

DOI: 10.11835/j.issn.2096-6717.2021.242



开放科学(资源服务)标识码(OSID):



# 基于多维度聚类算法的重庆住宅空调使用特征分析

薛凯, 刘猛, 晏璐, 何昱洁

(重庆大学 土木工程学院; 国家级低碳绿色建筑国际联合研究中心; 绿色建筑与人居环境营造  
国际合作联合实验室; 风工程及风资源利用重庆市重点实验室, 重庆 400045)

**摘要:**长江流域夏季炎热、冬季阴冷, 全年高湿, 室内热环境恶劣, 多样化的空调使用习惯对住宅供暖空调能耗有重要影响。大数据技术发展为更大样本、更高精度、更多维度的空调行为监测提供了基础, 弥补了现有研究方法误差大和分类指标单一的不足。选取重庆市作为长江流域典型城市的代表, 随机抽取2 000台住宅房间空调器样本, 从空调使用时长、温度需求及能耗角度, 构建空调运行的5个特征参数, 采用多维度聚类算法识别出重庆地区空调使用习惯的典型类别, 通过深入分析不同使用习惯类别的特征差异, 总结出三类典型群体。

**关键词:**聚类算法; 数据挖掘; 使用习惯; 房间空调器

中图分类号: TU831.2 文献标志码: A 文章编号: 2096-6717(2022)04-0167-09

## Characteristics of occupants' behavior in Chongqing residential air-conditioning based on multi-dimensional clustering algorithm

XUE Kai, LIU Meng, YAN Lu, HE Yujie

(School of Civil Engineering; National Centre for International Research of Low-Carbon and Green Building; Joint International Research Laboratory of Green Building and Built Environment; Chongqing Key Laboratory of Wind Engineering and Wind Energy Utilization, Chongqing University, Chongqing 400045, P. R. China)

**Abstract:** With a hot summer, cold winter and high humidity climate, residential energy consumption in the Yangtze River Basin is strongly affected by diverse air-conditioning behaviors in such a harsh indoor thermal environment. The development of big data technology provides a basis for larger samples, higher accuracy, and more dimensions of air-conditioning behavior monitoring, which can make up for the current situation of large errors in existing research methods and single classification indicators. By selecting 2 000 samples of residential room air conditioners (RACs) in Chongqing as the representative city: First, five characteristic parameters of air-conditioning operation are constructed from the perspective of air-conditioning using period, temperature demand and energy consumption; Then, a multi-dimensional clustering algorithm was used to identify the typical categories of air-conditioning behavior; Finally, through in-depth analysis of the characteristic differences among the clustering results, three typical air-conditioning

收稿日期: 2021-08-11

基金项目: 国家重点研发计划(2018YFD1100704)

作者简介: 薛凯(1990-), 男, 博士生, 主要从事室内环境与大数据应用研究, E-mail: xuek1213@126.com。

刘猛(通信作者), 男, 教授, 博士生导师, E-mail: liumeng2033@126.com。

Received: 2021-08-11

Foundation item: National Key Research and Development Program of China (No. 2018YFD1100704)

Author brief: XUE Kai (1990-), PhD candidate, main research interests: big data and indoor thermal environment, E-mail: xuek1213@126.com.

LIU Meng (corresponding author), professor, doctoral supervisor, E-mail: liumeng2033@126.com.

behavior groups are summarized for residential buildings in Chongqing.

**Keywords:** cluster algorithm; data mining; occupants' behavior; room air-conditioner (RACs)

长江流域地区占中国国土面积的 25%，人口占 55%，GDP 占近 45%，是一个人口密集，经济、产业高度发达的地区。该地区夏季炎热、冬季阴冷，全年高湿，室内热环境恶劣。近年来，多数家庭将房间空调器作为改善室内热环境的重要措施，根据国家统计局数据，目前中国住宅房间空调器拥有量达 3.6 亿台<sup>[1]</sup>，其中重庆地区已接近户均 2 台。龙惟定等<sup>[2]</sup>指出，随着每年 34 亿 m<sup>2</sup> 的城镇住宅面积的增长，维持住宅建筑环境的空调能耗仍会有明显上升，因此，控制房间空调器电耗是减少建筑间接碳排放的主要措施，对尽快实现“双碳”目标也有积极意义。近年来的研究表明，住宅空调使用模式是影响供暖空调能耗的主要因素<sup>[3-4]</sup>；中国经济发展水平不均衡，住户的年龄、家庭条件等有很大差异，造成居民在空调使用习惯上有很大不同。多样化的居民使用习惯直接影响空调的运行功率与运行时长，进而影响房间空调器能耗。为总结出不同使用习惯的典型、剖析使用习惯之间的关系，对空调用户使用习惯的分类研究势在必行。

