

ICS 号

CCS 号

团体标准

团体标准编号

代替团体标准编号

建筑能耗模拟用人为基础数据采集 技术导则

Technical guidelines for basic data acquisition of
occupant behavior in building energy consumption
simulation

(征求意见稿)

XXXX-XX-XX 发布

XXXX-XX-XX 实施

中国建筑节能协会 发布

前 言

根据《中国建筑节能协会团体标准管理办法（试行）》（国建节协（2017）40号）及《关于印发〈2019年度第一批团体标准制修订计划〉的通知》（国建节协[2019]5号）的要求，由主编单位会同有关单位组建编制组，经广泛的调查研究，认证总结实践经验，考察有关国内外标准和先进经验，并在广泛征求意见的基础上，共同编制了本导则。

本导则的主要内容包括：1.总则；2.术语；3建筑用能行为描述方法；4.建筑模拟用人为描述参数及基础数据；5. 数据获取方法；6. 人行为数据质量要求与预处理；7. 数据分析方法；附录；本导则用词说明；引用标准名录。

本导则由中国建筑节能协会标准化管理办公室负责管理（联系电话：010-57811483，邮箱：biaoban@cabee.org），由主编单位负责具体内容的解释。执行过程中如有意见或建议，请寄送至主编单位（地址：上海市嘉定区曹安公路4800号，邮编：200092）。

本导则主编单位：同济大学、清华大学

本导则参编单位：浙江大学，北京工业大学，北京建筑大学，天津城建大学，天津大学，湖南大学，上海建科院，华建集团，太古地产，上海华电源信息技术有限公司，洛阳优创信息科技有限公司，广东美的白色家电技术创新中心有限公司，浙江大华技术股份有限公司，宁波方太厨具有限公司

本导则主要起草人员：潘毅群，燕达，李峥嵘，李翠，王闯，陈淑琴，杨斌，周翔，张静思，许鹏，俞准，简毅文，潘嵩，丁研，季亮，张颖

本导则主要审查人员：

目 次

1 总则	1
2 术语与符号	2
2.1 术语.....	2
2.2 符号.....	4
3 建筑用能行为描述方法	5
3.1 人员移动.....	5
3.2 人员动作.....	5
4 描述参数及基础数据定义.....	8
4.1 固定作息表及所需的基础数据.....	8
4.2 位移模型参数及所需的基础数据.....	9
4.3 动作模型参数及所需的基础数据.....	10
5 基础数据获取方法	11
5.1 获取方式.....	11
5.2 采集方法.....	12
6 基础数据质量与预处理	14
6.1 基础数据质量要求.....	14
6.2 基础数据预处理方法.....	14
7 基础数据分析方法	16
7.1 环境因素的影响分析.....	16
7.2 非环境因素的影响分析.....	17
7.3 建筑用能行为特征分析.....	17
附录 A 适用于隐私要求较高的基础数据采集设备.....	21
附录 B 数据预处理方法.....	23
附录 C 各用能行为的指标分析推荐用图/表.....	26
本导则用词说明	29
引用标准名录	30
附：条文说明	31

Contents

1 general provisions.....	1
2 terms and symbols	2
2.1 terms	2
2.2 symbols.....	4
3 description method of energy-related occupant behavior in buildings.....	5
3.1 occupant movement.....	5
3.2 occupant action	5
4 Description Parameters and Basic Data Definition	8
4.1 Fixed Schedule and Required Basic Data	8
4.2 Movement Model Parameters and Required Basic Data.....	9
4.3 Action Model Parameters and Required Basic Data	10
5 Basic Data Collection Method.....	11
5.1 Collection Method	11
5.2 Measuring Method.....	12
6 Basic Data Quality and Preprocessing.....	14
6.1 Basic Data Quality Requirements.....	14
6.2 Basic Data Preprocessing Methods	14
7 Basic Data Analysis Method.....	16
7.1 Impact Analysis of Environmental Factors.....	16
7.2 Impact Analysis of Non Environmental Factors.....	17
7.3 Analysis on Behavior Characteristics	17
Appendix A Data Collection Equipment for High Privacy Requirements	21
Appendix B Data Preprocessing Methods.....	23
Appendix C Recommended Diagram / Table for Indicator Analysis	26
Explanation of Wording in This Guide.....	29
List of Quoted Standards	30
Addition: Explanation of Provisions	31

1 总则

1.0.1 为指导和规范建筑能耗模拟相关的人行为基础数据采集与应用，建立符合我国国情的各类民用建筑人行为标准数据集，特制定本导则。

1.0.2 本导则适用于办公、酒店、商场、住宅等民用建筑全性能模拟过程中，生成各种人行为描述参数所需的基础数据的采集、处理和分析。

1.0.3 建筑能耗模拟中人行为基础数据的采集、处理和分析方法除应符合本导则外，尚应符合国家现行标准的规定。

2 术语与符号

2.1 术语

2.1.1 建筑用能行为 occupant energy-related behavior in buildings

与建筑环境和能耗直接且密切相关的使用模式、使用状况等，如人员移动、开窗、遮阳调节、照明、空调供暖设备使用、电器设备使用等行为。

2.1.2 人员移动 occupant movement

人在建筑及房间位置的变化，即房间内的人员数量与人员位置分布。包括人员进入或离开某个房间，以及人员移动到房间内的某个位置。

2.1.2 人员动作 occupant action

人员基于感知到的周围环境与以往经验的总和进行比较，从而有意识或无意识控制周围人工环境物理参数的动作，主要体现在房间设备对象的状态变化上。

2.1.3 建筑模拟用人行为 occupant behavior in building energy simulation

建筑性能模拟中用于预测人员所在位置、建筑内各类设备对象的运行状态和人员对于设备控制使用特征的模型和方法。

2.1.4 位移模型 movement model

刻画和描述室内人员数量与位置特征的分布，追踪人员在建筑中的移动，解决房间有没有人、有多少人等基本问题。

2.1.5 动作模型 action model

刻画和描述室内人员对某设备控制使用特征的分布，以及动作发生与时间或环境等因素的关联性。

2.1.6 人行为数据 occupant behavior data

通过各类传感器、移动互联网大数据、问卷调查、环境参数测试装置等方式获取的与人员行为活动相关的数据。

2.1.7 人员移动数据 occupant movement data

用于表示人员的移动规律，可根据不同的模拟目的选择对应的采集方法。

2.1.8 人员动作数据 occupant action data

用于表示人对设备的调控规律，可根据不同的设备选择对应的采集方法。

2.1.9 接触式测量 contact measurement

使用直接接触人员的方式采集人员位移及动作，应以不对室内人员形成负担为原则。

2.1.10 非接触式测量 non-contact measurement

通过不直接接触人员的设备采集人员位移及动作，应充分考虑室内人员隐私问题。

2.1.11 可穿戴设备 wearable sensor

直接将传感器与人员身上的配件相结合的一种便携式设备，可与各类终端及互联网连接从而采集多种人员信息，如人体代谢率、人员位置等。

2.1.12 异常数据 abnormal data

不满足相关国家标准中规定的正常取值范围的可疑数据。

2.1.13 缺失数据 missing data

由于传感器数据缺失、调查类数据缺失和数据后处理缺失导致数据的不完整性。

2.2 符号

$E(x_i)$	——	平均值
T	——	温度, °C
u	——	阈值参数, °C
l	——	尺度参数, °C
k	——	形状参数
c	——	事件对于随环境因素变化的基本函数形式产生的附加影响
α	——	常参数
$\Delta\tau$	——	模拟步长
r	——	开关比
VIF	——	方差膨胀因子
TOL	——	容差
Y	——	遮阳设备使用状态特征值
E	——	室内照度水平, lux
D	——	灯具开关状态
$L_2(x_i, x_j)$	——	待填补样本已知属性与完整样本的欧式距离
σ	——	标准差
μ	——	均值
v_i	——	残差
IQR	——	四分位间距
x^*	——	待转换变量的转换值
x_{min}	——	最小值
x_{max}	——	最大值

3 建筑用能行为描述方法

3.1 人员移动

3.1.1 宜使用固定作息表和基于马氏链与事件的室内人员移动模型两种方式描述人员移动行为，其中固定作息表应符合本导则 3.1.2 的规定，基于马氏链与事件的室内人员移动模型应符合本导则 3.1.3~3.1.7 的规定。

3.1.2 固定作息表以房间人员逐时在室率表示。房间人员逐时在室率为房间逐时人数与房间最多人数的比值。根据房间设置的最多人数及逐时在室率，即可计算逐时房间人数。人员作息表的模型通式按下式计算：

$$x_i = f(t_i) \quad (3.1.2)$$

式中： x_i —— 房间逐时在室率， $0 \leq x_i \leq 1$ ；

t_i —— 第 i 个小时。

3.1.3 基于马氏链与事件的室内人员移动模型可用来描述人员个体的随机移动过程。其基本想法是将人员个体的日常活动与位置移动用移动事件和马尔可夫随机过程（Markov Chain，简称马氏链）来近似。模型建立应符合以下规定：

- 1 以建筑房间作为位置单元，人员位置以房间编号标识；
- 2 基于个体的移动过程描述，不同人员个体的移动过程相互独立；
- 3 采用马尔可夫链表示人员移动随机过程；
- 4 为日常活动定义移动事件，各个事件通过修改马尔可夫转移矩阵来驱动人员移动过程；
- 5 根据人员位置初始状态，进行马尔可夫链数值模拟，生成人员逐时位置、统计各房间人员状况等。

3.1.4 人员的位置移动伴随或包含在一系列日常活动和事件之中，称为事件机制，例如上班、下班、开会、起床、睡觉等，都会引起人员位置的变化。宜引入移动事件反映这些日常活动在人员移动过程中的作用。

3.1.5 随机走动的主要数值特征应包括人员在各个房间停留的时间比例和人员在各个房间平均每次逗留的时间。

3.1.6 办公建筑中的典型事件包括上班、下班、午餐、会议等，这些典型事件的特征参数应包括起始时间、结束时间、平均上班/下班/出发/返回时间等。

3.1.7 住宅建筑中的典型事件包括起床、睡觉、上班、下班等，这些典型事件的特征参数应包括起始时间、结束时间、平均起床/睡觉/上班/下班时间等。

3.2 人员动作

3.2.1 人员动作包括人员对房间内部空调、照明、设备及窗户的开关等使用或控制行为。

3.2.2 人员动作可采用固定作息表、条件概率模型、Logistic 模型等三种方法进行描述，

其中固定作息表法应符合本导则 3.2.3 的规定，条件概率模型法应符合本导则 3.2.4 的规定，Logistic 模型法应符合本导则 3.2.5 的规定。

3.2.3 人员动作的固定作息表即房间照明、设备、空调、供暖、通风等运行作息，用来描述房间各类装置的逐时使用情况，应根据房间功能类型和典型日类型进行设置。房间各类装置使用作息表的模型应符合以下规定，并按下式计算：

$$x_i = f(t_i) \quad (3.2.3)$$

式中： x_i —— 作息表逐小时值， $0 \leq x_i \leq 1$ ；

t_i —— 第 i 个小时。

- 1 房间空调（供暖）作息表包括设定温湿度和启停状态两部分。空调（供暖）设定温湿度分别用温度值、相对湿度值表示。空调（供暖）启停状态用 0 和 1 来表示，0 表示关闭，1 表示开启。
- 2 房间照明或设备作息表以房间照明灯具或家用/办公设备的逐时使用率表示。房间照明（设备）逐时使用率是房间逐时照明（设备）功率与房间最大照明（设备）功率的比值。根据房间设置的照明（设备）最大功率及逐时使用率计算逐时房间照明（设备）功率。
- 3 房间通风作息表以房间的逐时通风换气次数表示。房间逐时通风换气次数是房间逐时通风量体积流量与房间体积的比值，应根据实际情况确定。

3.2.4 人员动作的条件概率模型将人员动作视为由特定的时间、事件或物理环境条件所触发、并按一定概率随机发生。条件概率模型应从开动作和关动作分别描述，分为环境反馈型和时间型两大类，应符合以下规定：

- 1 环境反馈型条件概率用于描述与室内外物理环境因素相关的动作，例如开暖气、开空调、开灯、开窗等动作，其发生受环境因素（如室内温度、照度、CO₂ 浓度等）刺激和影响。环境反馈型条件概率按下式计算：

- 1) 动作发生概率与环境因素正相关：

$$P = \begin{cases} 1 - e^{-\left(\frac{x-u}{l}\right)^k} \Delta\tau, & x > u \\ 0 & x \leq u \end{cases} \quad (3.2.4 - 1)$$

- 2) 动作发生概率与环境因素负相关：

$$P = \begin{cases} 1 - e^{-\left(\frac{u-x}{l}\right)^k} \Delta\tau, & x < u \\ 0 & x \geq u \end{cases} \quad (3.2.4 - 2)$$

式中： P ——动作发生的概率；

x ——环境变量，如温度、湿度、照度、CO₂ 浓度、太阳直射强度等各类环境变量；

$\Delta \tau$ ——模拟计算采用的时间步长；

u ——阈值参数，表征人体感受环境刺激的阈值特征，量纲与 x 相同； $x-u$ 表示环境变量 x 偏离阈值的大小；

l ——为尺度参数，表征环境刺激的比例因子；令 $\bar{l} = l(\Delta\tau)^{-\frac{1}{k}}$ ，则 \bar{l} 的量纲与 x 相同， $\frac{x-u}{\bar{l}}$ 是环境变量 x 的无量纲化；

k ——形状参数，表征动作随环境变化的敏感性， k 值越大，使用者对于该自变量因素越敏感，概率变化曲线越陡峭。

2 时间型条件概率用于与时间或事件相关的动作，即动作发生在某些特殊时刻，如进出门时，上下班时，起床或睡觉时。时间型条件概率按下式计算：

1) 动作发生概率为常数：

$$P = \begin{cases} p & , \tau = \tau_0 \\ 0 & , \tau \neq \tau_0 \end{cases} \quad (3.2.4 - 3)$$

2) 动作发生概率与时长（如出门离去）有关：

$$P = \begin{cases} 1 - e^{-\left(\frac{t_{leave}}{l}\right)^k} & , \tau = \tau_0 \\ 0 & , \tau \neq \tau_0 \end{cases} \quad (3.2.4 - 4)$$

3 人员动作的条件概率模型与人员移动随机模型类似，都是针对人员个体进行定义的。

对于多人房间，可结合个体行为独立性假设，模拟得到人员群体的行为结果。

3.2.5 人员动作的 Logistic 模型描述方法应符合下列规定：

1 Logistic 模型按下式计算（以开关空调为例）：

$$p = \frac{1}{1 + \exp^{-(bx+c)}} \times 100\% \quad (3.2.5)$$

式中： p ——行为状态的发生概率（空调开着或关着的概率）；

x ——热环境变量（如室内温度、室外温度等）（ $^{\circ}\text{C}$ ）；

b 、 c ——拟合系数，没有明确的物理意义。

2 Logistic 模型中的概率通常针对房间设备（如空调、窗户、照明灯具等）的运行状态进行定义，适合描述人员群体的行为规律。

4 描述参数及基础数据定义

4.1 固定作息表及所需的基础数据

4.1.1 固定作息表法所需采集的数据根据被测行为确定，可以是逐时状态数据，也可以是逐时动作数据。若所采集基础数据为动作数据，则应将动作及其对应时刻数据转化成为状态数据使用。

4.1.2 人员逐时在室状态固定作息表以典型日内的人员逐时在室率表示。所需采集的基础数据包括逐时在室人数、或者人员移动动作及其对应时刻。人员逐时在室率(%)按下式计算：

$$x_i = \frac{\text{第 } i \text{ 小时在室人员数量}}{\text{典型日在室人员最大数量}} \quad (4.1.2)$$

式中： x ——人员逐时在室率；

i ——典型日内 i 时。

4.1.3 窗户逐时使用状态固定作息表以典型日内的逐时开窗状态、逐时通风换气次数或逐时通风量表示。所需采集的基础数据包括逐时的窗户开关状态(窗户开关动作及其对应时刻)、逐时的通风换气次数或逐时的通风量。逐时开窗状态、逐时通风换气次数或逐时通风量的描述应符合以下规定：

1 逐时窗户开关状态由 0,1 表示，0 表示窗户关闭，1 表示窗户开启；

2 逐时通风换气次数宜用最大通风换气次数与逐时通风换气次数比例共同表示，其中逐时通风换气次数比例(%)按下式计算：

$$x_i = \frac{\text{第 } i \text{ 小时通风换气次数}}{\text{典型日最大通风换气次数}} \quad (4.1.3 - 1)$$

