

王乙竹,陶 伟,陆思宇.基于神经网络的“回南天”观测数据质量控制方法初探[J].气象研究与应用,2024,45(2):37-44.  
WANG Yizhu, TAO Wei, LU Siyu. Preliminary study on the quality control method for observation data of “Continuous Wet Weather” based on neural network[J]. Journal of Meteorological Research and Application, 2024, 45(2): 37-44.

## 基于神经网络的“回南天”观测数据质量控制方法初探

王乙竹<sup>1</sup>, 陶 伟<sup>1\*</sup>, 陆思宇<sup>2</sup>

(1. 广西壮族自治区气象技术装备中心, 南宁 530022; 2. 广西壮族自治区气象科学研究所, 南宁 530022)

**摘要:**为判别“回南天”观测设备数据可靠性,基于传统反向传播神经网络(BPNN),结合粒子群优化算法(PSO-BPNN),对广西“回南天”观测数据进行质量控制研究。结果表明:(1)在模型估算温度与实测温度对比验证中,与BPNN模型相比,PSO-BPNN模型精度更高,PSO-BPNN模型没有明显高估或低估,而BPNN模型在10℃附近出现较大偏差。(2)在使用测试集数据对模型进行测试中,瓷砖地面和墙面温度在10~30℃范围,模型的适用性更强,PSO-BPNN模型稳定性优于BPNN模型。(3)在随机添加人工误差进行的模型检验中,PSO-BPNN模型瓷砖地面、墙面、水泥地面温度的最佳质量控制参数分别为1.73、1.64、1.68,BPNN模型分别为1.82、1.83、1.78。

**关键词:**质量控制;反向传播神经网络;粒子群优化;“回南天”

**中图分类号:** P412 **文献标识码:** A **doi:** 10.19849/j.cnki.CN45-1356/P.2024.2.06

“回南天”是指经过长时间低温过程后气温快速回升,当暖空气露点温度大于等于物体表面温度时,水汽达到饱和状态,使得物体表面出现潮湿、小水珠、冒水等天气现象<sup>[1-2]</sup>,多出现于华南和江南地区。“回南天”易导致衣物难干、食物生霉、电器短路等,对人类身体健康、工业生产、交通出行方面造成诸多不便。

为更好监测、预报、服务“回南天”过程,广西壮族自治区气象局于2016年完成91个“回南天”观测系统布网<sup>[3]</sup>,常态化观测站内瓷砖地面温度、墙面温度、水泥地面温度和空气温度<sup>[4]</sup>。数据及时、时间间隔小、站点覆盖率高。但由于目前缺乏有效的质量控制手段,无法判断“回南天”观测数据可靠性。而高质量的实时“回南天”观测数据是运行监控、技术保障、过程预报的重要依据,因此对“回南天”观测数据进行质量控制势在必行。

传统的观测数据质量控制方法是以范围检查、极值检查、内部一致性检验相结合的综合检查

法<sup>[5-7]</sup>。针对传统质控算法时空维度检验不足问题,后续的研究中加入时间一致性检验、空间一致性检验和时变检验等方法,使观测资料判别精度得到提升<sup>[8-10]</sup>。然而该方法计算过程复杂,需大量参数信息,数据处理速度慢,容错率低。近年来,神经网络(NN)、支持向量机、随机森林和深度学习模型等被用于数据质量控制研究<sup>[11-13]</sup>。机器学习模型可以表达各种变量之间的非线性关系或复杂关系,有效提高数据处理能力。

反向传播神经网络(back propagation neural network, BPNN)简单、易实现,具有较强自学习和自适应能力,但其泛化能力差、且容易陷入局部极值<sup>[14]</sup>。为探求一种判别精度高、易实现、处理速度快的“回南天”观测数据质量控制方法,本文基于BPNN,结合粒子群优化(particle swarm optimization, PSO)算法对BPNN权值阈值进行优化,首次对广西全区“回南天”观测数据进行质量控制初步研究,以期后续技术落地提供方法指导,为监测、预报、服

收稿日期: 2023-11-02

基金项目: 广西气象科研计划项目(桂气科2021Z05)、广西壮族自治区气象技术装备中心自立项目(ZBKY202304)

第一作者: 王乙竹(1996-), 硕士, 助理工程师, 主要从事气象装备运行保障、质量控制研究。E-mail: 617575956@qq.com

\*通讯作者: 陶伟(1983-), 高级工程师, 主要从事综合气象观测研究。E-mail: 249525294@qq.com

务“回南天”过程提供有效数据支持。

## 1 数据来源与方法

### 1.1 数据来源

本研究使用的数据包括:1)“回南天”设备观测数据,2)国家地面气象观测资料,3)辅助数据(海拔、经度、纬度等)。其中,地面气象观测资料和辅助数据作为模型输入,“回南天”设备观测数据作为模型输出,搭建神经网络模型。