住宅空调使用习惯的研究方法常有调研法、实测法、模拟法等，然而，由于现有方法本身的局限性，均无法总结出典型用户群体及其使用习惯。对于调研实测法，研究机构开展了大量研究，其中，Guo 等<sup>[5]</sup>通过调研总结出该地区人员的主观特点，但由于使用习惯具有多样性，调研法很难通过单一指标总结出一两种典型使用模式。何玥儿<sup>[6]</sup>发放了 800 余份问卷对重庆地区空调使用模式进行广泛调研，并将空调使用时刻信息作为分类依据，构建了老幼群体的全天使用、上班族的早、中、晚间歇运行以及早、晚间歇运行的 3 种典型使用模式。然而，基于调研的分类多是依据经验进行主观确定的<sup>[7]</sup>，与实际情况存在差异。所以，通过调研与实测，可以获取一定主客观数据，但是数据质量、数据维度不够全面，导致典型行为获取具有局限性。模拟方法通常采用特定的理想条件，而实际情况却往往有很大出入。实际运行时，模型输入往往不一定遵照设定的理想条件，存在诸多不确定性<sup>[8]</sup>。在目前的建筑能耗模拟软件中，多数采用固定作息的方法进行描述，也有一些采用理想化反馈模型或随机模型等简化处理方式<sup>[9]</sup>；还有大量学者采用验证模型的思路，如通过实测或调研辅助手段提升模型的精度。Knight 等<sup>[10]</sup>

采用 ECOTECH 软件模拟学校建筑的能耗，并与实际监测数据进行对比，发现模拟方法明显低估了建筑物的电耗情况，差异达到 43%，通过进一步探索发现，造成差异的主要原因是用户对自身设备作息过于主观<sup>[8, 10]</sup>。由于模拟方法对用户行为的设置一般简化为固定/静态方式<sup>[11]</sup>，实际运行情况下的特别复杂行为模拟方法往往很难描述<sup>[12]</sup>，导致模拟结果难以准确反映实际人的行为对建筑用能水平的影响。

从分类指标或特征来看，高岩等<sup>[3]</sup>、刘猛等<sup>[13]</sup>采用能耗模拟软件，结合标准与文献记录，将设置温度、空调作息作为分类指标，设置几种典型的空调使用模式，推广了“部分时间”、“部分空间”的住宅空调节能运行方式。这些研究的空调设置模式维度较为单一，指标具有局限性（如仅采用使用时刻作息等），仅仅实现对空调使用习惯模式的简单分类，常常与实测结果产生较大偏差。

由于影响用户使用行为的因素较为复杂，也就无法基于单一指标的简单量化进行分类（如平均值、分位数等），很多学者提出多指标聚类算法（或综合指标聚类）<sup>[14]</sup>。随着互联网、大数据、人工智能的兴起，智能空调不断涌现，更加客观、更大区域、更多样本、更高效率、更全方位的数据监测与获取方式成为了可能。陈焕新等<sup>[15-16]</sup>阐述了制冷空调行业中常用的聚类分析、关联规则分析等无监督学习算法的原理与应用，分析了大数据在空调系统优化、故障诊断、建筑能耗与维护预测等方面的用途，指出大数据在用于分析人的行为上具有明显的优势。Yu 等<sup>[17]</sup>基于室外环境、建筑信息、人员信息、设备功能等 4 个指标，应用该算法对家庭用户进行聚类分析，研究结果可较好地用于评估设备对能耗的影响以及建筑能耗模拟的优化，作者在处理聚类指标过程中采用分类变量进行量化预处理，同时指出，由于调研方法的限制，对于众多参数影响下的建筑物能耗进行全面调查是不切实际的，查找典型指标进行研究可行性更高。康旭源等<sup>[18]</sup>采用聚类分析获取人员作息，分析了人员行为对建筑能耗的影响特点，通过人员位置信息大数据库，重构人员在室率指标，基于 *k*-means 聚类方法，获取了 16 栋不同建筑的典型人员作息曲线，为人员行为研究提供了数据支撑。

综上所述，因样本量以及研究方法的限制，现有

多数研究采用的指标不够全面,且主观性强,无法真实反映用户的使用习惯特点;此外,对指标的预处理方式过于简单,导致与实际情况存在较大差异,多维度聚类分析则为人员行为分类研究提供了参考。笔者基于房间空调器监控平台,获取了近2000台房间空调器详细、客观的运行数据,从空调使用时长、温度需求及能耗角度构建影响空调使用习惯的特征参数,提取多维度聚类特征,并通过数据清洗、衍生、标准化等预处理方法,对空调样本进行分群,提出房间空调器典型群体,并定性描述其使用特点。

## 1 研究方法

在机器学习中,聚类算法属于“无监督学习”算法的一种,目标是通过无标记训练样本的学习来揭示数据的内在性质及规律<sup>[19]</sup>。笔者对空调行为习惯采用常用的  $k$ -means 算法,在大样本的快速计算方面,比模糊划分算法、期望最大化算法等其他算法更具优势<sup>[20]</sup>。

### 1.1 聚类算法

聚类根据计算自身的距离或相似度将数据划分为若干组,划分的原则是组内距离最小而组间距离最大。 $k$ -means 采用距离作为相似性评价指标,即认为两个样本距离越近,其相似度就越大<sup>[21]</sup>。采用欧氏距离度量样本间相似性<sup>[17]</sup>。

$$d(k, l) = \sqrt{(x_{k1} - x_{l1})^2 + (x_{k2} - x_{l2})^2 + \dots + (x_{kn} - x_{ln})^2} \quad (1)$$

式中: $k, l$  分别代表两个样本; $x_{k1}, x_{k2}, \dots, x_{kn}$  分别为  $k$  样本的  $n$  个特征; $x_{l1}, x_{l2}, \dots, x_{ln}$  分别为  $l$  样本的  $n$  个特征。直观来看,式(1)刻画了组内样本围绕该组均值的紧密程度,距离  $d$  越小,则组内样本的相似性越高。

