

基于PCA的EC管外壁阻垢率预测模型研究

陈杰^{1,2}, 陈永保^{1,2}, 刘婧楠^{1,2}, 赵顺安^{1,2,3}, 章立新^{1,2*}

¹上海理工大学能源与动力工程学院, 上海

²上海市动力工程多相流动与传热重点实验室, 上海

³中国水利水电科学研究院, 北京

收稿日期: 2021年12月6日; 录用日期: 2021年12月31日; 发布日期: 2022年1月10日

摘要

由于蒸发式冷凝器(EC)换热管外阻垢率不易准确测试, 搭建了一个在喷淋水中加入柠檬酸(CA)阻垢的实验装置并获取了180组实验数据。本文利用BP、GRNN两种神经网络对阻垢率进行预测并结合主成分分析(PCA)对实验输入参数进行降维并对降维前后的预测结果进行对比。降维前两种模型的输入参数为实验进行时间、换热管外壁温度、喷淋水温度、pH值、电导率, 降维后两种模型的输入参数为 Y_1 和 Y_2 , 输出参数均为阻垢率。对比降维前后的预测结果得出: 经过PCA降维后, BP的网络运行时间从20s降为15s, 误差指标RMSE、MAPE分别从2.45%、3.6%降为1.44%、2.38%, 相关系数R从0.9745升高到0.9885; GRNN的运行时间从0.8s降为0.4s, 误差指标RMSE、MAPE分别从1.78%、2.85%降为1.04%、1.98%, 相关系数R从0.9853升高到0.9966, 并且降维后GRNN模型比BP模型预测时间超短、误差指标小、相关性更高, 更适合本领域。

关键词

蒸发式冷凝器, 阻垢, 主成分分析, 神经网络预测

Research on Prediction Model of Scale Resistance Rate of EC Pipe Outer Wall Based on PCA

Jie Chen^{1,2}, Yongbao Chen^{1,2}, Jingnan Liu^{1,2}, Shun'an Zhao^{1,2,3}, Lixin Zhang^{1,2*}

¹School of Energy and Power Engineering, University of Shanghai for Science and Technology, Shanghai

²Shanghai Key Laboratory of Multiphase Flow and Heat Transfer of Power Engineering, Shanghai

³China Institute of Water Resources and Hydropower Research, Beijing

Received: Dec. 6th, 2021; accepted: Dec. 31st, 2021; published: Jan. 10th, 2022

*通讯作者。

文章引用: 陈杰, 陈永保, 刘婧楠, 赵顺安, 章立新. 基于PCA的EC管外壁阻垢率预测模型研究[J]. 建模与仿真, 2022, 11(1): 76-87. DOI: 10.12677/mos.2022.111007

Abstract

Because it is difficult to accurately measure the external scale inhibition rate of evaporative condenser heat exchange tube, an experimental device for adding citric acid into spray water was built and 180 groups of experimental data were obtained. In this paper, BP and GRNN neural networks are used to predict the scale inhibition rate, and combined with PCA, dimension reduction of experimental input parameters is carried out, and the prediction results before and after dimension reduction are compared. Before dimensionality reduction, the input parameters of the two models are experiment time, outer wall temperature of heat exchange tube, spray water temperature, pH value and conductivity. After dimensionality reduction, the input parameters of the two models are Y_1 and Y_2 , and the output parameters are scale inhibition rate. By comparing the prediction results before and after dimension reduction, the running time of BP network decreased from 20s to 15s, error index RMSE and MAPE decreased from 2.45% and 3.6% to 1.44% and 2.38% respectively, and correlation coefficient R increased from 0.9745 to 0.9885. The running time of GRNN decreased from 0.8s to 0.4s, the error index RMSE and MAPE decreased from 1.78% and 2.85% to 1.04% and 1.98% respectively, and the correlation coefficient R increased from 0.9853 to 0.9966. Moreover, GRNN model after dimensionality reduction has shorter prediction time, smaller error index and higher correlation than BP model, which is more suitable for this field.

Keywords

Evaporative Condenser, Discaling, PCA, Neural Network Prediction

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

蒸发式冷凝器(EC)是一种常用于冷库降温的换热设备[1]。EC 在运行过程中,循环的喷淋水喷淋到换热管壁面,主要通过喷淋水的蒸发与对流换热将管内工艺系统的循环制冷工质的热量放散到管外空气中。在 EC 运行期间,喷淋水与管外壁发生热交换,使喷淋水不断蒸发,水中的盐类和杂质被浓缩,换热管外壁出现结垢,增加了管内外的传热热阻,缩短了换热管的使用寿命,增加了设备的运行成本[2]。结垢是工业界与学术界中十分关注并且是一个长期存在的问题,90%以上的工业换热器都存在结垢问题,因此人们越来越重视对阻垢的研究[3]。综合目前国内外研究结果得出:阻垢方法大致可分为化学阻垢、物理阻垢、绿色阻垢三类,其中使用化学阻垢剂最为普遍,但含氮和磷的传统化学阻垢剂会导致水体富营养化,且增加了水处理系统设计和污水回收或排放的困难,而绿色阻垢剂在水体中的存在不会对生态环境造成恶劣的影响[4]。

