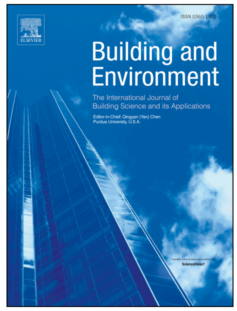
Accepted Manuscript



Thermal comfort, occupant control behaviour and performance gap – A study of office buildings in north-east China using data mining

Cheng Sun, Ran Zhang, Steve Sharples, Yunsong Han, Hongrui Zhang

PII: S0360-1323(18)30779-0

DOI: https://doi.org/10.1016/j.buildenv.2018.12.036

Reference: BAE 5870

To appear in: *Building and Environment*

Received Date: 2 September 2018

Revised Date: 26 November 2018

Accepted Date: 14 December 2018

Please cite this article as: Sun C, Zhang R, Sharples S, Han Y, Zhang H, Thermal comfort, occupant control behaviour and performance gap – A study of office buildings in north-east China using data mining, *Building and Environment* (2019), doi: https://doi.org/10.1016/j.buildenv.2018.12.036.

Thermal comfort, occupant control behaviour and performance gap – a study of office buildings in north-east China using data mining

Cheng Sun [a, b, *], Ran Zhang [a, b], Steve Sharples [c], Yunsong Han [a, b], Hongrui Zhang [a, b]

[a] School of Architecture, Harbin Institute of Technology, Harbin, 150001, China
[b] Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology, Ministry of Industry and Information Technology , Harbin, 150001, China
[c] School of Architecture, University of Liverpool, Leverhulme Building, Liverpool, L69 7ZN, United Kingdom

ABSTRACT

Simulation techniques have been increasingly applied to building performance evaluation and building environmental design. However, uncertain and random factors, such as occupant behaviour, can generate a performance gap between the results from computer simulations and real buildings. This study involved a longitudinal questionnaire survey conducted for one year, along with a continuous recording of environmental parameters and behaviour state changes, in ten offices located in the severe cold region of north-east China. The offices varied from private rooms to open-plan spaces. The thermal comfort experiences of the office workers and their environmental control behaviours were tracked and analysed during summer and winter seasons. The interaction of the thermal comfort experiences of the occupants and behaviour changes were analysed, and window-opening behaviour patterns were defined by applying data mining techniques. The results also generated window-opening behaviour working profiles to link to building performance simulation software. The aim was to apply these profiles to further study the discrepancies between simulation and monitored results that arise from real-world occupant behaviour patterns.

Key words: window-opening behaviour; office building; cold climate; cluster analysis; association rules mining.

1. Introduction

In the process of architectural design and building energy-efficiency-evaluation studies, various types of building performance simulation techniques have become basic tools for building energy calculation, design optimisation, operation management, and building energy-saving diagnosis [1-5]. Performance-based simulation analysis methods and evaluation indices are also widely used for the energy-efficient design of new buildings, energy-saving renovation of existing buildings, energy-efficient technology assessment, and formulation of energy-saving standards [6-9]. The process of architectural design is being transformed from result control to process control, and from separation of architectural geometric design and evaluation of building performance to the combination of those aspects [9,10].

Although the potential of building performance-based design has been widely recognised, it is still not possible to provide the best solution for designing energy-efficient buildings due to the gap between real results and those expected from architectural design schemes [11, 12]. This discrepancy also fails to give a true feedback on the impact of a building design on performance; therefore, it does not provide designers with a real perception of building system performance. The reason for this difference is the large number of input parameters in the simulation process and their non-linearity, discreteness, and uncertainty, represented by the user behaviour parameters in this study, which further increase the complexity of influencing elements [13].

Among these uncertain input parameters, occupant behaviour is a major factor affecting the thermal comfort and energy efficiency of buildings, indicating the importance of establishing a behaviour mode for modelling and predicting building performance [14-16]. The influence of occupant behaviour on buildings is greatly influenced by geographical regions and ethnic cultures [17, 18]. This is reflected in many case studies, especially for residential or office buildings, which tend to have more individual controls [19-21]. Presently, the exploration of the impact of different climates and cultural backgrounds on behaviour patterns still needs further development. An accurate model is based on a wide range of data collection. Behaviour-control data collection in recent research is derived from long-time recordings and transverse questionnaire surveys [22, 23]. Longitudinal questionnaires, with the characteristics of being time-consuming and labour-intensive, are relatively rare in occupant-behaviour studies, even though the results, when combined with measured data, can provide more opportunities for exploring changes in behaviour control.

There are four main methods to examine the behaviour mode: agent-based modelling, statistical analysis, machine learning, and stochastic modelling [24-29]. Zimmermann [30] first applied the agent-based modelling method to build simulation models for behaviour control and motivating factors; Haldi and Robinson [31] studied the numerical relationship between occupant behaviour and other information; D'Oca and Hong [32] used data mining to discover occupancy patterns in office spaces; Erickson et al [33] modelled and estimated occupancy status and related energy consumption. These classic studies, with their different approaches, focused on different aspects of behaviour, providing both theoretical support and application guidance for determining patterns of occupant behaviour.

Research relating to occupant behaviour in Chinese buildings has only been active in the last few years. For example, Yu [34] conducted a winter and summer survey amongst elderly occupants to investigate their thermal comfort and adaptive behaviour characteristics in a hot summer/cold winter area of China; Song et al [35] surveyed five office rooms located in a cold region of China to identify the influencing factors of window-opening behaviour; Xin [36] focused on summer window-opening behaviour triggers and classification in a hot summer/cold winter part of China. Due to China's large regional differences in climate, research on different climatic regions is imperative. In our previous studies, basic characteristics of the summer occupant behaviour were researched [37], and a comparison of the influencing factors and predictive models between different modes of occupant behaviour in offices were examined [38]. Furthermore, research on the simulation optimisation of building performance linked with an occupant behaviour configuration file is relatively scarce in the literature.

This study focuses on the interaction between thermal comfort and occupant behaviour in different-sized offices located in the north-eastern China city of Harbin, which experience a severe cold winter climate. The study involved a one-year longitudinal questionnaire survey and logging of occupant environmental control behaviours in winter and summer. Window-opening behavioural patterns were identified using data mining techniques, with an attempt at classifying the behaviour mode to reflect the characteristics of different behaviour categories. Next, efforts were made to try and reduce the gap between simulation results and real data by directly linking the behaviour modes to simulation software, to improve the accuracy of the simulation and reflect the real mechanism of the impact of occupant behaviour on building simulation in office buildings.

This study contributes to findings about thermal comfort and occupant behaviour in different-sized offices with and without air conditioning during the hot summer and cold winter in Harbin regarding the following:

- Long-term occupant thermal comfort and behaviour characteristics in private offices, shared-private offices, and open-plan offices;
- Influencing factors of adaptive behaviour for both summer and winter;
- Defining the window-opening behaviour duration patterns, window-opening behaviour classification, and behaviour profiles in the hot summer season and cold winter period via data mining techniques;
- Modifying the building thermal performance gap and verifying the window-opening behaviour profiles in selected offices.

2. Methodology

For the extreme Harbin climate of hot summer and cold winter, this study established a data set from a long-term survey, with the application of statistical analyses and data mining techniques, to define window-opening behaviour and attempted to fix the building performance simulation gap. A longitudinal survey was conducted for a one-year period, interviewing for both subjective and objective variables relating to occupant thermal comfort and adaptive behaviour. The basic characteristics of occupant thermal comfort experiences and behaviour in the summer and winter were obtained. Logistic regression was applied to analyse the parameters influencing window-opening behaviour. Data mining technology combed data, summarised rules, and classified categories of these data, obtained from the longitudinal questionnaire survey and field measurements in the summer and winter seasons. Finally, behaviour profiles were obtained and linked into DesignBuilder, and then, the performance simulation was optimised. Fig. 1 schematically shows the methodological approach.

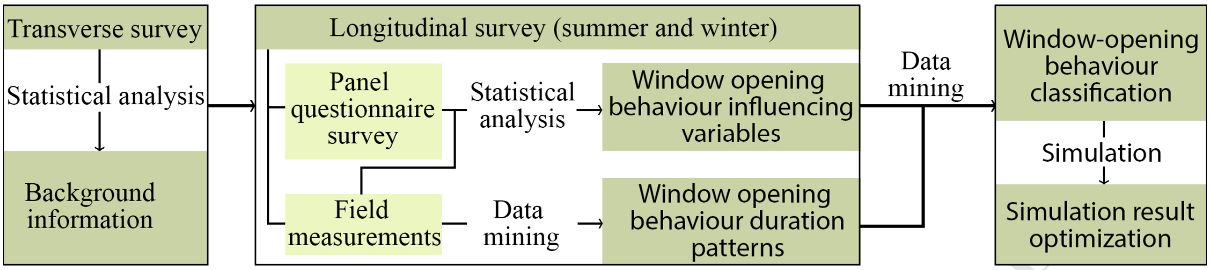


Fig. 1 Work flow of the behaviour classification and building performance simulation optimisation

2.1 Sample selection

Harbin is a typical city in north-eastern China. It experiences a temperate continental monsoon climate with four distinct seasons. The winter is long and cold, while the summer is short but hot. A district heating (DH) scheme is widely applied in Harbin, with six months of uninterrupted winter heating. Ten volunteer offices distributed around six office buildings in representative districts of Harbin were chosen from the samples of the transverse survey, including private offices, shared-private offices, and open-plan offices (Fig. 2)[38]. In summer, the background transverse survey revealed that it is uncommon for air conditioning (AC) to be used in small-scale offices but was more commonly employed in large open-plan offices. In winter, district heating is the most common heating method, but a few buildings still use electric heating (EH). Based on the characteristics of heating and AC systems, typical offices buildings were selected to give a range of different types and sizes.