式中： x ——逐时通风换气比例；

i ——典型日内 i 时。

3 逐时通风量宜用最大小时通风量与逐时通风量比例共同表示，其中逐时通风量比例(%)按下式计算：

$$x_i = \frac{\text{第 } i \text{ 小时通风量}}{\text{典型日最大小时通风量}} \quad (4.1.3 - 2)$$

式中： x ——逐时通风比例；

i ——典型日内 i 时。

4.1.4 照明逐时使用状态固定作息表以典型日内的逐时照明开关时间表示。所需采集的基础数据包括照明个数、各照明功率、各照明逐时使用状态、或者照明开关动作及其对应时刻。逐时照明开关时间的描述应符合以下规定：

1 照明逐时使用状态由 0,1 表示, 0 表示照明关闭, 1 表示照明开启;

2 逐时照明开关时间 (%) 按下式计算:

$$x_i = \frac{\text{第}i\text{小时处于开启状态的照明功率总和}}{\text{照明总功率}} \quad (4.1.4)$$

式中: x ——逐时照明开关时间;

i ——典型日内 i 时。

4.1.5 设备逐时使用状态固定作息表以典型日内的设备逐时使用率表示。所需采集的基础数据包括设备个数、各设备功率、各设备逐时使用状态、或者设备开关动作及其对应时刻。

设备逐时使用率的描述应符合以下规定:

1 设备逐时使用状态由 0,1 表示, 0 表示设备关闭, 1 表示设备开启;

2 逐时设备使用率 (%) 按下式计算:

$$x_i = \frac{\text{第}i\text{小时处于开启状态的设备功率总和}}{\text{设备总功率}} \quad (4.1.5)$$

式中: x ——逐时设备使用率;

i ——典型日内 i 时。

4.1.6 遮阳逐时使用状态固定作息表以典型日内的逐时遮阳使用状态表示, 0 表示遮阳关闭, 1 表示遮阳开启。所需采集的基础数据包括遮阳逐时使用状态、或者遮阳开关动作及其对应时刻。

4.1.7 空调逐时使用状态固定作息表由逐时空调运行状态、空调逐时温度两部分来描述。逐时空调运行状态以典型日内的逐时空调启停状态表示, 0 表示空调关闭, 1 表示空调开启。所需采集的基础数据包括逐时空调启停状态、或者空调开关动作及其对应时刻。空调逐时温度所需采集的基础数据为空调逐时设定温度。空调逐时温度 (°C) 按下式计算:

$$t_i = \frac{\text{第}i\text{小时空调设定温度为 } t_1 \text{ 的频次}}{\text{第}i\text{小时各空调设定温度总频次}} * t_1 + \dots + \frac{\text{第}i\text{小时空调设定温度为 } t_n \text{ 的频次}}{\text{第}i\text{小时各空调设定温度总频次}} * t_n \quad (4.1.7)$$

式中: t ——空调设定温度;

i ——典型日内 i 时;

n ——第 i 小时采集到的不同空调设定温度数。

4.2 位移模型参数及所需的基础数据

4.2.1 人员位移模型数据主要分为人员移动数据和人员在室率数据。

4.2.2 应根据事件发生机制定义工作日内会导致人员移动行为的典型事件, 根据人员移动的时间和空间特征, 获得人员在典型事件之间的转移概率。

4.2.3 人员在建筑中的分布情况可通过热图的形式获取或表达。

4.3 动作模型参数及所需的基础数据

4.3.1 影响人员动作的参数可根据所需的动作模型需求设定，行为数据可利用智能插座或者摄像头等方式获取，物理环境因素可利用传感器设备获取，生理、心理等主观因素可通过问卷或生理监测设备获取。

4.3.2 动作模型所需的参数采集时长和周期需要根据模型而定。

4.3.3 用能设备的动作模型应符合下列规定：

- 1 建筑中的设备散热量考虑受设备种类、功率和开启比例等因素影响。
- 2 办公设备以计算机为例，建立人员对该类设备的动作模型需要确定的影响因素有人员作息时间表、计算机配置比例、计算机运行状态、人员使用行为习惯和人员不同离开时长关闭计算机的概率。
- 3 计算机影响因素的数据采集周期和间隔应根据动作模型而定。
- 4 生活设备的使用应基于人员对该类设备的动作模型需要，确定关键影响因素，如人员作息时间表和人员用能的消耗水平等。

4.3.4 照明的动作模型应符合下列规定：

- 1 影响建筑中照明负荷的因素主要是照明的功率和照明开启数量。
- 2 人员产生开灯行为的室内工作面照度范围应根据案例建筑中人员的光环境需求和偏好而定。
- 3 单人办公房间内，建立人员对照明设备的动作模型需要确定的影响因素除了环境因素，还包括基于特定事件的行为习惯和节能意识。
- 4 多人办公房间内，建立人员对照明设备的动作模型需要确定的影响因素除了以上所述的因素外，还应考虑人员在室率对照明开启数量的分布影响。

4.3.5 开关窗动作模型应符合下列规定：

- 1 影响开关窗动作的因素主要是开关窗动作的发生时刻和与开关窗行为有显著相关的环境参数，通常为室内温度、室外温度、室内湿度、室外湿度、室内 CO₂ 浓度等。
- 2 建立室内人员对开关窗的动作模型所需的基础数据主要包括开关窗的动作时刻、动作时刻对应的环境参数、以及该时刻对应的开关窗动作。

4.3.6 空调启停的动作模型应符合下列规定：

- 1 影响建筑中空调启停动作的因素主要是其相关的热环境参数，通常为室内温度、室外温度等。
- 2 建立人员对空调的动作模型所需的基础数据主要包括开关空调的动作时刻、动作时刻对应的热环境参数、以及该时刻对应的空调开关动作等。

5 基础数据获取方法

5.1 获取方式

5.1.1 人行为数据包括人员移动数据和人员动作数据。人员移动数据可采用接触式设备、非接触式设备、移动互联网大数据、问卷调查方式获取，人员动作数据可采用设备调控行为测试装置、环境参数测试装置获取。在测试过程中，为了便于分析人行为数据，各测试设备的时间应同步，宜具备时钟校正功能，宜具备相同的测试起始时间和采样时间步长。

5.1.2 使用接触式设备采集人员位移及动作应以不对用户形成负担为原则，可使用基于Wifi型号的室内人员识别、佩戴可穿戴设备、佩戴射频识别传感器、三轴加速器或座椅压力传感器等设备。

5.1.3 使用非接触式设备采集人员的位移及动作应考虑室内人员隐私，可通过摄像头、深度相机、热成像仪、毫米波雷达、被动式红外传感器、CO₂传感器等设备，结合统计分析或机器学习算法获取人员在室和活动的情况，从而获得室内人员数量及移动规律，宜采用传感器融合技术提高检测准确性。非接触式数据获取方式可根据表 5.1.3 中各类型检测方法的特点进行选取。

表 5.1.3 不同类型非接触式检测方法

非接触式数据获取方式	采集数据类型	现有基础设施	隐私问题	数据处理复杂程度
摄像头、深度相机	是否有人、人员数目、人员轨迹、人员姿态	无	有	高
热成像仪	是否有人、人员数目、人员轨迹、人员姿态	部分	部分	高
毫米波雷达	是否有人、人员数目、人员轨迹、人员姿态	无	无	中
被动式红外传感器（PIR）	是否有人、人员数目	无	无	低
CO ₂ 传感器	是否有人、人员数目	部分	无	低

5.1.4 建筑内总人数可通过手机、电脑等移动终端采集的定位大数据获取。采用此方法获取建筑内总人数时，应考虑建筑类型及建筑尺寸。数据需注明建筑类型、采集时间以及采集样本占总体的比例。

5.1.5 可通过问卷调查的方法获取建筑人员位置信息以及行为习惯。问卷调查过程需根据

调查目的以及研究内容确定问卷调查内容及对象，样本数可参考表 5.1.5。

表 5.1.5 不同规模问卷调查样本容量范围

总体规模	抽样比
100 人以下	50%以上
100-999 人	20%-50%
1000-4999 人	10%-30%
5000-9999 人	3%-15%
10000-100000 人	1%-5%
100000 人以上	1%以下

5.1.6 人员对设备的调控行为应使用设备调控行为测试装置进行获取，包括用电设备开关及档位、开关窗行为及窗户开度、遮阳装置开度等，相关仪器设备如电量记录仪、磁开关记录仪、摄像头、超声波记录仪、拉绳位移记录仪、手机 APP 等。

5.1.7 记录人员对设备的调控行为的同时，应测试相对应的环境参数，包括温湿度、照度、CO₂、VOC、PM_{2.5} 等，进而分析人员对设备调控行为的触发条件模型。

5.2 采集方法

5.2.1 基础数据的采集包括人员移动数据的采集和人员动作数据的采集两个部分，其中人员移动数据的采集应符合本导则 5.2.2~5.2.7 的规定，人员动作数据的采集应符合本导则 5.2.8~5.2.12 的规定。

5.2.2 室内人员移动数据用于获取人员的移动规律，可根据不同的模拟目的选择对应的采集方法，包括人移动轨迹、房间是否有人、房间人员数量、建筑内人员密度和人员在室率等数据。

5.2.3 房间是否有人可用于判定房间内的设备是否发生调控行为的可能性，宜使用一个或多个被动式红外传感器、毫米波雷达布置在房间内，仪器测量范围应能覆盖整个房间，测量步长宜不低于 1min 一次。

5.2.4 房间人员数量能够可用于确定室内人员发热量、室内人员行为交互等信息，在没有隐私要求的房间宜采用摄像头配合人脸识别的方法进行采集，在隐私要求较高的房间宜采用座椅压力、人员计数器等设备，数据记录时间步长宜不低于 1min 一次。

5.2.5 人移动轨迹用来描述人在室外和室内、不同房间之间、同一房间内部的移动规律，没有隐私要求情况下宜使用摄像头采集，在隐私要求较高时宜使用基于 Wifi 型号的室内人员识别、佩戴可穿戴设备、佩戴射频识别传感器采集，可用的数据采集设备参见附录 A。

5.2.6 建筑内人员在室情况可以通过人员密度和人员在室率反映。应根据建筑能耗模拟目的，对不同空间、时间尺度选择不同的人员密度和人员在室率获取方式。

5.2.7 室内人员的动作数据用于获取人对设备的调控规律，可根据不同的设备选择对应的采集方法，设备调控行为、环境参数、人员位移的采集时间应保持一致。

5.2.8 室内人员的动作数据用于获取人对设备的调控规律，可根据不同的设备选择对应的采集方法，设备调控行为、环境参数、人员位移的采集时间应保持一致。

5.2.9 采暖/空调行为应采集空调器开关情况、环境温度和人员是否在室状况，可采集空调器设定温度、环境湿度、人员进入室内时间和停留时间，测量步长不低于 1min 一次。

5.2.10 开/关窗行为应采集窗的开启和关闭状态、人员是否在室状况，可采集窗的开度、室内外温湿度、空气品质等参数，净化器行为应采集净化器的开启和关闭状态、人员是否在室状况，可采集净化器开启档位、室内外空气品质等参数，测量步长不低于 1min 一次。

5.2.11 遮阳行为应采集遮阳装置的开启和关闭状态、人员是否在室状况，可采集遮阳装置的开度、室内照度、室外太阳辐射照度、眩光等参数，照明行为应采集灯具的开启、关闭和档位状态、人员是否在室状况，可采集室内照度、照度均匀度等参数，测量步长不低于 1min 一次。

5.2.12 个体舒适装置调控行为应采集装置的开启和关闭状态、人是否在个体舒适装置调控的微环境中，可采集不同种类个体舒适装置的档位、对应的环境状态参数等，测量步长不低于 1min 一次。

6 基础数据质量与预处理

6.1 基础数据质量要求

6.1.1 人行为基础数据的质量要求应符合以下规定：

- 1 应对第 5 章所述方法获取的人行为基础数据进行质量控制。质量控制过程中应包含对数据的准确性、完整性和一致性三个要素的检验。
- 2 应根据模拟用途以及所选择的行为描述方法，从数据采集范围、周期以及采集频率三个层面分析并检验人行为基础数据的完整性要素。
- 3 人行为基础数据的准确性应根据所获取的数据类型，参照现行国家和行业相关标准进行测量精度的检验。
- 4 应根据不同获取方法所采集的数据之间的相关性，对数据进行一致性检验。
- 5 在对人行为基础数据进行质量控制时，宜采用计算机自动处理和人工审核校验相结合的方法进行。

6.1.2 典型时间表的样本采集数量和采集周期应根据其数据获取方法和行为类别进行确定。

6.1.3 动作与位移模型应符合以下规定：

- 1 对于获取人员动作和位移的部分传感器，其检测量程和精度要求应符合表 6.1.3 的规定。

表 6.1.3 部分传感器的量程以及精度要求

测试参数	量程	精度
干球温度	-10~50°C	±0.5°C
相对湿度	10~100%	±5%
空气流速	0~5m/s	±(0.05+5%读数) m/s
黑球温度	0~60°C	±0.5°C
定向辐射热	-2kW/m ² ~2kW/m ²	±5 W/m ²
表面温度	~10~60°C	±1°C
PIR	半径 3~5 m	/
超声波	185 平方米	/

- 2 数据采集范围应涵盖人员使用和活动的主要区域，被采集对象应具有代表性。
- 3 数据的采集周期以及采集频率，宜根据建筑类型以及具体行为类别进行确定。

6.2 基础数据预处理方法

6.2.1 数据预处理应符合以下一般规定：

1 在数据质量控制过程中，宜按照数据清洗、数据集成与规约、数据转换与离散化三个基本步骤对人行为基础数据进行预处理。

2 应根据获取的数据类别和模拟应用场景等要求选择不同的数据清洗方法。

3 当存在多种数据源时，宜采用数据集成与规约对人行为基础数据进行预处理。

6.2.2 数据清洗应符合以下规定：

1 缺失数据处理应符合以下规定：

1) 人行为基础数据的缺失通常包含传感器数据缺失、调查类数据缺失和数据后处理缺失。在对人行行为数据进行数据清洗时，应对上述数据缺失进行处理。

2) 缺失数据处理方法主要包括保留缺失项、直接删除和缺失值填补。应根据人行行为数据类别以及缺失情况进行选择。

3) 可采用均值填补法、线性插值法、回归填补法和 KNN 填补法等相关方法对人行行为数据缺失值进行填补，具体缺失值处理方法见附录 B.1-B.3。

2 异常数据处理应符合以下规定：

1) 应根据 6.1.2 至 6.1.4 节中的规定，将不满足相关国家标准中规定的正常取值范围的数据视作可疑数据。

2) 可采用莱茵达准则、箱线法和聚类法等方法判断数据的正确性，并对可疑数据进行标识，具体方法见附录 B.4-B.6。

3) 宜进一步对异常数据进行人工校核以避免误删正确的异常数据。

4) 对具有高频噪音的时序类数据（如 CO₂ 浓度等），宜采用小波算法等时频分析方法对其进行去噪处理，具体方法及步骤见附录 B.7。

6.2.3 数据集成与规约应符合以下规定：

1 人行为基础数据之间通常存在较大的相关性，宜采用相关分析等方法对人行行为数据进行一致性检验后，将多源数据进行统一集成和存储。

2 人行行为相关数据通常具有特征多和数据量大等特点，宜采用主成分分析等数据规约方法实现降低数据维度的目的，具体方法及步骤见附录 B.8。

6.2.4 数据转换与离散化应符合以下规定：

1 在对具有不同量纲的人行行为相关数据进行处理和分析时，为避免量级较小数据的效果被掩盖，可采用 z-score 和 min-max 等标准化方法将所采集的数据转化为无量纲的数值，具体方法及步骤见附录 B.9-B.10。

2 在采用基于离散型变量的数学方法建立人行行为相关模型时，宜采用分箱法和聚类法等方法对连续型变量进行数据离散化处理。

7 基础数据分析方法

7.1 环境因素的影响分析

7.1.1 分析环境因素对用能行为的影响时，可用描述性统计，单因素分析，聚类分析，独立样本 t 检验，多因素方差分析，皮尔逊 (Pearson) 相关系数，斯皮尔曼(Spearman)相关系数，加权法，共线性诊断检验，Kaplan-Meier 生存分析等方法。