在传统研究中,影响阻垢率的因素是多方面的,采用传统手段进行数据处理非常困难且效率低下[2]。值得注意的是,人工神经网络技术经过几十年的发展已具备优异的非线性映射和泛化能力,只要选择合适输入变量及网络参数,其输出值可以任意精度逼近期望值[5]。早在 20 世纪 50 年代,第一代神经网络就被提出,它的算法只有两层,主要解决线性问题,当函数稍微复杂一些,该神经网络无法得到有效解[6]。经过后人的努力,解决了第一代神经网络所遇到的问题,为后面的研究提供了极大的帮助,大量神

神经网络得以构建, 比如多层感知器[7]、BP神经网络[2]、PT神经网络[8]、DNN神经网络[9]、CNN神经网络[10]、RBF神经网络[11]、GRNN神经网络[12]等。目前处理非线性问题最常用的神经网络有BP与GRNN[13], 本文采用PCA对网络输入数据进行降维并通过预测结果选取合适的网络模型。

为此搭建了模拟蒸发式冷凝器运行的喷淋水系统, 将柠檬酸(CA)加入喷淋水中并计算出阻垢率, 采用各模型进行预测并选取最优的模型。

2. 实验研究

2.1. 阻垢机理

CA是一种通过微生物发酵而产生的有机酸, 分子式为 $C_6H_8O_7$, 结构如图1所示。

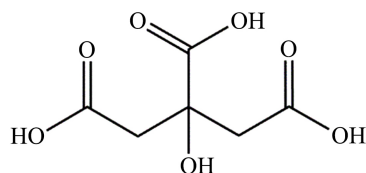


Figure 1. Molecular structural formula of CA
图 1. CA 分子结构式

由于CA有着自身的特异性以及螯合作用, 在工业中应用比较广泛[14]。换热管壁与喷淋水换热生成的水垢主要成分是碳酸钙($CaCO_3$), 柠檬酸会与碳酸钙在一定温度下发生化学反应, 水垢就会溶于酸, 在柠檬酸含量充足的情况下, 水垢的溶解度会随着温度的升高而增大。

CA在水中电离过程如式(1)~式(3)所示:



2.2. 实验系统与实验原理

2.2.1. 实验系统

实验系统如图2所示, 主要由循环水泵、增压水泵、模拟蒸发式冷凝器喷淋结垢实验装置、铂电阻、温度采集仪、pH采集仪、电导率采集仪、浮子流量计、电加热、补水箱等装置组成。



Figure 2. Experimental system
图 2. 实验系统

2.2.2. EC 模拟装置

图 3 为模拟 EC 的实验装置，主要由风机、收水器、喷淋管、水槽、模拟换热管、壳体组成。

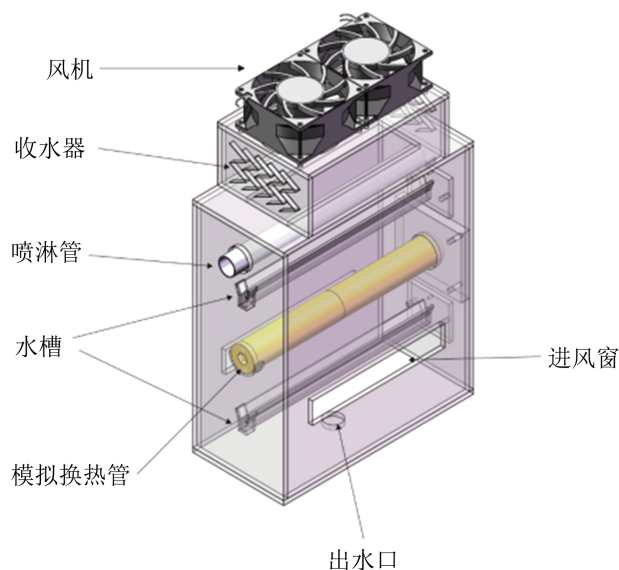


Figure 3. Simulation of countercurrent closed tower experimental device
图 3. 模拟逆流闭塔实验装置

模拟换热管相关参数列于表 1。

Table 1. Parameters of simulated heating tube
表 1. 模拟换热管相关参数

参数	数值
材质	黄铜
外径/mm	28
壁厚/mm	4
总长/mm	160
导热系数/ $\text{W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$	381
热电偶布置数量	6
最大加热功率/W	100

实验装置的壳体用有机玻璃板制成，上有与风机和喷淋管的接口，两侧有进风窗，底部有出水口。进风窗长度为 160 mm，宽度为 25 mm，进风口顶端和换热管底部的高度差为 40 mm。