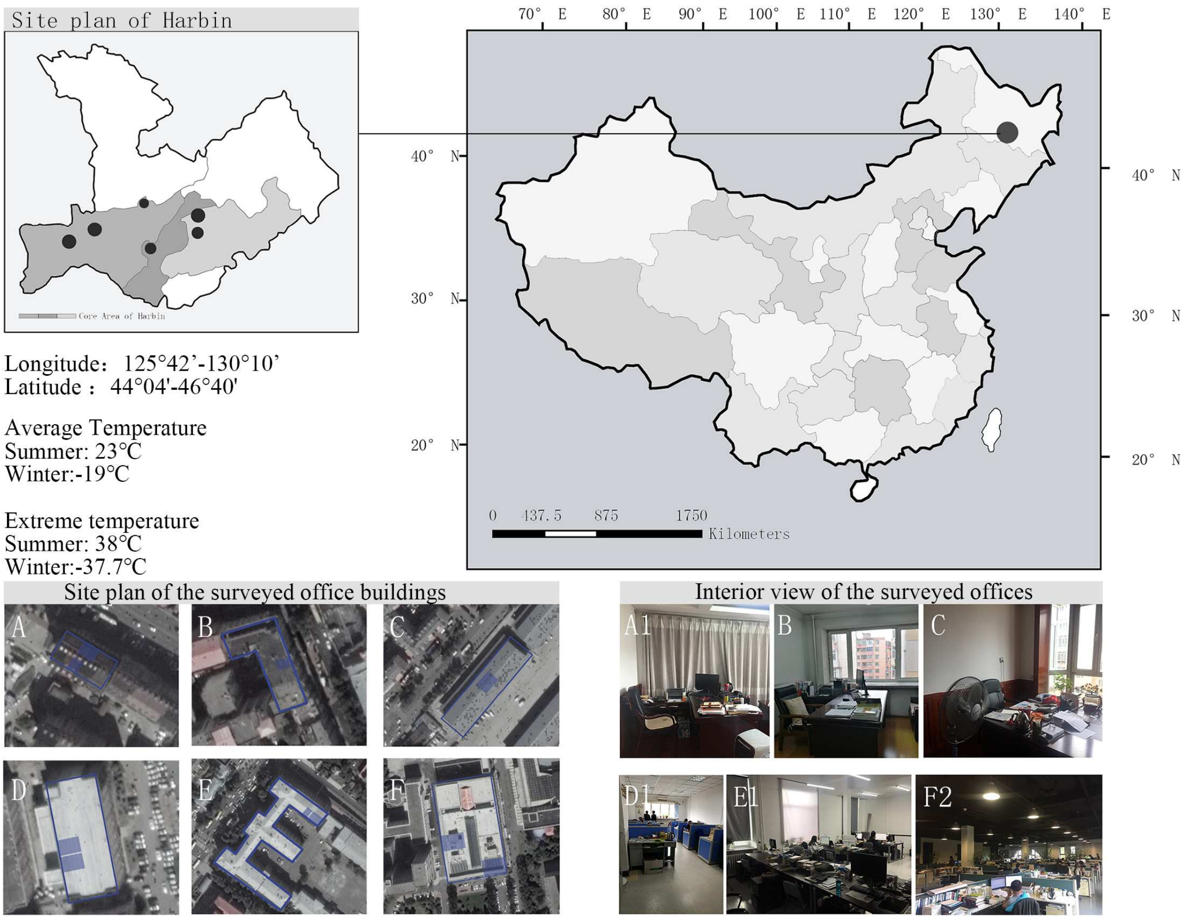


Fig. 2 Site and location of the surveyed offices in Harbin, North-eastern China [38]

All basic building information, including the characteristics of subjects, geometric parameters, and the available environmental equipment controls, are shown in Table 1. The surveyed offices include four private and shared-private offices, two open-plan offices with 3-10 occupants, two open-plan offices with 11-20 occupants, and two open-plan offices with more than 20 occupants. In summer, occupants in offices D1 and D2 were able to control single-unit AC, and those in office D1 chose to switch-off the AC when feeling cold. The AC in building D was removed in the second week of the summer survey for equipment replacement. It should be noted that all the offices in this study are in buildings with east- or west-facing main façades, due to the limitations of the urban layout of the available buildings. However, according to background research, in Harbin, the main façade in most office buildings are oriented east-west, rather than north-south. There were 80 occupants who completed the questionnaire survey. The number of the occupants in each surveyed office was defined as per ASHRAE Standard 55-2013 [39]. The ratio of male to female was close to 1, similar to the results of transverse surveys.

Table 1 Basic information for the occupants of the surveyed offices and available facility controls

Office No.	Subjects (Surveyed)	No. Male	Office type	Room size (m^2)	Orientation	Available Control Summer	Winter
A1	1(1)	0	Private	25.62	Northeast	Fan	EH[c]
A2	2(2)	0	Shared-private	15.47	Southwest	Fan	EH
B	1(1)	1	Private	21.74	Northeast	Fan	DH[d]
C	2(2)	0	Shared-private	18.6	Northwest	Fan	DH
D1	5(5)	2	Open plan	40.34	West	AC[a] + fan	DH
D2	5(5)	4	Open plan	40.34	West	AC[a]+ fan	DH
E1	15(15)	6	Open plan	66.2	Southwest	Fan	DH
E2	11(6)	6	Open plan	37.66	Southwest	Fan	DH
F1	50(22)	15	Open plan	380	West	AC[b] + fan	DH
F2	50(21)	9	Open plan	380	East	AC[b] + fan	DH

Notes: a. AC is single unit air conditioning;
b. AC is central air conditioning;
c. EH is electric heating; d. DH is district heating.

2.2 Longitudinal survey

2.2.1. Panel Questionnaire survey

The panel questionnaire applied in this study was a survey that used the same subjects from the same office environments to track their changes in thermal comfort and adaptive behaviour during the year. This panel questionnaire survey was designed so that the results could be integrated with long-term monitoring behaviour data, so that the interactive relationship between naturally ventilated behaviour, occupant experience, and physical environment parameters could be obtained.

Fig. 3 shows the framework of the methodology. A pilot questionnaire was conducted for a week among workers in similar office environments; then, the official questionnaire survey was conducted. The pilot questionnaires helped to improve the clarity of presentation and integrity of the survey content, according to the opinions and feedback from the subjects. Following the pilot survey, a three-part questionnaire was conducted, consisting of a start survey, a daily survey, and a final survey. The questionnaires were sent via WeChat, the most commonly used social media software in China, at 10:00 am in the morning and 3:00 pm in the afternoon for two weeks to help the subjects develop the habit of answering the questions on time. The purpose of the start questionnaire was to obtain some basic information from the subjects, their overall feelings of their office environment, and the range of adaptive control behaviours available to them under restrictive office conditions. The daily survey sought to obtain the occupant clothing level, thermal comfort experiences on a seven-point scale [39], and different behavioural status at the time of the questionnaire, involving the most concise and accurate questions. The final survey focused on summary questions about the overall thermal comfort experience during the two-week survey and the satisfaction with the questionnaire.

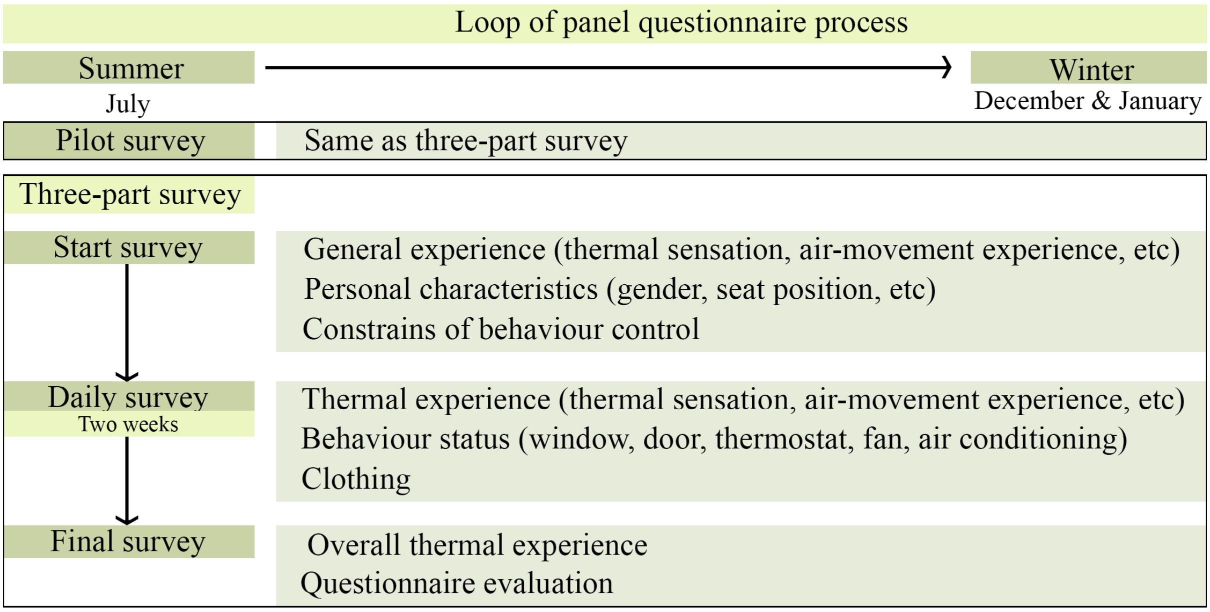


Fig. 3 Framework and content of the three-step panel questionnaire survey [38]

2.2.2. Field measurements

The spatial organisation of the office buildings, along with the geometric design parameters of the monitored buildings, were measured in detail to facilitate data analysis and simulation modelling, using an infrared rangefinder (model UT392). Indoor and outdoor physical parameters were continuously recorded by a weather station (E-Log environmental data logger) and HOBO U12 data loggers (air temperature and relative humidity) at 30-min and 15-min intervals, respectively (these intervals are also used by the dynamic building energy modelling software DesignBuilder that was used in another part of this study, to be described later).

Occupant adaptive control, including the use of fans, AC use in summer, and heating facilities in winter, were recorded by the panel questionnaires at the time the questionnaires were answered, together with the continuous measurement of window status using the Hobo UX 90-001 state/event data loggers. The number and duration of the status changes were recorded. Due to equipment limitations, the size of opening could not be recorded, but the windows in the surveyed buildings were all casement windows. This is the most common window type for office buildings in Harbin, and based on the data statistics of the background transverse survey, they are usually opened fully in most cases.

2.3 Data analysis

Data mining generally refers to the process of searching for hidden information in a large amount of data using algorithms. Fig. 4 shows the work flow of the data mining in the classification and characterisation of the window-opening behaviour in the different sized Harbin offices.

In this research logistic analysis was applied to analyse the influence factors of window-opening behaviour. The degree of association between changes in window status and each parameter, including non-nominal and nominal variables obtained from the panel questionnaires and measured datasets, was examined by binary logistic regression. Logic regression analysis results were combined with the measured distribution characteristics of long-term behaviour to determine the influencing factors of window-opening behaviour in north-east China.

Cluster analysis was used to obtain the window-opening duration patterns of the office occupants via monitoring data over one year. To form continuous and operational working user profiles, the time of the day was divided into six periods, early morning, morning, noon, afternoon, evening, and night. The average performance of window-opening behaviour duration in these time periods of summer and winter was grouped using cluster analysis. The grouping results of weekends and workdays were separately considered. The summer and winter window-opening duration patterns were then obtained via cluster analysis.

Association rules mining was then used to classify the behaviour with the results of logistic analysis and cluster mining. Each office was classified into its own type, according to the influencing factors of window-opening behaviour and the window-opening duration patterns in summer and winter. Finally, occupant window-opening behaviour profiles were formed and then linked to the modelling of building thermal performance, using the dynamic analysis software DesignBuilder.

The cluster analysis and association rule mining were employed, along with the open source data mining program Rapid Miner, to mine the classification of the window-opening behaviour.