国家地面气象观测资料经过严格质量控制可用率达99.82%<sup>[15]</sup>。本研究使用地面观测资料提供的小时降雨、气温、气压、湿度、海平面气压、露点温度、12 h露点温度变化值、水汽压、2 min平均风向、2 min平均风速和能见度数据,逐小时间隔观测。数据来源于气象大数据云平台。

“回南天”观测站与国家气象观测站同址建设,布设在自然通风独立房间内,分别在室内瓷砖地面、墙面、水泥地面凹槽安装传感器,观测不同材质物体表面的实时温度,此外该装备还可提供室内气温观测。91个观测站中四温度要素站(观测瓷砖地面温度、墙面温度、水泥地面温度、空气温度)9个(图1红点),两温度要素站(观测瓷砖地面温度、墙面温度)82个(图1黑点)。由于“回南天”气温观测数据可根据国家地面观测气温进行简单质控,本研究在此不做讨论。数据来源于省级“回南天”观测数据库,选用逐小时观测数据。

此外,站点海拔(5.5 m到1030.9 m)、经度、纬度、儒略日(Julian Day, JD)、时间等信息作为模型的辅助输入数据。

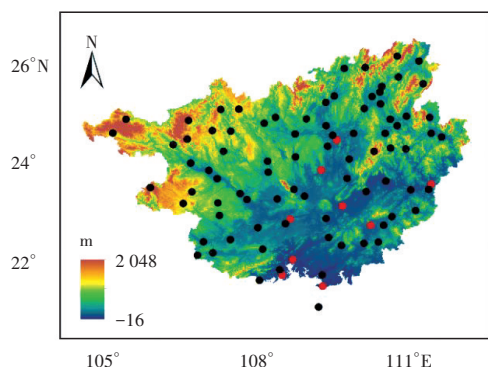


图1 研究区域海拔及国家级自动观测站/“回南天”观测站(91个)空间分布。红点为“回南天”四温度要素站观测站(瓷砖地面温度、墙面温度、水泥地面温度、空气温度),黑点为“回南天”两温度要素站(瓷砖地面温度、墙面温度)

### 1.2 研究方法

BPNN本质是多层前馈网络,通过数据正向和误差方向传播不断学习更新网络来提高准确性,其结构简单,广泛应用于气象领域<sup>[16-17]</sup>。但在数据相关因素过多的情况下,BPNN依赖于设计结构,易存在陷入局部极值,有时无法找到全局最优值。“回南天”数据质控模型输入包括经纬、时间、降雨、气温等14类相关要素,PSO算法具有较强的全局收敛能力<sup>[18]</sup>,利用PSO算法对神经网络进行优化(即PSO-BPNN),发挥BPNN非线性映射能力和PSO算法的全局寻优能力,以此提高模型判别准确性。

图2显示 PSO-BPNN“回南天”观测数据质量控制方法的流程。首先建立以国家地面气象观测资料、辅助数据为输入,“回南天”设备观测数据为输出的BPNN,进行初始训练。利用PSO算法全局性搜索的特点,寻找合适的神经网络阈值和网络初始连接权值。进行新一轮BPNN训练,提高局部收敛速度和系统预测的准确性。最终通过多次迭代优化调整,获得最佳预测模型。

基于 MATLAB 2016b 中的 Neural Network Toolbox 工具箱进行基本三层网络的构建,考虑样本集的处理和网络结构的设计<sup>[19]</sup>,采用广泛的预试验分析法进行拓扑结构设计。本研究BPNN模型使用试凑法确定网络结构为三层(输入层、隐藏层、输出层),输入层为14,隐藏层神经元数为20,输出层为1,设置瓷砖地面、墙面温度数据模型最大迭代次数300次,目标误差0.001,水泥地面温度模型最大迭代次数99次,目标误差0.001。PSO算法中,三个模型惯性因子为0.42,群粒子个数为10,学习因子为2,最大限制速度设为1。

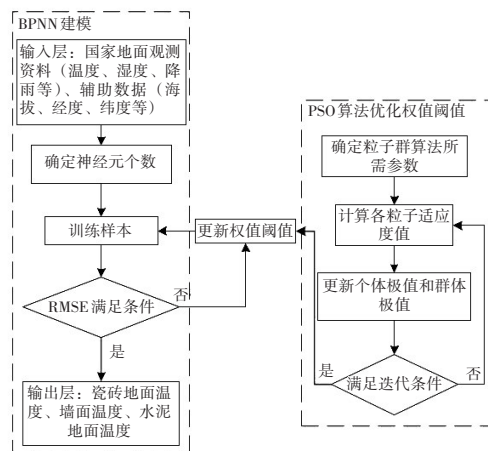


图2 PSO-BPNN“回南天”观测数据质量控制方法流程图

为提高模型精确性,分别搭建瓷砖地面、墙面、水泥地面温度三个 PSO-BPNN 模型。剔除地面观测要素未通过质控的数据或“回南天”观测设备因故障、更换、调试等出现的明显异常数据,2021 年获取有效地面气象观测数据和对应时间瓷砖地面、墙面温度数据共 448 508 组,有效地面气象观测数据和对应水泥地面温度有效数据共 63 632 组。随机选择一半数据(瓷砖地面和墙面 224 254 组,水泥地面