2.2.3. 测试环境

实验相关条件如表 2 所示。

Table 2. Experimental conditions
表 2. 实验相关条件

条件	数值
测试环境	室内
环境温度/°C	15.2 ± 1
相对湿度/%	78 ± 3
初始内壁温度/°C	36 ± 0.2
自来水硬度(以 CaCO ₃ 记)/mg·L ⁻¹	80
CaCl ₂ 和 NaHCO ₃ 的摩尔比	1:2
实验进行时间/h	30

在表 2 条件满足之后,把喷淋水的硬度提升为 1000 mg/L,并同时加入 CA,此时水质波动剧烈,待稳定(0.25 h)之后每隔 10 min 采集一组数据。因喷淋水蒸发和垢在换热表面的析出,喷淋水的硬度会产生变化,故采用补水箱自动补水,并适时加入配好的 CaCl₂ 和 NaHCO₃ 溶液,通过监控 pH 值和电导率 σ 以使水质基本保持稳定。

阻垢率 η 按式(4)定义,单位为%。

$$\eta = \frac{R_f - R_{f(\text{CA})}}{R_f} \quad (4)$$

式中: R_f 与 $R_{f(\text{CA})}$ 分别为相同时刻未加 CA 与加 CA 两组实验计算所得的污垢热阻,该时刻污垢热阻按式(5)计算,单位为(m²·°C)/W。

$$R_{f(t)} = \frac{1}{K(t)} - \frac{1}{K(t=0)} \quad (5)$$

式中: $K(t)$ 为 t 时刻的传热系数, $K(t=0)$ 为实验起始时的传热系数。

传热系数 K 通过采集到的模拟换热管的电加热功率 P 、模拟换热管内壁面温度 T_{bi} 、喷淋水在模拟换热管外壁面上的平均温度 T_s , 结合模拟换热管几何尺寸(内外直径及管长)和导热系数,计算得到。详细的热阻测量原理参考文献[15]。实验共获得了 180 组数据。

3. BP 神经网络与 GRNN 神经网络预测结果分析

3.1. 主成分分析

水质因素(电导率 σ 、pH 值)、实验工况因素(模拟换热管外壁面温度 T_{bo} 、喷淋水在模拟换热管外壁面上的平均温度 T_s)、实验持续时间 τ , 均对阻垢率 η 有影响,各个因素间具有强非线性耦合,模型计算复杂度高。主成分分析法的思想是将高维度的相关变量转变为低维度互不相关且能代表原始数据大部分信息的新变量,降低模型计算复杂度,其流程如图 4 所示。

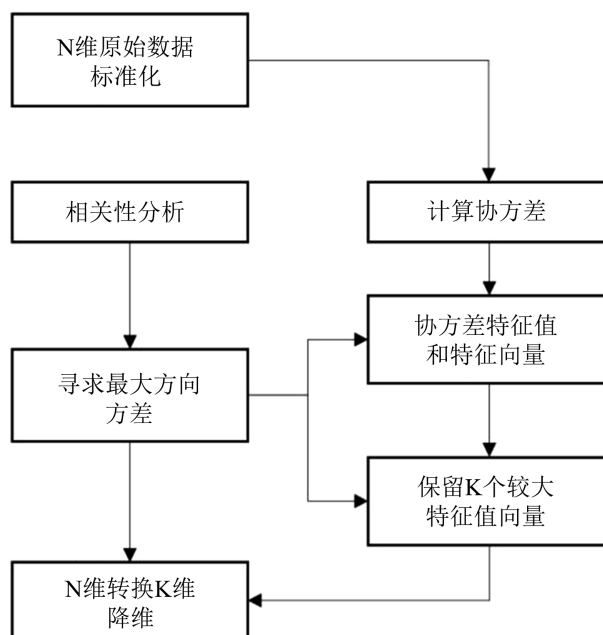


Figure 4. Flow diagram of principal component analysis
图 4. 主成分分析流程图

3.1.1. 各影响因素之间的相关分析

各影响因素之间的相关关系如表 3 所示。

Table 3. Correlation coefficient matrix of influencing factors
表 3. 影响因素间相关系数矩阵

影响因素	pH 值	σ	τ	T_{bo}	T_s
pH 值	1	-0.42	0.95	0.46	0.19
σ	-0.42	1	-0.12	0.47	0.73
τ	0.95	-0.12	1	0.70	0.48
T_{bo}	0.46	0.47	0.70	1	0.92
T_s	0.19	0.73	0.48	0.92	1