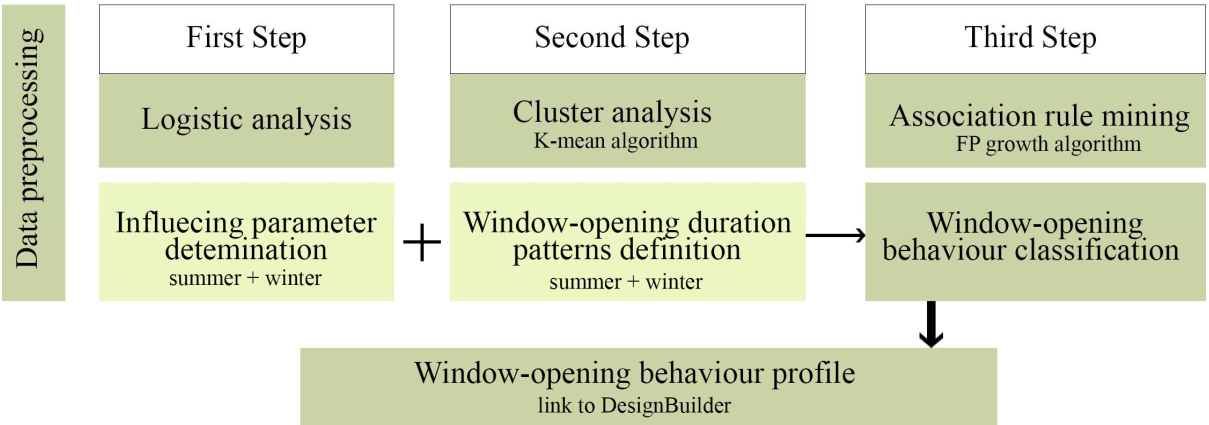


Fig. 4 Method for window-opening behaviour pattern and working profile definition

2.4 Statistical analysis technique

Logistic regression analysis is a generalised linear regression analysis model, an algorithm used for classification and prediction, which characterises the influencing factors of nominal variables and the predictive probability of the occurrence of events. To solve a problem of regression or classification, a cost function is established; the optimal model parameters are iteratively solved by an optimisation method and, finally, the quality of the model is verified. For binary logistic regression, when there are only two dependent variables (e.g. happen or not happen), a regression analysis between conditional probability $P\{Y=1|x\}$ and x is used, substituting the difficult method by attempting to build the relationship between independent and dependent variables directly, which is equivalent to looking at a value in the domain of a continuous function from 0 to 1. Equations (1) and (2) describe this relationship of P and x:

$$Logit(P) = \ln P / (1 - P) \quad (1)$$

$$P = \frac{\exp(\beta_0 + \beta_1 x_1 + ... + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + ... + \beta_k x_k)} \quad (2)$$

where:

P is the probability

$P / (1 - P)$ is the odds ratio

The changes in behaviour status often correspond to the categorical variables, such as window open and closed. In this study, binary logistic regression was applied to define the relationship between the related variables and window-opening probabilities. The significance of the variables, based on a likelihood ratio test, using a 5% significance level, was tested to estimate the regression coefficients.

2.5 Data mining techniques

2.5.1. Cluster analysis

Cluster analysis is processed to classify similar objects into different groups or subsets by statistical classification, so that all the member objects in the same subset have similar attributes.

In this study, the window-opening duration modes in the observed offices were analysed using the K-means clustering approach [40]. This method involves a vector quantisation of clusters and is the most commonly used algorithm for basic clustering. For a data set D, K-means clustering initially distributes the n data points in D into k random clusters. Each cluster is associated with a centroid (centre point), and the distance from each data point to all k centroids is calculated. A data point is then assigned to the cluster whose centroid is closest to it so that similar data points can be gathered together. It is an iterative method, and the next iteration calculates the new centroids of these new clusters by calculating the average of the distances between the points and the centroid. This iteration continues until convergence is achieved.

The similarity between clusters is usually evaluated via the distance between groups, and the distance is obtained through the measured Euclidean distance (Equation (3)).

$$d(a,b) = d(b,a) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + \cdots (b_n - a_n)^2} \quad (3)$$

where,

$a = (a_1, a_2, \cdots a_n)$,

$b = (b_1, b_2, \cdots b_n)$,

and a and b are two points in Euclidean space.

The performance of clustering was evaluated using the Davies-Bouldin Index (DBI) index. The DBI index refers to the ratio between the average distance in the group and between the groups (Equation (4))

$$E = \frac{1}{n}\sum_{i=1}^{n} \max_{i \neq j}\left[\frac{R_i + R_j}{M_{ij}}\right] \quad (4)$$

where:

n is the group number,

R_j is the average distance inside group j found by averaging the distance between each cluster data point and the cluster centre,

and M_{ij} is the distance between the centre of each group.

According to Equation (3), a smaller DBI value indicates better performance for the cluster algorithm result. Groups with low DBI indicators represent clusters of low internal distances (i.e., high cluster similarity) while high DBI indicators represent clusters of high internal distance (i.e., low cluster similarity).

2.5.2. Association Rules Mining

The purpose of applying association rules mining is to find the relationship between variables in large data sets and reveal the implicitly related features in the data [41]. The general form of association rules mining can be presented by Equation (5).

$$X \Rightarrow Y \quad (5)$$

where:

X is the preceding item of the rule,

Y is the latter item of the rule,

X and Y can be a project or an item set from the data set.

Although many association rules relationships can be identified via the method, only a few of the relationships may be valid. There are two values for evaluating the validity of the mining results: *Confidence* and *Support*.

Confidence is the measurement of the accuracy of the association rules. It describes the probability of item Y containing item X, and reflects the possibility of Y appearing under the condition of X. If the confidence level is high, the possibility of the emergence of X is high, reflecting the conditional probability of Y under a given X. Its formula can be described as Equation (5)

$$C_{X \to Y} = \frac{|T(X \cap Y)|}{|T(X)|} \tag{5}$$

where:

$|T(X)|$ represents the number of transactions that contain the project X,

$|T(X \cap Y)|$ means the number of transactions that contain both the project X and the project Y.

Support measures the universality of the association rules and represents the probability of the concurrent occurrence of project X and project Y, and the formula is

$$S_{X \to Y} = \frac{|T(X \cap Y)|}{|T|} \tag{6}$$

where:

$|T|$ represents the total number of transactions

Confidence and *Support* can only measure the validity of the results of association rules, but they fail to measure whether the results are practical. Therefore, the index of lift is applied to measure whether the appearance of X can motivate Y. The index of lift, shown in Equation (7), is the ratio of *Confidence* to later *Support*, and the greater the value is, the better are the results.

$$L_{X \to Y} = \frac{C_{X \to Y}}{S_Y} = \frac{|T(X \cap Y)|}{|T(X)|} / \frac{|T(Y)|}{|T|} \tag{7}$$

In this study, the frequent pattern growth algorithm (FP-Growth algorithm) was applied to define the classifications of window-opening behaviour with the results of influencing factors and the duration modes of window-opening.

3. Results and Discussion

The results were analysed to define the window-opening behaviour type for modifying the simulation gap. The main outcomes are summarised as follows:

3.1. Thermal comfort characteristics

During the summer season in Harbin, the indoor temperature T_{in} was maintained at around 30°C in the surveyed natural ventilation buildings, and around 27.5 °C in the AC offices. The Chinese evaluation standards for indoor thermal environments in civil buildings (GB/T 50785-2012) [42] limit the range of T_{in} to between 18°C to 28°C, which means that the indoor temperature of all naturally ventilated offices were in the uncomfortable range. In winter, the indoor temperature of all surveyed offices was in the comfortable temperature range.

During the summer survey, the outdoor temperatures were generally lower for the second half of the questionnaire, with the average value dropping from 30.2°C to 26.9°C. In winter, the change was upwards, from -17.2°C to -13.5°C. Fig. 5 to Fig. 8 shows the scatter plots of average thermal sensation and thermal satisfaction votes for each surveyed office building, and the line chart shows the mean value of votes for occupants from offices of different sizes and layouts.

The summer thermal sensation votes show that all kinds of office buildings with the same size had a certain degree of consistency (Fig. 5). Occupant sensations in offices with 1-2 persons were very hot in the first week and neutral in the second week due to the decline in outdoor temperature. After the removal of AC equipment, although the outdoor temperature cooled down, the occupants of offices with 3 to 10 persons (offices D1 and D2) felt very hot. Occupants felt the same level of warmth in offices of 10 to 20 people during all the survey runs, indicating that the average value of the thermal sensation vote changed little when the outdoor temperature varied between 26.9°C to 30.2°C. The thermal sensation vote of office occupants using AC equipment was neutral because of the stable indoor environment.

In winter, most of the mean thermal sensation vote results were in the level between cool and neutral, except those from the open-plan offices with 3 to 10 occupants in which the value was between cool and cold (Fig. 6). According to the line chart of the mean thermal sensation vote, divided by office size, all occupants' experiences in the winter were between cool and warm.

From the thermal satisfaction vote results in Fig. 7, only occupants in offices with three to ten persons experienced low satisfaction below 'dissatisfied' in most cases in summer. The average level of thermal satisfaction corresponded to the level of thermal sensation in offices with 3–20 and > 20 occupants. Private and shared-private offices had a neutral assessment about the indoor environment, despite the hot experience of the thermal sensation, which may be because the occupants working in the more independent office environments had more control over their behaviour. In winter, the occupants in offices with 3 to 10 persons also presented a low level of satisfaction, while others in the range were 'a little dissatisfied' to 'a little satisfied', which is consistent with the mean thermal sensation voting (Fig. 8).