采用  $k$ -means 算法,对于每个簇而言,要求其簇内差异小,而簇外差异大,轮廓系数(Silhouette)正

是描述簇内外差异的关键指标<sup>[22]</sup>,见式(2)。

$$S = \frac{b - a}{\max(a, b)} \quad (2)$$

式中: $a$  为某个样本与其所在簇内其他样本的平均距离; $b$  为某个样本与其他簇样本的平均距离;由式(2)可知, $S$  取值范围为  $(-1, 1)$ ,当  $S$  越接近于 1,则聚类效果越好,越接近  $-1$ ,聚类效果越差,0 代表聚类重叠,没有很好地划分聚类。轮廓系数指标针对样本空间中的一个特定样本,计算它与所在聚类其他样本的平均距离  $a$  以及该样本与距离最近的另一个聚类中所有样本的平均距离  $b$ ,将整个样本空间中所有样本的轮廓系数取算数平均值,作为聚类划分的性能指标<sup>[23]</sup>。

聚类算法在训练之前需要对聚类各特征进行预处理,预处理方法主要有离群值检验、正态化以及标准化,具体处理方法见文献<sup>[24]</sup>,其中,离群值检测采用分位数方法,剔除数据中距离上下分位数较远的值,一般取 1.5 倍的分位间距;正态化的作用使得各特征近似呈现正态分布,消除数据分布中“尾迹”的影响;标准化的作用是消除不同特征量纲的影响,使数据标准化为均值为 0、方差为 1 的标准正态分布。

笔者基于机器学习工具 Scikit-learn 的  $k$ -means 包进行训练,该工具基于 Python 语言,在处理大数据集时具有编程简单、运算速度快的优势。

### 1.2 数据描述

针对重庆地区空调使用行为习惯进行研究,从某空调制造企业大数据平台随机抽取重庆主城区六区共计 2000 台住宅空调样本,从 2016 年到 2017 年共计一整年的空调运行数据集,数据集参数范围、分辨率、采集方式、传感元件、采集频率等见表 1。由于原始数据规模大,监测维度多,造成聚类分析计算距离失效,为避免“维数灾难”,需要基于行为研究对聚类参数进行重构与衍生<sup>[20]</sup>。

表 1 数据采集参数

Table 1 Data collection parameters

编号	采集参数	分辨率	范围	采集方式	传感元件	采集频率
1	能耗	0.01 kW·h	0~2 000 kW·h	直接	电量记录	
2	开机室内温度	0.1 °C	5~35 °C	直接	温度记录	15 min
3	设置温度	1 °C	16~30 °C	直接	状态记录	
4	运行小时数	1	0~3 000	间接	时间记录	
5	累积时长	0.25 h	0~3 000 h	间接	时间记录	1 min
6	开机次数	1	0~400	间接	时间记录	



聚类参数(特征)的多维度衍生应遵从3个原则:1)聚类参数尽量反映原始参数的研究意义,并覆盖更多维度;2)聚类参数尽可能接近正态分布,由于原始数据在收集过程中会产生诸多“坏值”,导致原始数据往往会呈现偏态,会大大影响聚类用户群的识别;3)参数应该进行相关性检验,由于高度相关的特征/参数会影响聚类分析的结果,在开始聚类分析之前需要先对冗余变量进行筛选和剔除。通过使用习惯特点进行筛选与衍生,形成与使用习惯相关的5个维度参数。

## 2 结果

### 2.1 特征分析

#### 1) 运行参数

长江流域地区空调运行具有“部分时间”、“部分空间”的使用特点<sup>[25]</sup>,其中“部分时间”即反映间歇运行的特性,采用单次开机运行时长反映该运行特点。如图1(a)所示,单位次数运行时长集中分布在0~10 h之间,根据多项夏热冬冷地区空调间歇运行研究,当单次运行时长分布在0~5 h之间时,可认为该样本偏好间歇运行<sup>[26-28]</sup>。

#### 2) 温度设置参数

将用户温度设置作为空调能耗影响的重要参数,由于夏热冬冷地区冬、夏季设置温度存在很大区别,为了描述季节性影响,构建综合性指标 $\alpha$ 来描述温度季节差异,衡量用户的综合温度设置偏好,见式(3)。

$$\alpha = \frac{ST_{\text{weight,w}}/T_{\text{comfort,w}}}{ST_{\text{weight,s}}/T_{\text{comfort,s}}} \quad (3)$$

式中: $T_{\text{comfort,w}}$ 与 $T_{\text{comfort,s}}$ 分别为冬、夏季舒适温度,按照暖通设计规范<sup>[29]</sup>,分别取20℃<sup>[13]</sup>与26℃<sup>[30]</sup>。 $ST_{\text{weight}}$ 代表某样本设置温度按时长加权的平均值。图1(b)为设置温度参数分布,该指标呈明显的双峰特点。总体上看,当 $\alpha=1$ 时,代表该用户偏好冬、夏季舒适温度设置;当 $\alpha>1$ 时,代表用户偏好冬季高温、夏季低温的设置方式,空调行为偏耗能;当 $\alpha<1$ 时,代表用户偏好冬季低温、夏季高温的设置方式,空调行为偏节能,从分布图可以看出,采用该设置方式的样本较少。

