

文章编号: 1671-6612 (2023) 02-225-06

# 基于 PCA-BP 神经网络的智慧建筑暖通空调能耗预测

社会军

(中铁建设集团华北工程有限公司 天津 300000)

**【摘要】** 智慧暖通空调受到通风、温度和湿度高维数据影响, 存在数据高维、冗余、异常问题, 影响能耗预测精度。为了获取精准预测结果, 结合 PCA-BP 神经网络方法预测智慧建筑暖通空调能耗。采用 PCA (principal components analysis) 方法对通风、温度和湿度能耗数据降维处理, 将三维特征样本点映射到二维平面, 通过特征归一化处理, 获取保留原特征低维能耗数据。加权平均处理能源数据, 解决数据缺失、冗余等问题。对 PCA 降维矩阵做归一化处理, 查找出异常数据, 通过线性标量假设, 去除其贡献率低的数据。利用 BP 神经网络 (Back propagation neural network) 非线性逼近特性, 平滑处理能耗数据, 取代异常点数据。构建能耗预测目标函数, 避免 BP 神经网络陷入局部最优解, 完成 BP 神经网络的动态预测。由实验结果可知, 该方法制冷能耗波动范围为 500-3750kW/月, 制热能耗波动范围为 650-3450kW/月, 与实际数据一致, 说明使用该方法具有精准预测结果。

**【关键词】** PCA 主成分分析; BP 神经网络; 智慧建筑; 暖通空调; 能耗预测  
中图分类号 TU831 文献标识码 A

## Prediction of HVAC Energy Consumption in Smart Buildings Based on PCA-BP Neural Network

Du Huijun

(China Railway Construction Group North China Engineering Co., Tianjin, 300000)

**【Abstract】** Intelligent HVAC is affected by high-dimensional data of ventilation, temperature and humidity, and there are high-dimensional, redundant and abnormal data problems, which affect the accuracy of energy consumption prediction. In order to obtain accurate prediction results, the PCA-BP neural network method is used to predict the energy consumption of HVAC in smart buildings. The PCA (principal components analysis) method is used to reduce the dimension of ventilation, temperature and humidity energy consumption data, map the three-dimensional feature sample points to the two-dimensional plane, and obtain the low dimensional energy consumption data that retains the original features through feature normalization. Weighted average processing of energy data to solve problems such as data loss and redundancy. Normalize PCA dimension reduction matrix, find out abnormal data, and remove the data with low contribution rate through linear scalar assumption. Using the nonlinear approximation characteristics of BP(Back propagation neural network) neural network, the energy consumption data is smoothed to replace the abnormal point data. The objective function of energy consumption prediction is constructed to avoid BP neural network falling into local optimal solution and complete the dynamic prediction of BP neural network. It can be seen from the experimental results that the fluctuation range of cooling energy consumption of this method is 500-3750kW/month, and the fluctuation range of heating energy consumption is 650-3450kW/month, which is consistent with the actual data, indicating that this method has accurate

作者 (通讯作者) 简介: 社会军 (1981.05-), 男, 硕士, 高级工程师, E-mail: quanshou0475021686@163.com  
收稿日期: 2022-12-16

prediction results.