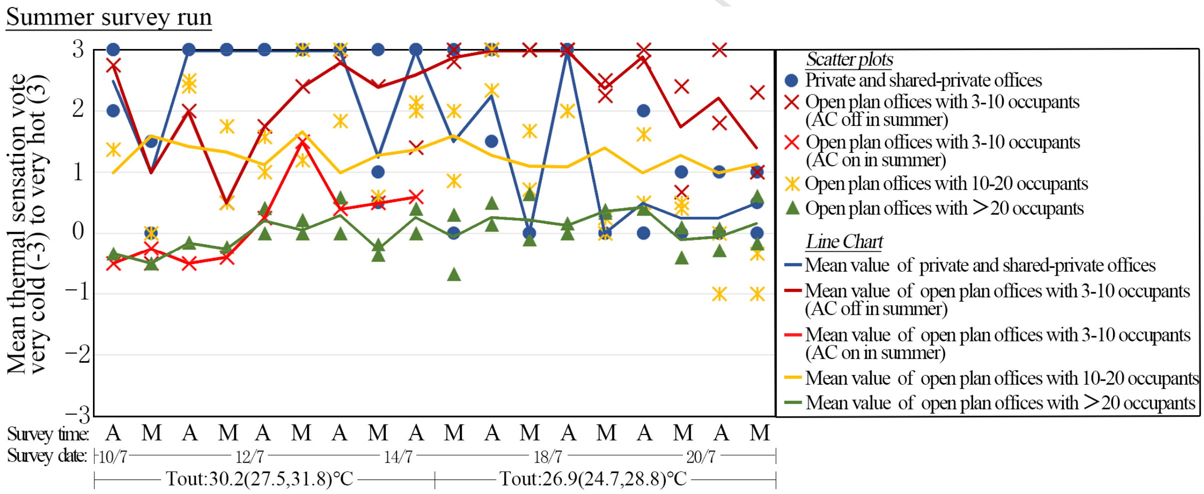


Fig. 5 Scatter plots and line chart of average thermal sensation vote of occupants in different-sized offices in summer

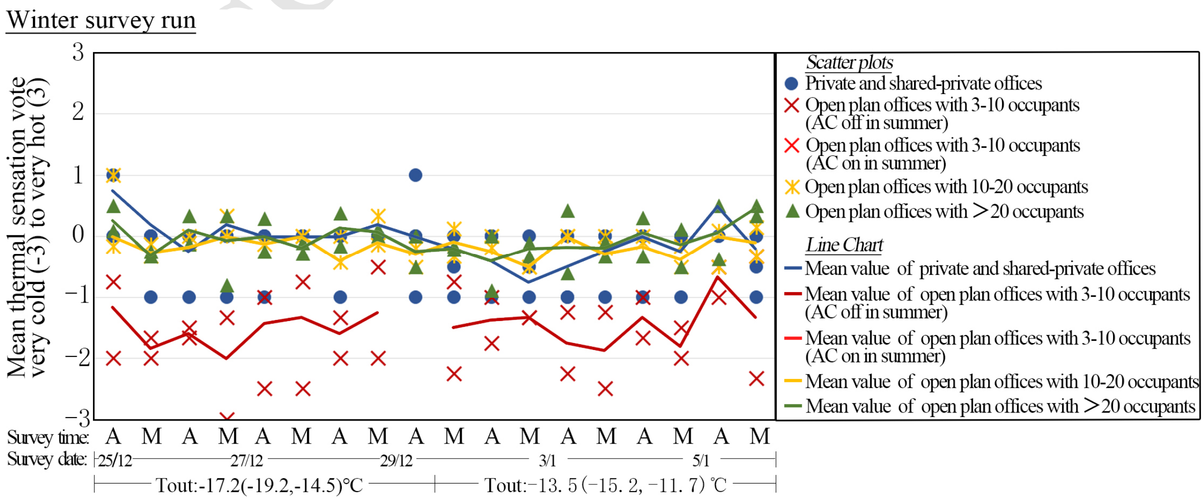


Fig. 6 Scatter plots and line chart of average thermal sensation vote of occupants in different-sized offices in winter

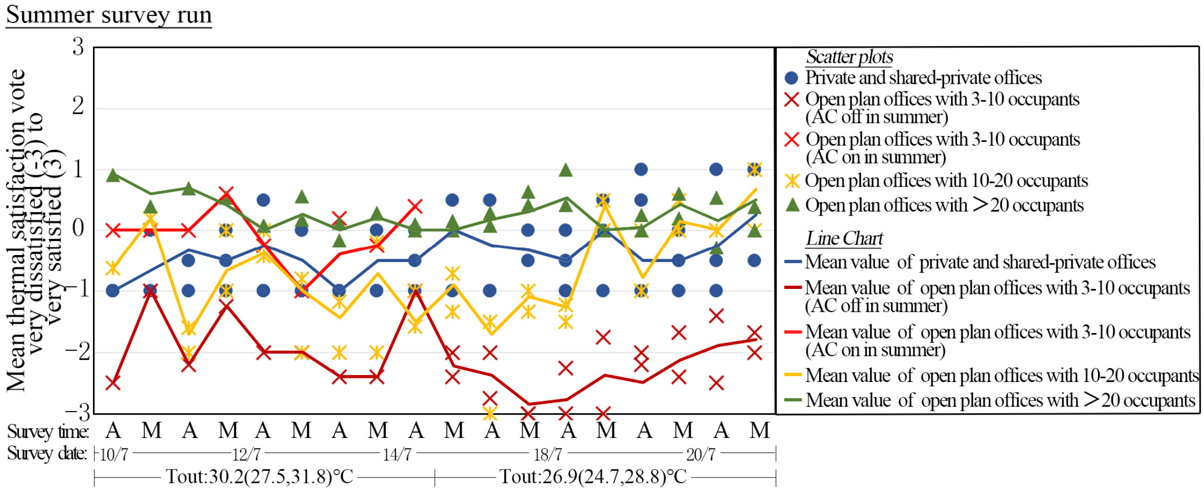


Fig. 7 Scatter plots and line chart of average thermal satisfaction vote of occupants in different-sized offices in summer

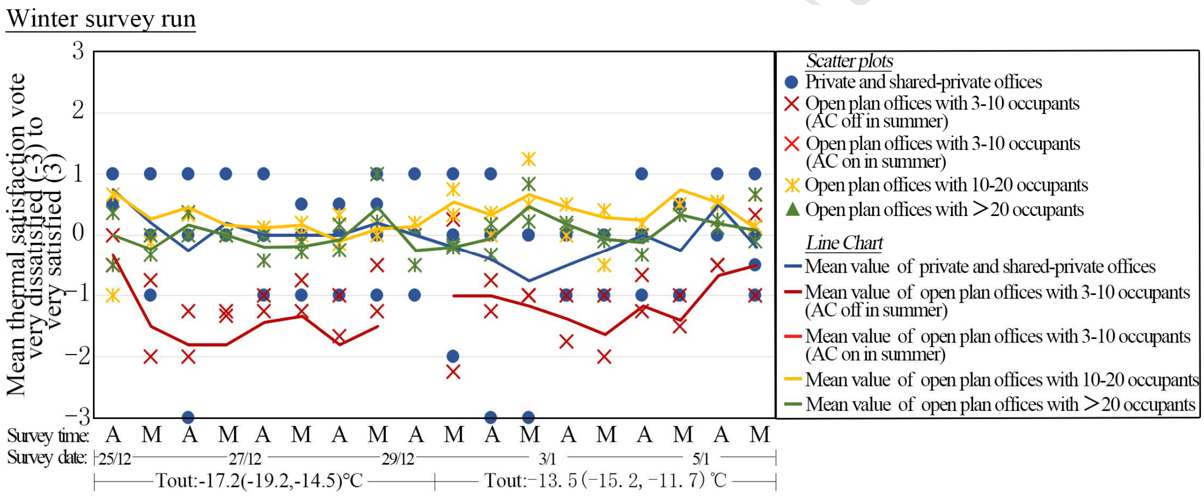


Fig. 8 Scatter plots and line chart of average thermal satisfaction vote of occupants in different-sized offices in winter

3.2. Behaviour control characteristics

One day time was divided into six intervals to obtain the formation of a continuous window-opening behaviour profile, containing early morning time, morning time, noon time, afternoon time, evening time, and night time (Table 2).

Table 2 presents the average value of window-opening duration in each surveyed office on workdays and weekends in summer and winter, with the value of the variance measuring the dispersion of the recording data. Around 50% of the surveyed offices from small- to large-scale exhibited a window-opening time of no closures across the entire summer typical season in July during day and night, which means the occupants never closed the window during this period. It is worth noting that most of these buildings also showed the extremely opposite performance of having totally closed windows during the days and nights on weekdays and weekends in December and January.

In summer, four rooms (A2, D1, D2, D2', and F2) kept their windows in an open state during the work time, with most of the occupants keeping to a routine of opening the window when they arrived and closing the window when they left. In winter, among these surveyed offices, D1, D2, and F2 rooms also showed a short duration of opening windows at the time of people's arrival or their lunch break.

In Harbin, it is a very common phenomenon that people work overtime on weekends. In summer, the windows in the offices with day-night window-opening behaviour were open on weekends throughout the summer. The other offices of the 'routine type' presented a greater dispersion of window-opening behaviours, which may be due to increased randomness of the overtime work period on the weekends. In winter, the window-opening behaviour of occupants of all the surveyed office buildings was remarkably consistent, that is, no window-opening at all during the winter season, which may be due to the fact that the cold winter reduced the overtime hours or the overtime on weekends in winter was not too long. Due to the characteristics of window-opening behaviour, most surveyed office rooms showed a correlation with habit in winter and summer.

Table 2 Window-opening duration of weekdays and weekends in each surveyed office in summer and winter with average and variance value

Summer results												
	Window-opening duration on work days (hours) (average value (variance))						Window-opening duration on weekends (hours) (average value (variance))					
	6 am-9 am	9am-12am	12am-3pm	3pm-6pm	6pm-9pm	9pm-12pm	6-9am	9am-12am	12am-3pm	3pm-6pm	6pm-9pm	9pm-12pm
A1	3(0)	3(0)	3(0)	3(0)	3(0)	6(0)	3(0)	3(0)	3(0)	3(0)	3(0)	6(0)
B	3(0)	3(0)	3(0)	3(0)	3(0)	6(0)	3(0)	3(0)	3(0)	3(0)	3(0)	6(0)
A2	0.45(0.2)	2(0.1)	2(0.6)	0.25(0.1)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
C	3(0)	3(0)	3(0)	3(0)	3(0)	6(0)	3(0)	3(0)	3(0)	3(0)	3(0)	6(0)
D1	1.25(0.1)	3(0.1)	2.45(0.2)	2(0.2)	0.25(0)	0(0)	0(0)	1(0.9)	1(0.2)	0.5(0.1)	0(0)	0(0)
D2	1(0.3)	1(0.4)	0.25(0.2)	0.25(0.2)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
D2’	1(0.4)	2(0.2)	2(0.2)	1.25(0.1)	0(0)	0(0)	0.5(1)	2(0.2)	2(0.4)	0.75(0.8)	0(0)	0(0)
E1	3(0)	3(0)	3(0)	3(0)	3(0)	6(0)	3(0)	3(0)	3(0)	3(0)	3(0)	6(0)
E2	1.5(1.2)	2.5(1.4)	2.5(0.4)	2.5(0.4)	2.25(1)	5(2)	2.5(1)	3(0.1)	3(0.1)	3(0.6)	2.5(1)	6(1)
F1	3(0)	3(0)	3(0)	3(0)	3(0)	6(0)	3(0)	3(0)	3(0)	3(0)	3(0)	6(0)
F2	1.25(0.2)	3(0.1)	2.45(0.3)	2(0.1)	0.25(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
Winter results												
	Window-opening duration on work days (hours) (average value (variance))						Window-opening duration on weekends (hours) (average value (variance))					
	6am-9am	9am12am	12am-3pm	3pm-6pm	6pm-9pm	9pm-12pm	6-9am	9am-12am	12am-3pm	3pm-6pm	6pm-9pm	9pm-12pm
A1	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
B	0.02(0)	0.05(0)	0.1(0.1)	0.1(0.1)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
A2	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
C	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
D1	0.25(0)	0(0)	0.05(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
D2	0.1(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
E1	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
E2	0.04(0)	0.07(0)	0.2(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
F1	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
F2	0(0)	0.16(0)	0.04(0)	0.02(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)

Notes: a. A1 and B are private offices, A2 and C are shared-private offices, D1, D2 (3-10 persons, AC was closed or removed for equipment update), D2’ (with AC on in summer), E1 and E2 (11-20 persons) are open-plan offices, F1 and F2 (> 20 persons) are open-plan offices (with AC in summer).

b. The offices on bold were those open all the time in summer and closed all the time in winter.

