

基于 IAGA-BP 神经网络改进的降水测量

蔡辉[†], 李寅, 唐乃乔

(南京信息工程大学自动化学院, 江苏 南京 210044)

摘要:针对降水的小型化的精准测量的需求,采用了多元压电陶瓷片采集降水信号作为神经网络模型输入,并且对 AGA-BP 神经网络算法进行改进。首先使用该雨量计和标准雨量桶同时采集降水数据获得后期建模所需数据样本,再使用 GA-BP 神经网络对数据样本进行建模。融入自适应遗传算法对模型进行优化后,又针对传统 AGA-BP 神经网络进化后期个体适应度差异减小选择过程弱的缺点,使用了非线性函数对变异和交叉因子进行自适应调节。实验结果与数据分析显示:使用 IAGA-BP 神经网络拟合的降水模型的回归系数高达 0.96, MSE 为 0.014627, 决定系数 R^2 为 0.88, 精度得到明显提升,为全自动化测量降水优化手段提供了新的参考。

关键词:降水;压电陶瓷;神经网络;自适应遗传算法

中图分类号:TP39

文献标识码:A

Improved Precipitation Measurement Based on IAGA-BP Neural Network

CAI Hui[†], LI Yin, TANG Naiqiao

(School of Automation, Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China)

Abstract: In order to meet the demand of miniaturized and accurate measurement of precipitation, this study uses multi-component piezoelectric ceramic pieces to collect precipitation signal as the input of neural network model, and improves the AGA-BP neural network algorithm. Firstly, the rain gauge and the standard rain barrel were used to collect precipitation data at the same time to obtain the data samples required for later modeling, and then GA-BP neural network was used to model the data samples, and adaptive genetic algorithm was incorporated to optimize the model, and the weak selection process was reduced due to the individual fitness difference in the late evolution of the traditional AGA-BP neural network. The non-linear function is used to adjust the variation and crossover factors. Experimental results and data analysis show that the regression coefficient of the precipitation model fitted with the IAGA-BP neural network is as high as 0.96, the MSE is 0.014627, and the determination coefficient R^2 is 0.88, which significantly improves the accuracy and provides a new reference for the fully automated precipitation measurement optimization method.

Key words: precipitation; piezoelectric ceramics; neural network; adaptive genetic algorithm

作为传统气象六要素之一的降水是全球气候变暖大条件下着重关注的要素,其与人们的生活息息相关。由于中国的地理位置处于陆地季

风气候的敏感区域,人们时常被极端降水威胁到生命财产安全,农、林、牧业每年也因其出现巨大的财产损失。雨量计作为降雨监测的主要手段,其测量

精度是否达标对于天气预报、预防气象地质灾害、指示农牧业生产以及地质土壤研究等场合是一个极其重要的衡量标准。

文献[1]中阐述了传统的人工观测雨量桶到翻斗式雨量计、虹吸式雨量计,再逐渐发展到基于电子技术和敏感器件的超声波雨量计、压电式雨量计、光学雨量计等,人们针对不同的场合拥有了更多的选择。上述各雨量计技术虽然已经比较成熟但是仍然有各自的缺陷,文献[2]指出了虹吸式雨量计需要定期进行人工记录与维护,并且虹吸管容易被污染影响测量精度;文献[3]中提到了翻斗式雨量计在发生暴雨时存在翻转不及时的弊端,也存在需要定期维护清洗、体积大、不方便搬移等缺点;光学雨量计本质使用的是光散射的原理,但实际使用过程中容易发生镜头污染,对实际测量造成影响,并且成本过于昂贵,无法进行大面积布设;文献[4]中使用压电式传感器经过机理建模的方法设计压电式雨量计,虽然体积更小,但建模过程烦琐;文献[5]中的压电式雨量计运用了神经网络的方法,但对网络的参数缺少对应优化。

针对压电雨量计的神经网络参数优化问题,对压电雨量计神经网络模型进行适当的改进,提出了基于 AGA-BP 神经网络改进的压电雨量计相关设计方法,提升了整个模型的收敛速度和测量精度。本研究为降雨检测数据优化提供了新的手段,提升国产仪器设备的研发能力。

1 雨量计原理及系统结构

1.1 压电传感器测量原理

文献[6]介绍了压电陶瓷是一种拥有压电效应的介电材料,通过对该材料施加压力,该种材料会在内部产生正负电荷的极化现象,从两端连接导线后便可以获取到材料受压而产生的电荷。材料的极化强度和应力位线性关系如下公式所示:

$$p = d \times \sigma$$
(1)