31 816 组)用于构建神经网络模型,剩余数据(瓷砖地面和墙面 224 254 组,水泥地面 31 816 组)用于验证模型(表 1)。此外,广西全区发生“回南天”天气现象常出现于一月到四月<sup>[2]</sup>,为测试模型适用性,选择 2020 年 1 月至 4 月观测数据对模型进行测试应用。训练集、验证集和测试集“回南天”数据温度分布基本一致,主要集中在 10℃到 30℃之间。

表 1 训练集、验证集、测试集所使用数据组的瓷砖地面、墙面、水泥地面温度分布情况表

温度区间(℃)	训练集(组)			验证集(组)			测试集(组)		
	瓷砖地面	墙面	水泥地面	瓷砖地面	墙面	水泥地面	瓷砖地面	墙面	水泥地面
<10	3 885	4 739	87	3 906	4 747	85	3 908	4 777	29
10 ~ 20	85 463	81 918	9 199	85 548	81 967	9 283	121 135	115 717	10 434
20 ~ 30	120 469	116 405	20 871	120 391	116 341	20 896	50 884	55 284	10 663
>30	14 437	21 192	1 659	14 409	21 199	1 552	357	506	615
合计	224 254	224 254	31 816	224 254	224 254	31 816	176 284	176 284	21 741

### 1.3 评估方法

利用相关系数(correlation coefficients,  $R$ ),均方根误差(root mean square error,  $RMSE$ ),偏差(mean bias,  $MB$ )3个统计指标对神经网络模型精度进行评估和误差分析。 $R$ 用于评估 PSO-BPNN 估算温度( $T_{NN}$ )与“回南天”设备观测温度( $T_{OBS}$ )之间的相关性, $RMSE$ 反映  $T_{NN}$ 与真实  $T_{OBS}$ 之间的差异, $MB$ 评估  $T_{NN}$ 是否高估或低估。

$$R = \frac{\sum_{i=1}^n (T_{NN} - \overline{T_{NN}})(T_{OBS} - \overline{T_{OBS}})}{\sqrt{\sum_{i=1}^n (T_{NN} - \overline{T_{NN}})^2 \sum_{i=1}^n (T_{OBS} - \overline{T_{OBS}})^2}} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (T_{NN} - T_{OBS})^2} \quad (2)$$

$$MB = \frac{1}{N} \sum_{i=1}^n (T_{NN} - T_{OBS}) \quad (3)$$

式中  $T_{NN}$ 为神经网络输出的估算温度, $T_{OBS}$ 为“回南天”设备观测温度, $N$ 为样本量。

$$|T_{NN} - T_{OBS}| \leq fS \quad (4)$$

记  $f$  为质量控制参数,用于判断数据是否可疑。若  $T_{OBS}$  满足(4)式认为数据通过检验,否则认为数据可疑, $S$  为  $T_{OBS}$  的标准差。

## 2 结果与分析

### 2.1 验证模型性能评估

分别将 PSO-BPNN 和 BPNN 估算温度与“回南

天”设备观测温度进行对比验证。图 3 显示 PSO-BPNN(图 3(a)—3(c))和 BPNN(图 3(d)—3(f))估算温度与真实观测温度的对比散点图。相比而言,PSO-BPNN 模型比 BPNN 模型具有更高的精度, $R$  值高, $RMSE$  值小。特别是观测值在 10℃附近时,BPNN 模型个别样本偏移较大,最大差值可达 23.25℃。而 PSO-BPNN 模型样本基本分布于 1:1 线附近。

对比三种不同材质(瓷砖地面、瓷砖墙面、水泥地面)模型情况,其中水泥地面模型精度(PSO-BPNN:  $R=0.987$ ,  $RMSE=0.822^\circ\text{C}$ ; BPNN:  $R=0.985$ ,  $RMSE=0.872^\circ\text{C}$ )略高于瓷砖地面(PSO-BPNN:  $R=0.981$ ,  $RMSE=1.121^\circ\text{C}$ ; BPNN:  $R=0.977$ ,  $RMSE=1.220^\circ\text{C}$ )和墙面模型(PSO-BPNN:  $R=0.978$ ,  $RMSE=1.260^\circ\text{C}$ ; BPNN:  $R=0.973$ ,  $RMSE=1.410^\circ\text{C}$ )。三个模型的平均偏差均在  $\pm 0.045^\circ\text{C}$  内,说明估算值相比观测值整体而言没有高估或低估。