3.3. Behaviour influencing factors

The correlation between potential influencing factors and window-opening behaviour was analysed in summer, winter, and two quarters to assess the main factors affecting behaviour in a single season and across different seasons (Table 3). The correlation of physical parameters, consisting of indoor and outdoor temperature and relative humidity, with the window state was determined using logistic analysis. The occupants' experience of the thermal environment were also included in the discussion for further understanding of the interaction between the occupants' thermal comfort experience and the window-opening behaviour control. Nominal variables, e.g. season and morning/afternoon, were tested for correlation with the window-opening behaviour.

According to the basic features of window-opening of office buildings in different scales, there are significant differences in the opening duration in winter and summer, and this was also verified by the result of correlation analysis. From the perspective of data analysis throughout the year, temperature and relative humidity indoors and outdoors, as well as the corresponding thermal sensation, temperature preference, and humidity feelings, were related. Meanwhile, the specific statistics from the summer and winter analysis showed that the window-opening behaviour in the surveyed office rooms did not show high correlation to the temperature or humidity change in an individual season.

In summer, the occupant window status varied with outdoor temperature only in offices F1 and F2. Correspondingly, the thermal sensation feeling influenced the window-opening behaviour of occupants in F1 and F2, and temperance preference in F2. The window-opening changes were also influenced by indoor relative humidity and humidity feelings for the occupants in F1 and F2. Occupants in F1 also thought the air movement and overall satisfaction were the reasons for their behavioural changes towards the window.

In winter, there were only two surveyed offices presenting correlation of environmental physical parameters and thermal sensation evaluation vote. There was significant correlation for the temperature preference, air movement, and overall satisfaction with the window status in office B but no physical factors, and in addition to the temperature preference, users of office F2 were affected by outdoor temperature, indoor humidity, and the corresponding thermal sensation and humidity feeling.

Table 3 Influencing factors of window-opening behaviour in summer, winter, and across the entire year

	Nominal variables	Non-Nominal variables														
	Season	Tin			Tout			RHin			RHout			Clo		
		S	W	H	S	W	H	S	W	H	S	W	H	S	W	H
A1	√	×	×	√	×	×	√	×	×	√	×	×	×	×	×	√
B	√	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
A2	√	×	×	√	×	×	√	×	×	√	×	×	√	×	×	√
C	√	×	×	×	×	×	×	×	×	×	×	×	×	-	×	×
D1	√	×	×	√	×	×	√	×	×	√	×	×	√	×	×	×
D2	√	×	×	√	×	×	√	×	×	√	×	×	√	×	×	×
E1	√	×	×	√	×	×	√	×	×	×	×	×	√	×	×	√
E2	√	×	×	√	×	×	√	×	×	√	×	×	×	×	×	√
F1	√	×	×	√	√	×	√	√	×	√	×	×	×	×	×	√
F2	√	×	×	√	√	√	√	√	√	√	×	×	×	×	×	√
	Interval variables															
		Thermal sensation feeling			Temperature preference			Humidity feeling			Air movement			Overall satisfaction		
Office No.		S	W	H	S	W	H	S	W	H	S	W	H	S	W	H
A1		×	×	√	×	×	√	×	×	√	×	×	×	×	×	√
B		×	×	√	×	√	√	×	×	×	×	√	×	×	√	√
A2		×	×	√	×	×	√	×	×	×	×	×	×	×	×	×
C		-	×	√	-	×	√	-	×	√	-	×	×	-	×	√
D1		×	×	√	×	×	√	×	×	×	×	×	×	×	×	√
D2		×	×	√	×	×	√	×	×	×	×	×	×	×	×	√
E1		×	×	√	×	×	√	×	×	√	×	×	√	×	×	√
E2		×	×	√	×	×	√	×	×	√	×	×	×	×	×	√
F1		√	×	×	×	×	×	√	×	√	√	×	√	√	×	×
F2		√	√	√	√	√	√	√	×	√	×	×	×	×	×	×
Tested but not relevant parameters					Smell; Outdoor noise; Outdoor Air Quality; Time: (Morning or afternoon); fan use; AC use											

Notes: √ means $p < 0.05$, the correlation is significant; × means $p > 0.05$, the correlation is not significant.
- means no data was collected because of holidays and other reasons.
S is summer; W is winter; and H is the results considering winter and summer.
Tin is indoor temperature; Tout is outdoor temperature; RHout is outdoor relative humidity; Rhin is indoor relative humidity; Clo is thermal resistance of clothing; No. is number.

In summary, for all surveyed buildings, the window status change had the strongest correlation with the seasons (Table 4). Changes in physical parameters within a certain threshold in summer and winter only affected the window-opening behaviour of users in a few buildings. In Section 3.2, the statistical results of window-opening duration showed that some of the occupants exhibited the habit of opening all windows in summer and closing all windows in winter. Some other offices kept the routine of opening the window during work time in summer and winter. Office E2 was the only office with random window-opening behaviour. Combining with the analysis of the basic features of the window-opening behaviour in the previous section, it can be speculated that the main factors influencing behaviour included season, habit, and thermal comfort experiences, and this will be further tested in the next section as the premise input for association rules mining, via classifying the categories of behaviour. Other variables, including smell, outdoor noise, outdoor air quality, time (morning or afternoon), and other behaviour were also tested but showed no correlation with the behaviour.

Table 4 Statistics of window-opening behaviour influencing factors across the whole year

Influencing factors	Offices No.									
	A1	B	A2	C	D1	D2	E1	E2	F1	F2
Season	√									
Thermal comfort	×	√[a]	×	×	×	×	×	×	√[b]	√[c]
Habit	√							×	√	

Notes: a.√ means the factor affect window-opening behaviour in winter;
b.√ means the factor affect window-opening behaviour in summer;
c.√ means the factor affect window-opening behaviour in summer and winter.

3.4. Window opening duration patterns

With the application of cluster analysis, three window-opening duration patterns in summer and four patterns in winter were obtained via the data of occupant performance of window-opening behaviours in offices from private offices to open-plan offices in north-east China.

3.4.1 Summer window-opening duration patterns

In the typical summer month of July, according to the statistics in Table 2, some of the surveyed offices exhibited a window open time of the entire duration of the month, while others showed a close connection between the working hours and window open time. The cluster results (Fig. 9) agree with these basic statistics of the window-opening behaviour, and three types are defined.

The first type was where windows were open the entire time on weekdays, which consisted of private offices A1 and B, shared-private office C, and open-plan offices E1 and F1. The offices belonging to this type were also those with the continuous window-opening behaviour during the weekends. Type two was the “working time opening routine” example, which was also consistent with the statistics of Table 2, comprising shared-private office A2 and open-plan offices D2 (with and without single unit AC) and F2 (with central AC). For this type, the occupants used the window during the work time period. This pattern of behaviour similarly occurred during the weekends. After cluster mining, E2 belonged to a separate category, and it also had the phenomenon of open window during the day and night, but its duration was shorter than the duration of type 1 in each period of time. It can be seen from Table 2 that the duration of the E2 open window in each part of the day has great discreteness. The duration of occupants working in E2 had a larger randomness by the monitoring results from the space occupancy sensor HOBO UX90-006x, which may lead to this discreteness.

3.4.2 Winter window-opening duration patterns

Data from the typical two winter months, December and January, were included in the data mining. Four types of window-opening behaviour duration were defined (Fig. 10). Table 2 also reveals that, on the weekends, all windows of the monitored offices were in the state of being totally closed during these two months.

Type 1 was windows opened for a short time of 1 min or less during the early morning and morning time, which included five situations in which offices had entirely closed windows, except office D2. The other three types of window-opening behaviour in winter involved a relatively longer open window time of about 15 min, but each of the open times were concentrated at different time periods: type two was in the early morning, type three was in the morning time, and type four was in the noon time. These three types of windows also had a very short time of open window, 1–3 min during other periods of the day.