7.1.2 对调查总体所有变量的有关数据的描述性统计包括数据的频数分析、集中趋势分析、离散程度分析、分布以及一些基本的统计图形。

7.1.3 单因素分析是在一个时间点上对某一变量的分析，目的在于描述事实。

7.1.4 聚类分析是将物理或抽象对象的集合分组为由类似的对象组成的多个类的分析过程，其目标就是在相似的基础上收集数据来分类。

7.1.5 独立样本 t 检验是对于相互独立的两个来自正态总体的样本，利用独立样本的 T 检验来检验这两个样本的均值和方差是否来源于同一总体。

7.1.6 多因素方差分析是对一个独立变量是否受多个因素或变量影响而进行的方差分析。

7.1.7 皮尔逊 (Pearson) 相关系数可用于检验两个连续变量之间是否存在线性相关，相关系数在-1 到 1 之间，越接近 1 或-1，相关性越强。

7.1.8 斯皮尔曼(Spearman)相关系数可用于分析不服从正态分布的数据可采用斯皮尔曼相关系数，即描述两个变量之间的关联程度与方向。斯皮尔曼相关主要用于解决分类变量和连续变量相关的问题。

7.1.9 加权法指相应参数乘以权重系数的分析方法。可用于分析某些参数对空调开启率的综合影响。例，当室内温度和室外温度综合影响人行为时，采用加权分析这两个参数的影响程度。

7.1.10 共线性诊断是检验自变量间是否存在近似的线性关系，即某个自变量是否能近似的用其他自变量的线性函数来描述。

7.1.11 Kaplan-Meier 生存分析用于研究单一变量对生存的影响。

7.1.12 环境因素对建筑用能行为的影响的分析可按照表 7.1.11 选择分析方法。

表 7.1.11 建筑用能行为分析方法-环境因素

分析目的		分析方法
证明研究普遍性		独立样本 t 检验
相关性分析-环境因素	两个连续变量	单因素分析
		加权分析
		皮尔逊相关系数
	两个及以上连续变量	多因素方差分析
		共线性诊断
分类变量	斯皮尔曼相关系数	

	连续变量	
使用模式		聚类分析

7.2 非环境因素的影响分析

7.2.1 分析非环境因素对用能行为的影响时，可考虑一天中不同时间段、工作日/周末、季节、建筑类型、建筑高度、建筑朝向等因素。常用的方法有单因素 ANOVA 检验、独立样本 t 检验。

7.2.2 非环境因素的描述应符合以下规定：

- 1 对于一天中不同时间段的影响分析，应将一天中的时间按照需要划分为相同的区间，计算在每个区间内对应的开窗/空调/照明/遮阳概率，描述时间对开窗行为的影响。
- 2 对于工作日/周末的影响分析，应通过汇总被测建筑分别在工作日和周末的日均开窗/空调/照明/遮阳时长，对比分析出室内人员在工作日和周末的开窗/空调/照明/遮阳行为不同之处。
- 3 对于不同季节的影响分析应对比不同季节下，开窗/空调/照明/遮阳时长、日均开窗/空调/照明/遮阳次数等参数的差别。
- 4 对于不同建筑类型的影响分析：对比在相同因素下，不同建筑类型的开窗/空调/照明/遮阳概率，得出其对应的开窗行为特征，并通过对比得出几种类型建筑的区别。
- 5 对于建筑物高度的影响分析应符合以下规定：
 - 1) 对于独栋建筑，应将超高层建筑按照避难层划分为低层、中层及高层区三部分，记录的数据同时按朝向划分。记录该建筑不同朝向、不同高度的开窗/遮阳比例。
 - 2) 对于多栋建筑，应通过对比每栋建筑在同一时刻开窗/遮阳比例的不同，得出其不同建筑高度的开窗/遮阳行为特征。
- 6 对于建筑物朝向的影响分析，应记录一栋建筑不同朝向在相同时刻的开窗/遮阳比例，分析其不同朝向对开窗/遮阳行为的影响。

7.2.3 单因素 ANOVA 检验用于确定三个及其以上的数据组之间的均值是否具有统计差异。

7.2.4 独立样本 t 检验的具体原理及应用方法见 7.1.4。

7.2.5 非环境因素对建筑用能行为的影响的分析按照表 7.2.5 选择分析方法。

表 7.2.5 建筑用能行为分析方法-非环境因素

分析目的		分析方法
证明研究普遍性		独立样本 t 检验
相关性分析-非环境因素	两个分类变量	独立样本 t 检验
	两个以上分类变量	单因素 ANOVA 检验
使用模式		聚类分析

7.3 建筑用能行为特征分析

7.3.1 室内人员位移特征可使用以下指标分析，每个指标对应的推荐用图/表见附录 C.0.1。

- 1 占用率/在室率（%），即人员在某时刻在室时长与总测试时长比值。
- 2 房间人数（人），即室内人员数量。
- 3 工作日与周末在室比例（%），即工作日在室总时长与周末在室总时长之比。
- 4 转移概率（%），即人员在时刻 τ 处于位置 i 时，在时刻 $\tau+1$ 处于位置 j 的概率，亦即人员在时刻 τ 从子空间 i 出发，下一时刻移动到子空间 j 的概率。
- 5 在室/空置概率（%），即反映在一定时间下人员在室/不在室内的可能性大小。
- 6 占用/空置时长（h），即综合反映建筑中在测试阶段人员在室内/房间空置的时间长度。
- 7 人员每日活动时长（min），即在每天的活动中，第一次“到达”和最后一次“离开”的时间的差值对应于“每日活动”的持续时间。
- 8 每日/每周累计在室时长（min），即在测试期间内，人员每天/每周在室内的总时长。
- 9 各时刻在室时长（min），即一天 24 小时中，每一时刻对应的人员在室时长。
- 10 人员日均离开/到达房间次数（次），即房间每日已占用的活动数目与空置活动数目。

7.3.2 窗户使用行为特征可使用以下指标分析，每个指标对应的推荐用图/表见附录 C.0.2。

- 1 单次开启时长（分钟）即某一次窗户保持打开状态的时长。
- 2 每小时开启时长（分钟/小时）即窗户保持开启状态的总小时数占总测试小时数的百分比。
- 3 日均开窗时长（分钟/天）即各被测房间窗户开启总时长与测试天数的比值。
- 4 总开窗时长（分钟）即在测试期间，各被测房间的窗户开启总时长，通常采用柱状图来表示，被测房间作为横坐标，窗户开启时长作为纵坐标。
- 5 日均开窗次数（次/天）即测试期间窗户状态由关闭至开启的次数与总测试天数的比值。
- 6 开/关窗次数（次）即将各测试因素划分为不同的区间，在某区间内，窗户由关闭状态到打开状态（由打开状态到关闭状态）的次数。
- 7 窗户状态改变概率（%）即在某一时间段内，窗户状态改变的次数占总次数的比值。
- 8 开/关窗频率（%）即在某一时间段内的窗户状态由关闭到打开（由打开到关闭）这一动作发生的次数除以总开/关窗次数。
- 9 开窗概率/开窗频率/开启率（%）即将各测试因素划分为不同的区间，在某区间内，窗户处于开启状态的数量占窗户状态记录总和的百分比即为对应的开窗概率，通常使用点线图来表示，各因素划分的区间作为横坐标，开窗概率作为纵坐标。
- 10 开窗比例（%）即被测建筑在某因素的影响下开启的窗扇数占所有的窗扇数的比值。
- 11 开关比即所收集到的各房间的数据中，每个房间窗户打开状态与关闭状态的比值，理想中的开关比的等于或接近于 1，除可描述开窗特征外，还可用于对建模数据进行筛选，按下式计算：

$$r = \frac{N_0}{N_1} \quad (7.3.2)$$

式中： N_0 ——窗户为“开状态”的数据量，

N_1 ——窗户为“关状态”的数据量，

r ——开关比，%。

7.3.3 空调使用行为特征可使用以下指标分析，每个指标对应的推荐用图/表见附录 C.0.3。

1 逐时开启率（%）应符合以下规定：

- 1) 对于单台设备，逐时开启率（%）为每小时运行的累积时长与一小时之比。
- 2) 对于多台设备，逐时开启率（%）为每小时开启空调器数量与空调器总样本量之比。

2 每日开启率（%）应符合以下规定：

- 1) 对于单台设备，每日开启率（%）为当日空调器运行的累积时长与 24 小时之比。
- 2) 对于多台设备，每日开启率（%）为当日开启空调器数量与空调器总样本量之比。

3 基于影响因素的开启率（%），即在影响因素的单位范围内，空调器开启时间与监测时间之比。影响因素指环境因素或非环境因素，其可为连续变量或分类变量。

4 基于影响因素的累计开启率（%），即在影响因素的指定范围内，空调器开启率之和。

5 单次运行时长（h），即空调器在每一次相邻的开启动作与关闭动作之间所运行的时长。

6 单次运行时长 t 小时的占比（%），即单次运行时长 t 的数量与单次运行时长总样本量之比。

7 单次运行时长累积概率（%），即在单次运行时长的指定范围内，空调器开启率之和。

8 日均开启时长（小时），即空调器在监测时期的总运行时长与运行天数之比。

9 在开启日的运行次数（次），即若某天空调器开启，空调器在该开启日所运行的次数（每一次相邻的开启与关闭动作记为一次运行）。

10 在第 t 时刻的开启比例（%），即空调器在第 t 时刻运行的累积时长与空调器在所有时刻运行的累积时长之比。

11 开启/关闭动作发生在 t 时刻的占比（%），即开启/关闭动作发生在第 t 时刻的运行次数与空调器总运行次数之比。

12 温度设定值（℃），即非连续性操作时设定的空调器运行温度；连续性操作时最后设定的空调器运行温度。

13 日均温度设定动作频次（次），即对空调器进行温度设定动作的次数与运行天数之比。

14 日均开启频次（次），即空调器在监测时期的总开启次数与运行天数之比。

15 日均能耗（kWh/day），即空调器在监测时期的总能耗与运行天数之比。

16 整夜运行率（%），即空调器整晚运行的天数与总运行天数之比（若空调器在 23 时至次日 6 时之间的 95% 时间内处于开启状态，则定义为整晚运行）。

7.3.4 遮阳行为特征可使用以下指标分析，每个指标对应的推荐用图/表见附录 C.0.4。

- 1 遮阳调节概率（%），即遮阳位置被调整的次数与总实测次数之比。
- 2 遮阳上调比例（%），即遮阳设备上调（增加开启面积）的次数与总实测次数之比。
- 3 遮阳下调比例（%），即遮阳设备下调（增加关闭面积）的次数与总实测次数之比。
- 4 遮挡率（%），即遮阳设备关闭面积与总面积之比。
- 5 日均调节次数（次），即调整遮阳设备动作的次数与监测天数之比。
- 6 遮阳设备使用状态特征值 Y ，按下式取值：

$$Y = \begin{cases} 0, & \text{遮阳设备完全打开} \\ X_i, & \text{遮阳设备关闭面积与完全关闭面积之比} \\ 1, & \text{遮阳设备完全关闭} \end{cases} \quad (7.3.4)$$

- 7 遮阳特征值 Y 的比例（%），即相应 Y 取值的数量与所有取值数量之比。

7.3.5 室内照明行为特征可使用以下指标分析，每个指标对应的推荐用图/表见附录 C.0.5。

- 1 累计开灯时长（h），即测试（计算）期间灯具开启的总时长。
- 2 累计开灯次数（h），即测试（计算）期间灯具开启的总次数。
- 3 单次开灯时长（h），即每次开灯之后（直至关闭）灯具所开启的时长。
- 4 单次关灯时长（h），即每次关灯之后（直至打开）灯具所关闭的时长。
- 5 每次开灯最大时长（h），即每次开灯之后（直至关闭）灯具所开启的最大时长。
- 6 每次开灯最短时长（h），即每次开灯之后（直至关闭）灯具所开启的最短时长。
- 7 开灯照度阈值（lux），即当室内照度高于一定水平后，开灯动作发生的概率恒为 0，则称该照度水平为开灯照度阈值。
- 8 平均开灯时数（h），即测试（计算）期间灯具开启的平均时数。
- 9 最低开灯照度（lux），即室内要求达到的最低照度水平。
- 10 关灯延迟时间（h），即人员离开到灯关闭所延迟的时间。
- 11 开（关）灯概率（%），即反映在一定条件下开（关）灯动作发生的可能性大小。
- 12 平均每日人员在室时间（h），即测试（计算）期间人员在室的平均时长。
- 13 平均每日照明使用时间（h），即测试（计算）期间室内照明的平均时长。
- 14 照明开启比率（%），即在某一时间段，开启的灯具数占灯具总数的百分比。

附录 A 适用于隐私要求较高的基础数据采集设备

设备名称	精度范围	使用介绍
三轴加速器	$\pm 1.0g \sim \pm 5.0g$	可长时间持续监测人员身体活动情况，其体积小、重量轻、功耗低、成本低，非常适合应用于收集人员信息
无线路由器		对于用户的活动范围可以实现全方位覆盖，无需佩戴任何设备，不侵犯用户个人隐私
智能手环		易于佩戴，实时监控人员心率、代谢率等生理参数，配有三轴加速剂、陀螺测试仪、GPS 等装置
Kinect DK 深度相机	人数检测准确率： 98.12% 姿态检测准确率： 83%	Kinect 摄像机与计算机中 SDK 和 API 相结合可实现人员检测。
红外夜视人体测温热成像仪	姿态检测准确率： 87.44%	安装相机对应的计算机平台可得到人员数目和轨迹，与深度学习算法结合可得到人员姿态。
智能毫米波传感器标准天线插件模块	姿态检测准确率： 96.76%	传感器模块在软件开发套件 (SDK) 上进行应用评估和开发，执行与人行行为检测相应的程序代码。
路由器	人数检测归一化均方根误差：0.096 人员轨迹检测精度： 1.385 m	分析移动设备和 WiFi 路由器之间的流量，并通过在线顺序极限学习机算法实现人员数据采集。

<p>RFID 电子标签</p> <p>RFID 传感器</p> <p>RFID 协调器</p>	<p>人员轨迹检测精度： 0.799m</p>	<p>传感器通过无线传感器网络将所有跟踪标签的信息发送给 RFID 协调器，协调器发送给位置服务器，通过机器学习算法采集位置数据。</p>
<p>标准式 PIR 传感器</p> <p>长距离式 PIR 传感器</p> <p>墙装式 PIR 传感器</p>	<p>是否有人准确率： 92.75%</p>	<p>PIR 传感器的信号传输至微控制器，再传输至无线通信模块，最终传输到远程服务器，通过计数估计算法得到人员数目。</p>
<p>CO₂ 温湿度记录器</p>	<p>是否有人准确率： 94%</p> <p>人员数目准确率： 77%</p>	<p>CO₂ 浓度传输至计算机中，与最大似然法、卡尔曼滤波、优化算法、机器学习算法结合分析出是否有人以及人员数目。</p>

附录 B 数据预处理方法

B.1 缺失值处理方法

B.1.1 线性插值法利用待填补值相邻两点的已知数据,通过构造一次多项式计算待填补值。

线性插值法应按下式计算:

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1} \quad (\text{B.1.1})$$

式中: $(x_1, y_1), (x_2, y_2)$ ——已知点对应数值

(x, y) ——待填补值

B.1.2 回归填补法将待填补的缺失变量和其他相关变量分别作为因变量和自变量,建立多元线性回归模型,根据此模型可利用其他相关变量的数值计算缺失变量的填补值。回归填补法应按下式计算:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad (\text{B.1.2})$$

B.1.3 KNN 填补法为基于距离加权为待填补样本寻找距其最近的 K 个完整样本,将这 K 个样本在相应属性上的均值或加权平均值作为填补值。KNN 填补法应按以下步骤进行:

1 按如下公式计算待填补样本已知属性与完整样本的欧式距离:

$$L_2(x_i, x_j) = \left(\sum_{l=1}^n |x_i^{(l)} - x_j^{(l)}|^2 \right)^{\frac{1}{2}} \quad (\text{B.1.3})$$

2 根据欧式距离排序选取 K 个距离最近的完整样本, K 值可采用交叉验证法选取。

3 根据欧式距离计算加权平均值作为填补值。

B.2 异常数据处理方法

B.2.1 莱茵达准则规定变量的观测值 x_i 与其均值 μ 的残差 v_i 应不超过 3σ , σ 为变量的标准差,若超过此范围则视为异常数据,按下式计算:

$$\sigma = \frac{\sqrt{\sum_{i=1}^N (x_i - \mu)^2}}{N - 1} \quad (\text{B.2.1 - 1})$$