其中, σ 和 d 分别为压电材料所受的压力和压电张力常数, p 为极化强度。小信号检测领域使用最为广泛的是铌酸钾钠压电陶瓷,因其具有较高的谐振频率且压电应变常数 d_{33} 大。文献[7]使用该材料制作了一款声速度计,所以压电陶瓷也适合用来检测雨水击打产生的微小振动信号。针对上述特性可以用来表征降水击打物体强度的能力,且高精度放大器拥有使其信号线性放大的能力,便有了本研究采集数据及拟合降水强度的实验及相关

数据优化分析。压电陶瓷的实物以及其极化产生的电荷波形示意图如图 1 和图 2 所示。

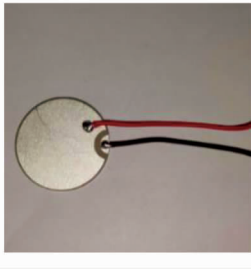


图 1 压电陶瓷实物图

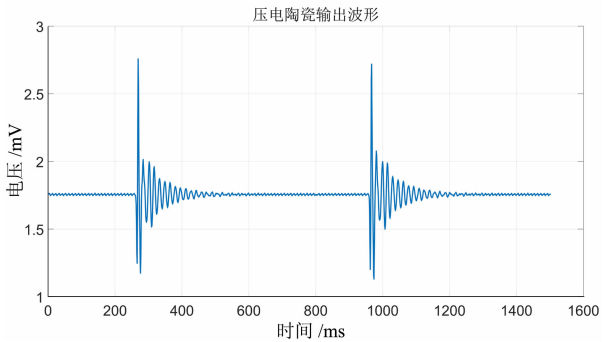


图 2 压电陶瓷波形图

1.2 雨量计结构及数据处理

该雨量计的结构如图 3 所示,整个雨量计外壳由工业级 PLA 材料制成,其收缩率低、韧性好、强度高,因而可以适应降雨监测的恶劣环境。降雨时雨水先进入偏心漏斗结构的盛雨器,经过漏斗流入疏水槽后进入隔离槽,最后通过出水口排出。

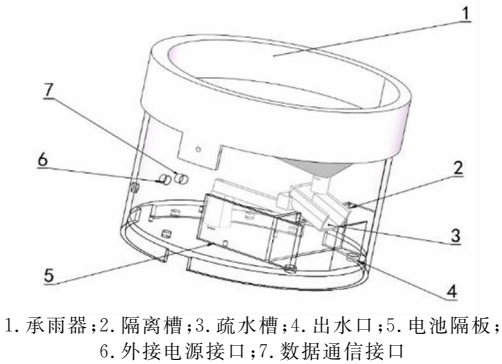


图 3 压电式雨量计结构设计

为了去除传感器的边界效应,相对于传统机理建模方法不同的地方在于同时使用 4 个压电陶瓷片均匀放置在传感器顶部内侧进行数据采集并传入神经网络,该雨量计的信号处理总体示意图如图 4 所示。

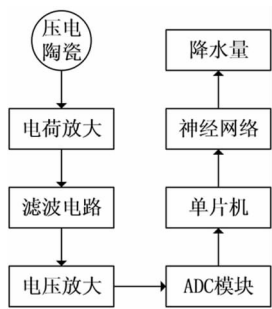


图 4 信号测量电路原理图

图 5 将 1 个压电陶瓷每分钟内测得电压累计值作为一次测量数据,数据采样率为 2.5 kHz,后期将 4 个压电传感器测量值传入神经网络后与每分钟降水量进行非线性拟合,得出两者的输入输出映射关系模型。

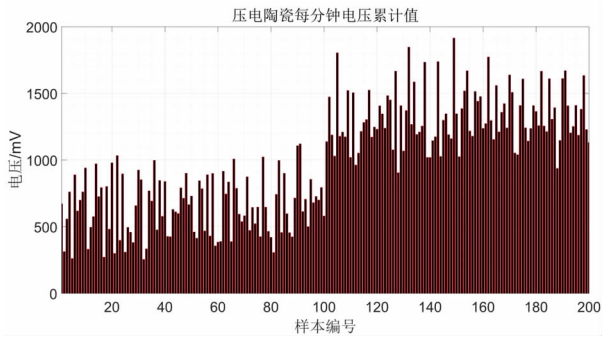


图 5 压电传感器电压累计值

2 降水模型建立及相关优化

2.1 GA-BP 神经网络模型建立

BP 神经网络是一种基于梯度下降法的多层前馈神经网络,以目标函数的负梯度方向对神经网络的参数进行调整,根据文献[8]中使用全连接神经网络进行红外降水测量的思路,本研究的整个 BP 神经网络模型的拓扑结构如图 6 所示。

