
Fig. 1 Design process of optimized BP neural network prediction model with improved genetic algorithm

2.2 基于改进遗传算法优化 BP 神经网络预测模型设计

获得输入数据后,对数据进行预处理,去除脏数据,确定网络拓扑结构,对 BP 神经网络进行初始化操作,获得初始权值和阈值后,对初始值进行编码,将模型输出误差平方根的倒数作为适应度,进行选

择、交叉、变异操作,通过选择、交叉、变异操作计算适应度的值,并且对适应度的值进行判断,确定是否满足终止条件,若不满足则重复以上操作,若满足则可得到最优的权值和阈值,计算误差,并对权值和阈值进行更新,当满足终止条件时,得到结果,若不满足则重复以上操作。

将改进遗传算法与 BP 神经网络结合,构建出了基于改进遗传算法优化 BP 神经网络的预测模型,提升了预测模型的准确性。基于改进遗传算法优化 BP 神经网络预测模型设计流程图如图 2 所示,神经网络的结构如图 3 所示。

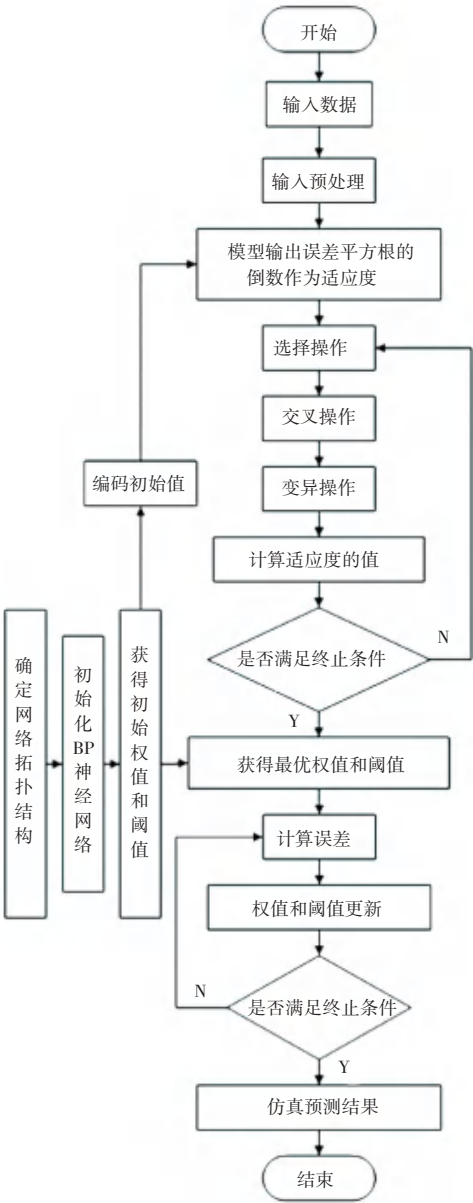


图 2 基于改进遗传算法优化 BP 神经网络预测模型设计流程图
Fig. 2 Design flowchart of optimized BP neural network prediction model based on improved genetic algorithm