PSO-BPNN 和 BPNN 模型估算的瓷砖地面、瓷砖墙面、水泥地面温度值与设备观测值差异均呈现正态分布特征(图 4)。其中 PSO-BPNN 差值在  $0^\circ\text{C}$  附近的峰值明显高于 BPNN,说明 PSO-BPNN 模型偏差较小的点多。瓷砖地面温度、墙面温度的估算值与观测值的差异大部分在  $\pm 3^\circ\text{C}$  内(PSO-BPNN: 98.5%、98.3%; BPNN: 98.3%、96.9%),水泥地面温度估算值与观测值的差异大部分在  $\pm 2^\circ\text{C}$  内(PSO-BPNN:



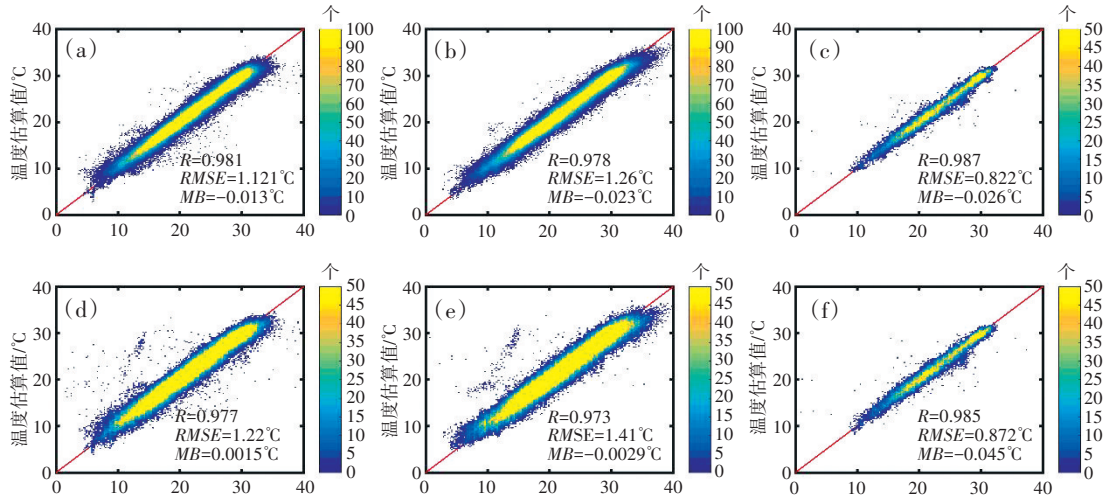


图3 模型估算值与设备观测值对比散点图,红实线为1:1线,色柱表示重叠点个数。

(a)PSO-BPNN估算瓷砖地面温度;(b)PSO-BPNN估算瓷砖墙面温度;(c)PSO-BPNN估算水泥地面温度;  
(d)BPNN估算瓷砖地面温度;(e)BPNN估算瓷砖墙面温度;(f)BPNN估算水泥地面温度

99.0%;BPNN:97.8%)。而 $\pm 1^{\circ}\text{C}$ 范围内,PSO-BPNN模型估算的瓷砖地面、墙面、水泥地面温度误差比例分别为67.2%、63.6%、82.1%,而BPNN分别为65.4%、