**【Keywords】** PCA principal component analysis; BP neural network; Smart buildings; HVAC; Energy consumption prediction

## 0 引言

暖通空调在提高室内空气品质、提高室内居住舒适性等诸多功能的同时，也产生了大量能耗问题，制约我国生态环境稳定发展。其中一个重要的问题在于，常规的暖通空调和监测设备不能及时调整最佳的送风量，以适应不断改变的室内环境的各种参数，从而造成了空调的能源消耗过剩。另外，“非均匀非正常”的各种因素也与暖通设施的安全风险和防治有着很大关系，比如在火灾中，烟雾等的扩散速率可以高达 3m/s，因此有必要开发一种新型的监测技术。路凯文等人结合实验研究，证明了利用暂态区模式进行的室内热状况预报是可行的，其结果表明：大气热层化和暂态壁温的计算是正确的。但是，多区模式的建立都是以室内空气的均匀性为基础，而在快速水力计算中，忽视了水流的重要影响。结果表明，这些方法的预报结果仍然有很大的偏差，不能达到工程中对预报的准确性要求<sup>[1]</sup>；曾剑锐等人利用 OpenFOAM 软件建立了一种新的计算模型，并通过实验证明了其在房间中的流动特性。但是，由于传感器的数量不多，不能很好地描述整体/非均匀分布（或内部人群）的环境参数，因此不能保证对最好的空气质量进行评价。如果进行大规模的监控，代价高昂，而且存在信息重复情况<sup>[2]</sup>。本文为优化空调控制策略，达到节能减排目的，提出了基于 PCA-BP 神经网络的智慧建筑暖通空调能耗预测方法。

## 1 智慧建筑暖通空调能耗数据 PCA 降维处理

PCA 是目前应用最为普遍的一种降维算法，PCA 的基本思路就是把多个维度的特性映射到一个维度上，这个维度是一个新的正交特性，也就是所谓的主分量，它是在原来多个维度上再构建一个维度的特性<sup>[3]</sup>。PCA 的任务是从原来的空间中依次找到一套互相垂直的轴系，而这些新的轴线的选取与数据自身有着紧密联系。

降维的线性建模的主要作用是对基于 PCA 的初步预报数据库进行快速地扩展和数据的降维，从而大大减少了系统建设费用。此方法分为两种：一

种是降维模式，另一种是直线模式<sup>[4]</sup>。在此基础上，采用体积平均法对数据进行了压缩，达到了数据精确性要求（例如，在实际应用中，该算法的要求不超过 10%）。图 1 显示了降维的基本理论，它采用了基于容积均值的离散和数据的降维重建方式。

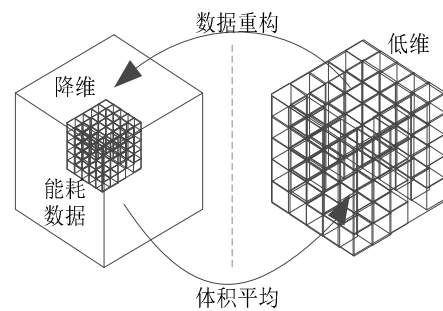


图 1 能耗数据 PCA 降维模型原理图

Fig.1 Schematic diagram of the PCA dimension reduction model for energy consumption data

由图 1 可知，线性（标量）模式分为线性通风模式、温度模式和湿度模式，它可以将各种不同的环境指标（例如浓度、温度、湿度）线性叠加，可以迅速地获得大量资源（如污染源、热源、湿度源）的影响，从而达到数据库扩展的目的<sup>[5,6]</sup>。

假设智慧建筑暖通空调能耗数据中存在  $m$  个样本和  $n$  个特征，可表示为：

$$\begin{cases} \mathbf{x}^{(1)}, \mathbf{x}^{(2)} \dots \mathbf{x}^{(m)} \\ \mathbf{x}^{(i)} = [\mathbf{x}_1^{(i)}, \mathbf{x}_1^{(i)} \dots \mathbf{x}_n^{(i)}]^T \end{cases} \quad (1)$$

由于在数据降维时会产生映射错误，因此需先对特征数据进行归一化处理，这样才能达到对能耗数据降维的目的。在整个过程中，应保证每个特性都具有相同的比例，否则会出现特征过大或特征过小的问题，影响整个降维效果<sup>[7]</sup>。从三维向二维的降维，可以理解为找到两个向量  $\bar{x}_1$ 、 $\bar{x}_2$ ，通过这两个向量分别构造出一种新的特征，并将其与原有的采样点相对应<sup>[8,9]</sup>。将三个特征的采样点投影到二维平面上会出现一些错误，因此必须对每个图像进行精确地测量。错误幅度的计算公式，如下所示：

$$e = \frac{1}{m} \sum_{n=1}^m \left\| \mathbf{x}_n^{(m)} - \mathbf{x}_a^{(m)} \right\|^2 \quad (2)$$