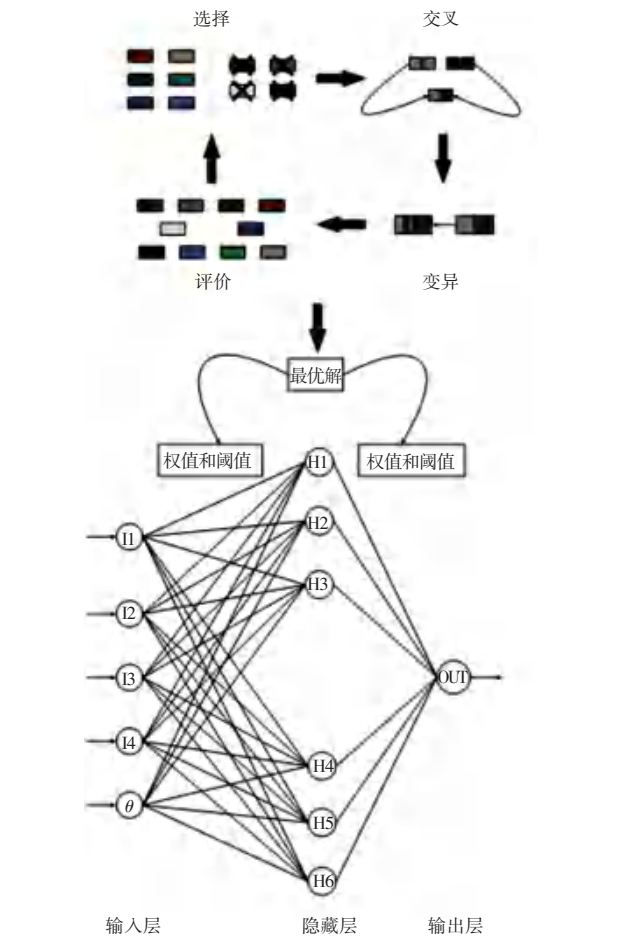


图 3 基于改进遗传算法优化 BP 神经网络预测模型结构图

Fig. 3 Structure schematic diagram of optimized BP neural network prediction model based on improved genetic algorithm

3 实验仿真分析

3.1 设置模型的参数

本文使用的特征变量分别为年降雨量、年日照时数、平均气温、农作物受灾面积、水稻播种面积,输入变量有 5 个,因此输入层的节点数设置为 5,隐含层的节点数设置为 9 个。为了实现预测功能,只需要输出一个实数作为预测结果,因此将输出层设置为 1。选择 Sigmoid 函数为激活函数,迭代次数设定为 1 000 次。

3.2 模型评价指标

本文采用平均绝对误差 (MAE)、平均绝对百分比误差 (MAPE) 和均方根误差 (RMSE) 来评价模型的预测准确性,公式如下:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - m_i| \quad (11)$$

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - m_i|}{y_i} \quad (12)$$

$$\text{RMSE} = \sqrt{\sum_{i=1}^N (y_i - m_i)^2} \quad (13)$$

其中, y_i 表示第 i 个样本的真实值; m_i 表示第 i 个样本的预测值; N 表示样本数量。

3.3 数据来源

本文的数据来源于国家统计局,共有 73 个样本,每个样本包含年降雨量、年日照时数、平均气温、农作物受灾面积、水稻播种面积 5 个特征值和水稻产量 1 个目标值,前 58 个样本作为训练集数据,剩余 15 个样本作为测试集数据。

3.4 仿真实验结果与分析

将经过处理的数据放入训练集,对模型进行训练,使用测试集对模型的性能进行评估。最后,将 3 个模型的预测值与真实值进行对比分析。

由 BP 神经网络所得到的产量预测结果与真实产量值的对比图以及产量误差分析图如图 4 和图 5 所示,输出结果与真实值吻合程度较差,预测误差波动变化剧烈,平均相对误差达到了 4.02%。