结果表明：众多因素之间有强烈的相关性。因此在建立 η 预测模型之前，需要采用 PCA 降维处理消除各影响因素间的相关性。

3.1.2. 各影响因素的主成分分析

主成分分析结果如表 4 所示，一共有 5 个主成分，主成分 Y_1 和 Y_2 的特征值分别为 3.937 和 1.910，贡献率分别为 58.738% 和 38.203%，累计贡献率达到 96.94%，累计贡献率超过 95% 满足作为主成分的条件，说明前两个主成分能够很好代替各影响因素的主要信息。因此，选取前两个主成分 Y_1 、 Y_2 作为 η 预测模型的输入节点。

Table 4. Principal component analysis of each influencing factor
表 4. 各影响因素的主成分分析

主成分	特征值	贡献率(%)	累计贡献率(%)
Y ₁	2.937	58.738	58.738
Y ₂	1.91	38.203	96.94
Y ₃	0.127	2.54	99.481
Y ₄	0.023	0.465	99.946
Y ₅	0.003	0.054	100

3.2. 神经网络结构及其相关评价指标

BP 和 GRNN 神经网络的基本结构都是由输入层、隐含层和输出层组成。

相关评价指标

用于表示神经网络预测结果的相关参数主要是：相关系数(R)、均方根误差(RMSE)、平均绝对百分比误差(MAPE)，它们的表达式为式(6)~(8)。

$$R = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \cdot \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (6)$$

其中：N 为样本个数、x 为输入值、y 为输出值。

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (a_i - p_i)^2} \quad (7)$$

式中：a_i 为实际输出值，p_i 为预测值。

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{a_i - p_i}{p_i} \right| \quad (8)$$

3.3. 神经网络的构建

选择 180 组数据中随机的 144 组数据作为训练样本，余下 36 组数据作为验证样本，分为以下几个步骤对数据进行处理：

1) 确定输入节点。未经 PCA 处理的参数为 τ 、 T_{bo} 、 T_s 、 σ 、pH 值；PCA 处理后的输入节点为 Y₁ 和 Y₂。

2) 确定隐含层节点。BP 神经网络隐含层神经元个数 n 通常由黄金分割法按式(9)确定。GRNN 神经网络隐含层节点数采用 newgrnn 函数设置。

$$n = \sqrt{a+b} + k \quad (9)$$

式中：a 为输入节点个数、b 为输出节点个数、k 为常数，k 的范围为[1~10]。

3) 确定输出节点参数为 η 。

4) 归一化处理。为了使数据统一、可靠性更高、便于计算，采用式(10)的归一化函数 mapminmax 进行数据处理。

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (10)$$

式中： x 为被归一化的值， y 为归一化之后的值。

3.4. 网络模型的训练与优化

3.4.1. BP 神经网络运行结果

根据式(9)得出未经 PCA 处理的隐含层节点数范围为[4~14]，经 PCA 处理后的隐含层节点数范围为[2~12]，PCA 处理前后的网络运行时间随着隐含层节点数的变化规律如图 5 所示，经过 PCA 处理后，相同隐含层节点下神经网络的运行时间降低 3s。

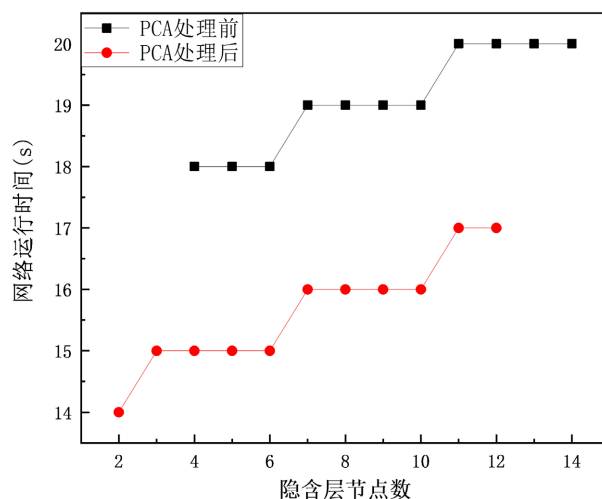
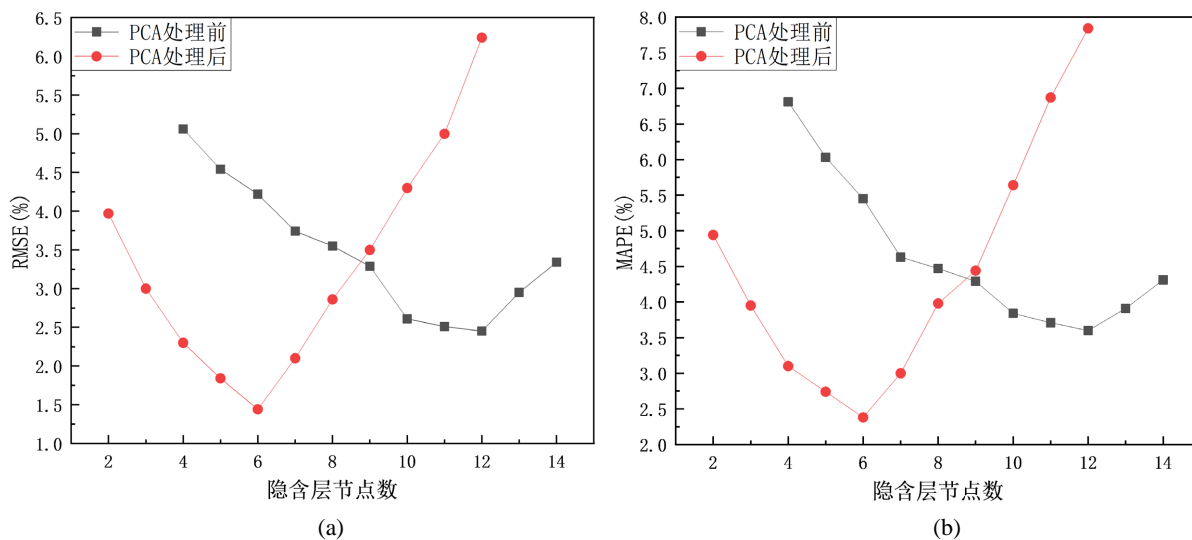