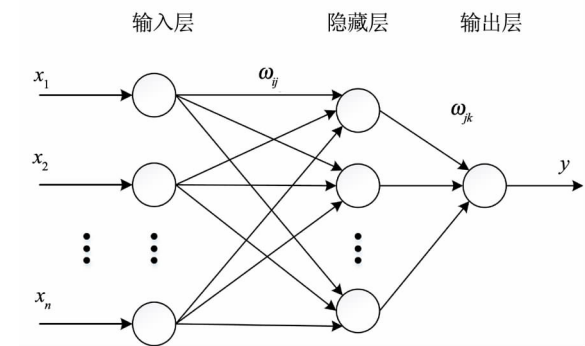


图 6 BP 神经网络拓扑结构图

信号从输入层进入传递到隐藏层,经过隐藏层的激活函数 $g(t)$ 后再传递到输出层,其中 $g(t)$ 为双曲正切函数 $\text{tansig}(t)$,输出层根据输出 Y_k 和预测值 P_k 和的均方误差(MSE)进行梯度下降法对整个网络的权值和阈值进行迭代更新。BP 神经网络经过上述过程不断迭代更新后使损失函数 $L = \frac{1}{2} \sum_{k=1}^q (Y_k - P_k)^2$ 值最小,从而得到压电传感器电

压累计值 x_i 与降水量 Y_k 的关系为:

$$Y_k = \sum_{j=1}^p [g(\sum_{i=1}^n \omega_{ij} x_i + \alpha_j)] \omega_{jk} + \beta_k \tag{2}$$

为克服 BP 神经网络在训练的过程中初始权值与阈值选取不当陷入局部最优解的缺陷,文献[9]和文献[10]运用了遗传算法分别解决了材料晶体特征参数寻优和裂纹智能识别优化。因而本研究还使用遗传算法对 BP 神经网络进行相应的模型优化,具体的遗传算法优化 BP 神经网络参数的步骤为编码、选择、交叉、变异四个阶段,整个优化流程如图 7 所示。遗传算法的初期先随机产生神经网络解的个体作为初始种群,并使用 $[-2, 2]$ 区间内的实数串编码,再使用初始种群计算每个个体解的适应度,通过适应度的不同来对下一代种群进行选择,自适应度函数 F_i 的计算公式如下:

$$F_i = \sqrt{\sum_{i=1}^M (P_i - Y_i)^2} \tag{3}$$

其中, P_i 为预测值, Y_i 为实际待测值。

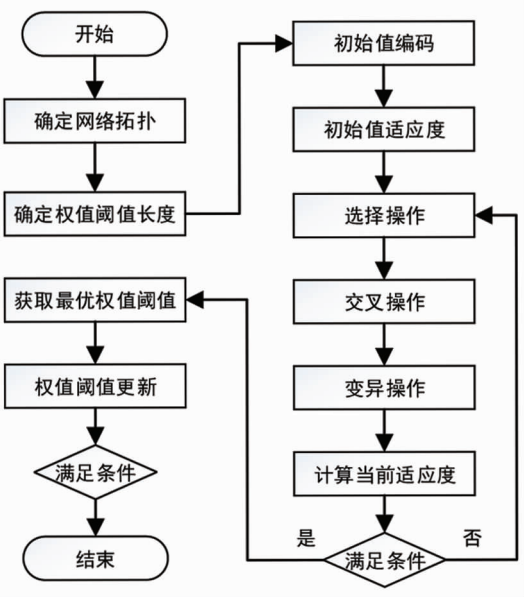


图 7 遗传算法优化 BP 神经网络流程图

2.2 自适应遗传算法及其优化

由于传统遗传的交叉算子 P_c 和变异算子 P_m 是常数,因而针对不同的优化问题需要通过反复实验来确定这两个参数,因而有了文献[11]提出的根据染色体适应度值不同自行调整 P_c 和 P_m 的自适应遗传算法:

$$P_c = \begin{cases} k_1 (f_{\max} - f') / (f_{\max} - f_{\text{ave}}) & , f' \geq f_{\text{ave}} \\ k_1 & , f' < f_{\text{ave}} \end{cases}$$
$$P_m = \begin{cases} k_2 (f_{\max} - f') / (f_{\max} - f_{\text{ave}}) & , f' \geq f_{\text{ave}} \\ k_2 & , f' < f_{\text{ave}} \end{cases}$$

(4)