公式 (2) 中,  $x_a^{(n)}$  表示映射后坐标。若用目前取样空间来表示该坐标, 则该降维矩阵的尺寸与取样值相符, 基于此构建降维矩阵:

$$x = \begin{bmatrix} (x^{(1)})^T \\ \vdots \\ (x^{(m)})^T \end{bmatrix} = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \cdots & x_n^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ x_1^{(m)} & x_2^{(m)} & \cdots & x_n^{(m)} \end{bmatrix} \quad (3)$$

在获取降维效果之后, 用降维矩阵将该模型映射到低维空间上<sup>[10]</sup>。降维公式如下所示:

$$x' = (n \times c)^T \times (n \times 1) \quad (4)$$

公式 (4) 中,  $c$  表示数据降低的维度。 $c$  值越大, 低维空间上特征矢量也就越多, 降维误差也就越小, 保留原来特征特性。

## 2 基于 PCA-BP 神经网络的能耗预测

采用 PCA 和 BP 神经网络相结合的方法, 将空调节能控制中的初始数据进行了精简, 剔除了对其影响较小的部分, 同时也充分发挥了 BP 神经网络的非线性及平行运算的特点, 进一步改进了 BP 神经网络的算法架构, 从而使其具有较好的学习性能和较好稳定性。对于基于 PCA-BP 神经网络的能耗预测, 首先, 将空调节能控制信息进行分析, 利用 PCA 对其进行预处理, 得出累积贡献度, 然后由最开始向后依次抽取出小于特征总量的主成分。将 PCA 所得的主成分数据引入 BP 神经网络的输入级, 初始化 BP 神经网络权值, 获取精准预测结果。

### 2.1 能耗数据加权归一化处理

由于不同负载比例下能耗数据存在数据缺失、冗余等问题, 对能源消耗的预测效果产生较大不利影响。针对有问题的能源消耗数据, 采取有目标的方法进行处理<sup>[11]</sup>。能源数据加权平均处理公式, 可表示为:

$$\lambda_x(d, t) = \omega_1 \lambda_x(d, t+1) + \omega_2 \lambda_x(d, t-1) + \cdots + \omega_{m-1} \lambda_x(d, t+1) + \omega_m \lambda_x(d, t-1) \quad (5)$$

公式 (5) 中,  $w$  表示加权权重;  $d$  表示日期;  $t$  表示时间, 利用此方程对能耗数据进行处理填补。填补完毕后, 将 PCA 降维模型做标准化处理, 获得的数据参量, 获得的所有参量再进行训练采

样, 将训练采样结果与所获得的参量分别进行归一化处理, 由此查找出异常数据<sup>[12]</sup>。计算公式如下:

$$\hat{x}(n) = \frac{x(n) - x_{\min}}{x_{\max} - x_{\min}} \quad (6)$$

公式 (6) 中,  $x_{\max}$ 、 $x_{\min}$  分别表示待处理能耗数据最大和最小值。将能耗变化率上限 5% 作为临界点, 如果能耗预测测量结果与实际测量结果的偏差大于设定的临界点, 则表明所预测到的能量消耗数据是异常数据。

### 2.2 累计贡献率计算

根据线性标量假定 (参照以上的直线模式原则), 可以将内部环境贡献系数解释为各个来源 (例如: 房间内的热能和湿气来源) 对诸如温度、湿度等的整体环境参数的影响因素<sup>[13]</sup>。利用热源或湿气来源的影响因素 (建立在降维线性建模基础上的资料库), 再加上源能强度的改变, 可以有效实现“超实时”地预测温度和湿度。在热、湿两个方面, 其房间环境贡献系数计算公式为:

$$R_s(w_i) = \frac{\Delta x_s(x_a^{(i)})}{\Delta x_{s,0}} = \frac{x_s(x_a^{(i)}) - x_n}{x_{s,0} - x_n} \quad (7)$$