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (\text{B.2.1 - 2})$$

式中: N ——观测样本数量

x_i ——变量的观测值

μ ——均值

v_i ——残差

B.2.2 箱线法根据数据分布特征的四分位间距 IQR（即上四分位数 Q_1 和下四分位数 Q_3 之差）定义数据的取值边界值，若数据大于上边界值或小于下边界值则被视为异常点。上下边界值分别为 $Q_3+a \times IQR$ 和 $Q_1-a \times IQR$ ，其中 $IQR=Q_3-Q_1$ 。a 的取值通常为 1.5 或 3。

B.2.3 K-Means 聚类法将距离聚类中心较远或属于数据量较少、较稀疏的集群的数据视作异常值。K-Means 聚类法应按以下步骤进行：

- 1 任意选择 K 个样本作为初始的聚类中心；
- 2 计算每个样本与 K 个聚类中心的欧式距离；
- 3 根据每个样本与聚类中心的距离，将其划分至距离最近的集群；
- 4 重新计算每个集群的平均值并将其作为新的聚类中心；
- 5 重复上述步骤，直至聚类中心不再更新。

B.2.4 小波去噪利用小波变换对原始信号进行多尺度分解，并通过设定合理阈值去除原始信号中的高频噪音成分，应符合以下规定：

- 1 应按下式计算：

$$W_{\psi}x(a,b) = \langle x(t), \psi_{a,b}(t) \rangle \geq \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt \quad (\text{B.2.4})$$

式中：a——尺度参数，控制小波函数的伸缩，对应于频率，并与频率成反比；

b——平移参数，控制小波函数的平移，对应于时间。

- 2 应按以下步骤进行：

- 1) 设定分解层数，并计算小波系数（包括近似系数和细节系数）。为达到良好的去噪效果，宜根据需求选择适合的分解层数；
- 2) 合理确定小波各层细节的阈值，并采用阈值法去除集中在高频的噪音成分；
- 3) 利用经过阈值量化处理的各层高频系数和最高分解层次的低频系数进行信号重构，得到剔除噪声后的信号。

B.3 数据规约方法

B.3.1 主成分分析通过降维的思想将多个可能存在相关性的变量，经过正交变换转化为少数几个互相无关的综合变量的一种统计分析方法，应按以下步骤进行：

- 1 对原始数据矩阵 X 进行归一化处理；
- 2 根据公式 $C = \frac{1}{m} XX^T$ 计算特征的协方差矩阵，并计算特征值和特征向量；
- 3 确定主成分个数 k，主成分个数可根据累积方差贡献率确定。累积方差贡献率宜设置在 75%~95% 范围内，此时对应的 k 个主成分可包含原始变量中的绝大部分信息；
- 4 构造主成分矩阵，即将前 k 个成分对应的特征向量作为降维后的样本特征。

B.4 数据标准化方法

B.4.1 Z 分数标准化 (z-score Normalization) 对数据进行转换使其均值为 0、标准差为 1, 按下式计算:

$$x^* = \frac{x - \bar{x}}{\sigma} \quad (\text{B.4.1})$$

式中: x^* ——待转换变量的转换值

x ——待转换变量的原始值

\bar{x} ——待转换变量的均值

σ ——待转换变量的标准差

B.4.2 min-max 标准化 (min-max normalization) 将数值型数据转换至[0,1]范围内, 按下式计算:

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (\text{B.4.2})$$

式中: x^* ——待转换变量的转换值

x ——待转换变量的原始值

x_{\min} ——待转换变量的最小值

x_{\max} ——待转换变量的最大值

附录 C 各用能行为的指标分析推荐用图/表

C.0.1 室内人员位移特征指标分析推荐用图/表应按表 C.0.1 确定。

表 C.0.1 人员位移的指标分析推荐用图/表

统计指标	点线图	箱型图	折线图	柱形图	拟合曲线+散点	制表
占用率/在室率			√			
工作日与周末在室比例				√		
在室/空置概率					√	
占用/空置时长			√	√	√	√
人员每日活动时长			√			
每日/每周累计在室时长			√			
每时刻在室时长		√				
人员日均离开/到达房间次数				√		
房间人员位移信息						√

C.0.2 窗户使用行为指标分析推荐用图/表应按表 C.0.2 确定。

表 C.0.2 窗户使用行为的指标分析推荐用图/表

特征描述方法	饼图	折线图	点线图	柱状图	散点图	散点图+拟合曲线图	箱型图
窗户状态改变概率	√						
总开窗时长			√	√	√		√
日均开启时长			√	√	√		√
单次开窗时长					√		
每小时开窗时长					√		√
开窗概率		√	√	√	√	√	√
日均开窗次数							√
开窗次数				√			√
开/关窗比例		√	√				
开关窗频率	√		√			√	

C.0.3 空调使用行为指标分析推荐用图/表应按表 C.0.3 确定。

表 C.0.3 空调操作行为的指标分析推荐用图/表

行为特征	行为指标	核概率密度	散点图+拟合曲线	点线图	曲线图	柱状图	箱型图	制表
时刻特	逐时开启率			√	√			

征	在第 t 时刻开启比例							√	
	单次运行时长 t 小时的占比		√						
	开启/关闭动作发生在 t 时刻的占比					√	√	√	
动作特征	日均温度设定动作频次					√			√
	在开启日的运行次数					√		√	√
	日均开启频次					√			√
	每日开启率						√		√
	整夜运行率								√
运行时间	单次运行时长							√	√
	日均开启时长					√			√
	单次运行时长累积概率		√						
驱动因素	基于 S 因素的开启率				√				
	基于 S 因素的累计开启率		√						
操作情况	温度设定值								√
能耗	日均能耗					√	√		√
特征描述方法	饼图	折线图	点线图	柱状图	散点图	散点图+拟合		曲线图	箱型图
窗户状态改变概率	√								
总开窗时长			√	√	√				√
日均开启时长			√	√	√			√	√
单次开窗时长					√			√	
每小时开窗时长					√				√
开窗概率		√	√	√	√	√			√
日均开窗次数									√
开窗次数				√					√
开/关窗比例		√	√						
开关窗频率	√		√					√	

C.0.4 遮阳行为指标分析推荐用图/表应按表 C.0.4 确定。

表 C.0.4 遮阳行为的指标分析推荐用图/表

行为指标	核密度	柱状图	曲线图	点线图	散点图+拟合曲线	直方图	饼状图	箱型图	制表
遮阳调节概率	√	√		√					√
遮阳上调比例		√	√	√	√				
遮阳下调比例		√	√	√	√				
遮挡率			√	√					√
日均调节次数									√

遮阳设备使用 状态特征值 Y	√				√		√	
遮阳特征值 Y 的比例		√		√	√			
特征描述方法	饼图	折线图	点线图	柱状图	散点图	散点图+拟合	曲线图	箱型图
窗户状态改变概率	√							
总开窗时长			√	√	√			√
日均开启时长			√	√	√		√	√
单次开窗时长					√		√	
每小时开窗时长					√			√
开窗概率		√	√	√	√	√		√
日均开窗次数								√
开窗次数				√				√
开/关窗比例		√	√					
开关窗频率	√		√				√	

C.0.5 照明行为指标分析推荐用图/表应按表 C.0.5 确定。

表 C.0.5 照明行为的指标分析推荐用图/表

统计指标	散点图	点线图	直方图	制表
累计开灯时长	√		√	√
累计开灯次数	√		√	
单次开灯时长	√			
单次关灯时长	√			
每次开灯最大时长				√
每次开灯最短时长				√
开灯照度阈值	√			
平均开灯时数				√
最低开灯照度				√
关灯延迟时间				√
开灯（关）概率	√	√		√
平均每日人员在室时间			√	√
平均每日照明使用时间				√

统计指标	散点图	点线图	直方图	制表
照明开启比率				√

本导则用词说明

- 1 为便于在执行本导则条文时区别对待，对要求严格程度不同的用词说明如下：
- 1) 表示很严格，非这样做不可的用词：
正面词采用“必须”，反面词采用“严禁”；
 - 2) 表示严格，在正常情况下均应这样做的用词：
正面词采用“应”，反面词采用“不应”或“不得”；
 - 3) 表示允许有选择，在条件许可时首先应这样做的用词：
正面词采用“宜”，反面词采用“不宜”；
 - 4) 表示有选择，在一定条件下可以这样做的，采用“可”。
- 2 条文中指明应按其他有关标准执行的写法为：“应符合……的规定”或“应按……执行”。

引用标准名录

- 1 《公共建筑节能设计标准》 GB 50189-2015
- 2 《信息技术数据质量评价指标》 GB/T 33850-2017
- 3 《工业数据质量 通用技术规范》 GB/T 39400-2020
- 4 《建筑热环境测试方法标准》 JGJ/T 347-2014

中国建筑节能协会团体标准

建筑能耗模拟用人为基础数据采集技术导则

T/CABEE XXX-2020

条文说明

编制说明

《建筑能耗模拟用人为基础数据采集技术导则》T/CABEE 00X-2021 经中国建筑节能协会 2021 年 X 月 XX 日以第 X 号公告批准发布。

本导则的主要内容包括：1.总则；2.术语；3 建筑用能行为描述方法；4.建筑模拟用人为描述参数及基础数据；5. 数据获取方法；6. 人行为数据质量要求与预处理；7. 数据分析方法；附录；本导则用词说明；引用标准名录。

为了便于读者在使用本导则时能正确理解和执行条文规定，《建筑能耗模拟用人为基础数据采集技术导则》编制组按章、节、条顺序编制了本导则的条文说明，对条文规定的目的、依据及执行中需注意的有关事项进行了说明。但是，条文说明不具备与标准正文同等的法律效力，仅供使用者作为理解和把握标准规定的参考。

目 次

1 总 则	34
2 术语与符号	35
3 建筑用能行为描述方法	36
3.1 人员移动.....	36
3.2 人员动作.....	44
4 描述参数及基础数据定义	52
4.1 固定作息表及所需的基础数据.....	52
4.2 位移模型参数及所需的基础数据.....	52
4.3 动作模型参数及所需的基础数据.....	52
5 基础数据获取方法	55
5.1 获取方式.....	55
5.2 采集方法.....	58
6 基础数据质量与预处理	61
6.1 基础数据质量要求.....	61
6.2 基础数据预处理方法	62
7 基础数据分析方法	64
7.1 环境因素的影响分析.....	64
7.2 非环境因素的影响分析	64

1 总 则

1.0.1 本导则只关注与建筑用能有关的人行为，包括人员移动行为和人员动作行为。动作行为主要包括各种对建筑环境进行控制的行为，如开窗行为、空调操作行为、遮阳调节行为、照明开关与调节行为等。

2 术语与符号

2.1.4 位移模型应适当简化降低复杂度，以方便进行人行为调研分析且易于应用于建筑模拟计算并保证模拟结果的精度。

2.1.5 动作模型应适当简化降低复杂度，以方便进行人行为调研分析且易于应用于建筑模拟计算并保证模拟结果的精度。

2.1.9 常用的接触式测量装置包括三轴加速器或座椅压力传感器、室内人员佩戴可穿戴设备或射频识别传感器，以及基于 Wifi 信号的室内人员识别等。

2.1.10 非接触测量设备一般指摄像头、深度相机、热成像仪、毫米波雷达、被动式红外传感器、CO₂ 传感器等设备，结合统计分析或机器学习算法获取人员在室和活动的情况，从而获得室内人员数量及移动规律，宜采用传感器融合技术提高检测准确性。

3 建筑用能行为描述方法

3.1 人员移动

3.1.1 固定作息表方法用于描述房间人员在室状况，方法简单，属于建筑能耗模拟软件中的基本方法。这种方法需要预先设置每类房间的最大人数及逐小时在室率，不能描述房间人数在各个时刻的随机变化以及不同房间在同一时刻的人数差异与关联，很难保证建筑总人数的动态守恒。基于马氏链与事件的室内人员移动模型是一种新的描述人员个体随机移动过程的方法，能够较好地描述人员随机性、时相关性和空间互相关性，具有简单清晰的特征参数，同时便于程序实现人员移动过程和室内人员在室状况。

3.1.2 房间人员作息表一般按典型房间功能类型和典型日（工作日、节假日等）进行预先设置。以 DeST 默认设定为例，见表 1，其中民用住宅以起居室和卧室为例，商用建筑以办公室和会议室为例，运行时段分为工作日和节假日。

表 1 房间人员逐时在室率

建筑类别	房间类型	运行时段	下列计算时刻 (h) 房间人员逐时在室率 (%)												
			1	2	3	4	5	6	7	8	9	10	11	12	
民用住宅	起居室	工作日	0	0	0	0	0	0	0	0	33	33	0	0	0
		节假日	0	0	0	0	0	0	0	0	0	33	33	66	66
	卧室	工作日	100	100	100	100	100	100	100	100	0	0	0	0	0
		节假日	100	100	100	100	100	100	100	100	100	50	50	0	0
商用建筑	办公室	工作日	0	0	0	0	0	0	0	0	0	100	100	100	100
		节假日	0	0	0	0	0	0	0	0	0	30	30	30	30
	会议室	工作日	0	0	0	0	0	0	0	0	7	100	100	100	100
		节假日	0	0	0	0	0	0	0	0	0	38	38	38	38
建筑类别	房间类型	运行时段	下列计算时刻 (h) 房间人员逐时在室率 (%)												
			13	14	15	16	17	18	19	20	21	22	23	24	
民用住宅	起居室	工作日	0	0	0	0	0	33	33	100	100	100	33	0	
		节假日	66	33	33	33	66	66	66	100	100	100	100	33	
	卧室	工作日	0	0	0	0	0	0	0	0	0	0	50	100	
		节假日	0	0	50	50	0	0	0	0	0	0	0	50	
商用	办公室	工作日	30	100	100	100	100	50	10	10	10	0	0	0	

建筑	节假日	10	30	30	30	30	30	10	10	10	0	0	0
	工作日	25	25	100	100	100	100	25	25	25	25	7	0
	节假日	13	13	25	25	25	25	13	13	13	13	0	0

3.1.3 基于马尔可夫链及事件机制的人员移动过程:假设一个建筑物内部有 n 个房间（内部子空间），每个房间编号依次为 $1, 2, \dots, n$ ，同时把外界（建筑物外部）当作一个特殊子空间，编号为 0 。这些子空间就构成一个具有 $n + 1$ 个节点的封闭拓扑网络（闭图），每个节点代表一个子空间，人员位置用子空间节点编号进行标识。人员在建筑内外各个子空间之间移动，其位置状态可视为随机变量。如果他的移动范围覆盖了所有子空间（也可以是其中一个子集），则其位置状态可能的全部取值为 $\{0 = \text{outside}, 1 = \text{room1}, 2 = \text{room2}, \dots, N = \text{roomN}\}$ 。人员在各个时刻的位置 X_τ ，就构成一个随机时间序列 $\{X_\tau\}$ 。用马氏链近似表示这个位置序列，即在任意时刻 $\tau + 1$ ，人员位置 $X(\tau + 1)$ 仅与上一时刻的位置 $X(\tau)$ 有关。转移概率 $p_{ij}(\tau) = P\{X_{\tau+1} = j | X_\tau = i\}$ 表示人员在时刻 τ 处于位置 i 时，在时刻 $\tau + 1$ 处于位置 j 的概率，亦即人员在时刻 τ 从子空间 i 出发，下一时刻移动到子空间 j 的概率，如图 1 所示。由全部转移概率 p_{ij} 所组成的矩阵为马氏链 $\{X_\tau\}$ 的转移矩阵，见式（1）。

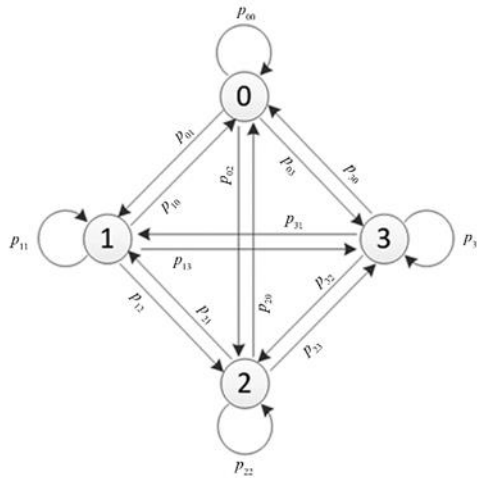


图 1 人员位置的转移概率示意图

$$P_\tau = (p_{ij})_{(n+1) \times (n+1)} = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \dots & p_{0n} \\ p_{10} & p_{11} & p_{12} & \dots & p_{1n} \\ p_{20} & p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n0} & p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix} \quad (1)$$