#### 3) 季节参数

为了衡量用户对季节的偏好特点,采用季节使用率之差 $\beta$ 进行衡量,如式(4)。

$$\beta = C_{\text{summer}}/C_{\text{s,total}} - C_{\text{winter}}/C_{\text{w,total}} \quad (4)$$

式中: $\beta$ 为季节使用率之差,%; $C_{\text{summer}}$ 与 $C_{\text{winter}}$ 分别为夏季与冬季开机运行小时数; $C_{\text{s,total}}$ 与 $C_{\text{w,total}}$ 分别为夏季与冬季总小时数。图1(c)为季节参数分布。当 $\beta>0$ 时,夏季运行率高于冬季,表示该用户偏好夏季运行;当 $\beta=0$ 时,表示该用户冬、夏季运行率相差不大,季节差异不明显;当 $\beta<0$ 时,表示冬季运行率更高,该用户偏好冬季运行,从分布图可以看出,该地区偏好冬季的样本少,也符合该地区夏季为主的特点。

#### 4) 环境参数

为了综合反映冬、夏季用户对舒适温度的容忍情况,采用开机运行期冬、夏季室内温度与舒适温度的差值之和(运行期室内温度 $\gamma$ )来描述用户对室内

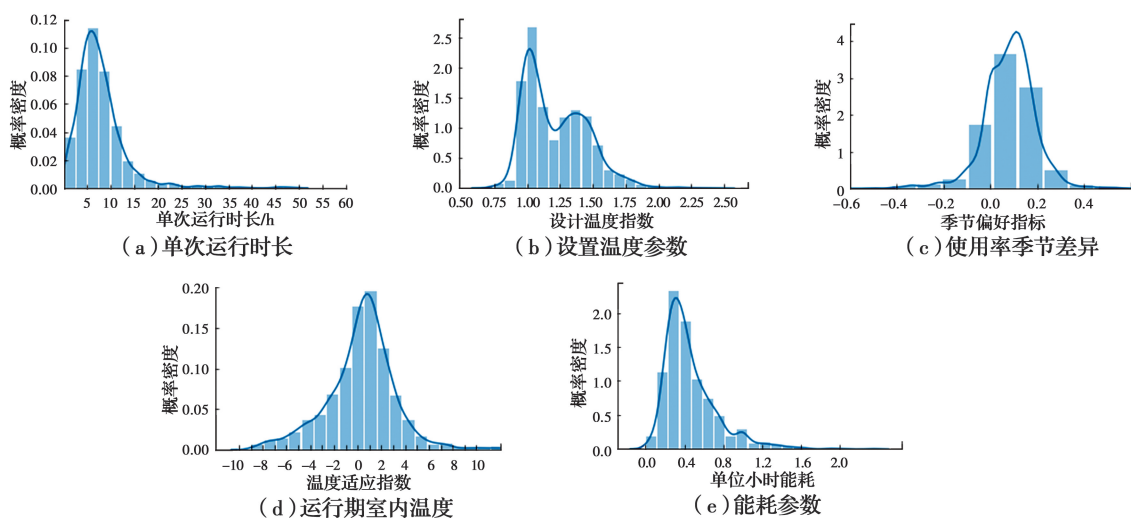


图1 分布密度图

Fig. 1 Distribution density diagram

环境的需求,见式(5)。

$$\gamma = (T_{in,s} - T_{comfort,s}) + (T_{comfort,w} - T_{in,w}) \quad (5)$$

式中: $T_{in,s}$ 、 $T_{in,w}$ 分别为冬、夏季空调开机时室内温度平均值,依照规范,夏热冬冷地区冬季与夏季舒适温度分别可取 20、26 °C。由式(5)可知, $\gamma$ 越大,表示对室内冬季偏冷与夏季偏热环境的适应能力越强; $\gamma$ 越小,表示对室内冬季过热与夏季过冷的适应能力越强。图 1(d)为环境参数分布。

### 5) 能耗参数

单位小时能耗可反映样本空调的制冷能力,间接反映用户的室内负荷特点,由图 1(e)可知,单位小时能耗分布更接近正态分布,且“尾迹”不明显。

## 2.2 聚类可视化

为了得到聚类最优数目,选择聚类数目从 2~9 分别建模,并计算各类别下的平均轮廓系数,如图 2 所示。当聚类数目为 2 与 3 时,平均轮廓系数最高,聚类数目越多,轮廓系数越小,表示聚类效果越差,所以,主要讨论类别为 2 与 3 时的分类特点。

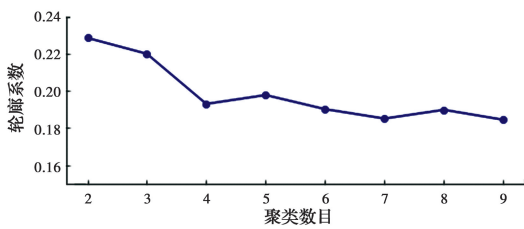


图 2 不同聚类数目轮廓系数图

Fig. 2 Silhouette value diagram of different clusters

为了在二维坐标系清晰展示样本分布,将 5 维数据转化为 2 维显示,如图 3(a)所示,采用 TSNE 降维方法<sup>[22]</sup>并单独计算每个样本的轮廓系数,再进行排序,如图 3(b)所示,其中红色虚线表示整体样本的平均轮廓系数。