公式 (7) 中,  $x_s$  表示湿度、温度源  $s$  的能耗值;  $\Delta x_{s,0}$  表示湿度、温度源  $s$  的湿度和热量发生变化后扩散到室内平均能耗<sup>[14]</sup>。利用室内环境贡献系数来快速地实现对温度和湿度能耗的快速预测, 剔除其所占比例小的部分, 利用 PCA 特征值来描述高维数据, 尽量保留“距离”, 减少各种运算过程中所需运算量, 提高了数据提取效率。

### 2.3 基于 BP 神经网络能耗预测及误差修正

在能耗数据中会出现与空调系统历史工作流程不一致的情况, 在相同异常点, 由于负载比例作用, 使数据噪音加剧, 增加了能耗预测难度。因此, 利用历史数据变异规则, 对离群值进行垂直处理, 并对其进行校正。假定预测空调系统能耗时, 在计算偏差值大于容许能量偏差值情况下, 以能量消耗数据为“非正常值”, 而在偏差值低于容许能量偏差值情况下, 针对非正常能量消耗, 调整能量偏差值, 以校正偏差值<sup>[15]</sup>。若能量消耗值为不正常值, 则以下列公式代替:

$$\ddot{x}(d, t) = \frac{\dot{x}(d+1, t) + \dot{x}(d, t-1)}{2} \quad (8)$$

公式 (8) 中,  $\dot{x}(d,t)$  表示异常数据。通过该公式能够对能耗连续过程进行一定平滑处理。通过对空调系统动态工况分析, 得出了各时段的能耗变化规律, 假设数据最大变动范围为  $f(t)$ , 可表示为:

$$|\dot{x}(d,t) - \dot{x}(d,t-1)| > f(t) \quad (9)$$

将能耗数据作为不良数据时, 利用 (9) 式进行了消减, 其计算公式为:

$$\dot{x}(d,t) = \frac{\dot{x}(d,t+1) + \dot{x}(d,t-1)}{2} \quad (10)$$

通过上述过程处理后, 得到了能源消耗的全面、可信数据。采用 BP 神经网络对空调系统能耗进行了预测。基于此, 构建的能耗预测目标函数可描述为:

$$y(x) = Z([y_t, v_t]) \quad (11)$$

公式 (11) 中,  $Z(\cdot)$  表示非线性函数;  $y_t$  表示 BP 神经网络  $t$  时刻输出值;  $v_t$  表示  $t$  时刻输入值。四层网络的结论与三层网络相比更容易陷入局部最小点, 同时网络中的网络数量越多, 网络泛化性能越差, 网络预测性能也越差, 因此只要有足够多的网络节点, 就能在任意地点逼近非线性函数, 逼近公式可表示为:

$$y'(x) = Z'_i([y_t, v_t]^T) \quad (12)$$

公式 (12) 中,  $Z'_i$  表示经过  $i$  次逼近的 BP 神经网络。BP 神经网络从输入级节点开始, 由输入级节点通过各个隐含层节点, 再由各个节点向外节点传递, 每个节点的输出量都会受到节点的影响。在 BP 神经网络中, 利用动态因素降低了神经网络的振动倾向, 提高了 BP 的收敛率, 从而得到了更优的解, 避免陷入局部最优解。训练 BP 神经网络, 使得  $y'(x)$  逼近  $y_t$ , 完成 BP 神经网络的动态预测。

### 3 实验研究

为了验证基于 PCA-BP 神经网络的智慧建筑暖通空调能耗预测方法的合理性, 将其与实际数据、Block-Gebhart 模型、OpenFOAM 预测方法进行对比。