3.1.4 移动事件（下文统一简称为事件）具有以下几个基本属性：

- 1 起始时间和结束时间：表示事件的发生和作用时段。在这个时间段内事件得到激活，并在某一时刻随机发生。例如，上班事件在早上某个时段内随机发生、下班事件在下午某个时段内随机发生等。

- 2 起始位置和目标位置：表示事件所对应的人员房间位置变化，而人员转移矩阵 P 中对应于这些房间位置的的概率元素，会受到事件的影响和控制。例如，上班事件对应着人员从室外进入办公室、下班事件对应着人员从办公室进入室外，而上下班事件只影响转移矩阵中与这两个位置相对应的概率元素。
- 3 特征参数：用于表征人员日常活动的数字统计特性，例如平均上班时间、平均下班时间等。这些特征参数的优点在于相对直观、易于理解，并且可以与人员转移矩阵中受事件控制的概率元素建立联系，实现数值上的相互转换。
- 4 作用方式：各个事件对人员移动过程的控制方式与作用途径是一致的，即事件在其作用时段内激活，根据其特征参数对人员转移矩阵中对应元素进行设置与修改，从而确定各个时刻人员的转移概率，从而预测下一时刻的人员位置。
- 5 优先级：用于解决多个事件之间的冲突问题。当不同事件的作用时段有重叠时，比较其优先级次序，只有具备最高优先级的事件才能发生，同一时刻只有唯一事件设置和修改人员转移矩阵。

3.1.5 人员在各个房间停留的时间比例和人员在各个房间平均每次逗留的时间与 P 矩阵的联系，以及设置 P 矩阵的方法基于以下几个数学关系式。

1 极限分布（平稳分布）

将人员随机走动过程的转移矩阵记为 P 。在不考虑其他事件影响的情况下， P 矩阵是时不变的。在我们的问题中，代表人员随机走动的马氏链是不可约和遍历的，具有唯一平稳分布，记为 S ， $S = (S_0, S_1, \dots, S_n)$ ，其中 S_i 表示时间充分长以后马氏链处于状态 i 的概率，也就是长期来看，人员在子空间 i 中停留的时间比例。而且有：

$$\sum_{i=0}^n S_i = 1 \quad (2)$$

$$S = SP \quad (3)$$

如果已知 P ，联立上述两式，即可求解向量 S ，可表示为：

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 \\ p_{01} & p_{11} - 1 & \dots & p_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ p_{0n} & p_{1n} & \dots & p_{nn} - 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (4)$$

$$S^T = A^{-1}b \quad (5)$$

其中 S^T 为向量 S 的转置。

2 平均逗留时间

马氏链在状态 i 的停留时间记为 ST_i ，它是一个如式(6)所示的几何分布（在连续马氏链中它就是指数分布）。其中， $P\{ST_i = k\}$ 表示从马氏链进入状态 i 开始、 k 步之后离开状态 i 的概率。

$$P\{ST_i = k\} = p_{ii}^{k-1}(1 - p_{ii}) \quad (6)$$

ST_i 的期望值，即平均逗留时间 $E(ST_i)$ ，可表示为：

$$E(ST_i) = \sum_{k=1}^{\infty} k \cdot P\{ST_i = k\} = \sum_{k=1}^{\infty} k \cdot p_{ii}^{k-1}(1 - p_{ii}) = \frac{1}{1 - p_{ii}} \quad (7)$$

$$p_{ii} = 1 - \frac{1}{E(ST_i)} \quad (8)$$

马氏链在各个状态的平均逗留时间，记为 Est ， $Est = (Est_0, Est_1, \dots, Est_n)$ ，其中 Est_i 表示马氏链在状态 i 的平均逗留时间，也就是人员每次在子空间 i 中停留的平均时间。式(6)和式(8)说明了 P 矩阵和停留时间比例和平均逗留时间的关系。如果给定人员在各个子空间的停留时间比例 π 和平均逗留时间 Est ， P 矩阵的设置就转化为如下最优化问题：

$$\begin{aligned} & \min \|(A^{-1}b)^T - \pi\|_2 \\ & \text{s. t. } p_{ij} \geq 0 \\ & \sum_j p_{ij} = 1 \\ & p_{ij} = 1 - \frac{1}{Est_i} \end{aligned} \quad (9)$$

其中 $\tilde{S} = (A^{-1}b)^T$ 表示对 S 的估计。这个最优化问题可使用 Matlab 软件自带的 `fmincond` 函数求解。这样，具有 $(n+1) \times (n+1)$ 个元素的 P 矩阵，通过 $2 \times (n+1)$ 个参数就能加以确定，从而大大降低了输入参数的复杂度，特别是建筑内有多个房间的情况。

3.1.6 条文说明如下：

1 上班:上班事件（即早晨的到达）对应于从室外到个人办公室的位置变化，它只涉及 P 矩阵中与这两个子空间相对应的行和列上的元素。其有效时段通常是工作日办公时间之前的某段时间，例如早上 7:00 到 8:30，分别对应着办公人员最早和最迟的上班（到达办公室）时间 FA_1 ， FA_2 。

上班到达的过程可以用一个具有吸收态的两状态马氏链表示：

$$P_{go_office} = \begin{matrix} & 0 & 1 \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{bmatrix} p_{00} & p_{01} \\ 0 & 1 \end{bmatrix} \end{matrix} \quad (10)$$

其中，0,1 分别表示室外子空间和个人办公室的编号，个人办公室是一个吸收态。这个吸收态意味着人员一定会在某个时间进入他的办公室，而首次进入吸收态的时间就是人员到达办公室的时间。在上班事件的有效时段内，人员 P 矩阵中对应于室外和个人办公室的元素将按式（10）进行设置。假设 p_{00}, p_{01} 不随时间发生变化，那么，从上班事件的起始时间 FA_1 （7:00）算起，早上到达的时间 $\frac{FA - FA_1}{\Delta t} + 1$ （即从 FA_1 开始经历多少步进入吸收态）是一个几何分布，平均到达时间（记为 $E(FA)$ ）可表示为：

$$E(FA) = \frac{1}{1 - p_{00}} \quad p_{00} = 1 - \frac{1}{E(FA)} \quad (11)$$

如果上班时间与正常办公时间相同（例如上班事件起始时间和结束时间都为 8:00），则有 $p_{00} = 0$ ， $E(FA) = 1$ 。此时，上班事件将不再随机，而是一个确定的时间，效果与固定作息方式一致。上班事件优先级为 2。上班事件发生后，人员将开始办公楼内的随机走动过程。

- 2 下班:与上班事件的定义类似，下班事件（即晚上的离开）对应于人员从办公室到室外的位置移动。它的有效时段通常是工作日办公时间之后的某段时间，例如晚上 17:00 到 21:00，分别对应着办公人员最早和最晚的下班（离开办公室）时间 LD_1 、 LD_2 。下班离开过程可以用一个具有吸收态的两状态马氏链表示：

$$P_{\text{offwork}} = \begin{matrix} & 0 & 1 \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{bmatrix} 1 & 0 \\ p_{10} & p_{11} \end{bmatrix} \end{matrix} \quad (12)$$

其中，0,1 分别表示室外子空间和个人办公室的编号，而室外子空间是一个吸收态，意味着人员一定会在某个时间离开他的办公室，而首次进入吸收态的时间就是人员离开办公室的时间。在下班事件的有效时段内，人员 P 矩阵中对应于室外和个人办公室的元素将按式 (12) 进行设置。与上班事件类似，从下班事件的起始时间 LD_1 （17:00）算起，下午离开的时间 $\frac{LD-LD_1}{\Delta\tau} + 1$ （即从 LD_1 开始经历多少步进入吸收态）是一个几何分布，平均离开时间（记为 $E(LD)$ ）可表示为：

$$E(LD) = \frac{1}{1 - p_{11}} \quad p_{11} = 1 - \frac{1}{E(LD)} \quad (13)$$

如果下班时间与正常办公时间相同（例如下班事件起始时间和结束时间都为 17:00），则有 $p_{11} = 0$ ， $E(LD) = 1$ 。此时，下班时间是一个确定的时间点。另外，下班事件可反映工作日的加班状况，平均加班时间=平均离开时间 $E(LD)$ -规定下班时间 17:00。下班事件优先级为 2。下班事件发生后，人员将结束办公楼内的随机走动过程。

- 3 午餐:办公建筑中的午餐（午休）时间可分为两个事件。一个是出去吃午饭，表示午餐时间的开始，另一个是吃完饭返回办公室，表示午餐时间的结束。假设午餐地点是室外。这两个事件的处理方式类似于上下班事件。出去吃午饭的过程可用式 (14) 转移矩阵表示，其中 0、1 分别表示室外子空间和个人办公室的编号。从去吃午饭事件的起始时间算起（即最早的出发时间），平均出发时间（记为 $E(LL)$ ）可用式 (15) 表示。本事件之后，人员暂时离开随机走动过程。吃完饭返回办公室的过程可用式 (16) 的转移矩阵表示，其中 0、1 分别表示室外子空间和个人办公室的编号。从饭后返回事件的起始时间算起（即最早的返回时间），平均返回时间（记为 $E(LR)$ ）可用式 (17) 表示。本事件之后，人员再次进入随机走动过程。

$$P_{\text{lunchout}} = \begin{matrix} 0 & 1 \\ 1 & \begin{bmatrix} 1 & 0 \\ p_{10} & p_{11} \end{bmatrix} \end{matrix} \quad (14)$$

$$E(\text{LL}) = \frac{1}{1 - p_{11}} \quad p_{11} = 1 - \frac{1}{E(\text{LL})} \quad (15)$$

$$P_{\text{lunchback}} = \begin{matrix} 0 & 1 \\ 1 & \begin{bmatrix} p_{00} & p_{01} \\ 0 & 1 \end{bmatrix} \end{matrix} \quad (16)$$

$$E(\text{LR}) = \frac{1}{1 - p_{00}} \quad p_{00} = 1 - \frac{1}{E(\text{LR})} \quad (17)$$

午餐两个事件的优先级均为 2。

- 4 会议:在办公建筑中还有一类典型事件——会议，它对人员位置移动有很大影响，是造成室内人员分布不均匀的重要原因。一般而言，会议事件往往具有很强的随机性，很难通过固定作息的方式进行描述。由于会议事件通常发生在会议室，所以会议事件发生的频次与时长可以通过会议室的使用状况进行刻画。因此，本文提出一种基于两状态马氏链的方法来模拟随机会议事件。将会议室的使用状态（是-1，否-0）作为随机变量，其状态变化用马氏链描述，转移概率矩阵可表示为：

$$P_{\text{meet}} = \begin{matrix} 0 & 1 \\ 1 & \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} \end{matrix} \quad (18)$$

这个 2×2 的矩阵可由两个特征参数唯一确定。设会议室使用时间比例为 α （即被使用状态最终占全部工作时间的比例， $0 < \alpha < 1$ ）、会议平均时长为 β （即会议在使用状态 1 下的平均逗留时间，以单位时间步长来计算，例如取会议平均时长 1h，时间步长 5min，则 $\beta = 12$ ），则有：

$$\begin{aligned} p_{11} &= 1 - \frac{1}{\beta} \\ p_{10} &= 1 - p_{11} \\ p_{00} &= 1 - \frac{\alpha(1 - p_{11})}{1 - \alpha} = 1 - \frac{\alpha}{\beta(1 - \alpha)} \\ p_{01} &= 1 - p_{00} \end{aligned} \quad (19)$$

为确定与会人数，还必须定义最少最多与会人数、会议类型（人员构成）等，如表 2 所示。会议类型中，“组会”是指由单个房间的人所举行的会议（组内会议），“内部交流”是指多个房间的人所举行的会议（组间交流，此时应满足最小人数限制），各自所占比例按均匀分布进行处理。会议事件的优先级为 3。会议事件发生后，人员结束随机走动过程；会议事件结束后，人员恢复随机走动过程。

- 5 对于其他的典型事件，例如节假日加班、出差等都可以参照会议事件进行定义，此处不再赘述。将上述办公建筑移动事件的属性参数做一个汇总，见表 2。

表 2 办公建筑典型移动事件

办公移动事件	描述	特征参数	对应的日常含义
随机走动	办公楼内外各个房间之间的走动	停留时间比例	在各个房间停留的时间比例
		平均逗留时间	在各个房间平均每次的逗留时间
上班	从室外到个人办公室的位移，发生后进入随机走动过程	起始时间	最早到达时间
		结束时间	最晚到达时间
		平均上班时间	平均到达时间
下班	从个人办公室到室外的位移，发生后结束随机走动过程	起始时间	最早离开时间
		结束时间	最晚离开时间
		平均下班时间	平均离开时间
出去吃午饭	从办公室到室外的位移，发生后停止随机走动过程	起始时间	最早出发时间
		结束时间	最晚出发时间
		平均出发时间	平均出发时间
吃完饭返回办公室	从室外到办公室的位移，发生后恢复随机走动过程	起始时间	最早返回时间
		结束时间	最晚返回时间
		平均返回时间	平均返回时间
开会	从办公室到会议室的位移，发生后停止随机走动过程。基于会议室进行定义。	使用时间比例	会议室处于使用状态的时间比例
		会议平均时长	每次开会所占用的时长
		最少与会人数	最小与会人数
		会议类型与人员构成	与会人员的构成方式，组会或内部交流

3.1.7 住宅建筑中事件的定义与办公建筑类似，主要的区别在于住宅中需要考虑人员的睡眠活动：在这个时段内，人员所在的房间位置就是卧室，人员处于非活动状态，不会对房间内的设备有任何操作。因此要为人员定义一个新的状态变量，其状态变化可用以下转移矩阵表示：

$$P_{\text{wake}} = \begin{matrix} 0 & 1 \\ 0 & [p_{00} & p_{01}] \\ 1 & [p_{10} & p_{11}] \end{matrix} \quad (20)$$

其中，0、1 分别表示人员处于睡眠状态（非活动状态）、清醒状态（活动状态）。

1 起床事件对应于人员从非活动状态转为活动状态，而位置不发生变化。它对 P_{wak} 矩阵进行控制。其有效时段通常是早上的某段时间，例如早上 7:00 到 8:00，分别对应着人员最早和最迟的起床时间 GU_1 ， GU_2 。起床事件发生的过程用一个具有吸收态的两状态马氏链表示：

$$P_{\text{get up}} = \begin{matrix} 0 & 1 \\ 0 & [p_{00} & p_{01}] \\ 1 & [0 & 1] \end{matrix} \quad (21)$$

其中，0,1 分别表示睡眠状态、清醒状态，清醒状态是一个吸收态。这个吸收态意味着人员一定会在某个时间起床，而首次进入吸收态的时间就是人员起床的时间。在起床事件的有效时段内，人员 P_{wak} 矩阵中的元素将按式（21）进行设置。假设 p_{00}, p_{01} 不随时间发生变化，那么，从起床事件的起始时间 GU_1 （7:00）算起，起床的时间 $\frac{GU-GU_1}{\Delta\tau} + 1$ （即从 GU_1 开始经历多少步进入吸收态）是一个几何分布，平均起床时间（记为 $E(GU)$ ）可表示为：

$$E(GU) = \frac{1}{1 - p_{00}} \quad p_{00} = 1 - \frac{1}{E(GU)} \quad (22)$$

如果起床事件的起始时间和结束时间相同（平均起床时间也必然相同，例如都为 8:00），则有 $p_{00} = 0$ ， $E(GU) = 1$ 。此时，起床时间将不再随机，而是一个确定的时间，效果与固定作息方式一致。起床事件的优先级为 1。起床事件发生后，人员将开始住宅内的随机走动过程。

2 与起床事件的定义类似，睡觉事件对应于人员从活动状态转为非活动状态，位置则从其他房间转到其卧室。它的有效时段通常是夜间某段时间，例如晚上 22:00 到 23:00，分别对应着人员最早和最晚的睡觉时间 GS_1 、 GS_2 。睡觉时间发生的过程用一个具有吸收态的两状态马氏链表示：

$$P_{\text{go sleep}} = \begin{matrix} 0 & 1 \\ 0 & [1 & 0] \\ 1 & [p_{10} & p_{11}] \end{matrix} \quad (23)$$