聚类样本数量显示,聚类数目为 2 时,类别 A 样本占比 57%,类别 B 占比 43%;轮廓系数的总体均值约为 0.22,两个类别轮廓系数分布合理。从图 3(a)可以看出,在类别 B 中仍然有少量样本分类效果较差,轮廓系数为负,但仅有 39 个样本,占比不高(约占 3.4%)。

如图 3(b)所示,当样本数目为 3 时,各类别样本数占比分别为:47%(a),28%(b)和 25%(c),轮廓系数为负的样本共计 49 个,约占 4.2%,分布在类别 b 与类别 c 中;从两种聚类结果来看,类别 c 近似从类别 A、B 中各提取部分样本进行重新组合形成第 3 个类别,分类更加细化。

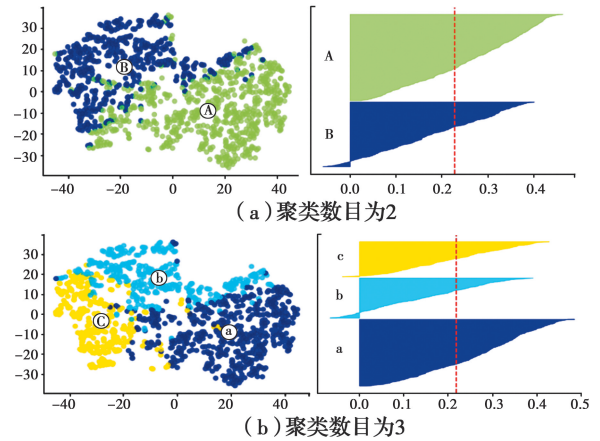


图 3 样本分布及轮廓系数排序

Fig. 3 Sample distribution and sorting of silhouette value

## 3 讨论

### 3.1 聚类特征差异性

为了进一步分析聚类结果下各个指标的分布情况以及不同类别之间的差异,采用雷达图显示特征分布,如图 4 所示。

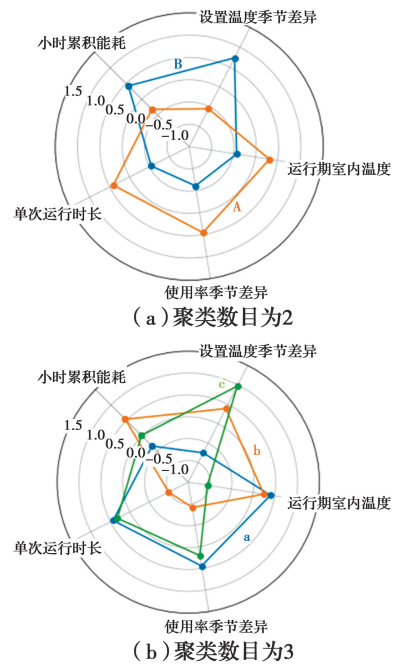


图 4 特征差异性图

Fig. 4 Characteristic difference diagram

当聚类类别为 2 时,从雷达图 4(a)可以看出,5 个特征在两个聚类类别中均表现出较为明显的差异,均可作为分离两个类别的优势特征;差异最大的特征为温度设置参数,A 类用户设置温度参数  $\alpha$  约为 1.0 左右,根据前文,该类用户偏好舒适温度设置(冬:20 °C,夏:26 °C);B 类用户均值约为 1.4 左右,

该类用户偏好冬季高温、夏季低温耗能设置行为。总体上, A 类用户较 B 类用户更偏向节能行为, 主要在夏季, 经过统计, 单位小时累积能耗类别 A 约为 0.2~0.4 kW·h, 类别 B 则略高, 处于 0.3~0.6 kW·h 之间, 类别 B 明显显示出高能耗的特点。

当聚类类别为 3 时, 同样采用雷达图分析各个指标特征, 如图 4(b) 所示, 其中类别 a 对应类别 A, 二者各项指标差异不大, 均反映夏季连续运行的偏好舒适温度运行的用户; 对于 b 类来说, 与 B 类各项指标差异不大, 均反映冬、夏季均有空调需求, 且多为间歇运行的用户群体; 对于 c 类来说, 指标设置温度、室内温度与 a 类差别大, 其他 3 个指标与 a 类近似相等, 两个类别均反映夏季连续运行的用户, 差异在于对热环境需求不同, c 类样本设置温度参数高、运行期室内温度低, 更偏好耗能设置的“过”舒适环境(冬季过热与夏季过冷), a 类倾向节能运行。

值得注意的是, 两个聚类方式均能较好地分出用户使用习惯类群, 聚为 3 类的分类方式在 2 类的基础上更加细化, 新用户群体是否具有实际意义仍需进一步讨论。

### 3.2 聚类数目讨论

聚为两个类别时, 样本分布较为均匀, 轮廓系数更高, 聚类错误率相对较低; 聚为 3 类时, 样本分布不均匀性提高, 轮廓系数低, 聚类错误率略高。但聚类数目的确定, 除了采用传统的评估指标(如距离、轮廓系数), 仍需结合实际具体问题具体探讨。样本分布占比如图 5 所示。



图 5 不同聚类数目样本占比

Fig. 5 Proportion of samples with different clusters

图 5 中外环代表聚类数目为 2(A 与 B), 内环代表聚类数目为 3(a、b 及 c)。通过对不同聚类数目的 5 个特征进行分析, A 与 a、B 与 b 的特点近似相同, 所以, 类别 c 可以看作是对聚 2 类时的进一步细分, 即类别 c 从类别 A 中分出 10% 的样本, 从类别 B 中分出 15% 的样本, 共计 25% 的样本形成 c 类, 为以