#### 3.1 实验标准

在制冷、制热能耗预测实验过程中, 对一年

12 个月空调系统制冷、制热进行计算, 制冷、制热原理如图 2 所示。

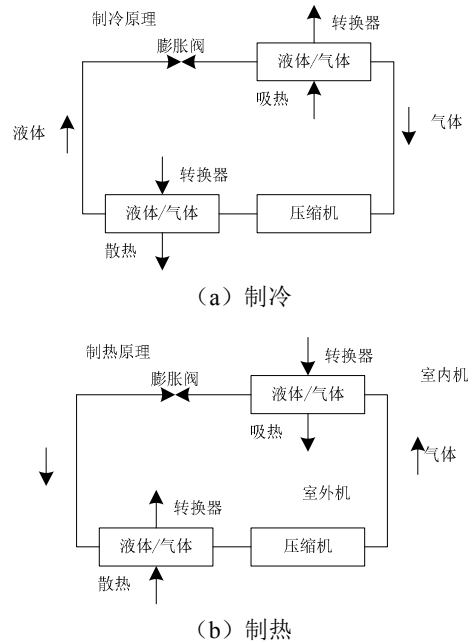


图 2 制冷、制热原理

Fig.2 Principle of refrigeration and heating

分析制冷、制热能耗与实际数值是否一致, 制冷、制热量计算公式为:

$$Q_{\text{制冷}} = \frac{C\rho V\Delta T_{\text{制冷}}}{12} \quad (13)$$

$$Q_{\text{制热}} = \frac{C\rho V\Delta T_{\text{制热}}}{12} \quad (14)$$

公式 (13)、(14) 中,  $C$  表示空调制冷或制热时水的比热容;  $\rho$  表示空调制冷或制热时水密度;  $V$  表示空调制冷或制热水体积流量;  $\Delta T_{\text{制冷}}$ 、 $\Delta T_{\text{制热}}$  分别表示制冷、制热进出水温度差。

#### 3.2 测试点布置

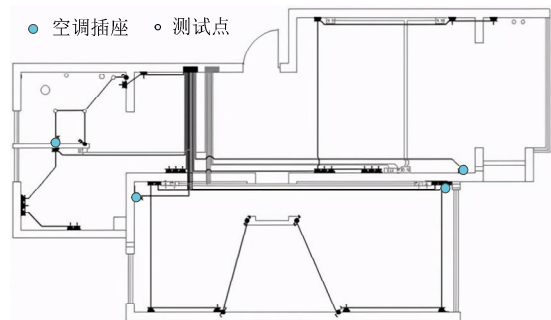


图 3 实验测试点布置

Fig.3 Arrangement of experimental test points

在智慧建筑暖通空调能耗预测过程中，以某写字楼建筑 1 匹空调为实验对象，依据空调系统实际运行能效，从空调系统制冷、制热量方面进行能耗分析。在无特殊情况下，设置的测试点如图 3 所示。

由图 3 可知，在该测试点布置下，获取的实际数据如表 1 所示。

表 1 实际制冷、制热能耗 kW/月

**Table 1 Actual cooling and heating energy consumption**

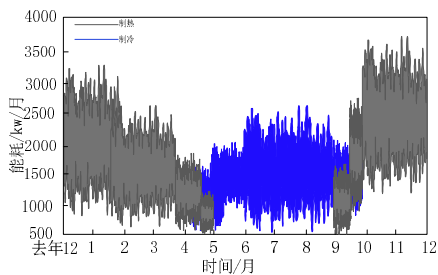
**kW/month**

| 时间/月    | 制冷       | 制热        |
|---------|----------|-----------|
| 去年 12-1 | 0        | 650-3000  |
| 1-2     | 0        | 650-3000  |
| 2-3     | 0        | 650-3000  |
| 3-4     | 0        | 1200-1850 |
| 4-5     | 500-1700 | 500-1500  |
| 5-6     | 900-2650 | 500-1250  |
| 6-7     | 500-3750 | 0         |
| 7-8     | 500-3750 | 0         |
| 8-9     | 500-3500 | 0         |
| 9-10    | 900-2350 | 500-1900  |
| 10-11   | 0        | 650-3450  |
| 11-12   | 0        | 650-3450  |