Figure 5. Variation of network running time with the number of hidden layer nodes before and after PCA processing

图 5. PCA 处理前后网络运行时间随隐含层节点数的变化情况

图 6(a)~(c)为降维前后 BP 模型运行之后各参数变化情况。



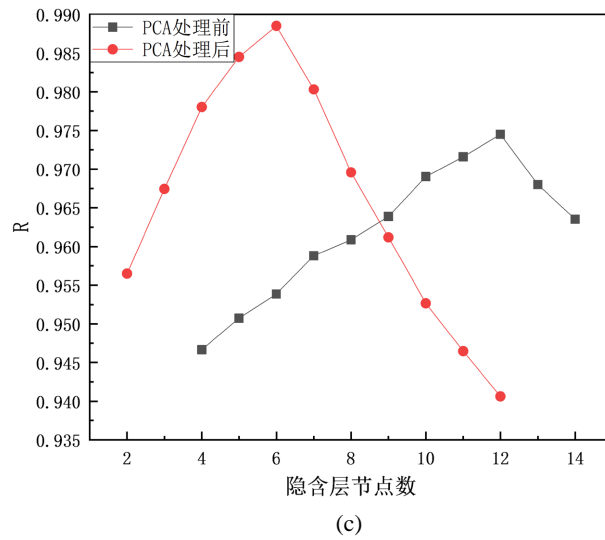


Figure 6. The variation of parameters of BP model with the number of hidden layer nodes before and after PCA processing
图 6. PCA 处理前后 BP 模型各参数随隐含层节点数的变化情况

综合图 5 与图 6(a)~(c)的结果得出：经过 PCA 降维后，BP 的网络运行时间从 20s 降为 15s，误差指标 RMSE、MAPE 分别从 2.45%、3.6%降为 1.44%、2.38%，相关系数 R 从 0.9745 升高到 0.9885。

为了选取最优的迭代次数，采用 MAPE 值作为评价指标，从迭代 1000 次开始并以 1000 为间隔递增，直到第 20,000 次时 MAPE 值趋于平稳，神经网络模型预测结果最佳，故选择 20,000 步作为最佳迭代次数。MAPE 随着迭代次数的变化关系如图 7 所示。

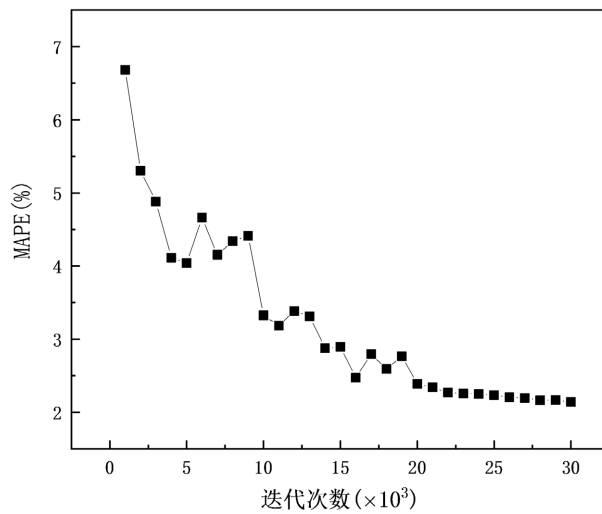


Figure 7. Changing trend of MAPE value with number of iterations
图 7. MAPE 值随着迭代次数的变化趋势

3.4.2. GRNN 神经网络运行结果

采用相同的数据用 GRNN 神经网络进行训练。spread 值是一个扩展常数，表示函数的光滑程度，对于数据的训练至关重要，spread 过大或过小都不能很好地预测模型。如果 spread 值过大，则会导致过拟合；如果 spread 过小，则会导致拟合结果远离实际值。spread 的范围为[0~2]，设定 spread 的初值为 0.1，

