

Analysis of Occupants' Air-Conditioning Opening Behaviour Based on the Survival Model

Jiawen Ren¹, Xin Zhou^{1,*}, Xing Shi², Xing Jin¹

¹School of Architecture, Southeast University, Nanjing, China

²Key Laboratory of Ecology and Energy-saving Study of Dense Habitat, Ministry of Education, College of Architecture and Urban Planning, Tongji University, Shanghai, China

Email address:

220190031@seu.edu.cn (Jiawen Ren), zhou-x06@seu.edu.cn (Xin Zhou)

*Corresponding author

To cite this article:

Jiawen Ren, Xin Zhou, Xing Shi, Xing Jin. Analysis of Occupants' Air-Conditioning Opening Behaviour Based on the Survival Model. *Urban and Regional Planning*. Vol. 6, No. 4, 2021, pp. 147-154. doi: 10.11648/j.urp.20210604.16

Received: October 26, 2021; **Accepted:** November 23, 2021; **Published:** November 24, 2021

Abstract: It is essential to predict building energy consumption through more accurate simulation of building energy consumption, and then put forward suggestions for building energy conservation. Therefore, it is a very important issue to study the variable, random and complex air conditioning usage mode. In the previous studies on air conditioning, it could be found that whether the air conditioning is on or off, it is only a mathematical function about environmental parameters. However, when we arrive at the office and feel uncomfortable, we don't open the window immediately. Instead, we put up with it for a while. In view of the above shortcomings, this study proposed a survival model based on Weibull function to predict the air-conditioning on behavior. Through the verification of the model, we found that the accuracy of the air-conditioning regulation model based on the survival model is more than 74%. We compared and verified the common three-parameter Weibull model with the survival model-based Weibull model, and found that the accuracy of the common three-parameter Weibull model was slightly higher than that of the survival model. At the same time, we analyzed the death event (tolerance temperature) of the survival model, and further improving the tolerance temperature is of great help to the accuracy of the model.

Keywords: Air-Conditioning Status, Air-Conditioning Opening Probability, Survival Models, Prediction Accuracy

1. Introduction

In 2018, building energy consumption accounted for about one-fifth of China's total energy consumption. [1, 2] Among them, air conditioning operation energy consumption accounts for more than 40% of the total building energy consumption [3]. Therefore, the energy saving of air conditioning is now a common concern of people. Researchers found that when air conditioning was used, occupants' behavior has a significant impact on air conditioning energy consumption, affecting 22 to 39 percent of a building's annual electricity consumption. Building energy simulation is usually used to study building energy consumption. However, air conditioning plans are usually predetermined with little variation between residents in such simulations, which is not quite accurate [4]. Therefore, better quantitative random usage models are needed to describe

variable, random and complex air conditioning usage patterns in buildings.

At present, many scholars have conducted a lot of research on the air-conditioning regulation behavior of occupants. It is mainly divided into influencing factors and model construction. In the aspect of influencing factors, the factors affecting the action of air conditioning switch are divided into subjective factors and objective factors. Objective factors include environmental factors (temperature, humidity, wind speed) [5-7] and non-environmental factors (users' income, age) [8-10]. Subjective factors are the influence of personal factors (such as life style and preference) on air conditioning operation behavior [11, 12]; In terms of model construction, scholars mainly explore the relationship between environmental factors and air conditioning switch probability, among which Logistic model [13] and Weibull model [14, 15] are the mainstream models at present. With the popularity of machine learning, more and more machine

learning algorithms are applied to this field [16].

In the previous studies on air conditioning, no matter the status of air conditioning is on or off, it is only the mathematical function about certain environmental parameter. But in the real situation, when we arrive at the office and reach the tolerance temperature, the air conditioning isn't opened immediately. Occupants often endure it for a period of time. In view of the above shortcomings, this study proposed a survival model to predict the air-conditioning on occupant behavior. Survival analysis is derived from medical science, which predicts the probability of event occurrence by combining the event endpoint and event occurrence time and comprehensively considering relevant factors. When survival analysis was applied to air conditioning regulation, the starting point of the event was reaching the tolerance temperature, and the

end point was the opening of the air conditioning.

2. Methodology

2.1. Overview

The flow of this work is shown in Figure 1. The relation of Window opening probability and temperature was explored at first. Then, the survival model was established by determining the initial event and parameters in three-parameters Weibull model. After establishing this, the model was validated by the testing method, namely switch curve accuracy and opening probability and the traditional three-parameters Weibull model was compared with the survival methods. Finally, the tolerance temperature of initial event was discussed.