夏季连续运行的用户群体, 空调行为偏好耗能设置, 对室内环境要求高, 因此, 聚 2 类中 A、B 的运行特点包含在聚 3 类时的运行特点中, 为此, 将聚类结果设置为 3 类进行后续分析, 见表 2。

表 2 不同类别用户群空调行为偏好

Table 2 Behavior preferences of air-conditioning with different clusters

聚 2 类(占比)	运行偏好	聚 3 类(占比)	运行偏好
A(57%)	夏季为主 连续运行 节能设置	a(47%)	夏季为主 连续运行 节能设置
		c(25%)	夏季为主 连续运行 耗能设置
		B(43%)	冬夏两用 间歇运行 耗能设置
b(28%)	冬夏两用 间歇运行 耗能设置		

由于夏热冬冷地区住宅空调夏季运行周期长, 对空调器依赖性高, 冬季采暖设备多样, 所以, 该地区空调仅采用夏季供冷单一方式的样本多。调研实测过程中, 大量学者也指出, 夏季存在低温设置的典型群体, 如陈颢等<sup>[31]</sup>通过对青年人热感觉调研指出, 夏季使用空调比例高的用户(依赖性更强)更适应较低的室温; 简毅文等<sup>[32]</sup>、谭晶月<sup>[33]</sup>、郭祺<sup>[34]</sup>采用现场询问方式对空调设置习惯的调研中发现, 有相当部分用户(超过 40%)在夏季会选择采用 26℃ 以下的低温设置, Haochen 等<sup>[35]</sup>也指出, 长江流域 20% 左右的用户会在夜间连续开启 18~24℃ 的设置温度; 柴盼等<sup>[36]</sup>通过调研也指出, 有约 30% 的用户夏季会通过设置低于标准规定的温度来提高舒适度。部分热舒适研究学者也指出, 体质较好的用户群体(如青年人等)会选择将室内环境设置较低温。

因此, 在夏季采用耗能的低温设置群体客观存在, 采用 3 个类别聚类更有实际意义。基于此, 总结聚类 3 个用户群体的特点:

1) 对于类别 a 用户群体, 运行特点为夏季为主、连续运行、节能设置, 空调使用过程具有较强的节能意识, 对空调耗电较为关注, 近似三代人家庭, 能耗最低, 可作为夏季空调运行的推荐模式。

2) 对于类别 b 用户群体, 运行偏好为冬、夏两用、间歇运行、耗能设置, 结合重庆地区调研结果, 重



点在于间歇运行特点,与上班族或两代人家庭的特征较为近似,能耗最高。该类高能耗人群应予以重点关注,可采用合理的规划及必要的节能技术,如自然通风技术等降低能源消耗。

3)对于类别c用户群体,运行特点为夏季为主、连续运行、耗能设置,重点在于夏季长期低温运行,多为身体适应性较强的人群,近似单身青年等家庭。该类人群需要在满足舒适性的情况下增加节能意识,减少不必要的能源浪费,避免长期运行引发的“空调病”。

## 4 结论

通过多维度聚类算法,结合重庆地区房间空调器数据监测平台,选择合适的评估指标对空调用户群体进行分类,识别出该地区空调使用的典型群体,主要研究结论如下:

1)从空调使用时长、温度需求及能耗角度,构建了空调运行的5个聚类参数,包括单次运行时长、运行期室内温度、使用率季节差异、设置温度季节差异、小时累积能耗参数,剖析重庆地区空调使用习惯的典型特征并识别典型群体,指出多维度聚类算法可用于多样化空调使用习惯的研究。

2)通过特征差异性分析,指出重庆地区空调用户运行方式多样、季节性明显的特点,并挖掘出该地区3类典型空调使用习惯,即以夏季为主,连续运行的低能耗群体;以夏季为主,连续运行的高能耗群体;冬、夏两用,间歇运行能耗较高的群体。