56.2%和80.4%。意味着PSO-BPNN在传统BPNN模型的基础上精度有所提高,大多数情况下可提供较为准确的温度估算。

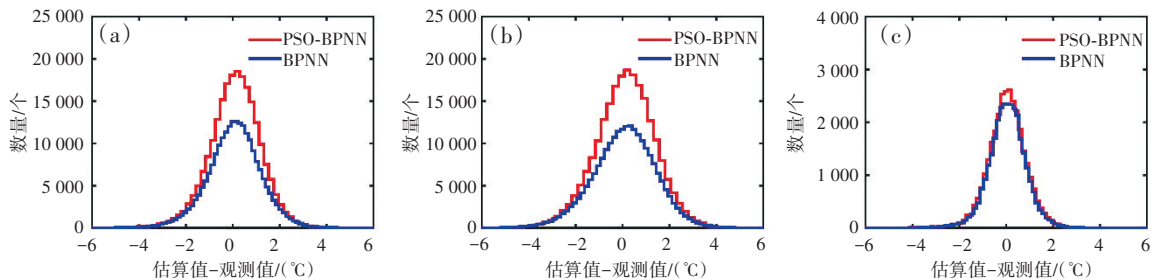


图4 PSO-BPNN和BPNN模型估算值与设备观测值之间的差异

(a)瓷砖地面;(b)瓷砖墙面;(c)水泥地面

## 2.2 测试模型性能评估

为测试PSO-BPNN和BPNN模型在其他数据集上的表现,使用测试集数据对模型进行测试,结果如图5所示。PSO-BPNN模型测试精度整体高于BPNN模型,PSO-BPNN模型估算值与观测值相关系数均大于0.819,瓷砖地面、墙面、水泥地面的 $RMSE$ 分别为 $2.220^{\circ}\text{C}$ 、 $2.018^{\circ}\text{C}$ 、 $1.780^{\circ}\text{C}$ 。而BPNN模型估算值与观测值 $R$ 大于0.741,瓷砖地面、墙面、水泥地面的 $RMSE$ 分别为 $2.630^{\circ}\text{C}$ 、 $2.925^{\circ}\text{C}$ 、 $1.900^{\circ}\text{C}$ 。与模型验证结果相比(图3),两个模型总体精度有所降低,主要是神经网络参数的调整作用于验证数据集上,与验证集更加拟合,当运用到实际数据时,会产生一定程度偏差。特别是当观测温度小于 $10^{\circ}\text{C}$ 或大于 $30^{\circ}\text{C}$ 时,模

型估算的瓷砖地面、墙面温度出现较大程度的偏差,这可能与训练集数据多集中于 $10^{\circ}\text{C}$ 到 $30^{\circ}\text{C}$ 之间有关(表1),如增加观测温度小于 $10^{\circ}\text{C}$ 或大于 $30^{\circ}\text{C}$ 训练样本,可有效提高模型精度,增强模型普适性。

2020年1月21—22日广西、广东多地气温回升,湿度增加,地面返潮,出现“回南天”现象。为测试本研究模型在“回南天”过程中的精度,使用该时间段(48 h)观测数据对模型进行精度评估。广西北海受“回南天”影响较大,图6显示北海站(59644)“回南天”观测数据和PSO-BPNN、BPNN模型的估算温度对比曲线。瓷砖地面温度模型偏向低估(PSO-BPNN  $MB=-1.01^{\circ}\text{C}$ , BPNN  $MB=-1.15$ ),而瓷砖墙面温度模型偏向高估(PSO-BPNN  $MB=0.926^{\circ}\text{C}$ ,

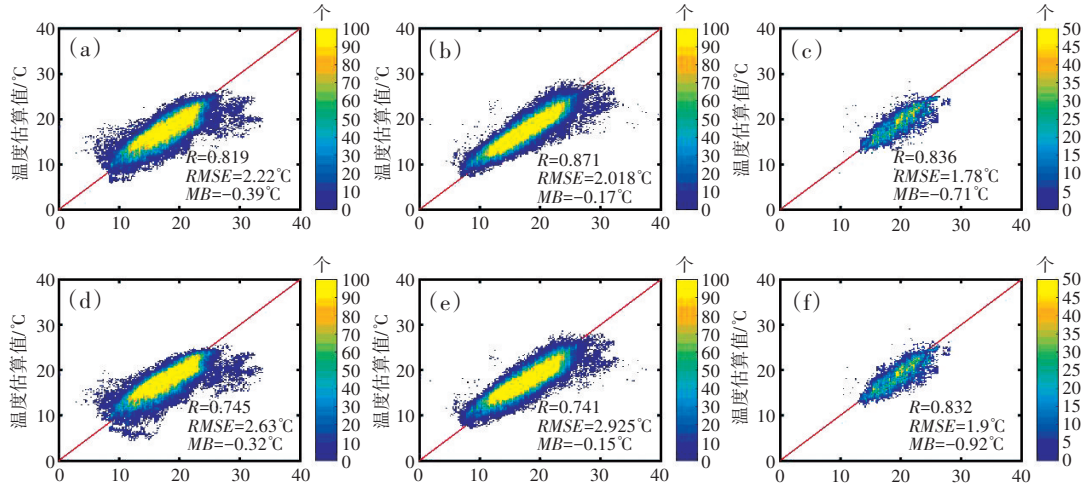


图5 测试集模型估算值与设备观测值对比散点图,红实线为1:1线,色柱表示重叠点个数。

(a)PSO-BPNN 估算瓷砖地面温度;(b)PSO-BPNN 估算瓷砖墙面温度;(b)PSO-BPNN 估算水泥地面温度;  
(d)BPNN 估算瓷砖地面温度;(e)BPNN 估算瓷砖墙面温度;(f)BPNN 估算水泥地面温度

BPNN  $MB=0.679$ ), 水泥地面模型有较小的高估 (PSO-BPNN  $MB=0.578^{\circ}\text{C}$ , BPNN  $MB=0.632$ )。其中 PSO-BPNN 估算的瓷砖地面温度和水泥地面温度偏差小于 BPNN 估算的瓷砖地面温度和水泥地面温