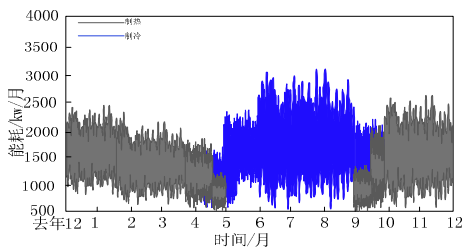
预测空调系统制冷、制热能耗，并将其与表 1 数据对比，分析能耗预测结果实际性能。

### 3.3 实验结果与分析

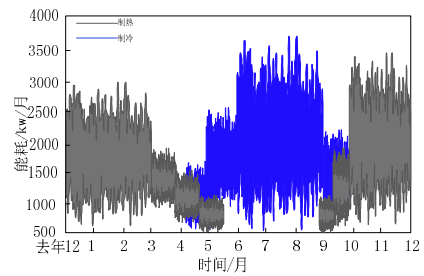
分别使用 Block-Gebhart 模型、OpenFOAM 预测方法与基于 PCA-BP 神经网络预测方法对比分析制冷、制热能耗预测结果，如图 4 所示。



(a) Block-Gebhart 模型



(b) OpenFOAM 预测方法



(c) 基于 PCA-BP 神经网络预测方法

图 4 不同方法能耗预测结果对比分析

**Fig.4 Comparative analysis of the energy consumption prediction results of different methods**

由图 4 (a) 可知，制冷时间主要是 4 月份到 10 月份，能耗波动范围是 500-2550kW/月；制热时间主要是去年 12 月到今年 5 月份，9 月份到 12 月份，能耗波动范围分别是 500-3400kW/月、500-3750kW/月。

由图 4 (b) 可知，制冷时间主要是 4 月份到 10 月份，能耗波动范围是 500-3150kW/月；制热时间主要是去年 12 月到今年 5 月份，9 月份到 12 月份，能耗波动范围分别是 500-2500kW/月、500-2550kW/月。

由图 4 (c) 可知，制冷主要是在 6-9 月份，能耗波动范围为 500-3750kW/月。在 4-6 月份、9-10 月份，能耗波动范围分别为 500-2650kW/月、900-2350kW/月；制热主要是在去年 12 月到今年 3 月份、10-12 月份，能耗波动范围分别为 650-3000kW/月、650-3450kW/月，其余月份能耗较小。

通过上述分析结果可知，使用 Block-Gebhart 模型、OpenFOAM 预测方法与表 1 数据不一致，而使用基于 PCA-BP 神经网络预测方法与表 1 数据一致，说明使用该方法具有精准预测结果。

## 4 结束语

针对目前我国智慧建筑暖通空调能耗占比高但能耗管理不佳的问题，提出了基于 PCA-BP 神经网络的智慧建筑暖通空调能耗预测方法，并得出如下结论：

(1) 利用 PCA 主成分分析法对通风、温度、湿度等能量进行降维处理，将三维特征样点投影到二维平面上，并对其进行特征标准化，得到了原始特征的低维能量消耗数据。

(2) 对能量数据进行加权平均, 解决数据缺失、冗余等问题。

(3) 采用线性标量假定, 对 PCA 的降维矩阵进行规格化, 并去除贡献率低的数据。

(4) 采用 BP 神经网络非线性并行处理能力, 将能量序列中的异常点数据进行平滑处理。根据 BP 神经网络的近似特征, 构造了能耗预测目标函数, 从而提高了网络的收敛性。