### 参考文献:

- [1] 国家统计局. 城镇居民平均每百户年末主要耐用消费品拥有量[Z]. 北京: 中国统计出版社, 2020.  
National Bureau of Statistics of China. Main durable goods owned per 100 urban households at year-end [Z]. Beijing: China Statistics Press, 2020. (in Chinese)
- [2] 龙惟定, 梁浩. 我国城市建筑碳达峰与碳中和路径探讨[J]. 暖通空调, 2021, 51(4): 1-17.  
LONG W D, LIANG H. Discussion on paths of carbon peak and carbon neutrality of urban buildings in China [J]. Heating Ventilating & Air Conditioning, 2021, 51(4): 1-17. (in Chinese)
- [3] 高岩, 安玉娇, 于喜哲, 等. 住宅使用模式对住宅建筑供暖空调能耗影响的模拟分析[J]. 建筑科学, 2010, 26(10): 287-291.  
GAO Y, AN Y J, YU X Z, et al. Simulation analysis of the influence of residential use pattern on heating and air-conditioning energy consumption of residential buildings [J]. Building Science, 2010, 26(10): 287-291. (in Chinese)
- [4] CHEN S, ZHANG G M, XIA X B, et al. A review of internal and external influencing factors on energy efficiency design of buildings [J]. Energy and Buildings, 2020, 216: 109944.
- [5] GUO S Y, YAN D, PENG C, et al. Investigation and analyses of residential heating in the HSCW climate zone of China: Status quo and key features [J]. Building and Environment, 2015, 94: 532-542.
- [6] 何玥儿. 基于量质分析的夏热冬冷地区住宅热环境营造技术优化研究[D]. 重庆: 重庆大学, 2017.  
HE Y E. Quantity-quality-based optimization of indoor thermal environment for residential buildings in hot summer and cold winter climate zones [D]. Chongqing: Chongqing University, 2017. (in Chinese)
- [7] 侯珊珊, 刘猛, 晏璐, 等. 运用社会化聆听方法的非集中供暖区住户对冬季室内热环境不满的分析[J]. 土木与环境工程学报(中英文), 2021, 43(1): 215-228.  
HOU S S, LIU M, YAN L, et al. Dissatisfaction expressed by occupants with winter indoor thermal environment in the non-district heating zone based on the social listening method full-text in English [J]. Journal of Civil and Environmental Engineering, 2021, 43(1): 215-228. (in Chinese)
- [8] RYAN E M, SANQUIST T F. Validation of building energy modeling tools under idealized and realistic conditions [J]. Energy and Buildings, 2012, 47: 375-382.
- [9] YAN D, XIA J J, TANG W Y, et al. DeST: An integrated building simulation toolkit. Part I: Fundamentals [J]. Building Simulation, 2008, 1(2): 95-110.
- [10] KNIGHT I, STRAVORAVDIS S, LASVAUX S. Assessing the operational energy profiles of UK educational buildings: Findings from detailed surveys compared to measured consumption [C]//2nd PALENC Conference and 28th AIVC Conference on Building Low Energy Cooling and Advanced Ventilation Technologies in the 21st Century. Crete, Greece. 2021: 531-536
- [11] 燕达, 丰晓航, 王闯, 等. 建筑中行人行为模拟研究现状和展望[J]. 建筑科学, 2015, 31(10): 178-187.  
YAN D, FENG X H, WANG C, et al. Current state

- and future perspective of occupant behavior simulation in buildings [J]. *Building Science*, 2015, 31(10): 178-187. (in Chinese)
- [12] HOES P, HENSEN J L M, LOOMANS M G L C, et al. User behavior in whole building simulation [J]. *Energy and Buildings*, 2009, 41(3): 295-302.
- [13] 刘猛, 薛凯, 袁逸群, 等. 典型城市居住建筑室内设置温度对供暖供冷能耗的影响[J]. *土木建筑与环境工程*, 2015, 37(Sup2): 204-210.  
LIU M, XUE K, ZHONG Y Q, et al. Impact of indoor setting temperature on heating and cooling energy consumption in selected cities [J]. *Journal of Civil, Architectural & Environmental Engineering*, 2015, 37(Sup2): 204-210. (in Chinese)
- [14] 涂正革, 谌仁俊. 中国碳排放区域划分与减排路径: 基于多指标面板数据的聚类分析[J]. *中国地质大学学报(社会科学版)*, 2012, 12(6): 7-13, 136.  
TU Z G, CHEN R J. Analysis on China's carbon emission division and reduction path: Based on multivariate panel data clustering analysis method [J]. *Journal of China University of Geosciences (Social Sciences Edition)*, 2012, 12(6): 7-13, 136. (in Chinese)
- [15] 陈焕新, 孙劭波, 刘江岩, 等. 数据挖掘技术在制冷空调行业的应用[J]. *暖通空调*, 2016, 46(3): 20-26.  
CHEN H X, SUN S B, LIU J Y, et al. Application of data mining technology to refrigeration and air conditioning industry [J]. *Heating Ventilating & Air Conditioning*, 2016, 46(3): 20-26. (in Chinese)
- [16] 陈焕新, 刘江岩, 胡云鹏, 等. 大数据在空调领域的应用[J]. *制冷学报*, 2015, 36(4): 16-22.  
CHEN H X, LIU J Y, HU Y P, et al. Application of big data in air-conditioning field [J]. *Journal of Refrigeration*, 2015, 36(4): 16-22. (in Chinese)
- [17] YU Z, FUNG B C M, HAGHIGHAT F, et al. A systematic procedure to study the influence of occupant behavior on building energy consumption [J]. *Energy and Buildings*, 2011, 43(6): 1409-1417.
- [18] 康旭源, 燕达, 孙红三, 等. 基于人员位置大数据的建筑人员作息模式研究[J]. *暖通空调*, 2020, 50(7): 1-10, 90.  
KANG X Y, YAN D, SUN H S, et al. Typical occupancy profiles in buildings based on big data of mobile positioning [J]. *Heating Ventilating & Air Conditioning*, 2020, 50(7): 1-10, 90. (in Chinese)
- [19] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.  
ZHOU Z H. *Machine learning* [M]. Beijing: Tsinghua University Press, 2016. (in Chinese)
- [20] 晋远, 燕达, 安晶晶, 等. 基于自编码算法聚类的城镇住宅建筑日用电典型模式分析[J]. *建筑科学*, 2020, 36(2): 1-7, 43.  
JIN Y, YAN D, AN J J, et al. Analysis on typical daily electricity consumption patterns of urban residential buildings from clustering analysis based on auto-encoder algorithm [J]. *Building Science*, 2020, 36(2): 1-7, 43. (in Chinese)
- [21] 张良均, 谭立云, 刘名军. Python 数据分析与挖掘实战[M]. 2版. 北京: 机械工业出版社, 2020.  
ZHANG L J, TAN L Y, LIU M J. *Hands-on data analysis and data mining with Python* [M]. Beijing: China Machine Press, 2020. (in Chinese)
- [22] YAN L, LIU M, XUE K, et al. A study on temperature-setting behavior for room air conditioners based on big data [J]. *Journal of Building Engineering*, 2020, 30: 101197.
- [23] ROUSSEEUW P J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis [J]. *Journal of Computational and Applied Mathematics*, 1987, 20: 53-65.
- [24] TARDIOLI G, KERRIGAN R, OATES M, et al. Identification of representative buildings and building groups in urban datasets using a novel pre-processing, classification, clustering and predictive modelling approach [J]. *Building and Environment*, 2018, 140: 90-106.
- [25] 程飞, 张旭, 张莉莉, 等. 夏热冬冷地区住宅分体式空调容量需求分析[J]. *建筑科学*, 2016, 32(8): 154-158.  
CHENG F, ZHANG X, ZHANG L L, et al. Analysis on capacity of the split type air conditioner in residential building in hot summer and cold winter area [J]. *Building Science*, 2016, 32(8): 154-158. (in Chinese)
- [26] 张旭, 程飞, 黄奕翔, 等. 变频空调器间歇供冷能耗特征测试分析[J]. *暖通空调*, 2020, 50(4): 8-13.  
ZHANG X, CHENG F, HUANG Y X, et al. Test and analysis of energy consumption characteristics of variable speed air conditioner intermittent cooling [J]. *Heating Ventilating & Air Conditioning*, 2020, 50(4): 8-13. (in Chinese)
- [27] 张晓洁. 长沙办公建筑间歇空调能耗模拟分析[D]. 长沙: 湖南大学, 2011.  
ZHANG X J. *Simulation and analysis on energy consumption of air-conditioned intermittent operation in*