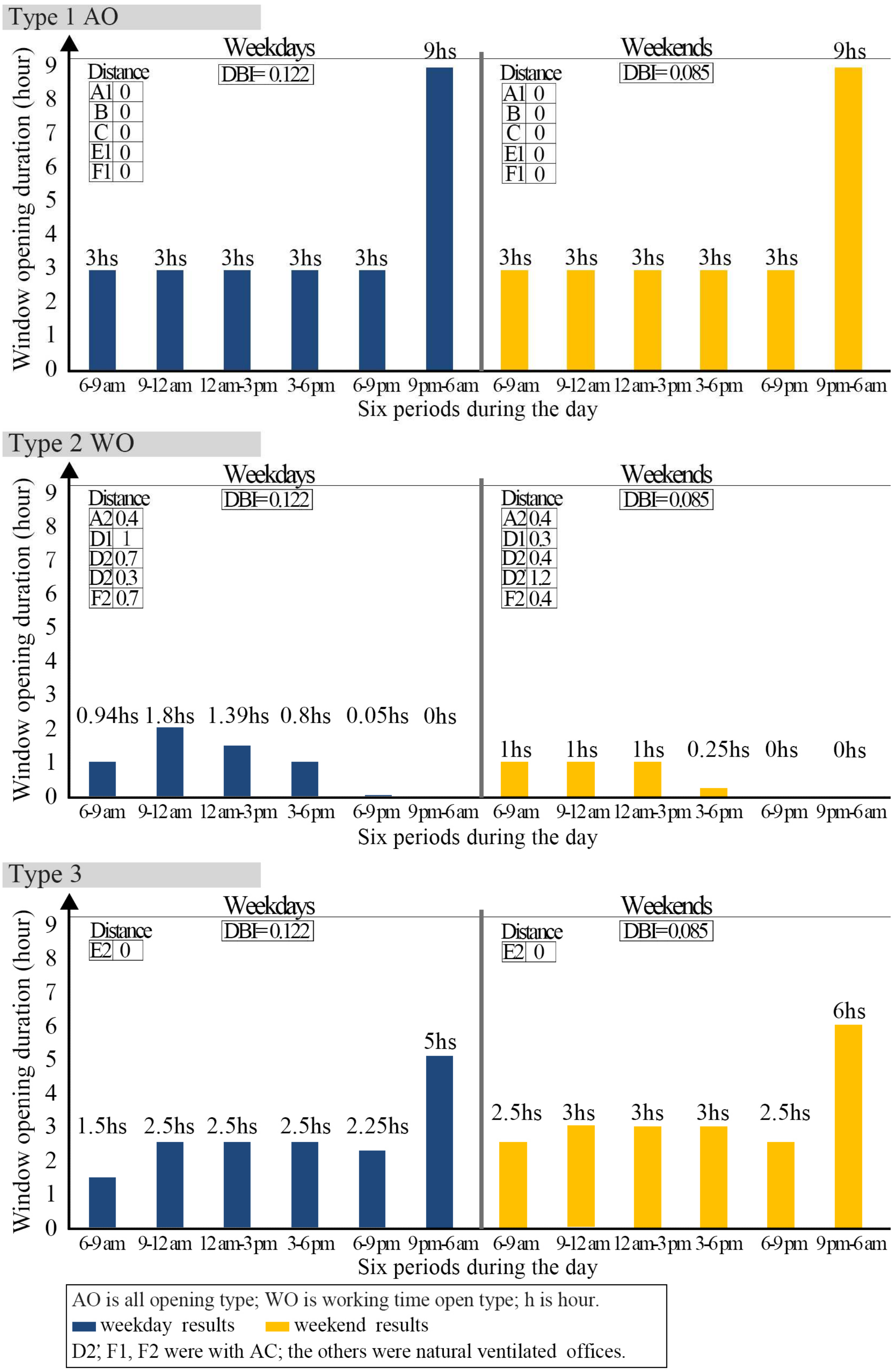


Fig. 9 Cluster-mining results of summer window-opening behaviour duration for the ten surveyed offices

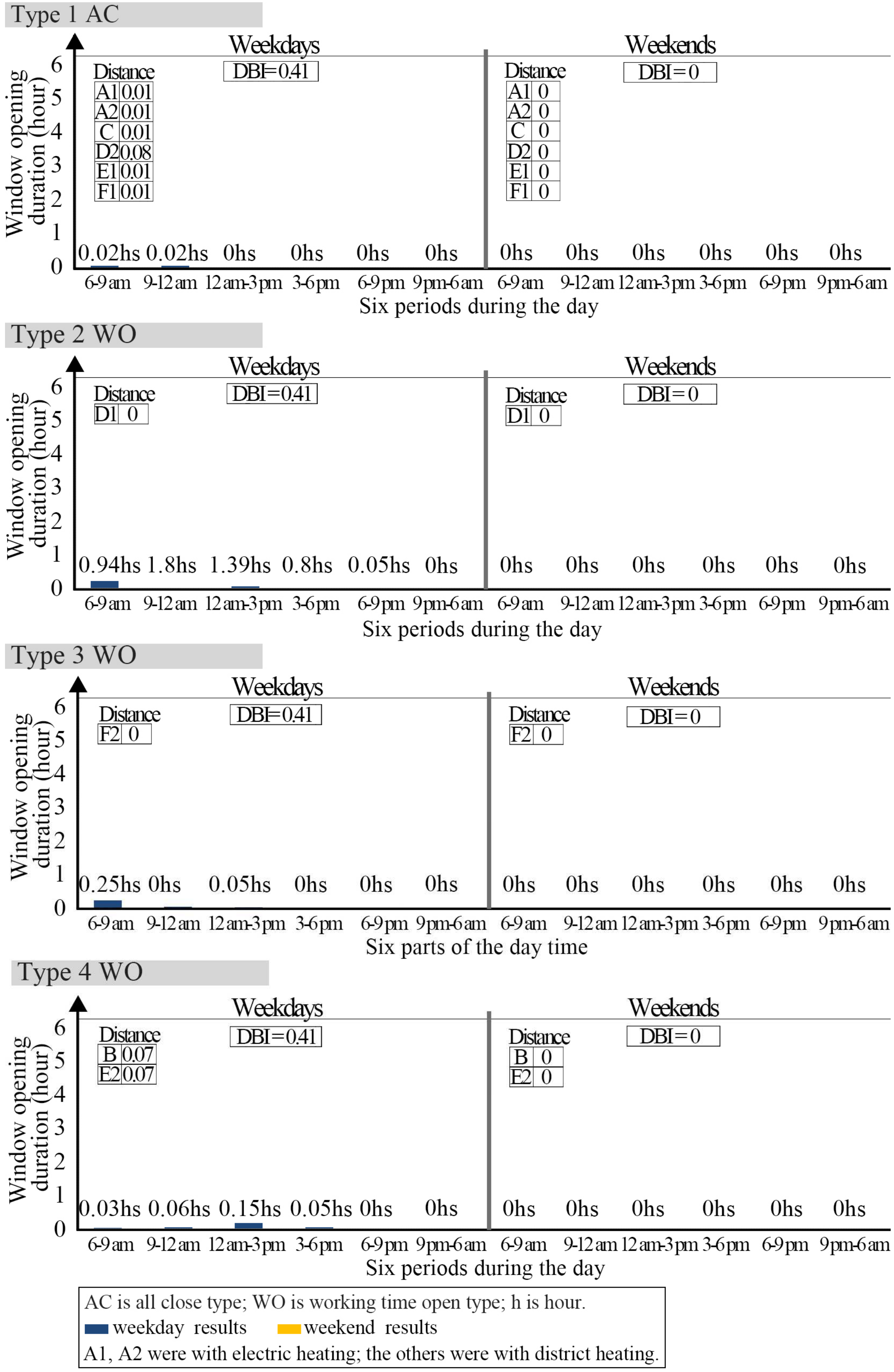


Fig. 10 Cluster-mining results of winter window-opening behaviour duration for the ten surveyed offices

3.5. Window-opening behaviour classification and profiles

Association rules mining was applied for classifying the window-opening behaviour together with the summer and winter opening duration and influencing factors. To get significant results from the association rules mining, the values of *Support*, *Confidence*, and lift were set at the minimum thresholds of 30%, 80%, and 1, respectively. The criteria were prescribed for each rule mined, which is that at least 30% of the data contained the premise and conclusion, in which the probability that a premise led to a conclusion was greater than 80%. Simultaneously, all the results mined were positively correlated with lift > 1. Finally, five types of window-opening modes for summer and winter seasons were obtained based on mining of the monitoring database for the entire year. The types and modes are summarised in Fig. 11.

Four of these types of window-opening behaviour modes were all-season and habit-motivated, except Type 5. The seasons are the most influential factors of the occupant window-opening behaviour in all types of office buildings. With the great changes in the physical data of temperature and relative humidity in different seasons, the behaviour of window-opening had an extremely large impact. Simultaneously, these types of behavioural modes were also significantly driven by behavioural habits, in which occupant behaviour during each of the seasons presented a stable window-opening duration.

Six of the surveyed offices showed an agreement between the summer and winter window-opening behaviour modes, which were all open in summer season and all closed in winter season (AO, AC) for Type 1, and work-time open for summer and winter for Type 3. For Type 3, the occupants opened the window when they arrived, whilst in summer the duration was close to full-time open during the working hours; in winter the duration was reduced to less than 15 min. This meant that the occupants of this type of office, who were executing the work-time window-opening behaviour mode (WO) still maintained this state in winter, with a significantly reduced duration of window-opening.

In Type 1, office F1 was also motivated by thermal comfort experiences, including thermal sensation, humidity feeling, and air movement in summer. In Type 3, only F2 showed similar window-opening behaviour factors, consisting of temperature preference and humidity feelings. The occupant behaviour of Type 2 exhibited work-time opening in summer and all-closed in winter with season and habit as the motivations. Office B in Type 4 exhibited an all-opening mode in summer and work-time opening in winter during the noon time, with a short open duration of about 15 min. The behaviour mode changed in winter due to the increasing need for thermal satisfaction and better air flow. Type 5, office E2, has also been described in the previous discussion, and it was noted that, because of the greater flexibility in working hours, its window-opening behaviour showed a strong correlation only with the seasons, and the window-opening duration did not demonstrate patterns consistent with other category types in the summer. E2 did show the same performance as office B with a short open duration of 15 min in the noon period in winter.

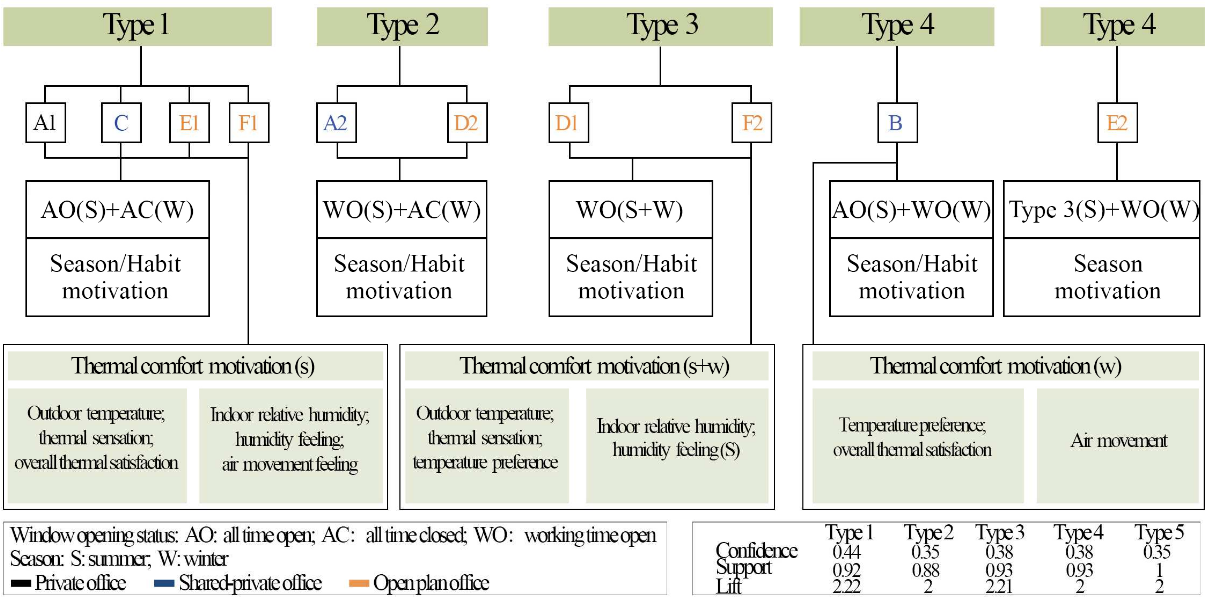


Fig. 11 Classification results of the window-opening behaviour with the duration patterns and influencing factors as premise conditions in summer and winter.