度,墙面温度 BPNN 偏差小于 PSO-BPNN。值得注意的是,三种材质模型中 PSO-BPNN 的稳定性更强,估算温度起伏波动相比 BPNN 较小,更有利于应用于观测数据质量评估判断。

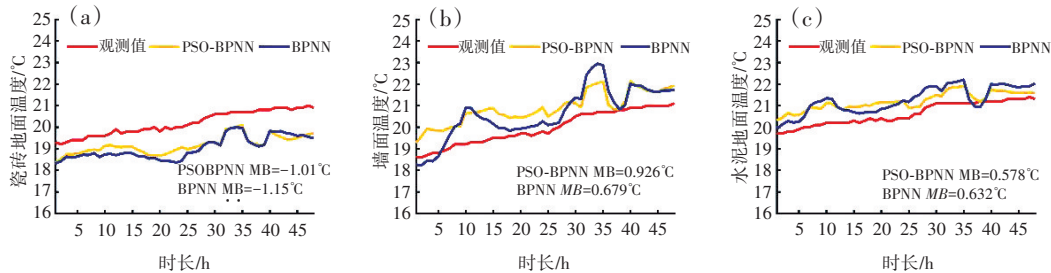


图6 北海站“回南天”过程观测温度和 PSO-BPNN、BPNN 模型的估算温度对比

质量控制中存在两类错误,第一类错误为原本正确的数据标定为错误,第二类错误为原本错误的的数据标定为正确。为达到质量控制的最好效果,在尽可能减少第一类错误发生的同时需保证第二类错误发生概率最小<sup>[20]</sup>。为检验 PSO-BPNN 和 BPNN 模型对“回南天”观测数据质量控制情况,在测试集上随机施加人工误差后对模型进行两类错误测试。图 7 分别为 PSO-BPNN 和 BPNN 模型瓷砖地面、墙面、水泥地面温度在不同质控参数下一二类错误率的变化曲线。PSO-BPNN 模型三种材质的最佳质控参数(红点)范围均在 1.5 到 2 之间,其中瓷砖地面温度模型最佳质控参数相对较高( $f=1.73$ ),对应一二类错误率约 23%,墙面温度模型最佳质控参数最低,值为

1.64,一二类错误率为 22.5%,水泥地面温度模型最佳质控参数为 1.68,对应错误率为 23.3%。BPNN 模型瓷砖地面、墙面、水泥地面温度最佳质控参数高于 PSO-BPNN 模型正负 0.19 内,分别为 1.82、1.83、1.78,对应的错误率也相对偏高,分别为 23.5%、23.7%、24.1%。因此,PSO-BPNN 模型比 BPNN 模型有更小的最佳质控参数和更低的错误率,质控效果更优。

最佳质控参数与气温、湿度、站点海拔等气象要素密切相关,同时模型训练精度、变温范围、季节也会影响其变动<sup>[11, 21]</sup>。为进一步检验模型在不同站点、不同区域的稳定性,对最佳质控参数( $f$ )和检错率进行空间分布分析。总的来说,PSO-BPNN 模型大部分站点的精度要优于 BPNN 模型。PSO-