#### 参考文献:

- [1] 路凯文, 黄晨, 李瑞彬, 等. Block-Gebhart 模型分区数对预测大空间热环境参数及负荷的影响[J]. 暖通空调, 2021, 51(3): 101-107.
- [2] 曾剑锐, 张尧立, 徐宏, 等. 基于 OpenFOAM 的超临界 CO<sub>2</sub> 流动与传热求解器的开发与验证[J]. 厦门大学学报(自然科学版), 2021, 60(1): 44-49.
- [3] 刘伟, 李怀, 黄巍, 等. 基于 TRNSYS 模拟的某近零能耗办公楼暖通空调系统优化配置分析[J]. 建筑科学, 2022, 38(4): 158-168.
- [4] 李玲, 齐兵, 李松. 基于 DeST 软件模拟的超低能耗建筑暖通空调系统适用性分析[J]. 暖通空调, 2022, 52(S1): 307-311.
- [5] 曹世杰, 任宸, 朱浩程. 基于有限监测与降维线性模型耦合预测的暖通空调系统在线监控方法与策略[J]. 建筑科学, 2021, 37(4): 83-92.
- [6] 刘新宇, 张晋, 季海波, 等. 有向图上暖通空调系统能耗博弈的分布式纳什均衡寻求设计(英文)[J]. 中国科学技术大学学报, 2022, 52(1): 39-46, 72.
- [7] 李卓阳, 靳小龙, 贾宏杰, 等. 考虑建筑物热动态特性的暖通空调模型预测控制方法[J]. 中国电机工程学报, 2020, 40(12): 3928-3940.
- [8] 吴剑林, 李怀, 于震, 等. 某近零能耗办公楼暖通空调系统优化运行分析[J]. 建筑科学, 2020, 36(6): 35-41.
- [9] 于震, 李立, 李怀, 等. 基于群智能的暖通空调输配系统拓扑描述方法和关键参数辨识[J]. 建筑科学, 2022, 38(2): 65-70, 81.
- [10] 董云霞. 建筑暖通空调设计中的节能减排技术研究[J]. 工业建筑, 2022, 52(4): 256.
- [11] 张诚. 变风量系统分区设计对某超高层办公建筑空调能耗的影响[J]. 暖通空调, 2020, 50(2): 46-52.
- [12] 李冠男, 姚庆, 毛前军, 等. 室温传感器偏差故障对武汉地区暖通空调系统的影响研究[J]. 可再生能源, 2022, 40(3): 332-337.
- [13] 谢谢, 周旭. 基于运行数据的暖通空调系统调适技术分析与实践[J]. 建筑技术, 2020, 51(9): 1098-1103.
- [14] 刘魁星, 兰博, 王进宇, 等. 暖通空调控制策略自动预检验方法[J]. 建筑科学, 2020, 36(12): 31-37, 125.
- [15] 柯尊友, 辛亚娟, 郑毅然, 等. 天津国家会展中心暖通空调系统冬季节能与防冻[J]. 建筑科学, 2020, 36(9): 92-98.
- [5] 彭光前, 王现林, 王喜成, 等. 空气源热泵除霜方式比较分析[J]. 制冷, 2021, 40(2): 34-37.
- [6] 孙家正. 空气源热泵除霜方法的研究现状及展望[J]. 建筑热能通风空调, 2017, 36(8): 42-46.
- [7] Zhang L, Dong J, Jiang Y, et al. A novel defrosting method using heat energy dissipated by the compressor of an air source heat pump[J]. Applied Energy, 2014, 133: 101-111.
- [8] 李延贺, 臧润清. 空气源热泵除霜方法的研究与发展[J]. 制冷技术, 2018, 46(9): 82-86.
- [9] 陈健勇, 李浩, 陈颖, 等. 空气源热泵空调技术应用现状及发展前景[J]. 华电技术, 2021, 43(11): 25-39.
- [10] 黄皓. 电子膨胀阀节流噪声数值模拟[D]. 杭州: 浙江理工大学, 2015.

(上接第 193 页)