The categorised types are classified by office scales from private office to open office. Table 5 shows the types of window-opening patterns and motivational factors of occupant window-opening behaviour in different-sized offices. For the private office, offices A1 and B had the same mode of full-time opening all summer, while there was a different behaviour in winter, with all the windows closed in office A1 and open for a short duration in office B. For the shared-private office, offices A2 and C showed the same performance of all windows being closed in winter, while A2 belonged to the all-open mode, and C exhibited the work-time opening mode in summer. For the open-plan offices with 3 to 20 occupants, offices D1, D2, E1, and E2 were not consistent with each other, displaying four different modes. It should be noted that D2 retained the habit of opening the window at the time of occupant arrival, but D2 was classified as a fully closed type due to its very short window-opening time when cluster analysis was performed. For the open-plan offices of more than 20 occupants, the two surveyed offices also showed inconsistent window-opening duration patterns, but their window-opening behaviour was affected by season, habit, and thermal comfort both in summer and by season, and by habit in winter. The offices with a thermal comfort experience motivation were all in the common range of indoor temperature and relative humidity, with no large difference as in the other surveyed offices.

Table 5 Summary of window-opening behaviour classification of office occupants

Office Type	Office number	Duration mode		Motivation factors
Private office	A1	Type 1	AO(S)+AC(W)	Season, habits
	B	Type 4	AO(S)+WO(W)	Season, habits, thermal comfort (in winter)
Shared-Private office	A2	Type 2	WO(S)+AC(W)	Season, habits
	C	Type 1	AO(S)+AC(W)	Season, habits
3-20 Open-plan office	D1	Type 3	WO(S+W)	Season, habits
	D2	Type 2	WO(S)+AC(W)	Season, habits
	E1	Type 1	AO(S)+AC(W)	Season, habits
	E2	Type 5	Type 3 + WO(W)	Season
> 20 Open-plan office	F1	Type 1	AO(S)+AC(W)	Season, habits, thermal comfort (in summer)
	F2	Type 3	WO(S+W)	Season, habits, thermal comfort

Notes: S is summer; W is winter.

After obtaining the window-opening behaviour classification, behavioural profiles were finally formed, which considered the results of the mode classification definitions and the temporal degree of subdivision of the DesignBuilder software. Usually, the time step was set as 15-min intervals to obtain a suitable simulation speed. The window-opening durations of less than 15 min cannot be calculated because it is too short a duration under this setting. The results are listed in Table 4.

Table 6 Window-opening behaviour profile of occupants in different-sized offices on weekdays and weekends during summer and winter season, divided into six periods

Duration of window-opening behaviour (hours)																							
Summer												Winter											
Weekdays						Weekends						Weekdays						Weekends					
Type1: A1, C, E1, F1																							
6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am
3	3	3	3	3	9	3	3	3	3	3	9	0	0	0	0	0	0	0	0	0	0	0	0
Type2: A2,D2																							
6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am
1	1.75	1.5	1	0	0	1	1	1	0.25	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Type3: D1																							
6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am
1	1.75	1.5	1	0	0	1	1	1	0.25	0	0	0.25	0	0	0	0	0	0	0	0	0	0	0
Type3: F2																							
6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am
1	1.75	1.5	1	0	0	1	1	1	0.25	0	0	0	0.25	0	0	0	0	0	0	0	0	0	0
Type4: B																							
6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am
3	3	3	3	3	9	3	3	3	3	3	9	0	0	0.25	0	0	0	0	0	0	0	0	0
Type5: E2																							
6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am	6-9am	9-12am	12am-3pm	3-6pm	6-9pm	9pm-6am
1.5	2.5	2.5	2.5	2.25	5	1	1	1	0.25	0	0	0	0	0.25	0	0	0	0	0	0	0	0	0

3.6. Building simulation optimisation

Four offices from two of the surveyed buildings were selected to investigate whether the application of behavioural models, by linking the new window-opening behavioural profile into DesignBuilder software, could be effective in reducing the differences between simulated and real environmental data. The weather file provided by the DesignBuilder resource platform was not a 2017 weather file (the year of this study). To overcome this, the outdoor weather station conditions measured in the study were matched with two days in summer and one day in winter that were very similar to the DesignBuilder weather file data. These days (July 10 and 26 and December 6) were used in the simulations. A calculation method for infiltration into DesignBuilder was used by entering the behaviour pattern code into the software. The calculated method needed to meet the conditions is shown in Equation (8)

$$T_{zone\_air} \rangle T_{setpoint}{}^{AND} T_{zone\_air} \rangle T_{outside\_air}{}^{AND} \text{ the schedule value.} \tag{8}$$

In summer, for an office with AC, for example D1 in the first half of July, the windows would be closed when the outdoor temperature was higher than the indoor one. Therefore, it was impossible for an office with AC to be simulated, as the real scenario involved the occupants using the windows and AC together. Therefore, for offices D1 and D2, only the data on July 26 were simulated for comparison with the template inside DesignBuilder and the behaviour modes of this research (offices D1 and D2 were natural ventilated rooms due to the AC being removed in the second half of July). In winter, the outdoor temperature was very low. When the DesignBuilder window-opening schedule template was used, a full-time window-opening mode during working hours, the windows were closed after the indoor temperature was reduced to a certain extent.

These four offices respectively belonged to Type 1, Type 2, Type 3, and Type 2. The original window-opening behaviour mode inside the software was due to the work time of the occupants. Fig. 12 and 13 present a comparison of the percentage difference of indoor temperature during the simulated work time with behaviour patterns of no behaviour (no opening), mode template inside DesignBuilder, and the real pattern from the data mining results in summer, with the real measured data as the baseline. The result of the winter real mode was very close to the no window-opening behaviour; therefore, only comparisons of the simulation results with the real mode and the template are shown.

In summer, the mode detail of work-time opening type (WO) obtained from the cluster analysis was very close to the DesignBuilder mode (Office A2, D1, and D2). The results of all behaviour types with the summer WO pattern were close to those of mode template inside DesignBuilder. The discrepancy between the real all-opening pattern (AO, office A1) and the DesignBuilder one was around 2.5% and 0.8°C. All the rooms with the no behaviour control showed a relatively high value, while the differences for office D1 and D2 were smaller, which may be because the nearby rooms were all offices with AC.

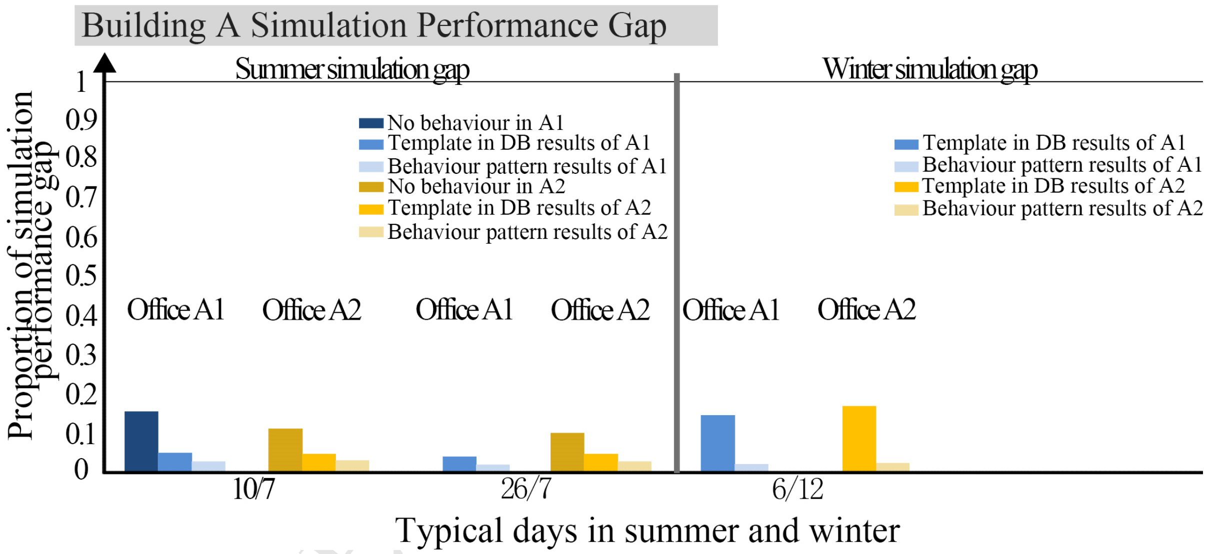


Fig. 12 Indoor temperature discrepancy applying no behaviour control, mode inside DesignBuilder, and real window-opening behaviour modes mined for private office A1 and shared-private office A2, with the measured temperature as the baseline during work time

In winter, due to the calculation method of DesignBuilder, the window always changed to closed status when the temperature goes down and, in particular, when the temperature was lower than the set point temperature for heating. For all the offices in building A (Fig. 13), with the AC mode in winter, the difference between the template in DB and the real mode result was around 13% with a temperature difference of 4.5°C. In building D (Fig. 13), the real pattern of D1 was with a

15-min opening in the early morning, and D2 was the all-closed mode. The discrepancy of D1 was 8% and 1.5°C, and for D2, it was 10% with a temperature difference of 1.8°C.

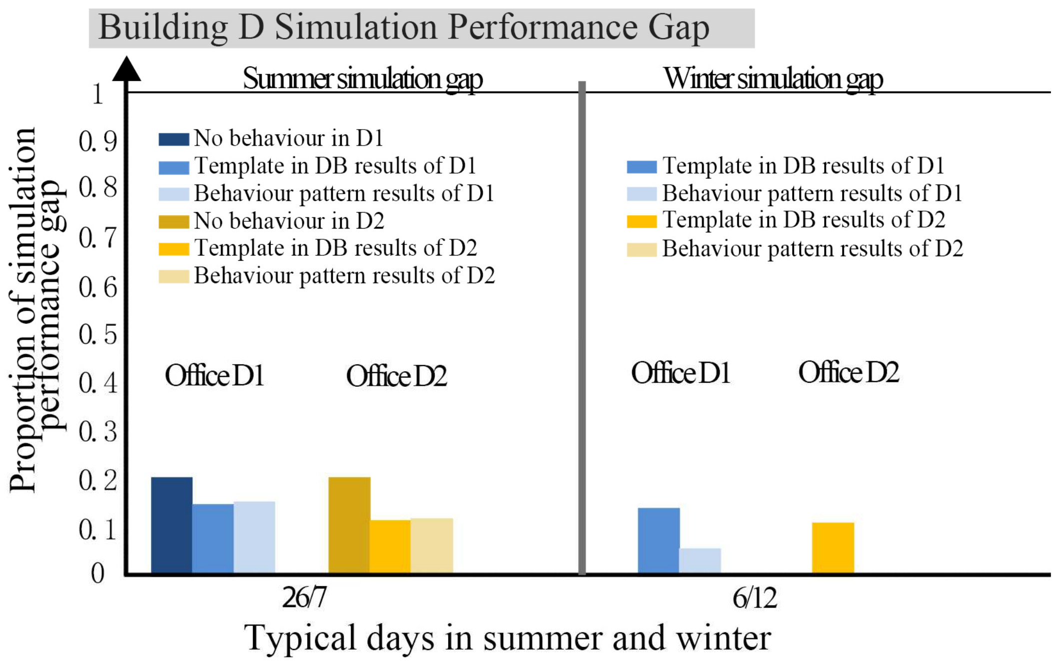


Fig. 13 Indoor temperature discrepancy applying no behaviour control, mode inside DesignBuilder, and real window-opening behaviour modes mined for open-plan offices D1 and D2, with the measured temperature as the baseline during work time