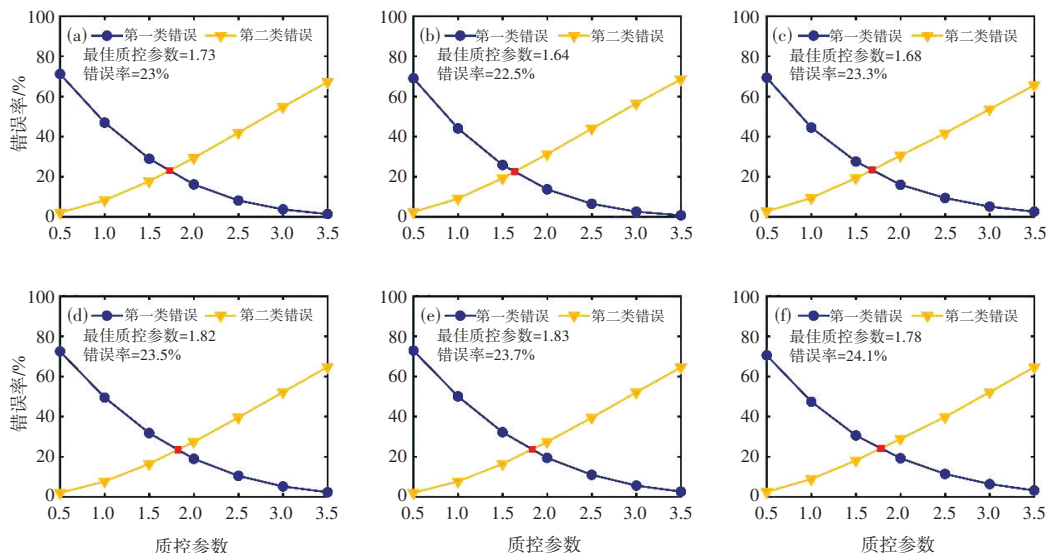


图7 PSO-BPNN模型瓷砖地面(a,d)、瓷砖墙面(b,e)、水泥地面(c,f)在不同质控参数下的两类错误率曲线,红点为最佳质控参数

BPNN 模型最佳质控参数在 1.2 到 2 之间,检错率 75% 以上,表明该方法对不同站点存疑的温度值有较好的识别率。BPNN 模型最佳质控参数在大于 1.4,最大可达 3,大部分站点检错率小于 79%。此外,不同站点最佳质控参数和检错率存在一定起伏,其中桂西北和桂东北的最佳质控参数高于其他区域(PSO-BPNN: $f > 1.4$ ,BPNN: $f > 1.6$ ),且检错率略低于其他区域(PSO-BPNN:检错率 $< 78\%$ ,BPNN:检错率 $< 77\%$ ),这可能与该区域的高海拔有一定关系(图1)。

### 3 结论与讨论

为判别广西全区“回南天”观测设备数据可靠性,本研究首次对该观测设备数据进行质量控制研究。基于传统 BP 神经网络,结合 PSO 算法搭建瓷砖地面、墙面和水泥地面温度估算模型。其中 2021 年一半数据用于模型训练,一半数据用于模型验证,2020 年 1 月到 4 月数据用于模型测试。主要取得如下结论:

(1) PSO-BPNN 模型估算的瓷砖地面、墙面、水泥地面的温度比传统 BPNN 模型估算的瓷砖地面、墙面、水泥地面温度更接近观测值。PSO-BPNN 模型相关性较高,均方根误差值较小,偏差均在  $0^{\circ}\text{C}$  附近,没有明显高估或低估。

(2) 对两种模型进行测试,PSO-BPNN 模型瓷砖地面、墙面、水泥地面温度模型估算值与观测值的  $R$  均 0.819 在以上, $RMSE$  小于  $2.221^{\circ}\text{C}$ ,BPNN 模型  $R$

在 0.741 在以上, $RMSE$  小于  $2.630^{\circ}\text{C}$ ,其中瓷砖地面和墙面  $10 \sim 30^{\circ}\text{C}$  范围适用性更强。此外,PSO-BPNN 模型稳定性优于 BPNN 模型。

(3) 测试集上随机添加人工误差来进行检错率分析,寻找到瓷砖地面、墙面、水泥地面网络的最佳质量控制参数分别为 1.73、1.64、1.68,错误率小于 23.3%,比 BPNN 模型最佳质量控制参数低(1.82、1.83、1.78),错误率小(小于 24.1%),不同站点最佳质控参数和对应检错率不同,表现为桂西北和桂东北最佳质控参数较高,检错率较低。

本研究使用 2021 年一半的数据集用于训练,其中具备水泥地面温度观测站点(9 个)相对较少,获取有效训练样本少,导致模型测试精度降低。同样,瓷砖地面和墙面温度在小于  $10^{\circ}\text{C}$  和大于  $30^{\circ}\text{C}$  的样本量相对较少,这也可能是测试时导致该温度段模型偏差较大的主要原因。后续研究将考虑加入长时间、多温段的数据集进行训练,可加强模型精度,有效应用于“回南天”监控系统中。对于模型测试,将考虑把该方法应用于多个“回南天”真实案例中,并讨论温度、降雨、季节对质控精度的影响。此外,广西壮族自治区气象局 2023 年增添“回南天”视频监控,通过图像识别技术判别物体表面潮湿情况,此后在“回南天”观测数据质量控制中可融入图像识别结果,为“回南天”过程提供更加有效数据支撑。

#### 参考文献:

[1] 罗小莉,古明悦,钟利华,等.近年广西典型“回南天”