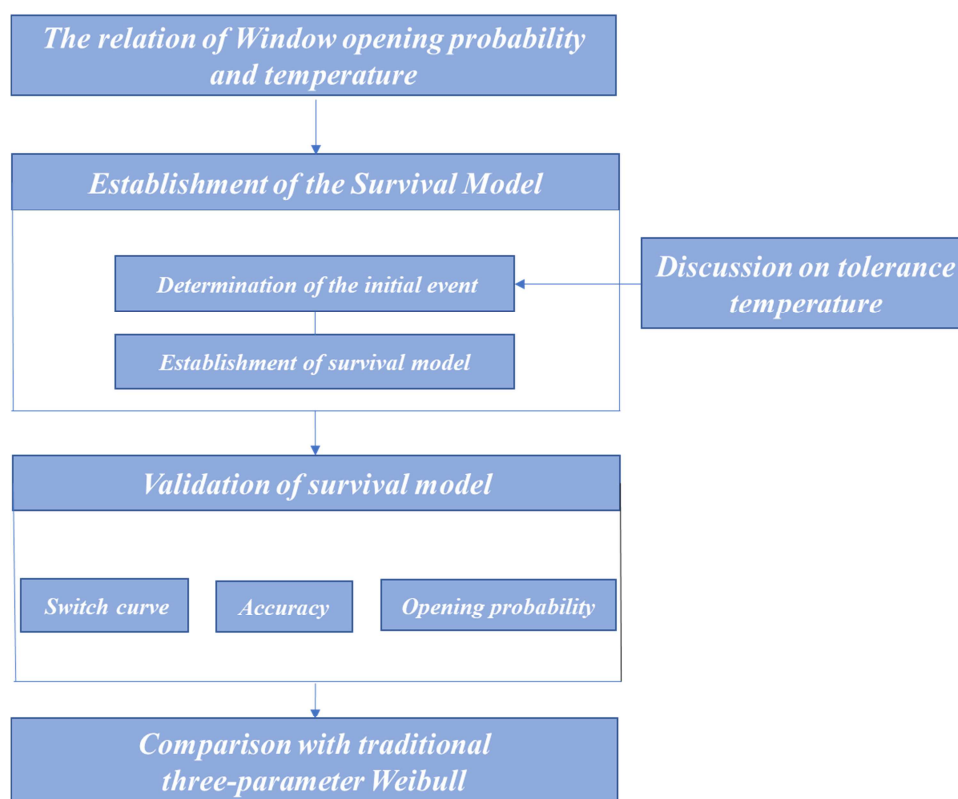


Figure 1. Flow chart of this work.

2.2. Data Source

Field measurements were made in an open-plan office in Nanjing, China. The indoor sensors are placed on the table, and other outdoor sensors are placed outside the window. All test instruments were recorded at an interval of 5 minutes. The test lasted from August to September 2016 for an open-plan Office. Therefore, in office, people use air conditioning for cooling.

2.3. Survival Model

Survival analysis method is a statistical method used to

study and observe the occurrence of specific events and their probability after a certain period of time in the process of existence or survival of individuals or groups. Among them, time often refers to the duration of a certain state in nature, human society and products and services, which is known as "life span". This life span can represent the survival time of organisms and the service time of products and services. In the survival analysis method, we call it "life time". Specific events, also known as Failure events, are survival outcomes set by researchers according to different research purposes. These specific events can be death, tumor metastasis, recurrence, or any clearly identifiable event, and different conditions are different grouping basis. In this study, the

operation of air conditioning was defined as the failure action in the survival model, and the grouping was based on the location of outdoor temperature.

There are many functions describing the rule of survival time, all of which are called survival time function. The general tools are survival function, cumulative distribution function, probability density function and risk function, which respectively illustrate different meanings of data, but are mathematically completely equivalent and can be derived from each other. Assuming that, the survival time is represented by a non-negative random variable T , the following function is used to describe the characteristics of the survival time T .

The survival function is also known as the cumulative survival probability, denoted as $S(t)$, and defined as the probability that the survival time of an individual exceeds the time point t , that is, the possibility that the research object still exists after the time point t . Its function expression is:

$$S(t) = \Pr(T > t) \quad (1)$$

When $t=0$, there is $S(0) = 1$. As time goes by, the survival function decreases monotonically, and there is $\lim_{t \rightarrow \infty} S(t) = 0$.

The curve drawn according to the $S(t)$ value is called the survival curve. A steep survival curve means that the individual or population under study has a shorter survival time or a lower survival rate, while a gentle survival curve means that the individual or population under study has a longer survival time or a higher survival rate.

2.4. The Methods of Survival Analysis

The statistical methods of survival analysis are usually divided into non-parametric method, parametric method and semi-parametric method. The nonparametric method does not make any assumptions about the distribution of survival time data and can analyze completely unknown survival functions. Kaplan-meier Method and life-table Method are commonly used, which are respectively applicable to different types of survival time data. In contrast, parametric method is to study the relationship between survival time and its influencing factors under the premise of making certain assumptions about survival function. The commonly used methods include exponential distribution, Weibull distribution and Logistic regression analysis. The semi-parametric rule has both characteristics. It can analyze the distribution of survival time and many factors of survival time through the model without assuming the distribution of survival time. Cox regression model is the most typical method. It only requires that the data meet the premise of proportional risk, so as to establish a multi-factor model of survival analysis, so as to study the effect of multiple factors on survival time.