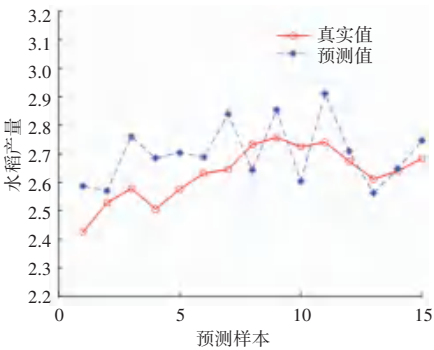


图 4 BP 神经网络产量预测对比图

Fig. 4 Comparison of BP neural network yield prediction

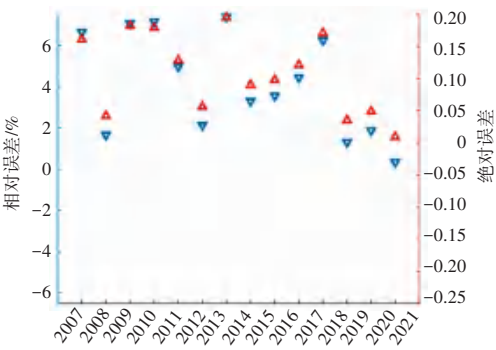


图 5 BP 神经网络产量误差分析图

Fig. 5 BP neural network yield error analysis diagram

经过遗传算法优化的 BP 神经网络所得到的产量预测结果与真实产量值的对比图以及产量误差分析图如图 6 和图 7 所示,输出结果与真实值吻合程度较单一 BP 神经网络略有提升,预测误差波动变化较小,平均相对误差为 3.33%,较单一 BP 神经网络

络降低了 0.69%。

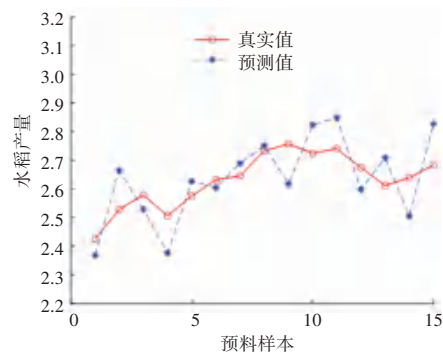


图 6 遗传算法优化的 BP 神经网络产量预测对比图

Fig. 6 Comparison of yield prediction of BP neural network optimized by genetic algorithm

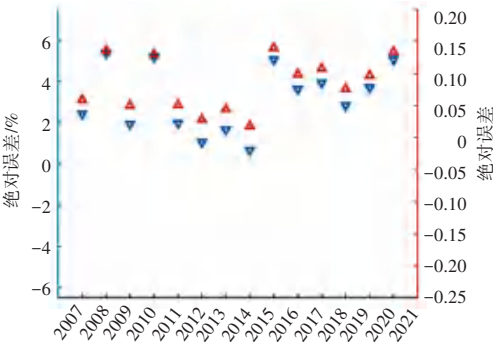


图 7 遗传算法优化的 BP 神经网络产量误差分析图

Fig. 7 Yield error analysis diagram of BP neural network optimized by genetic algorithm

改进后的遗传算法优化 BP 神经网络所得到的产量预测结果与实际产量值的对比图如图 8 和图 9 所示。改进后的遗传算法优化 BP 神经网络的 CGA-BP 神经网络模型在预测精度方面得到了很大的提高,产量预测模型的平均相对误差仅为 2.27%,较单一 BP 神经网络降低了 1.75%,较由遗

传算法优化的 BP 神经网络降低了 1.06%,输出结果与真实值吻合程度较高,变化趋势相同,且输出指标的预测误差变化波动较小,模型的预测效果相对较为稳定。

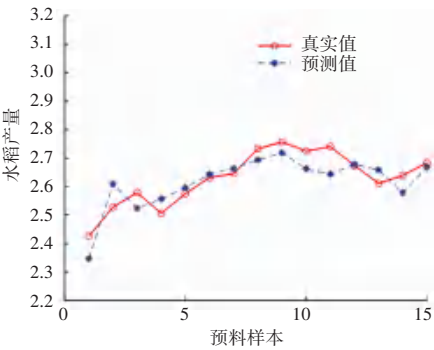


图 8 改进后的遗传算法优化 BP 神经网络产量预测对比图

Fig. 8 Comparison of BP neural network yield prediction optimized by improved genetic algorithm

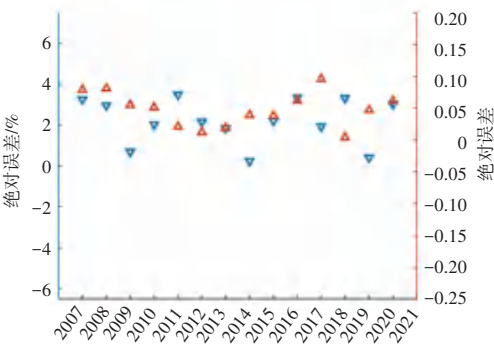


图 9 改进后的遗传算法优化 BP 神经网络产量误差分析图

Fig. 9 Yield error analysis diagram of BP neural network optimized by improved genetic algorithm