其中，0,1 分别表示睡眠状态、清醒状态，睡眠状态是一个吸收态，意味着人员一定会在某个时间睡觉，而首次进入吸收态的时间就是人员睡觉的时间。在睡觉事件的有效时段内，人员 P_{wak} 矩阵中的元素将按式（23）进行设置。与起床事件类似，从睡觉事件的起始时间 GS_1 （22:00）算起，睡觉的时间 $\frac{GS-GS_1}{\Delta\tau} + 1$ （即从 GS_1 开始经历多少步进入吸收态）是一个几何分布，平均睡觉时间（记为 $E(GS)$ ）可表示为：

$$E(GS) = \frac{1}{1 - p_{11}} \quad p_{11} = 1 - \frac{1}{E(GS)} \quad (24)$$

如果睡觉事件的起始时间和结束时间相同(平均睡觉时间也必然相同,例如都为 22:00),则有 $p_{11} = 0$, $E(GS) = 1$ 。此时,睡觉时间是一个确定的时间点。睡觉事件的优先级为 1。睡觉事件发生后,人员将结束住宅内的随机走动过程。

3 上班、下班事件与办公建筑类似,对于其他的典型事件,例如节假日出去玩、出差等都可以进行类似定义,此处不再赘述。将上述住宅移动事件的属性参数做一个汇总,见表 3。

表 3 住宅建筑典型移动事件

住宅移动事件	描述	特征参数	对应的日常含义
随机走动	住宅内外各个房间之间的走动	停留时间比例	在各个房间停留的时间比例
		平均逗留时间	在各个房间平均每次的逗留时间
起床	从睡眠状态到清醒状态,位置不变,发生后进入随机走动过程	起始时间	最早起床时间
		结束时间	最晚起床时间
		平均起床时间	平均起床时间
睡觉	从清醒状态到睡眠状态,转到卧室,发生后结束随机走动过程	起始时间	最早睡觉时间
		结束时间	最晚睡觉时间
		平均睡觉时间	平均睡觉时间
上班	从住宅到室外的位移,发生后停止随机走动过程	起始时间	最早上班时间
		结束时间	最晚上班时间
		平均上班时间	平均上班时间
下班	从室外到住宅的位移,发生后恢复随机走动过程	起始时间	最早下班时间
		结束时间	最晚下班时间
		平均下班时间	平均下班时间

3.2 人员动作

3.2.2 固定作息表方法用于描述房间设备使用状况,方法简单,属于目前建筑能耗模拟软件中使用最多的方法。但它很难体现空调、照明等设备使用受室内外物理环境条件的影响,也不能描述人员行为的随机性。针对这类与环境条件相关的行为,宜采用条件概率模型、logistic

模型等方法。

3.2.3 条文说明如下：

- 1 房间空调温度作息表以 DeST 中各类型房间的空调设定最高温度为例，见表 4。其中民用住宅以起居室和卧室为例，商用建筑以办公室和会议室为例，运行时段分为工作日和节假日。空调启停作息表同样以 DeST 默认设定为例，见表 5。其中民用住宅以起居室和卧室为例，商用建筑以办公室和会议室为例，运行时段分为工作日和节假日。

表 4 房间空调设定温度

建筑类别	房间类型	运行时段	下列计算时刻 (h) 空调设定最高温度 (°C)											
			1	2	3	4	5	6	7	8	9	10	11	12
民用住宅	起居室	工作日	26	26	26	26	26	26	26	26	26	26	26	26
		节假日	26	26	26	26	26	26	26	26	26	26	26	26
	卧室	工作日	26	26	26	26	26	26	26	26	26	26	26	26
		节假日	26	26	26	26	26	26	26	26	26	26	26	26
商用建筑	办公室	工作日	25	25	25	25	25	25	25	27	25	25	25	25
		节假日	25	25	25	25	25	25	25	27	25	25	25	25
	会议室	工作日	26	26	26	26	26	26	26	26	26	26	26	26
		节假日	26	26	26	26	26	26	26	26	26	26	26	26
建筑类别	房间类型	运行时段	下列计算时刻 (h) 空调设定最高温度 (°C)											
			13	14	15	16	17	18	19	20	21	22	23	24
民用住宅	起居室	工作日	26	26	26	26	26	26	26	26	26	26	26	26
		节假日	26	26	26	26	26	26	26	26	26	26	26	26
	卧室	工作日	26	26	26	26	26	26	26	26	26	26	26	26
		节假日	26	26	26	26	26	26	26	26	26	26	26	26
商用建筑	办公室	工作日	25	25	25	25	25	25	25	25	25	25	25	25
		节假日	25	25	25	25	25	25	25	25	25	25	25	25
	会议室	工作日	26	26	26	26	26	26	26	26	26	26	26	26
		节假日	26	26	26	26	26	26	26	26	26	26	26	26

表 5 空调启停状态 (0 表示空调关闭, 1 表示空调开启)

建筑类别	房间类型	运行时段	下列计算时刻 (h) 空调启停状态 (0 和 1)											
			1	2	3	4	5	6	7	8	9	10	11	12

			1	2	3	4	5	6	7	8	9	10	11	12
民用住宅	起居室	工作日	0	0	0	0	0	0	0	0	0	0	0	0
		节假日	0	0	0	0	0	0	0	0	1	1	1	1
	卧室	工作日	1	1	1	1	1	1	1	0	0	0	0	0
		节假日	1	1	1	1	1	1	1	1	1	1	1	1
商用建筑	办公室	工作日	0	0	0	0	0	0	0	1	1	1	1	1
		节假日	0	0	0	0	0	0	0	1	1	1	1	1
	会议室	工作日	0	0	0	0	0	0	1	1	1	1	1	1
		节假日	0	0	0	0	0	0	0	1	1	1	1	1
建筑类别	房间类型	运行时段	下列计算时刻 (h) 空调启停状态 (0 和 1)											
			13	14	15	16	17	18	19	20	21	22	23	24
民用住宅	起居室	工作日	0	0	0	0	0	0	1	1	1	1	1	1
		节假日	1	1	1	1	1	1	1	1	1	1	1	1
	卧室	工作日	0	0	0	0	0	0	0	0	0	0	1	1
		节假日	1	1	1	1	1	1	1	1	1	1	1	1
商用建筑	办公室	工作日	1	1	1	1	1	1	1	1	1	0	0	0
		节假日	1	1	1	1	1	1	1	1	1	0	0	0
	会议室	工作日	1	1	1	1	1	1	1	1	1	1	1	0
		节假日	1	1	1	1	1	1	1	1	1	1	0	0

2 房间照明或设备作息表以 DeST 默认设定为例，见表 6。逐时使用率基于对应房间的照明或设备的最大功率，其中民用住宅以起居室和卧室为例，商用建筑以办公室和会议室为例，运行时段分为工作日和节假日。

表 6 房间灯光逐时使用率

建筑类别	房间类型	运行时段	下列计算时刻 (h) 房间灯光逐时使用率 (%)											
			1	2	3	4	5	6	7	8	9	10	11	12
民用住宅	起居室	工作日	0	0	0	0	0	0	0	0	0	0	0	0
		节假日	0	0	0	0	0	0	0	0	0	0	0	0
	卧室	工作日	0	0	0	0	0	0	0	0	0	0	0	0

		节假日	0	0	0	0	0	0	0	0	0	0	0	0
商用建筑	办公室	工作日	0	0	0	0	0	0	0	0	10	10	10	10
		节假日	0	0	0	0	0	0	0	0	0	0	0	0
	会议室	工作日	0	0	0	0	0	0	0	7	10	10	10	10
		节假日	0	0	0	0	0	0	0	0	38	38	38	38
建筑类别	房间类型	运行时段	下列计算时刻(h)房间逐时使用率(%)											
			1 3	14	15	16	17	18	19	20	21	22	23	24
民用住宅	起居室	工作日	0	0	0	0	0	0	10	10	10	10	10	50
		节假日	0	0	0	0	0	0	10	10	10	10	10	50
	卧室	工作日	0	0	0	0	0	0	50	50	50	10	10	10
		节假日	0	0	0	0	0	0	50	50	50	10	10	10
商用建筑	办公室	工作日	3	10	10	10	10	50	10	10	10	0	0	0
		节假日	1	30	30	30	30	30	10	10	10	0	0	0
	会议室	工作日	2	25	10	10	10	10	25	25	25	25	7	0
		节假日	1	13	25	25	25	25	13	13	13	13	0	0

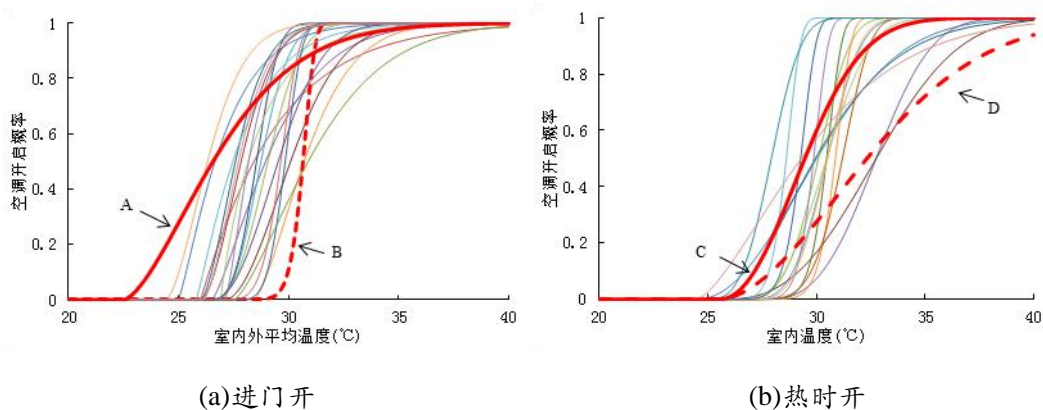
3 以办公建筑、教学楼等建筑为例，房间通风换气运行时段分为工作日和节假日，见表 7。

表 7 房间通风换气次数

建筑类型	运行时段	下列计算时刻 (h) 房间通风换气次数											
		1	2	3	4	5	6	7	8	9	10	11	12
办公建筑、教 学楼	工作日	0	0	0	0	0	0	7	7	7	7	7	7
	节假日	0	0	0	0	0	0	0	0	0	0	0	0
建筑类型	运行时段	下列计算时刻 (h) 房间通风换气次数											
		13	14	15	16	17	18	19	20	21	22	23	24
办公建筑、教 学楼	工作日	7	7	7	7	7	7	7	0	0	0	0	0
	节假日	0	0	0	0	0	0	0	0	0	0	0	0

3.2.4 条文说明如下:

1 环境反馈型条件概率模型用于描述基于环境条件触发的动作, 该模型以环境参数等纯物理量为自变量, 以 u 、 L 、 k 三个参数对自变量与动作发生概率之间的函数关系进行描述, 利用概率大小判断动作发生的可能性。该模型的 u 、 L 、 k 三个参数具有明确的物理意义, 可以体现人的个体差异, 因此主要用于针对个体的建筑用能行为描述。通常采用该模型描述人员个体的空调行为和照明行为。作为示例, 分别在进门开、热时开及睡前开的 3 种空调开启和冷时关、睡前关及入睡后关的 3 种空调关闭情况下, 针对北京地区的 28 间学生宿舍房间, 拟合确定出各个宿舍的空调开关行为概率曲线, 如下图所示。



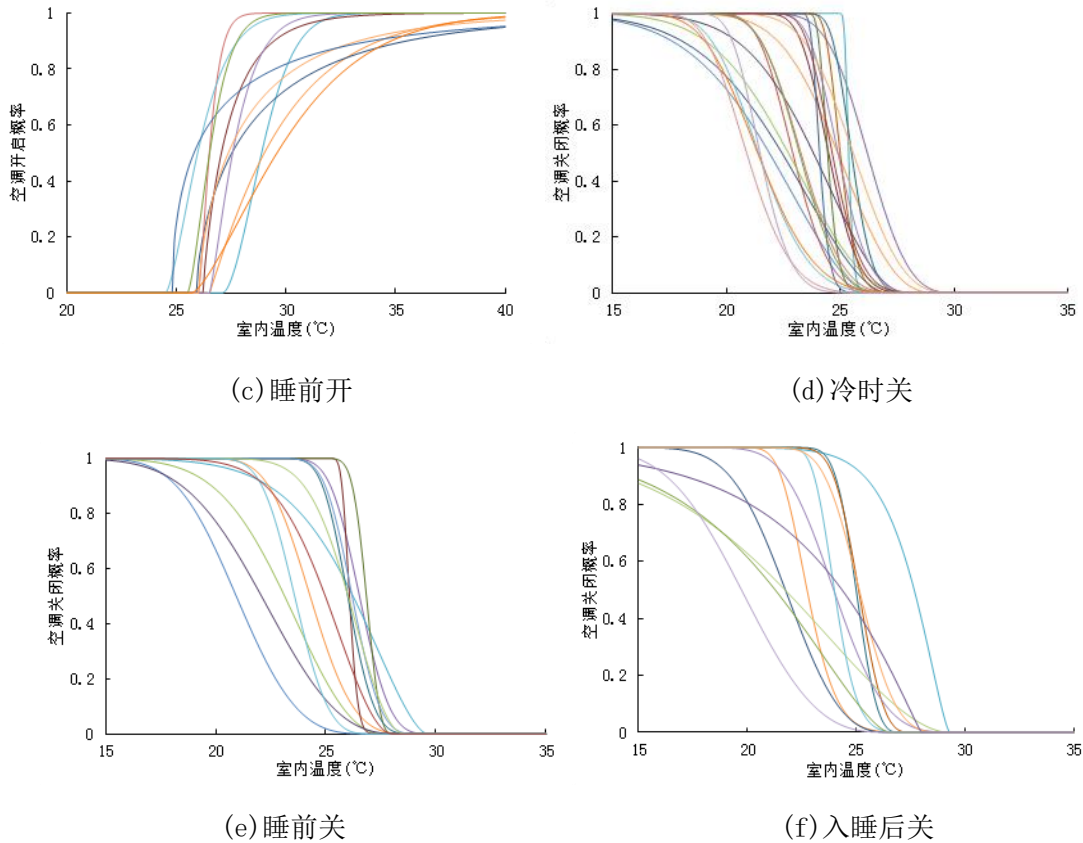
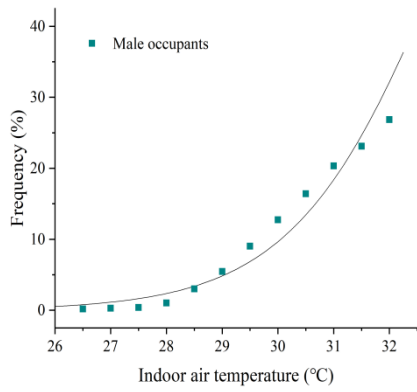


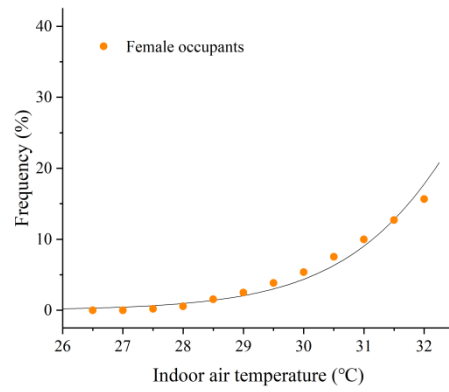
图 2 不同驱动类型下个体空调开关行为概率的变化曲线

2 一些动作发生在进出门、起床或睡觉等特殊时刻，往往是人的生活习惯所造成，不过由于各种随机因素的影响，其发生也是不确定的，例如随手关灯，可能绝大多数情况下如此，但偶尔也会忘记。一些关的动作还可能与人员离开房间的时长有关，比如照明关闭的概率与离开时长成正相关。

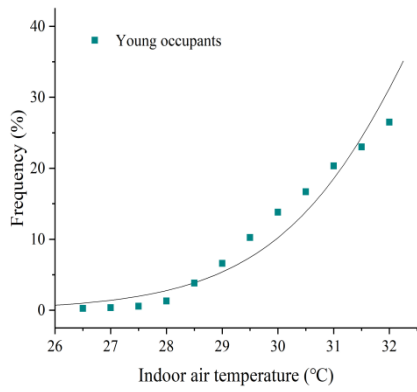
3.2.5 Logistic 模型主要基于 Logistic 回归分析，Logistic 回归分析源于描述生物种群增长规律，具有缜密的数学演绎与推导。logistic 回归的因变量可以是二分类的，也可以是多分类的，实际中最为常用的就是二分类的 logistic 回归，也更加容易解释。自变量既可以是连续的，也可以是离散的。因此，在建筑用能行为的描述中，Logistic 模型主要用来刻画设备（空调、窗、照明灯具）开、关的状态，又由于 Logistic 模型中各个系数没有明确的物理意义，因而 Logistic 模型更适宜对人群群体的行为状态进行描述。作为示例，分别针对北京地区的不同年龄的老年人群、中年人群和青年人群和不同性别的男性人群和女性人群，确定出各类人群空调开启行为的拟合概率曲线和拟合数据结果，如下图 3 和表 8 所示。