2.5. The Weibull Distribution

Weibull distribution is the most widely used statistical distribution model in the field of reliability research, and three-parameter distribution is the model with the strongest ability to adapt to data and the best fitting effect. However,

compared with normal distribution and two-parameter distribution, the structure of the distribution model is more complex. If the random variable follows the three-parameter Weibull distribution and its probability density function is shown in Formula (2), then its distribution function is (3):

$$f(X) = \frac{\beta}{\alpha} \left(\frac{X-\mu}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{X-\mu}{\alpha}\right)^\beta\right] \quad (2)$$

$$F(X) = P(X \leq x) = 1 - \exp\left[-\left(\frac{X-\mu}{\alpha}\right)^\beta\right] \quad (3)$$

If the survival time t satisfies the Weibull distribution, its probability density function is shown in Formula (4), and its distribution function is shown in Formula (5). Where X is the covariable value.

$$S(t, x)_{\text{survival}} = 1 - \exp[-\lambda(X)t^\nu] \quad (4)$$

$$S(t, x)_{\text{survival}} = P_{\text{survival}} = 1 - \exp[-\lambda(X)t^\nu] \quad (5)$$

$$\lambda(X) = \exp(\beta_0 + \beta_1 X + \dots + \beta_n) \quad (6)$$

2.6. The Prediction Model

In the prediction model, it is necessary to find the moment when the trigger condition is generated. If the trigger condition does not exist, the air conditioner remains in the state of the previous moment. This model is also composed of multiple discrete judgment processes based on the time step. Figure 2 shows the calculation logic within a time step. Whether the personnel are in the room or the outdoor temperature is taken as the input parameter for each step. It is important to note that, at the beginning of the simulation, it is necessary to determine whether there are two conditions: people arriving or people leaving. If no trigger condition exists, the air conditioner status is the same as the last time. If the occupants arrive, it will enter the random process based on the outdoor temperature and the $t=0$. The random number generated will be compared with the trigger probability under the trigger condition, and then determine whether the air conditioning state has changed. If the random number is smaller than or equal to the probability value of the trigger condition, the air conditioner is triggered. If the random number is larger than the probability value of the corresponding triggering condition, then the random process determines that the air conditioner is not turned on, which can be divided into two situations: The first is that the outdoor temperature when the personnel arrive indoors may not reach the critical temperature to trigger the opening of the air conditioner, so no change is made at this time, and the next time step continues to judge whether the outdoor temperature at that time triggers the opening of the air conditioner until the air conditioner is turned on or the personnel leave the room. In the second case, the tolerance temperature is reached, but with the increase of outdoor temperature, the air conditioning may be turned on after the occupants arrive indoors for some time, at which time, $t=t+1$. If the air conditioner is on when the occupant leaves, the air conditioner turns off when the occupant leaves. The final dynamic model outputs hourly air conditioning status data.