通过使用 BP、GA-BP、CGA-BP 模型对测试集进行预测后得到各模型的产量预测结果,并与真实值进行比较后得到各模型的相对误差,各模型的产量预测结果见表 1。

表 1 BP、GA-BP、CGA-BP 模型的产量预测结果

Table 1 Production prediction results of BP, GA-BP and CGA-BP models

年份/年	真实值	BP	BP 相对误差/%	GA-BP	GA-BP 相对误差/%	CGA-BP	CGA-BP 相对误差/%
2007	2.425 70	2.586 70	6.637 3	2.366 70	2.432 3	2.346 70	3.256 8
2008	2.528 00	2.570 00	1.661 4	2.664 00	5.379 7	2.609 00	2.966 8
2009	2.578 60	2.760 60	7.058 1	2.528 60	1.939 0	2.523 60	0.698 1
2010	2.506 00	2.685 00	7.142 9	2.376 00	5.187 5	2.557 00	2.035 1
2011	2.575 40	2.703 40	4.970 1	2.626 40	1.980 3	2.596 40	3.494 6
2012	2.631 63	2.687 63	2.128 0	2.603 63	1.064 0	2.643 63	2.166 0
2013	2.645 27	2.840 27	7.371 6	2.689 27	1.663 3	2.664 27	1.852 4
2014	2.732 68	2.642 68	3.293 5	2.750 68	0.658 7	2.693 68	0.256 2
2015	2.756 75	2.854 75	3.554 9	2.616 75	5.078 4	2.718 75	2.212 8
2016	2.724 61	2.603 61	4.441 0	2.823 61	3.633 5	2.662 61	3.376 6
2017	2.740 35	2.911 35	6.240 1	2.848 35	3.941 1	2.644 35	1.934 1
2018	2.674 01	2.709 01	1.308 9	2.598 01	2.842 2	2.678 01	3.328 3
2019	2.611 50	2.562 50	1.876 3	2.708 50	3.714 3	2.658 50	0.421 2
2020	2.638 94	2.647 94	0.341 0	2.504 94	5.077 8	2.577 94	3.031 5
2021	2.683 20	2.746 20	2.347 9	2.827 20	5.366 7	2.668 20	3.018 8

计算每个模型的均方根误差 (RMSE)、平均绝对误差 (MAE) 以及平均绝对百分比误差 (MAPE) 见表 2。可见 CGA-BP 模型表现最佳, 与 BP 模型相比, CGA-BP 模型的 RMSE、MAE 和 MAPE 分别提高了 6.84%、5.99% 和 2.29%; 相较于 GA-BP 模型, RMSE、MAE 和 MAPE 分别提高了 4.46%、4.23% 和 1.60%。说明 CGA-BP 预测模型具有更高的预测精度和更好的效果, 优于其他两种模型。

表 2 各个模型预测误差
Table 2 Prediction errors of each model

预测模型	RMSE	MAE	MAPE
BP	0.121 0	0.105 3	0.040 2
GA-BP	0.097 3	0.087 6	0.033 3
CGA-BP	0.052 6	0.045 3	0.017 3

通过深入分析预测结果, 不难发现改进的遗传算法优化 BP 神经网络模型 (CGA-BP) 在稳定性和准确性方面表现出了显著的优势。与传统的 BP 神经网络模型和 GA-BP 神经网络模型相比, CGA-BP 模型的预测误差普遍较小。全面评估这 3 种模型的 RMSE、MAE 和 MAPE 后, 可以看出 CGA-BP 模型在预测方面展现出了更高的准确性能。因此, 可以得出结论, CGA-BP 模型在预测方面具有优越的稳定性和准确性。