以 0.1 间隔递增, 降维前后 GRNN 模型运行之后各参数变化情况如图 8(a)~(c)所示。

经过 PCA 处理后, GRNN 的最优 spread 由 0.6 变为 0.2, 网络运行时间从 0.8s 降为 0.4s, 误差指标 RMSE、MAPE 分别从 1.78%、2.85%降为 1.04%、1.98%, 相关系数 R 从 0.9853 升高到 0.9966。

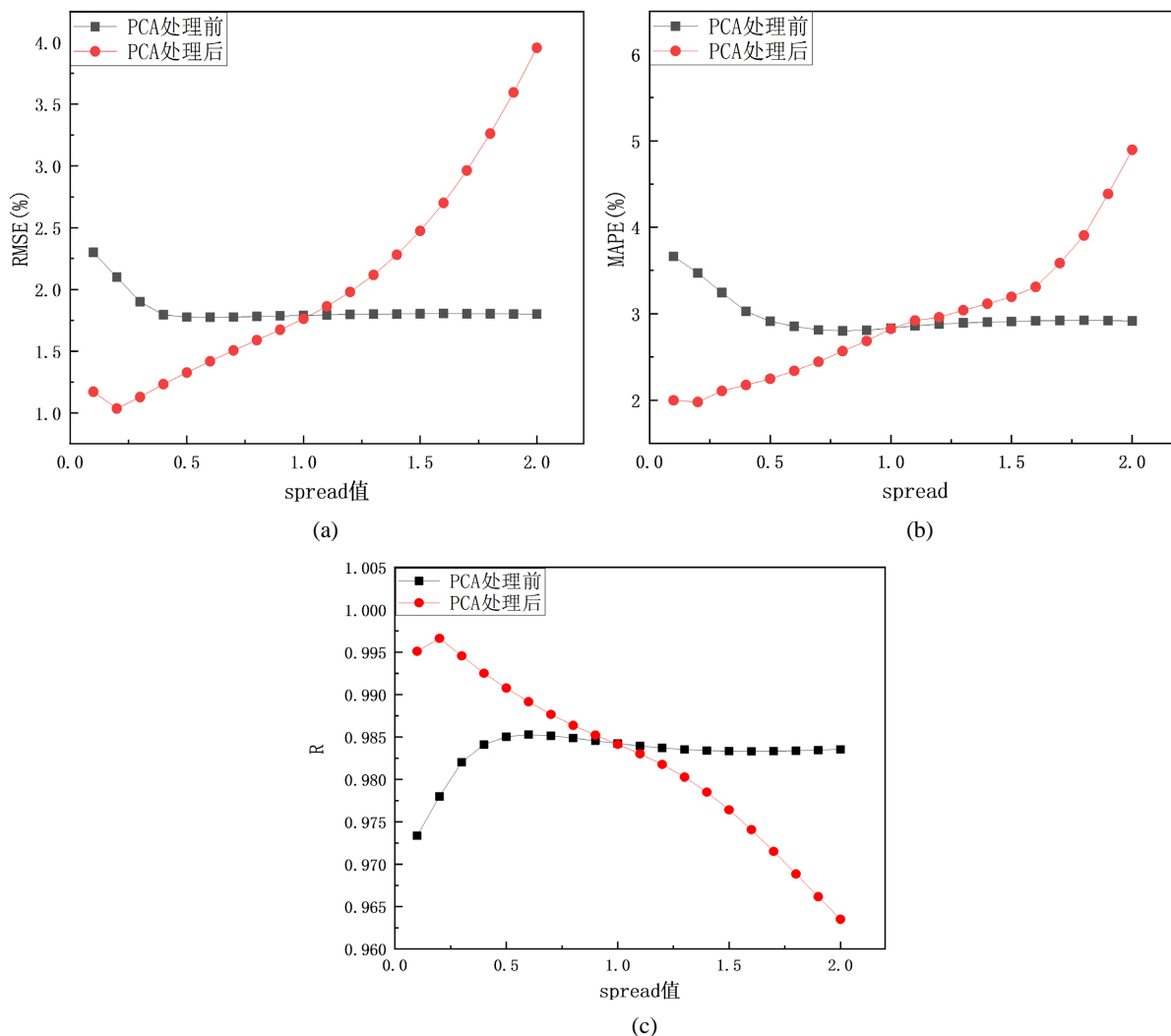


Figure 8. The variation of parameters of BP model with spread values before and after PCA processing

图 8. PCA 处理前后 GRNN 模型各参数随 spread 值的变化情况

3.4.3. 模型比较及分析

从 3.4.1 与 3.4.2 的结果分析得出: 经过 PCA 处理后的网络模型相较于未经过 PCA 处理的网络模型来说, 模型的运行时间更短, 在实际应用中能够明显的提高工作效率, RMSE 和 MAPE 更小, R 更大, 说明误差小, 预测精度更高。