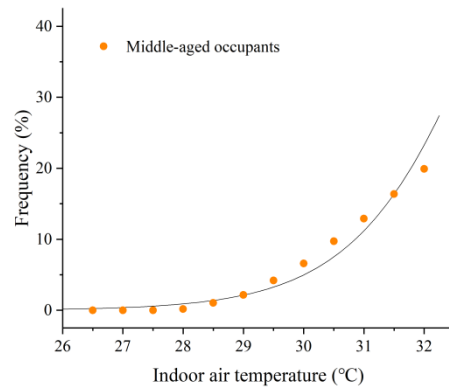
(a) 男性人群



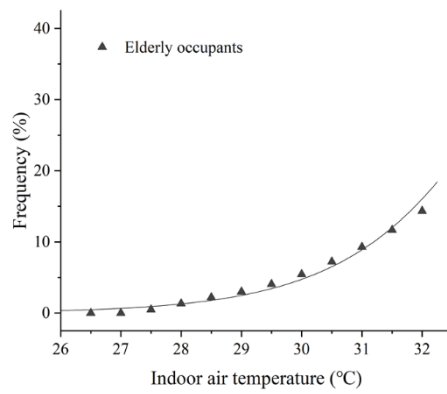
(b) 女性人群



(c) 青年人群



(d) 中年人群



(e) 老年人群

图 3 不同年龄人群的空调开启概率

表 8 Logistic 回归模型的拟合结果

Dependent variable	N	b	c	Exp (B)	Nagelkerke R ²	Hosmer-Lemeshow w	Classification accuracy (%)
Male	23654	0.745*	-24.586*	2.106	0.204	59.403*	90.1
Female	21025	0.778*	-26.427*	2.178	0.169	21.006	95.0
Young	13310	0.693*	-22.963*	2.001	0.188	66.946*	89.7
Middle-aged	19512	0.879*	-29.320*	2.410	0.210	33.816*	93.9
Elderly	11857	0.673*	-23.191*	1.960	0.148	19.661	95.1

其中，N: 样本数; b: 拟合回归系数; c: 拟合常熟; $\text{Exp}(\text{B}) > 1$: 正影响, $\text{Exp}(\text{B}) < 1$: 负影响; Nagelkerke R²: 拟合优度; Hosmer-Lemeshow > 0.05 ; * $p < 0.001$; classification accuracy: 计算精度。

4 描述参数及基础数据定义

4.1 固定作息表及所需的基础数据

4.1.1 如果所采集基础数据为动作数据，如窗户/照明/设备/遮阳/空调开关动作或人员移动及其对应时刻，则需转化成逐时的窗户/照明/设备/遮阳/空调使用状态或人员在室状态后使用。以窗户开关动作及其对应时刻为例，若采集的基础数据表示“A 窗户于典型日上午 8 时开启，下午 4 时关闭”，则典型日上午 8 时至下午 4 时的 A 窗户均为开启状态。由此将动作及其对应时刻数据转化成为状态数据使用。

4.2 位移模型参数及所需的基础数据

4.2.1 人员移动数据是基于人员层面，利用定位器获得人员在区域内的矢量轨迹记录，通过矢量图排布建立区域热点和移动轨迹，用于 netlogo 等模拟工具；人员在室数据是基于建筑层面，利用红外传感器等获取建筑在监测区间的逐时人数变化，为单体建筑的模拟进行数据累积。

4.2.2 定义一般办公环境导致人员移动行为的事件主要包括上班、上午去开会、上午开会后回办公室、去吃午饭、午餐后回办公室、下午去开会、下午开会后回办公室、下班。其中，上班、去吃午餐、午餐后回办公室和下班这 4 种事件为单位工作日内一定会发生的事件，区别在于不同办公要求或者不同人员发生的时间点或区间存在差异。具体的案例可根据作息规律对不同事件的发生进行判断。

4.2.3 对测试者佩戴定位设备进行位置追踪测试，通过上传人员经纬度坐标至地图生成人员轨迹图，同时结合 python 中的 heatmap 功能绘制出相应的热力轨迹图，可实现人员移动规律的可视化。

4.3 动作模型参数及所需的基础数据

4.3.1 人员动作描述和模拟所需数据不仅包括人员所处的物理环境因素，还包括生理、心理等主观因素。

4.3.3 条文说明如下：

1 公共办公建筑室内的能耗设备主要分为两大类。一类是办公设备，包括计算机、笔记本电脑、打印机、扫描仪等，另一类是生活设备，包括空气净化器、饮水机、加湿器等。不同类型和使用功能的办公建筑中的办公设备数量、能效和使用频率存在较大差异。设备的开启和关闭行为直接受人员行为习惯和规律影响，为建立人员动作模型，需要确定影响设备使用行为的因素。

2 条文说明如下：

1) 人员作息时间表根据 4.2 节人员位移数据和模型获得；

2) 计算机配置比例：同一类型建筑中不同房间人员计算机配置比例存在差异。通常一人

配置一台计算机，具有特别功能需求的人员一人配置 1~3 台计算机。(除了计算机外，通常一间小型办公室配备 1~2 台打印机，一间大型公共开放办公室按照每 20~25 人配备 1 台打印机)；

- 3) 计算机运行状态：主要有工作状态、待机状态、关机状态，对应的运行功率水平分别为高功率、中功率和低功率(0~10W)，不同品牌和功能的计算机的中功率和高功率阈值区间存在差异；
 - 4) 人员使用行为习惯：一般人员计算机的使用行为习惯分为“浪费型”、“适中型”和“节能型”。“浪费型”指人员保持计算机的长期开启，“适中型”指人员仅上班期间开启计算机，仅下班时间关闭计算机，“节能型”指人员在短暂离开房间也关闭计算机或改变运行状态；
 - 5) 人员不同离开时长关闭计算机的概率：离开时长分别按照 0~10 分钟、10~30 分钟、30~60 分钟、60~90 分钟、90 分钟以上进行设定，对人员不同离开时长下关闭计算机的概率进行统计，也可认为是不同离开时长会选择关闭计算机的人员在建筑中出现比例。
- 3 计算机配置比例、人员使用行为习惯和不同离开时长下关闭计算机的概率可通过问卷调研获取，计算机使用状态可通过在计算机插座上加装可监测用电功率的智能插座，通过测定单位时间的计算机运行功率判断运行状态。
- 4 以饮水机使用为例：人员作息时间表根据 4.2 节人员位移数据和模型获得；人员冷热水消耗水平取决于人员的取水渠道、水杯容量、每天的取水次数和偏好的饮水温度；取水渠道包括饮水机和电热水壶。

4.3.4 条文说明如下：

- 1 对于非传感照明建筑系统，确定照明开启数量所需的数据类型主要为室内外照度情况、人员对于自然采光或人工照明的使用偏好和习惯，其中室内外照度数据包括室内工作面照度、室外照度和室内的平均照度水平。在人员在室的前提下，工作面的照度是影响人员照明行为的主要因素，可根据灯具被打开前一刻的照度阈值确定照明开启或关闭的照度限制。
- 2 日常办公环境中人员开灯对应的室内工作面照度范围较大，在相关技术标准的推荐区间内人员开灯概率与照度联系密切，在照度低于下限时必会产生开灯行为，但并不排除照度高于上限时不存在开灯行为，与个人工作习惯或工作的特殊要求相关。
- 3 基于特定事件的行为习惯主要指人员趋于某种事件的发生而产生的固定行为，包括上班时刻达到房间，午饭或会议短暂离开房间和下班离开房间，需确定人员作息时间表，根据 4.2 节人员位移数据和模型获得；节能意识主要体现在人员离开多久后会选择关闭照明的概率，一般情况，照明关闭的概率与离开时长成正相关。
- 4 人员在室率的变化对照明开启数量的分布存在极限值，应确定这一极限值以及目标房间内的人员在室率时间表，该时间表根据 4.2 节人员位移数据和模型获得。

4.3.5 开关窗的动作模型用来描述窗户开闭的随机特征。模型的因变量为开窗或者关窗的概率，模型的自变量为开关窗的影响因素（通常为室内温度、室外温度等热环境参数）。开关窗的动作时刻、动作时刻对应的环境参数以及该时刻对应的开关窗动作等一组数据组成一个样本，宜采用 MATLAB、Python 等软件将所有样本进行拟合，从而得到开关窗动作模型。

4.3.6 空调启停的动作模型用来描述空调运行的随机特征。模型的因变量为开空调或者关空调的概率，模型的自变量为空调启停的影响因素（通常为室内温度、室外温度等热环境参数）。开关空调的动作时刻、动作时刻对应的热环境参数、以及该时刻对应的空调开关动作等一组数据可以组成一个样本，采用 MATLAB、Python 等软件可以将所有样本进行拟合，进而得到空调启停动作模型函数。

5 基础数据获取方法

5.1 获取方式

5.1.1 条文说明如下：

- 1 人员移动数据包括是否有人在室、室内人员密度、人员在不同房间的位移等，可根据需要测试的人员移动数据类型，选取不同的测试设备和测试方式，人员动作数据包括人员对房间内设备开关、调控行为以及产生该行为时室内的环境参数状态，可根据需要测试的设备调控行为类型，监测该设备的运行状态，选取适合的环境参数测试装置。
- 2 人行为数据获取，对于同时性有很高要求，通常人员在房间内移动、设备调控动作发生的时间比较短，各测试设备内置时钟在较长期监测过程中，容易产生系统偏移，引起行为判定的误差，甚至会导致测试数据无效，因此需要定期对各测试设备的时间进行同步，使用具备通信功能的设备进行内置时钟校正更为理想；在汇总数据时，采用相同的测试起始时间和采样时间步长，各测试设备的数据可直接进行合并，将大幅度提高数据分析的效率。

5.1.2 条文说明如下：

- 1 基于 Wi-Fi 信号的室内人员识别是通过利用计算机或用户携带的智能手机与安装在室内环境中的接收器和发射器共同实现的。由于实际测试环境的复杂性（房间结构复杂、物品放置自由），Wi-Fi 信号会出现多路径衰落和时间延迟，往往不能提供高精度的人员位置及行为识别。WiFi 基础设施、人员携带的支持 WiFi 的移动设备和机器学习算法（例如：在线顺序极限学习机算法）结合，可用于采集人员数量数据。使用过程中存在某种特殊情况，例如：忘记带走移动设备、一个人携带多个移动设备、睡眠模式 WiFi 关闭等，宜采用人员身份信息和历史居住习惯解决。
- 2 可穿戴设备（如，智能手环）可分析人员在室内空调环境下的运动，也可以利用 GPS 系统对人员进行定位。
- 3 射频识别传感器（RFID）包括多个 RFID 传感器、有源 RFID 标签、一个无线传感器网络、一个 RFID 协调器以及一个位置服务器，组成基于 Zigbee 协议的无线传感网络，由于 RFID 阅读器价格昂贵的原因，因此人员定位领域宜采用无线传感网络实现人员数目、人员位置的数据采集。
- 4 三轴加速器用于活动监测时，应依据具体需要监测的人员活动水平考虑设备精度、监测范围、摆放位置、使用数目等问题的选择。人员身体装配三轴加速器的数量应越少越好，如条件允许，宜在胸部、下背部、手腕、臀部、大腿、脚部皆放置三轴加速器，由此可监测人员大部分活动水平（如：上楼梯、跑步、步行、深蹲）。为了评估人员日常身体活动，三轴加速器的振幅范围应在 $\pm 12g$ ，记录频率应高于 20Hz。采样频率为 50Hz，足以捕捉人员将大多数日常活动。尽管人体活动幅度可达 12g，但躯干和臀部附

近的活动幅度仅达 6g。具体测试方案宜根据测试目的确定。g 是加速度的常用单位，表示三轴加速度传感器的量程。1g 代表一个重力加速度，即 9.8m/s^2 。除额外佩戴三轴加速度器外，还可使用智能手机中内置的三轴加速度、陀螺仪、动作感应器、GPS 系统识别人员位置及移动。

- 5 座椅压力传感器记录仪利用压力传感器感应座位上的压力来确定人员的出现或离开。但不能区分人员在室但离开座位的情况，适合房间中人数较多的办公建筑。

5.1.3 条文说明如下：

- 1 深度相机包含 3D 结构光相机、TOF 相机以及双目相机，Kinect 深度相机包含 ToF 摄像头、RGB 摄像头、IMU 传感器以及麦克风，。通过计算机中的软件开发工具包（SDK）和应用程序接口（API）非接触采集人员信息数据，可得到人员数目、人员身份、人员轨迹以及人员姿态。1 台 Kinect 设备最多同时跟踪 6 名人员，检测整个房间时，可放置在房间入口上方或者房间内部，
- 2 热成像仪主要由红外探测器、光学成像物镜、光机扫描系统三大部分构成，可识别图像的温度分布情况，实时跟踪人员轨迹并采集人数，宜用于保护人员隐私的场所。热成像仪与计算机的深度学习算法（例如：卷积神经网络算法）结合，进一步分析得到人员姿态。
- 3 毫米波雷达能够在复杂多变的环境中正常工作，如在光线不足、环境中烟尘较大等视线不良的环境中仍能保持较好的检测能力。可以实时采集到人员轨迹、生命体征、行为（手势和动作）、数目等数据，对人体的检测局限于外部轮廓，宜用于保护人员隐私的场所。
- 4 被动式红外传感器（PIR）与主动式红外传感器不同，它不依靠发射辐射能量来检测物体的存在，而是通过检测人或动物身体发射的红外辐射输出电信号，宜采集运动的人体数据，对于静止的人员检测精度较低，主要用于检测是否有人。传感器的有效感知区域宜覆盖人员活动的主要区域，有利于提高检测精度；同时，宜在房间入口处增设传感器，可辨别人员运动方向，同时实现人员数目的数据采集。
- 5 CO₂ 传感器检测的 CO₂ 浓度与室内人员密度之间存在较高的相关性，通过最大似然法、卡尔曼滤波、优化算法、机器学习算法（例如：极限学习机、支持向量回归算法）可分析出是否有人以及人员数目，不涉及用户信息保护用户隐私，但换气率、人数、门窗的开启等不可预知因素影响 CO₂ 浓度，通过数据预处理可减小这部分因素的影响。
- 6 使用非接触式设备采集人行为数据宜采用传感器融合技术，仅使用一种传感器很难提供可靠的数据信息，不同的传感器的功能可以相互补偿，在克服自身测量局限性的同时，充分利用其他传感器的优势，可提高检测准确性。对不同类型的数据进行收集和融合，选取重要特征，获得更准确的占用信息。例如：将摄像头、深度相机、热成像仪、毫米波雷达等成本和精度较高的传感器，与 WiFi 传感器、RFID 射频识别传感器、

被动式红外传感器、CO₂ 传感器等成本和精度较低的传感器结合，综合评判采集到的人员数据。

5.1.4 通过手机、电脑等移动终端的定位系统可以方便、快捷地获取建筑内总人数。但由于定位精度有限以及使用移动终端设备的人员，该方法适用建筑需满足以下条件：

- 1 建筑平面长、宽均大于定位系统的最小精度值（目前民用是 10m）；
- 2 建筑内超过 50%的人应具备手机、电脑移动终端，并开启定位服务。

对于商场、办公楼、医院（门诊）、交通枢纽、展览馆、体育馆等公共建筑，宜采用大数据方法获取总人数；对于居民楼、宾馆、医院（住院）等具有住宿功能建筑，应对夜间数据进行修正校核；对于幼儿园、小学、中学等主要人员不具备移动设备，不宜采用此方法获取人数。

5.1.5 问卷调查可获取人员的主观数据，如上下班习惯、设备调控喜好等，在测试的同时开展问卷调查，作为人行为分析的依据是非常有意义的。为保证问卷调查结果的科学合理性，问卷设计应遵循表述通俗易懂、没有歧义、选项互斥、避免回忆和估计的原则。问卷发放样本量需要满足抽样比，以满足统计分析的要求。