4 结束语

本文提出了一种水稻产量的预测模型, 通过将改进后的遗传算法与 BP 神经网络模型相结合, 得到了基于改进遗传算法优化 BP 神经网络预测模型 CGA-BP。该预测模型展现出优越的性能, 有效提升了水稻产量的预测准确度。

本文针对湖南省水稻产量预测问题, 选取了 5 个特征变量进行建模, 并使用改进的遗传算法优化 BP 神经网络预测模型 (CGA-BP) 进行预测。实验结果表明: CGA-BP 模型相较于传统的 BP 神经网络模型和 GA-BP 模型, 具有更好的预测效果和更高的预测精度。但 CGA-BP 模型仍然存在一些不足之处, 如模型的适应性仍有待提高, 环境中可能存

在其他因素导致预测结果产生误差。因此, 未来可以考虑结合环境中其他可能影响水稻产量的因素, 进一步优化和改进模型, 提高其预测精度和准确性, 还可以尝试将该模型应用于其他地区的水稻产量预测, 检验模型的普适性和泛化能力。

参考文献

[1] 王杰, 鄢卫东, 刘文超, 等. 中国水稻栽培现状、存在问题及改进措施[J]. 世界热带农业信息, 2021(7): 33-34.

[2] 谢琦, 王健, 肖向阳. 云南红河县水稻产量 ARIMA 模型分析预测[J]. 农村经济, 2019(6): 51-54.

[3] SUKARTINI N, ROVIK A H, PRAPTOMO B. Predicting rice yield in Thailand using random forest model [J]. Applied Mechanics and Materials, 2017, 859, 85-91.

[4] MO T, MYINT K T, LOW S C, et al. (2017). Prediction of rice yield in Myanmar using machine learning algorithms[J]. Air, Soil and Water Research, 2017, 10. DOI: 10.1177/1178622117704140

[5] ZHANG X, HUANG J, LI X, et al. Predicting rice yield in China at county scale by fusing multi-source data and models [J]. International Journal of Applied Earth Observation and Geoinformation, 2015, 35: 199-210.

[6] ZHANG Xin. A hybrid method for rice yield forecasting based on time series and artificial neural network[J]. Precision Agriculture, 2019, 20(4): 769-783.

[7] RABBI M F, BYUNG R C. Hybrid forecasting model for rice yield prediction using fuzzy cluster-art and time series[J]. The Scientific World Journal, 2015, 35:199-210.

[8] 张峰峰, 张欣, 陈龙, 等. 采用改进遗传算法优化神经网络的双目相机标定[J]. 中国机械工程, 2021, 32(12): 1423-1431.

[9] 周婧, 白云龙, 张小宝, 等. 基于 MEA-BP 神经网络的土壤养分评价模型[J]. 中国农机化学报, 2020, 41(2): 231-236.

[10] KUREMOTO T, KIMURA S, KOBAYASHI K, et al. Time series forecasting using a deep belief network with restricted Boltzmann machines[J]. Neurocomputing, 2014, 137: 47-56.

[11] 周游, 陆安江, 刘璇. 基于改进遗传算法的 BP 神经网络的水体叶绿素 a 含量预测[J]. 电子测试, 2022(15): 37-42.

[12] 肖宏峰, 谭冠政. 遗传算法在蚁群算法中的融合研究[J]. 小型微型计算机系统, 2009, 30(3): 512-517.

[13] HOLLAND J H. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence[M]. Cambridge, MA: MIT Press, 1992.

[14] 宋尧, 仰燕兰, 叶桦. 改进遗传算法在 MSPSP 问题中的验证[J]. 计算机系统应用, 2020, 29(10): 235-241.

[15] 王佳楠, 王玉莹, 何淑林, 等. 基于改进遗传算法优化 BP 神经网络的土壤湿度预测模型[J]. 计算机系统应用, 2022, 31(2): 273-278.