其中, $f_{\max}$ 为种群最大适应度; f_{ave} 为种群平均适应度; f' 为本轮迭代后两交叉个体中适应度较大的值。上述方法在参数寻优过程时中前期种群多样性差进化得快,并且随着优良种群的差异减小,适应度大的个体极可能出现变化太小甚至不变化情况,从而使整个优化算法出现“早熟”现象,对于早熟问题,文献[11]使用机器学习结合改进自适应遗传算法的方法对钻头选取参数进行了优化,而文献[12]也使用了非线性改变遗传算法参数方法提升了卡钻故障预测。为解决“早熟”现象,本研究也使用了指数量函数对传统自适应遗传算法的 P_c 和 P_m 进行优化,优化公式如下所示:

$$P_c = \begin{cases} k_1 \cdot e^{\frac{f_{\text{ave}} - f'}{f_{\max} - f_{\text{ave}}}} & , f' \geq f_{\text{ave}} \\ k_1 & , f' < f_{\text{ave}} \end{cases}$$
$$P_m = \begin{cases} k_2 \cdot e^{\frac{f_{\text{ave}} - f'}{f_{\max} - f_{\text{ave}}}} & , f' \geq f_{\text{ave}} \\ k_2 & , f' < f_{\text{ave}} \end{cases}$$

(5)

使用上述公式交叉算子 P_c 和变异算子 P_m 的优点在于后期随着优良种群的差异性减小后能够适当放大优良种群的差异性,便于优良种群朝更优良的方向进化,避免了“早熟”现象。

3 数据集获取与分析

3.1 数据获取与数据集划分

本研究采用的数据是由某电子科技有限公司的 OSA-1801 型雨量计和自制的压电式雨量计在空旷水平地带进行实际降水采集的,现场采样设备如图 8 所示,其中 OSA-1801 为精度 0.1 mm 的双翻斗式雨量计,原始数据一共 2500 组,每组数据为一分钟内 4 片压电陶瓷分别测量到的电压累计值

和实际降水量值,前者作为神经网络的 4 路输入,后者作为神经网络的输出进行相应模型训练。为方便后期的模型训练与评价,将所有数据按训练集、验证集、测试集分别按 60%、20% 和 20% 的比例进行相应的划分。



图 8 降水数据样本集采集

3.2 模型数据分析对比

(1)GA-BP 神经网络训练与拟合

传统 GA-BP 神经网络的拓扑结构为 4—15—1,所以该遗传神经网络的编码长度为 91,其中隐藏层的权值和阈值分别为 60 和 15 个,输出层有 15 个权值和 1 个阈值,遗传算法的初始化参数具体如下:整个种群的规模数量为 30,其中交叉概率设为 0.4,变异概率设置的相对较小为 0.05。该网络的收敛趋势和回归系数分布如图 9 所示,回归系数大约为 0.94,最佳 MSE 的值为 0.080524,相对较低,经计算后的决定系数 R^2 值为 0.84。

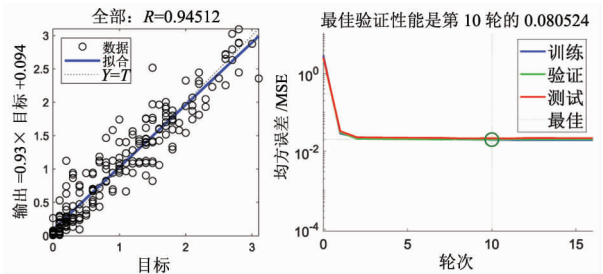


图 9 GA-BP 神经网络训练效果图

(2)AGA-BP 神经网络训练与拟合

在使用了拓扑结构为 4—15—1 的神经网络的情况下,使用传统 AGA-BP 神经网络预测的效果如图 10 所示,模型的回归系数相对于传统 GA-BP 神经网络提高到了 0.959 左右,MSE 也从 0.

080524 减少到了 0.02134,其中决定系数 R^2 约为 0.87,由此可见经过自适应算法的优化后对整个测试集与实际降水数据拟合度更高。

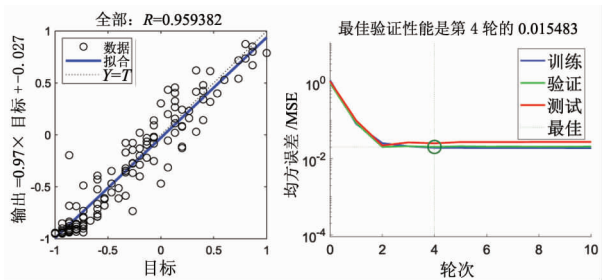


图 10 AGA-BP 神经网络训练效果图

(3)IAGA-BP 神经网络训练与拟合

使用改进后的非线性自适应改变交叉概率和遗传概率后,效果如图 11 所示,可以看出回归系数 R 的值从 0.959 左右提高到了 0.96 以上,但对于 MSE 的优化效果有所减缓,仅从 0.015 减小到 0.014 左右,决定系数 R^2 由 0.87 提升到 0.88 左右,所以改进的 AGA-BP 神经网络对模型优化虽然改善效果并不是那么明显,但是有优化效果,并且从训练的最佳 MSE 的训练速度来看,改进后的 IA-AGA-BP 神经网络的训练速度要明显优于其他两者。