4. Conclusion and limitation

This study applied data mining techniques to obtain the real occupant window-opening behaviour modes during a one-year period that involved longitudinal questionnaire surveys and behaviour state recording of different-sized offices in the severe cold winter climate of Harbin. Window-opening duration patterns using cluster analysis and influencing factors motivating behaviour via logistic analysis were defined for conditions to further classify the behaviour modes and form the behavioural profiles that were used in building performance simulation software. The findings of this study can be concluded as:

- In summer, the thermal sensation score was high, except from the offices with AC, during a certain range of outdoor temperature changes, while occupants from private and shared-private offices had better satisfaction with the same thermal sensation feeling. In winter, most of the occupants experienced a neutral thermal comfort level in offices of different sizes;

- Generally, season and habit were the major driving factors of window-opening behaviour. During the summer or winter runs, the behaviour tended to exhibit a stable change that was not influenced by physical environmental changes. Half of the surveyed office buildings showed the mode of window-opening day and night with no closing during summer, and all-closed in winter for the same occupants.
- Three summer patterns and four winter patterns of window-opening duration were obtained via cluster analysis. Five types of modes were classified and used to generate window-opening behaviour profiles. The types of patterns were not correlated with office size but were more related to habit.
- For the surveyed office, in the summer simulation, the difference between the simulated results via working-time open mode (WO) and the mode inside DesignBuilder was quite small, and thus, the mode profile inside the software could have replaced the WO mode. The AO pattern of window-opening behaviour suggested that the simulation performance gap could be fixed from 2.5% in summer. In the winter simulation, the difference between the real-mode calculated result and the template of DesignBuilder was more significant, from 10% to 13%. Due to the very short window-opening length of the winter WO mode, the correction level for the indoor thermal performance simulation of the two modes is very close.

Clearly, this study has imperfections with the limitation of the number of buildings surveyed and the number of recorded windows in the open-plan offices. Due to limited volunteer participation of the investigated offices, using offices from just one building to reduce the interference of other variables was difficult. This study considered ten offices of different sizes as the research objects and obtained some meaningful results. Extensive research of more building types is still necessary for further discussion.

The behaviour influencing factors shown in this study do not match the previous work by Xin [36], which found that in the summer season the environmental variables lose their predictive power of window-opening probability. This may be due to the distinct seasons with larger temperature differences in north-east China. Season had a great influence on window-opening behaviour in this study. In a single season, the change of window-opening behaviour of most occupants was very inactive. These inactive occupants did not change the state of windows with temperature. Some of the results from this study agree with the work of Song [35] about the influencing variables, while the findings of this study showed that season and habit are the major affecting parameters. The results of this study prove again the necessity of research on occupant behaviour to help revise and refine simulation results from building performance software.

5. Acknowledgements

The authors would like to acknowledge that this paper was financially supported by the China National Key R&D Program (Grant No. 2016YFC0700200), the National Natural Science Foundation of China (Grant No. 51708149), and the China Postdoctoral Science Foundation (Grant No. 2017M621276).

6. References

[1] G Augenbroe. Trends in building simulation. Build Environ. 37 (2002) 891-902.

[2] T Hong, SK Chou, TY Bong. Building simulation: an overview of developments and information sources. Build Environ. 35 (2000) 347-361.

[3] DB Crawley. Contrasting the capabilities of building energy performance simulation programs. Build Environ. 43 (2008) 661-673.

[4] JA Clarke. Energy Simulation in Building Design. Butterworth Heinemann. 2001.

[5] JLM Hensen and R Lamberts (eds.) Building Performance Simulation for Design and Operation. Routledge. 2011.

[6] S-H Lin, DJ Gerber. Evolutionary energy performance feedback for design: Multidisciplinary design optimization and performance boundaries for design decision support. Energ Build, 86 (2014) 416-441.

[7] C Sun, Y Han. Construction of Office Building Environment Information System in Severe Cold Area. New Architecture, 10 (2015) 4-9.

[8] YE Kalay. Performance-based design. Automation in Construction. 4 (1999) 395-409.

[9] Y Han, H Yu, C Sun. Simulation-Based Multiobjective Optimization of Timber-Glass Residential Buildings in Severe Cold Regions. Sustainability. 9(2017)2353-2371.

[10] C Sun, Y Han. Study on the theory of building digital energy saving design under the guidance of green performance. Architectural Journal, 11 (2016) 89-93.

[11] E Mantesi, C J Hopfe, MJ Cook, J Glass, P Strachan. The modelling gap: quantifying the discrepancy in the representation of thermal mass in building simulation. Build Environ. 131 (2018) 74-98.

[12] P Wilde. The gap between predicted and measured energy performance of buildings: a framework for investigation. Automation in Construction. 41 (2014) 40-69.

[13] CA Menezes, A Cripps, D. Bouchlaghem, R. Buswell. Predicted vs. actual energy performance of non-domestic buildings：using post-occupancy evaluation data to reduce the performance gap. Applied Energy, 97 (2012) 355-364.

[14] Da Yan et al. IEA EBC Annex 66: Definition and simulation of occupant behavior in buildings. Energ Build. 156 (2017) 258-270.

[15] M Bonte. Impact of occupant’s actions on energy building performance and thermal sensation. Energ Build. 76 (2014) 219.227.

[16] K-U Ahn, C-S Park. Correlation between occupants and energy consumption. Energ Build, 116 (2016) 420-430.

[17] D&R International Ltd. Buildings Energy Data Book. 2011.

[18] The Energy Data and Modeling Center，Handbook of Energy & Economic Statistics in Japan. 2011.

[19] Z Li, Y Jiang, WQ Fan. Environmental parameters and air conditioning behavior of residential air conditioning energy consumption impact analysis. Air Conditioning, 37 (2007) 67-71.

[20] Japan Institute of Architecture. Japan Residential Energy Consumption Online Database. http://tkkankyo.eng.niigata-.ac.jp/HP/HP/database/japan2/index.htm

[21] P Hoes，J Hensen，MGLC Loomans，B de Vries，D Bourgeois. User Behavior in Whole Building Simulation. Energ Build. 41 (2009) 295-302.

[22] M Jia, RS Srinivasan, AA Raheemb. From occupancy to occupant behavior: an analytical survey of data acquisition technologies, modeling methodologies and simulation coupling mechanisms for building energy efficiency. Renew Sust Energ Rev. 68 (2017) 525-540.

[23] R Li, G Dane, C Finck, W Zeiler. Are building users prepared for energy flexible buildings? - a large-scale survey in the Netherlands. Appl Energ. 203 (2017) 623-634.

[24] SP Corgnati, E Fabrizio, D. Raimondo, M Filippi. Categories of indoor environmental quality and building energy demand for heating and cooling. Build Simul, 4 (2011) 97-105.

[25] HB Gunay, W O'Brien, I Beausoleil-Morrison. A critical review of observation studies, modeling, and simulation of adaptive occupant behaviors in offices. Build Environ, 70 (2013) 31-47.

[26] CJ Andrews, H Chandra Putra, C Brennan. 2013. Simulation modeling of occupant behavior in commercial buildings. Prepared by the center for green building at rutgers university for the energy efficient buildings hub, Philadelphia, PA.

[27] A Mahdavi, C Pröglhöf. User behavior and energy performance in buildings. Wien, Austria: Internationalen Energiewirtschaftstagung an der TU Wien (IEWT), (2009) 1-13.

[28] J Zhao, B Lasternas, KP Lam, R Yun, V Loftness. Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining. Energy Build 82 (2014) 341–55.

[29] B Dong, B Andrews. Sensor-based occupancy behavioral pattern recognition for energy and comfort management in intelligent building. Proc Build Simul. (2009) 1444–51.

[30] G Zimmerman. Modeling and simulation of individual user behavior for building performance predictions. In: Proc. 2007 Summer Computer Simulation Conference, Society for Computer Simulation International, (2007) 913–920.

[31] F Haldi, D Robinson. On the behaviour and adaptation of office occupant. Build Environ 43 (2008) 2163–77.

[32] S D'Oca, T Hong. Occupancy schedules learning process through a data mining framework. Energy Build, 88 (2015) 395–408.

[33] VL Erickson, MÁ Carreira-Perpiñán, AE Cerpa. Observe: occupancy-based system for efficient reduction of HVAC energy. In: 10th International Conference on Information Processing in Sensor Networks (IPSN), IEEE, (2011) 258–269.

[34] Y Jiao, H Yu, T Wang, Y An, Y Yu. Thermal comfort and adaptation of the elderly in free-running environments in Shanghai, China. Build Environ, 118 (2017) 259-272.

[35] S Pan, Y Xiong, Y Han, Z Zhang, L Xia, S Wei, J Wu, M Han. A study on influential factors of occupant window-opening behaviour in an office building in China. Build Environ, 133 (2018) 41-50.

[36] X Zhou, T Liu, X Shi, X Jin. Case study of window operating behaviour patterns in an open-plan office in the summer. Energy Build, 165 (2018) 15-24.

[37] R Zhang, C Sun, S Sharples, HR Zhang. Thermal comfort and behaviour control in mix-mode office buildings in Harbin, China. Proc. Sustainable Design of the Built Environment 2017 Conference, 1(2017)318-330. London.

[38] C Sun, R Zhang, S Sharples, YS Han HR Zhang. A longitudinal study of summertime occupant behaviour and thermal comfort in office buildings in northern China. Building and Environment, 143(2018)404-420.

[39] ANSI/ASHRAE Standard 55-2013: Thermal Environmental Conditions for Human Occupancy. 2013：14. www.ashrae.org

[40] S Shukla, S Naganna. A review on K-means data clustering approach. International Journal of Information & Computation Technology, 4 (2014) 1847-1860

[41] S Kotsiantis, D Kanellopoulos. Association Rules Mining: a recent overview. GESTS International Transactions on Computer Science and Engineering, 32 (2006) 71-82

[42] Evaluation standard for indoor thermal environment in civil buildings. Ministry of Housing and Urban-Rural Development of the People's Republic of China & State Administration for Market Regulation. 2012.

Highlights:

· Long-term occupant thermal comfort and behaviour characteristics in different-sized offices in the severe cold region of China;

· Influencing factors of adaptive behaviour for both summer and winter;

· Defining the window-opening duration patterns;

. Defining the window-opening behaviour classification and profiles of office occupants via data mining techniques;

· Modifying the building thermal performance gap and verifying the window-opening behaviour profiles in selected offices.