图 9 为经过 PCA 处理后两种模型的线性回归分析对比图, 从图中可以发现经过 PCA 处理后的 GRNN 模型的相较于 BP 模型的线性拟合更加接近 45°线且数据更加集中, 表明该模型预测精度更高。

经过 PCA 处理后的两种模型相对误差如图 10 所示, BP 模型的最大相对误差为 8.74%, 主要集中在 4% 以内, 平均相对误差为 2.38%; GRNN 模型的最大相对误差为 -4.44%, 主要集中在 3% 以内, 平均相对误差为 1.98%。可见经 PCA 处理后的 GRNN 模型的误差更低。

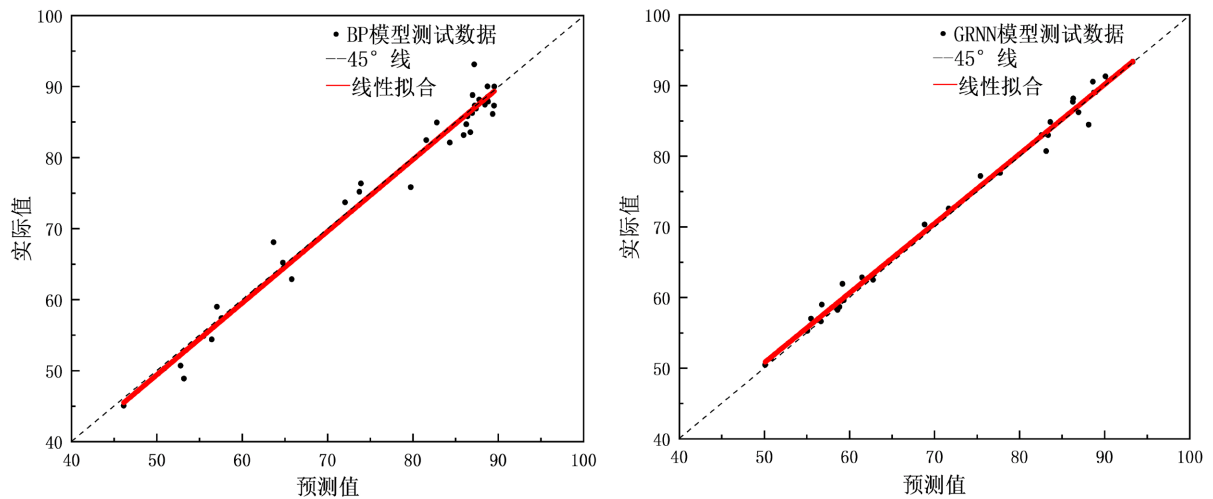


Figure 9. Linear regression analysis of predicted and actual values of the two models after PCA treatment
图 9. 经 PCA 处理后的两种模型的预测值与实际值的线性回归分析

对以上结果，从理论上分析是因为 GRNN 模型相较于 BP 模型来说，不需要进行权值的处理，以径向基网络为基础，非线性映射能力更强、收敛速度更快。GRNN 模型中人为调节参数很少，只有一个 spread 值，而 BP 模型需要人为调节的参数很多，有隐含层节点数、迭代步数等参数，这个特点决定 GRNN 模型得以最大可能地避免人为主观假定对预测结果的影响；而经过 PCA 处理后，原有的两两相关的变量合成了线性无关的低维变量，减少了模型的计算时间以及计算的复杂度。

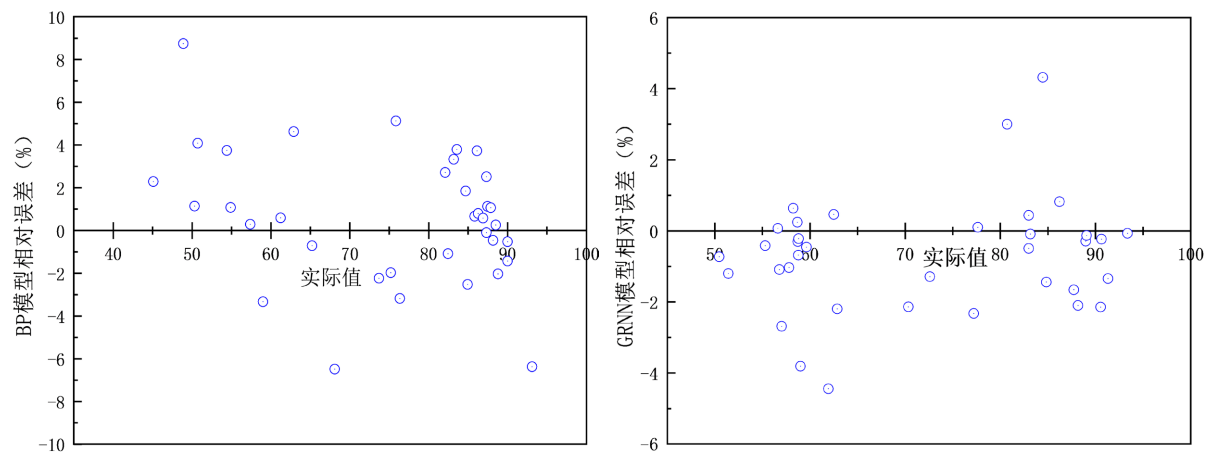


Figure 10. Relative errors of the two models after PCA treatment
图 10. 经 PCA 处理后的两种模型的相对误差