5.1.6 条文说明如下：

- 1 可使用电量记录仪进行测量分体式空调、风扇、照明设备的开关及档位情况，应具有实时测量功率和累计耗电量的功能。宜具有本地自动记录或网络远程上传到云端的功能。
- 2 可使用有磁开关记录仪和摄像头拍摄的方法测量开关窗行为和窗户开度，磁开关记录仪用于连续采集磁场变化的开关量，用户可以自行设计磁场环境并加以测量，在测试时分别将记录仪和永磁铁安装在窗户和对应的窗户边框上，当室内人员每开启或关闭窗户时，磁开关记录仪会记录一次数据。采用摄像头拍摄可以记录准确的窗户开度，但需人工手动处理视频数据，工作量较大。建议可采用位移记录仪或磁开关记录仪自动记录为主，摄像头拍摄为辅的方法。
- 3 常用的遮阳行为测试方法为（网络）摄像头或数码相机拍摄、感应自动记录（超声波测距仪、拉绳位移记录仪）、现场走访观察等。对于下拉型窗帘，可以使用超声波测距仪或者超声波测距传感器测量窗帘的开合程度。住宅的遮阳若是窗帘形式，可以采用布置多个测点方式进行测量，具体方式如图 4。

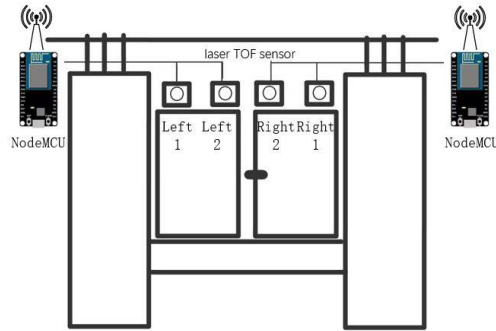


图 4 窗帘形式布点方案

4 使用手机终端 APP 进行控制的智能设备，可采用后台记录的功能获取设备的调控行为。

5.1.7 条文说明如下：

- 1 使用温湿度传感器测量房间内的温湿度，应具有实时测量温度和湿度的功能。宜具有自动记到本地或网络远程上传到云端的功能。
- 2 使用照度传感器获取室内工作面的照度，应尽量在人员位置附近布点。
- 3 一般通过测量室内污染物的含量来评价室内空气品质，可使用单一或多个传感器的空气监测仪，测量 CO₂，VOC，PM_{2.5} 等参数。

5.2 采集方法

5.2.2 室内人员状况（即房间里什么时候有人、有多少人等）是建筑能耗模拟计算的基本输入参数，而室内人员状况是由建筑中人员的移动过程决定的。因此，为准确模拟建筑热环境及室内设备的运行调节行为，必须首先对人员位移进行描述。根据不同的模拟目的，获取包括人移动轨迹、房间是否有人、房间人员数量、建筑内人员密度和人员在室率等数据。

5.2.3 房间是否有人，是人员移动数据采集的最基本要求，当房间内有人在室时，才有可能对房间内的设备进行调控。测量房间是否有人方法较简单，宜使用一个或多个被动式红外传感器、毫米波雷达覆盖房间，但房间有人员活动时，仪器显示为 1，判定室内有人。受到测量原理的限制，被动式红外传感器仅能感应到探测范围内有移动的人体。由于人员在室时很可能不产生移动，如静坐，睡觉等，将人体感应仪数据转换为人员在室状态时会产生一定错误，在数据处理过程中应将其修正。包括但不限于以下两种常见情况（为描述方便，“0”代表房间没人，“1”代表房间有人）：

- 1 当室内人员在 1min 内没有动作时，仪器可能将在室状态误判为“0”。因此，如果两次人体感应仪测得的数据为“1”，而中间一次测得数据为“0”，可能是房间内有人但未活动，需要对数据进行识别和修正。
- 2 在卧室的测量中，需对人员睡觉期间的在室状态进行修正，将仪器测量为“0”的数据改为“1”，以获得人员睡觉时在室状态的真实数据。若根据问卷调研结果，测试人员

有午睡习惯，则需对中午时间段内的数据进行类似的修正。

5.2.4 房间人员数量在建筑能耗模拟计算中，可以用来确定室内人员的发热量，另外当房间人数增加时，由于多人行为的交互作用，人员对设备调控趋向于统一，随机性减少。在没有隐私要求的情况下，宜使用摄像头配合人脸识别方法对人员数量进行检测，在摄像头无遮挡情况下精度较高；对于隐私要求较高时，测量人员数量宜采用座椅压力传感器，测量座位上的人员数量，或在房间门附近设置人员计数器。数据记录时间步长应不低于 1min 一次。

5.2.5 人员在室内的移动轨迹，是对于人员移动的精确描述，数据采集对于设备要求也相应较高。对于办公环境中，在没有隐私要求的情况下，宜使用摄像头配合人脸识别方法对人的移动轨迹进行检测。对于隐私要求较高时，宜使用基于 Wifi 型号的室内人员识别、佩戴可穿戴设备、佩戴射频识别传感器采集。

5.2.6 对于新建建筑或设计阶段的建筑，可参考《公共建筑节能设计标准》（GB50189）表 B.0.4-5 和 B.0.4-6 获取不同类型房间人员密度及在室率；对于投入使用的建筑，可通过红外传感器、Wi-Fi、蓝牙、摄像机等方式获取指定区域内总人数；也可通过门禁或者手机、电脑等移动终端获得建筑尺度的总人数，经式（25）计算可得到人员密度。人员在室率则根据模拟时间尺度，采用式（26）计算。此条文旨在说明人员在室情况的一种表述方法，由于人员密度和人员在室率均非直接测量得到，而是计算得到，因此在条文说明中加入了计算公式。而能被采集的实际上还是人数，人数根据不同空间尺度应选用不同获取方法。

$$\text{人员密度} = \frac{\text{区域内人数最大值}}{\text{区域面积}} \quad (25)$$

$$\text{人员在室率} = \frac{\text{某时刻区域内人数}}{\text{指定时间段内、区域内人数最大值}} \times 100\% \quad (26)$$

5.2.8 室内人员的动作数据用于分析在怎样的前提下，人会对设备发生调控动作，包括环境参数条件、事件触发条件等，根据不同的设备可选择对应的数据采集方法，如采暖/空调设备调控、开关窗、遮阳、采光、照明等，包括对设备调控行为的采集和对环境参数的采集，并结合人员的位移进行分析，如人员是否在室、人员进入/离开室内的时刻、人员在房间停留的时长等，从而获得设备调控行为的触发条件模型，各测试设备的采集时间应相一致以便于数据的对照和分析。

5.2.9 采暖空调行为包括人员对空调器的开关的触发条件模型、空调器设定温度等。空调开关情况可使用电量记录仪、在空调送风口安装温度自记仪、空调控制面板通讯等方式获取，环境温湿度可使用温湿度自记仪，测量步长应不低于 1 min 一次。

5.2.12 条文说明如下：

- 1 个体风扇，通过增加人员周围局部微环境的流速来调节人体热舒适，有效地降低人员头部及上半身的皮肤温度，因此采集到人员调控风扇的行为数据，将相应地改变空调系统的运行能耗。
- 2 加热/冷却键盘，通过键盘旁边的电热曲面加热手掌表面，通过键盘底部风扇冷却手掌

表面，在人员使用的过程中，人员调控设备的动作数据（例如：手掌放在加热/冷却键盘表面），导致模拟的建筑能耗发生变化。

- 3 个性化送风末端与人员周围的设备（例如：桌子、椅子、电脑等）结合，调节人体局部微环境，通过调控风口处温度和风速可改善人体热舒适，将其与混合送风结合，系统的运行能耗将发生变化，则检测人员调控送风末端的行为使模拟的能耗发生相应的变化。
- 4 因为加热（冷却）座椅的使用行为取决于房间的实际情况，故个性化行为应根据每一位使用者的位置进行调整。在现有能耗模拟工具应用中，人员应位于座位上才可以操控个性化热舒适装置。反之，关闭个性化热舒适装置。
- 5 下肢(脚和腿)可以通过辐射面板或反射灯泡等（暖脚器、暖腿器等）加热，也可以通过脚底下的导电垫加热。使用了局部加热设备的房间宜适当降低背景空调温度设定值。
- 6 个体热舒适装置可以在很宽的室温（18-30°C）范围内，将热舒适保持在可接受水平（热舒适高于 1），因此影响建筑系统能耗。由于个体热舒适装置的使用，暖通空调系统存在更宽的房间温度死区（不需要系统供热或者制冷的室内温度区间），暖通空调系统每年可节省 30%-40%的能源。

6 基础数据质量与预处理

6.1 基础数据质量要求

6.1.1 条文说明如下：

- 1 对人行为基础数据进行质量控制，是采用一定方法和模型判断数据质量的可靠程度，旨在保证所获取的数据能够满足特定的人行为模拟需求。数据质量的准确性、完整性和一致性的基本要素应符合现行《信息技术数据质量评价指标》和《GB/T 39400-2020 工业数据质量 通用技术规范》等相关国家标准的规定。通过检验准确性要素能够确定获取的人行为基础数据是否存在冗余、重复和异常数据。完整性要素用于确定所获数据的完整程度以及是否存在缺失数据。通过一致性检验确定不同数据所观测的人行为变化规律是否存在明显差异。
- 2 测量范围宜根据建筑功能、地域特征、模拟应用场景和具体行为类别进行确定，其基本要求应符合本导则 6.1.2 至 6.1.4 节中的规定。由于人行为变化存在一定周期性，其测量周期应包含不同作息规律（如工作日与非工作日）。对季节差异性较大的行为（如开关窗行为），通常需要检验其测量周期是否包含多个季节；对仅包含一个季节的数据，应添加相应标识。采集频率决定了所获取的数据描述人行为的精细程度，宜根据能耗模拟目的进行确定。在条件允许的情况下，宜增大采集频率以获取更多相关信息。
- 4 由于不同数据源在表征人行为变化规律时存在一定相关性，应根据不同数据源的相关性变化规律检验数据的一致性，对存在关系异常的数据应进一步进行人工校核，确定数据符合人行为变化的基本特征。
- 5 利用计算机软件采用不同质量控制方法对获取的数据进行质量控制，判断正确数据和可疑数据，并对可疑数据进行提取、显示或标识。同时结合相应知识进行人工质量审核，在部分情况下通过人工判断最终确定数据的正确性。

6.1.2 为保证所采集时间数据的典型性和完整性，当采用问卷调查等方法进行采集时，其样本数量宜保持在 100 户以上；当采用大数据在线监测等方法进行采集时，应注重保持样本的完整性（如完整住宅或办公区域）；就采集周期而言，对于受季节显著影响的适应性行为和用能行为时间数据（如开关窗行为和空调行为等），其采集周期应包含多个季节；对于受工作日和非工作日显著影响的行为数据（如人员在室占用率），其采集周期不宜低于一个星期；对于其他不受周期性和季节性影响的相关数据，其采集周期不宜低于一天。

6.1.3 条文说明如下：

- 1 上述参数主要参照国际标准 ISO 7730、现行国家标准《建筑热环境测试方法标准》以及人行为相关研究进行确定。

3 条文说明如下：

- 1) 对于位移模型，在采集周期方面，测试应至少涵盖人员在建筑环境内活动的一个完整周期，

如办公室的测试应从上班时间持续到下班时间，而住宅的测试时间应从起床时间持续到入睡时间。考虑到工作日和周末作息规律的区别，采集周期推荐为一周以上。在采集频率方面，人员在室内的活动变化会较为频繁，采集频率范围宜为 5 分钟-30 分钟。在测量条件允许的情况下，可进一步提高采集频率。

- 2) 对于动作模型，在采集周期方面，考虑人员动作具有较大的随机性，需要大量的数据挖掘其中的规律，因此采集周期宜持续超过一个月，且可通过具体行为类别考虑确定测量周期是否需要涵盖不同季节。在采集频率方面，考虑到一般人员动作变化频率往往相对较低，采集频率宜为 15 分钟-1 小时。在测量条件允许的情况下，可进一步提高采集频率。

6.2 基础数据预处理方法

6.2.2 条文说明如下：

1 条文说明如下：

- 1) 传感器数据缺失通常由传感器损坏、断电、存储空间有限等原因导致。调查类数据缺失通常由调查对象与调查者的自身原因，如调查对象回避敏感问题、调查者疏忽漏填错填等导致。数据后处理缺失通常由系统硬件及人为失误，如数据录入期间系统传输故障及操作人员错录数据等原因导致。
- 2) 保留缺失项是指对数据集中的缺失值不进行任何处理，直接应用于分析过程。在缺失项对人行行为相关模型构建不产生影响或影响较小的情况下可采取保留缺失项。直接删除是指通过剔除原始数据中不完整样本或者不完整属性，从而得到一个规模缩减但完备的数据集。在人行行为相关数据有多个属性缺失、被删除的含缺失值的对象与信息表中数据量相比非常小的情况下可采取直接删除。缺失值填补是指通过研究现有数据为每个缺失值估算一个尽可能合理的填补值，从而构造一个和原始数据集规模相同的完整数据集。
- 3) 样本量较为充足且缺失比例较小的情况下，可采用简单缺失值填补方法，如均值填补法和线性插值法。样本缺失量比较大且模拟精度要求比较高的情况下，可采用将数据集中的数据质量相对较高的样本当作训练集来建立预测模型，并根据训练出的预测模型估计缺失值，如回归填补法和 KNN 填补法等。

2 条文说明如下：

- 1) 相关国家标准包括《建筑热环境测试方法标准》等，应参照其中对于检测数据正常取值范围的规定。同时，还可结合实际场景或以相关文献中获得的取值范围作为参考，如果超出该范围认为数据异常。
- 2) 对于数值变化在一定范围内的时序类数据（如 CO₂、温湿度和照度），如出现较大突变，应视作可疑数据并采用莱茵达准则和箱线法等方法判断其正确性。对于发生

次数在一定范围内或满足一定分布规律的事件类数据，如发生次数过大应视作可疑数据，并采用箱线法和聚类等方法判断其正确性。对于用于分析人员移动行为和建立位移模型的时空类数据，考虑到人员在建筑中的空间移动距离以及移动间隔时间的变化范围通常有限，应通过检验其时间间隔和空间间隔以确定是否可疑，并采用箱线等方法判断其正确性。

7 基础数据分析方法

7.1 环境因素的影响分析

7.1.2 可以通过此方法分析影响人员行为的各参数的数据集中趋势。常用的指标有平均值、中位数、众数以及标准差等。

7.1.3 例如将室外温度按 2℃ 划分区间，统计各区间内的开窗/空调概率并绘制成折线图，通过观察曲线的波动趋势描述其对开窗/空调行为的影响。

7.1.4 在开窗/空调行为分析中，为了获得不同的行为模式，使用了聚类分析方法。可用于研究在人员为什么(动机模式)，多长时间(开放持续时间模式)，什么时候(互动模式)和多少(位置模式)。

7.1.4 可检验所测得环境参数与往年/月相同时间的环境参数是否有显著性差异，如果不存在显著性差异则说明所测数据并不具有“典型”特征，得到的研究结果可普遍应用于其他年/月份。在应用时要求数据满足三个条件：正态性、方差齐性、独立性。

7.1.7 皮尔逊相关系数可用于分析连续的环境变量（如：室外温度）与其对应温度下开窗/空调概率之间的相关性。应用该系数需满足以下条件：两个变量的标准差都不为零；两个变量之间是线性关系，都是连续数据；两个变量的总体是正态分布，或接近正态的单峰分布；两个变量的观测值是成对的，每对观测值之间相互独立。

7.1.10 在人员用能行为的研究中若两个变量之间存在明显的共线性则会影响回归分析的结果。可使用方差膨胀因子（VIF）和容差（TOL）来判定，当 $VIF < 5$ 或者 $TOL > 0.2$ ，可以说明基本不存在共线性。

7.1.11 在人员用能行为分析中可用于研究单一环境变量对动作持续时长的影响。曲线为连接人员发生动作的比例与动作持续时长所对应的下降阶梯曲线。曲线越平缓，发生动作的概率越大，动作持续时长越长。反之，曲线越陡，发生动作的概率越低，动作持续时长越短。为了排除其他变量对目标变量的影响，将其他变量控制在不会触发动作改变的状态范围内。

7.2 非环境因素的影响分析

7.2.3 以楼层/朝向为单因素变量，分析各组数据之间是否有显著性差异。如果产生显著性差异，则说明开窗/遮阳行为会受到其影响。