- an office building in Changsha [D]. Changsha: Hunan University, 2011. (in Chinese)
- [28] 刘晓庆. 间歇空调运行模式下住宅墙体热工性能研究 [D]. 重庆: 重庆大学, 2011.
- LIU X Q. Study on the thermal performance of residential walls in intermittent air conditioning run mode [D]. Chongqing: Chongqing University, 2011. (in Chinese)
- [29] 民用建筑供暖通风与空气调节设计规范: GB 50736—2012 [S]. 北京: 中国建筑工业出版社, 2012: 288.
- Design code for heating ventilation and air conditioning of civil buildings: GB 50736-2012 [S]. Beijing: China Architecture & Building Press, 2012: 288. (in Chinese)
- [30] 刘猛, 晏璐, 李金波, 等. 大数据监测平台下的长江流域典型城市房间空调器温度设置分析[J]. 土木与环境工程学报(中英文), 2019, 41(5): 164-172.
- LIU M, YAN L, LI J B, et al. Analysis of temperature setting habits of room air conditioners in the typical cities in Yangtze River Basin under the big data monitoring platform [J]. Journal of Civil and Environmental Engineering, 2019, 41(5): 164-172. (in Chinese)
- [31] 陈颢, 熊宸, 程楠, 等. 空调与自然通风工况下用户个性化特征对热舒适影响分析[J]. 制冷与空调, 2021, 35(2): 182-188.
- CHEN X, XIONG C, CHENG N, et al. Impact of the occupants' personalized characters on thermal comfort in air-conditioned and natural ventilation environment [J]. Refrigeration & Air Conditioning, 2021, 35(2): 182-188. (in Chinese)
- [32] 简毅文, 李清瑞, 白贞, 等. 住宅夏季空调行为对空调能耗的影响研究[J]. 建筑科学, 2011, 27(12): 16-19, 86.
- JIAN Y W, LI Q R, BAI Z, et al. Study on influences of usage behavior of residential air handling unit on energy consumption in summer [J]. Building Science, 2011, 27(12): 16-19, 86. (in Chinese)
- [33] 谭晶月. 基于大数据的重庆地区住宅建筑房间空调器使用特征研究[D]. 重庆: 重庆大学, 2018.
- TAN J Y. Research on the use characteristics of room air conditioners in residential buildings in Chongqing based on big data [D]. Chongqing: Chongqing University, 2018. (in Chinese)
- [34] 郭祺. 通风对重庆地区卧室房间空调器用能的影响研究[D]. 重庆: 重庆大学, 2018.
- GUO Q. Influence of ventilation on the energy consumption of air conditioning in the bedroom in Chongqing [D]. Chongqing: Chongqing University, 2018. (in Chinese)
- [35] JIANG H C, YAO R M, HAN S Y, et al. How do urban residents use energy for winter heating at home? A large-scale survey in the hot summer and cold winter climate zone in the Yangtze River region [J]. Energy and Buildings, 2020, 223: 110131.
- [36] 柴盼, 刘旭, 翁庙成, 等. 重庆某住宅小区能耗特征与节能潜力研究[J]. 制冷与空调, 2014, 28(6): 694-697.
- CHAI P, LIU X, WENG M C, et al. Research on energy consumption characteristics and energy saving potential of a residential area in Chongqing [J]. Refrigeration & Air Conditioning, 2014, 28(6): 694-697. (in Chinese)

(编辑 王秀玲)