- 现象成因分析[J]. 气象科技, 2015, 43(4):659-665, 674.
- [2] 陈见, 李佳颖, 高安宁, 等. 广西“回南天”发生特征及预报着眼点[J]. 气象, 2015, 41(3): 372-379.
- [3] 黄剑钊, 李艳萍, 陶伟, 等. 广西回南天观测系统装备与数据分析[J]. 气象研究与应用, 2020, 41(2):89-92.
- [4] 张新甲, 甘晓英, 刘艳辉, 等. “回南天”自动监测仪的研发与应用前景[C]//中国气象学会. 第28届中国气象学会年会——S1第四届气象综合探测技术研讨会. 台山市气象局, 2011:5.
- [5] JIMENEZ P A, GONZALEZ R J F, NAVARRO J, et al. Quality assurance of surface wind observations from automated weather stations [J]. *Journal of Atmospheric and Oceanic Technology*, 2010, 27(7): 1101-1122.
- [6] QI Y C, MARTINAITIS S, ZHANG J, et al. A real-time automated quality control of hourly rain gauge data based on multiple sensors in MRMS system [J]. *Journal of Hydrometeorology*, 2016, 17(6): 1675-1691.
- [7] BEGES G, DRNOVSEK J, BOJKOVSKI J, et al. Automatic weather stations and the quality function deployment method [J]. *Meteorological Applications*, 2015(22): 861-866.
- [8] RIENZNER M, GANDOLFI C. A procedure for the detection of undocumented multiple abrupt changes in the mean value of daily temperature time series of a regional network [J]. *International journal of climatology*, 2013, 33(5): 1107-1120.
- [9] 闵锦忠, 王晨珏, 贾瑞怡. 苏皖地面自动站资料的质量控制及结果分析 [J]. *大气科学学报*, 2018, 41(5): 637-646.
- [10] HUDDARD K G, YOU J. Sensitivity analysis of quality assurance using the spatial regression approach—A case study of the maximum/minimum air temperature [J]. *Journal of Atmospheric and Oceanic Technology*, 2005, 22(10): 1520-1530.
- [11] 殷利平, 刘宵瑜, 盛绍学, 等. 基于粒子群优化-BP神经网络-马尔科夫链的地面能见度观测资料质量控制 [J]. *科学技术与工程*, 2022 (13): 5125-5133.
- [12] SHA Y K, GAGNE L, DAVID J, et al. Deep-learning-based precipitation observation quality control [J]. *Journal of Atmospheric and Oceanic Technology*, 2021, 38(5): 1075-1091.
- [13] YE X L, YANG X, XIONG X, et al. A quality control method based on an improved random forest algorithm for surface air temperature observations [J]. *Advances in Meteorology*, 2017 :1-15.
- [14] 闵晶晶, 孙景荣, 刘还珠, 等. 一种改进的BP算法及在降水预报中的应用 [J]. *应用气象学报*, 2010, 21(1): 55-62.
- [15] 张志富, 任芝花, 张强, 等. 自动站小时气温数据质量控制系统研究 [J]. *气象与环境学报*, 2013, 29(4): 64-70.
- [16] 杨娜, 刘良明, 向大享, 等. 利用 BP 神经网络由特征气象要素预测土壤湿度 [J]. *土壤通报*, 2011, 42(6): 1324-1329.
- [17] 张天虎, 鲍艳松, 钱芝颖, 等. 基于 BP 神经网络与遗传算法反演大气温湿廓线 [J]. *热带气象学报*, 2020, 36(1): 97-107.
- [18] COUCEIRO M, SIVASUNDARAM S. Novel fractional order particle swarm optimization [J]. *Applied Mathematics & Computation*, 2016(283): 36-54.
- [19] 韩力群. 人工神经网络教程[M]. 北京: 北京邮电大学出版社, 2006:1-330.
- [20] BAYRAM S, GEZICI S. Stochastic resonance in binary composite hypothesis-testing problems in the Neyman-Pearson framework [J]. *Digital Signal Processing*, 2012, 22(3): 391-406.
- [21] 叶小岭, 沈云培, 熊雄. 一种基于改进克里金法的地面气温质量控制算法 [J]. *气候与环境研究*, 2016, 21(5): 614-620.



## Preliminary study on the quality control method for observation data of “Continuous Wet Weather” based on neural network

WANG Yizhu<sup>1</sup>, TAO Wei<sup>1\*</sup>, LU Siyu<sup>2</sup>

(1.Guangxi Meteorological Technical Equipment Center, Nanning 530022, China;

2.Guangxi Institute of Meteorological Sciences, Nanning 530022, China)

**Abstract:** In order to determine the reliability of the observation data of “Continuous Wet Weather”, a quality control study on the observation data of “Continuous Wet Weather” in Guangxi was carried out based on the traditional back-propagation neural network (BPNN), combined with particle swarm optimization (PSO) algorithm (i. e. PSO-BPNN). The results show that: (1) compared with the traditional BPNN model, the accuracy of the PSO-BPNN model is higher in comparing the model-estimated temperature with the measured temperature, without any significant overestimation or underestimation in the PSO-BPNN model, while the BPNN model shows a large deviation around 10 °C. (2) In the tests of PSO-BPNN and BPNN model, tile floor and wall temperatures in the range of 10 ~ 30 °C show greater applicability of the models, and the PSO-BPNN model is more stable than the BPNN model. (3) Randomly adding artificial errors for model validation, the optimal quality control parameters for the temperatures of tile ground, wall, and cement ground in the PSO-BPNN model are 1.73, 1.64, and 1.68, respectively, and 1.82, 1.83, and 1.78 for the BPNN model, respectively.

**Key words:** quality control; back propagation neural network; particle swarm optimization; Continuous Wet Weather