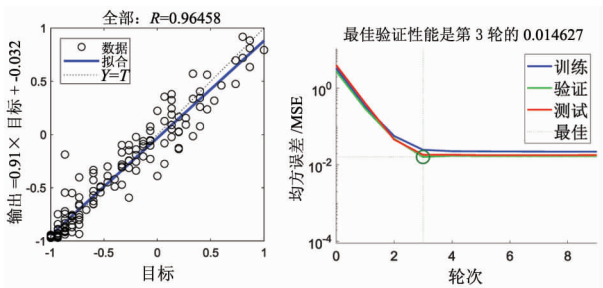


图 11 IAGA-BP 神经网络训练效果图

(4) 适应度曲线分析

三种算法的适应度寻优曲线如图 12 所示,传统遗传算法和自适应遗传算法在前期进化速度比较快,适应度迅速下降,但是到了后期明显适应度收敛速度明显下降,整个模型出现了“早熟”现象,IAGA-BP 神经网络前期迭代虽然比前两种优化算法慢,但是引入了非线性函数改变变异因子和交叉因子后加强了整个种群的进化,且优化效果更好。

3.3 数据拟合分析

经上述模型的相关模型性能参数分析后本研究使用实际降水采集数据进行拟合,效果如图 13 所示,三个模型对实际的数据拟合趋势基本相同,但对应的实际精度有仍然存在不同误差,其中 IA-

GA-BP 神经网络的预测效果最好,其次为 AGA-BP 和 GA-BP。

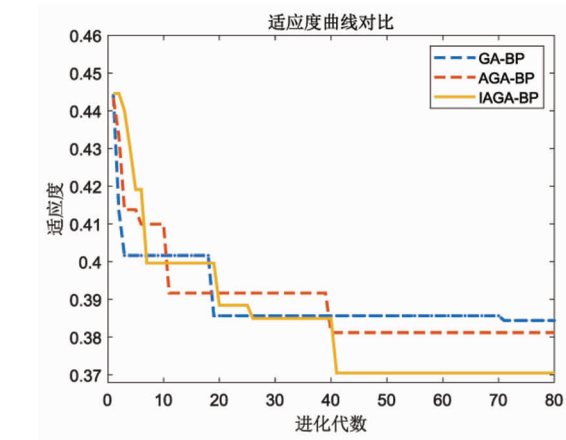


图 12 适应度寻优曲线

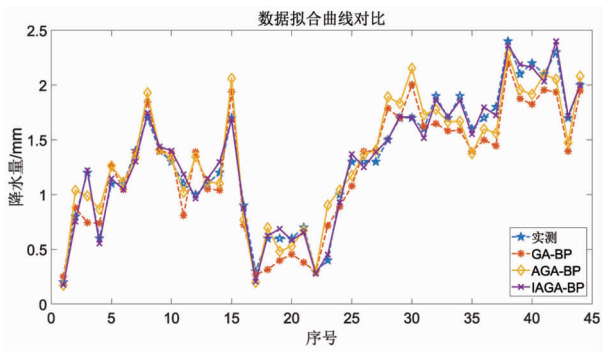


图 13 降水量预测值

为此从中随机抽取 10 组数据对具体的误差与分析,具体如表 1 所示。从表中数据也可以看出实际拟合过程中降水强度越小,拟合效果越差相对精度越低,对强度越大的降水拟合时的相对精度越高。

表 1 部分数据预测误差对比

序号	雨量桶 数据/mm	GA-BP 误差/mm	AGA-BP 误差/mm	IAGA-BP 误/mm
1	0.6	0.189	-0.126	-0.093
2	1.1	0.121	-0.117	-0.082
3	1.2	-0.101	0.124	0.094
4	0.3	0.189	0.131	0.097
5	0.7	0.164	-0.130	0.078
6	2.4	0.131	0.119	0.085
7	0.9	-0.252	0.176	-0.092
8	0.4	0.173	-0.115	0.103
9	2.3	0.152	0.165	0.089
10	1.6	0.143	0.116	0.081