4. 结论

经过 PCA 处理之后，模型的输入参数由 τ 、 T_{bo} 、 T_s 、 σ 、pH 值变成 Y_1 和 Y_2 ，且两者的累计贡献率达到了 96.94%。

经 PCA 处理后的 GRNN 模型预测的 R 为 0.9966，平均相对误差为 1.98%；而经 PCA 处理后的 BP 模型预测的 R 为 0.9885，平均相对误差为 2.38%。由此可见，前者的 R 更接近 1，而平均相对误差也更小。

因此, 在经过 PCA 处理后, BP 模型与 GRNN 模型的预测结果都有所提升, 并且经过 PCA 处理后的 GRNN 模型对 EC 管外壁阻垢率的预测性能明显优于 BP 模型。

基金项目

中国博士后科学基金(No. 2020M681347)。

参考文献

- [1] 张聪. 家用空调器用蒸发式冷凝器研究进展及应用前景[C]//中国制冷空调工业协会. 第七届中国制冷空调行业信息大会论文集. 北京出版社: 中国制冷空调工业协会, 2014: 4.
- [2] 刘津, 章立新, 沈艳, 等. 基于 BP 神经网络的闭塔换热管壁污垢热阻预测[J]. 热能动力工程, 2020, 35(12): 66-71.
- [3] Miller, R.F. (1986) Method for Minimizing Fouling of Heat Exchanger. *Journal of Heat Recovery Systems*, **6**, 8. [https://doi.org/10.1016/0198-7593\(86\)90197-9](https://doi.org/10.1016/0198-7593(86)90197-9)
- [4] 赵彦, 章立新, 高明, 等. 循环冷却水系统除碳酸钙污垢的研究进展[J]. 精细化工, 2020, 37(12): 2447-2456.
- [5] 孙辉, 章立新, 杨其国, 等. 基于 GA-BP 神经网络的超临界二氧化碳折射率及密度预测[J]. 热力发电, 2020, 49(10): 59-64.
- [6] McCulloch, W.S. and Pitts, W. (1943) A Logical Calculus of the Ideas Immanent in Nervous Activity. *The Bulletin of Mathematical Biophysics*, **5**, 115-133. <https://doi.org/10.1007/BF02478259>
- [7] Kia, A., Sensoy, S. and Xie, X. (2014) Classification of Earthquake-Induced Damage for R/C Slab Column Frames Using Multiclass SVM and Its Combination with MLP Neural Network. *Mathematical Problems in Engineering*, **2014**, Article ID: 734072. <https://doi.org/10.1155/2014/734072>
- [8] Jou, I.-C., You, S.-S. and Chang, L.-W. (1994) Analysis of Hidden Nodes for Multi-Layer Perceptron Neural Networks. *Pattern Recognition*, **27**, 859-864. [https://doi.org/10.1016/0031-3203\(94\)90170-8](https://doi.org/10.1016/0031-3203(94)90170-8)
- [9] 祝嘉声. 基于 DNN 的汉语语音识别声学模型的研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2015.
- [10] 许可. 卷积神经网络在图像识别上的应用的研究[D]: [硕士学位论文]. 杭州: 浙江大学, 2012.
- [11] Chang, W.-Y. (2013) Estimation of the State of Charge for a LFP Battery Using a Hybrid Method That Combines a RBF Neural Network, an OLS Algorithm and AGA. *International Journal of Electrical Power and Energy Systems*, **53**, 603-611. <https://doi.org/10.1016/j.ijepes.2013.05.038>
- [12] Sridharan, M. (2020) Application of Generalized Regression Neural Network in Predicting the Performance of Solar Photovoltaic Thermal Water Collector. *Annals of Data Science*, **1**, 1-23. <https://doi.org/10.1007/s40745-020-00273-1>
- [13] 何汉林, 孟爱华, 祝甲明, 等. 基于优化的 GRNN 和 BP 神经网络的磁滞曲线拟合对比分析[J]. 机电工程, 2013(1): 122-126.
- [14] 刘杨. 新型亲和反胶团系统及其蛋白质萃取特性研究[D]: [博士学位论文]. 天津: 天津大学, 2006.
- [15] 赵彦. 高频电磁场中柠檬酸对管外侧除垢特性的研究[D]: [硕士学位论文]. 上海: 上海理工大学, 2021.