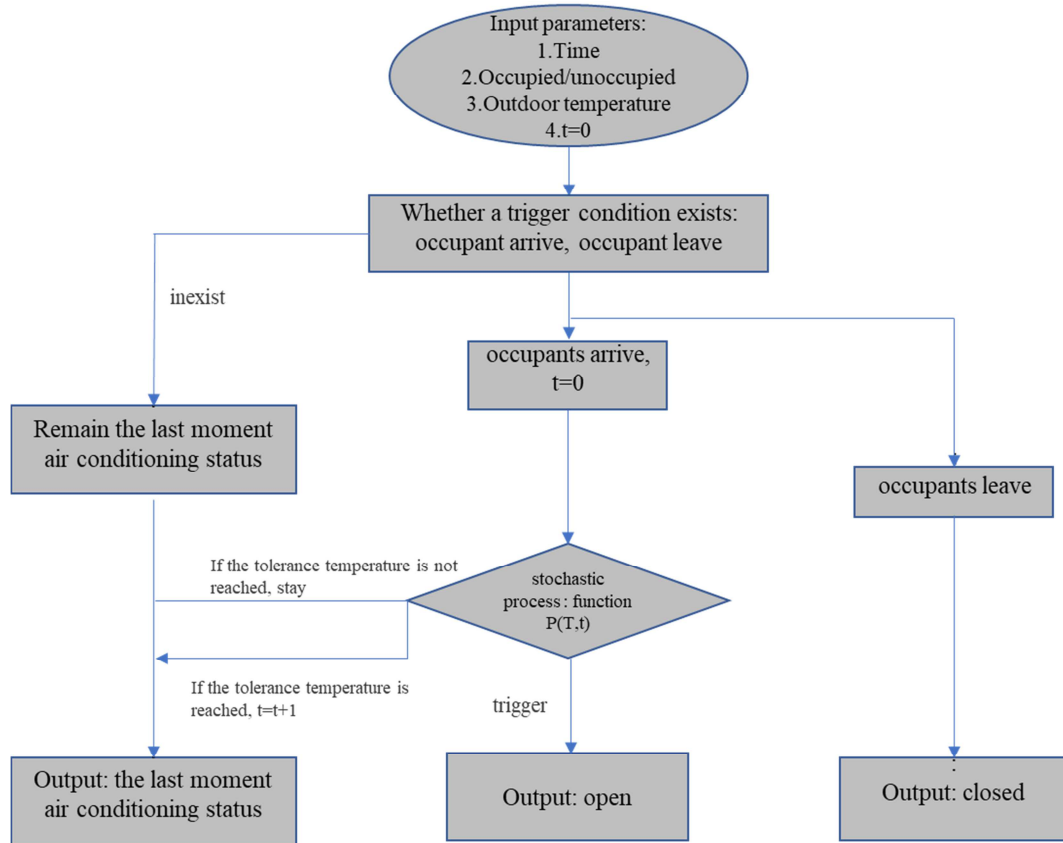


Figure 2. Model logic diagram.

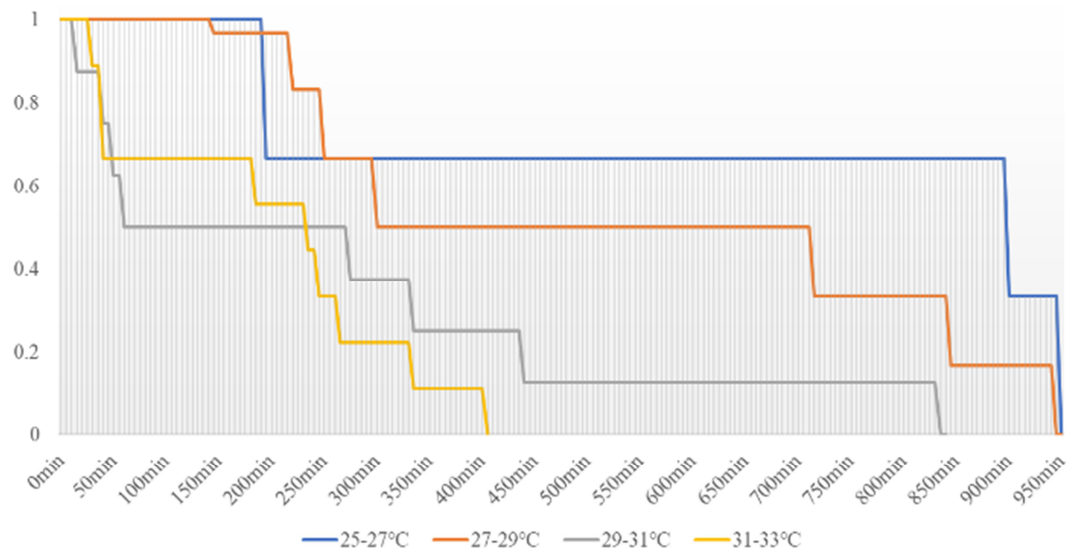


Figure 3. The probability of not turning air conditioning and survival time.

3. Results

3.1. The Relation of Window Opening Probability and Temperature

For traditional air conditioning, the probability is considered in the analysis of single or multiple environmental factors influence on air conditioning open action. In addition,

in order to determine whether air conditioning open actions related to the time of arrival, we adopt the method of survival analysis. We determine the getting to the office as the event starting, opening air conditioning action as the end of the event. The survival curve was used to study the effect of a single environmental variable on the endurance of air conditioning. This curve is a ladder descending curve drawn by the proportion of air conditioning turned on and the

appearance of endurance. The flatter the curve, the lower the probability of the air conditioner turning on, and the longer the time of not turning on the air conditioner. Conversely, the steeper the curve, the higher the probability of air conditioning, the shorter the time of not turning on air conditioning.

Figure 3 shows the ladder curve of survival ratio and survival time obtained by kaplan-Meier method. It can be seen from the curve that the higher the temperature, the shorter the time of not turning air conditioning, which is also consistent with our previous thought. As can be seen in Figure 3 when the outdoor temperature is 25-27°C, the probability of not opening air conditioning stay in the 1.0 for more than 3h, and 70% of the samples even exceed 15h. When the outdoor temperature ranges from 31°C to 33°C, the probability that the air conditioner is not turned on stays in the 1.0 for only 30 minutes. It can be seen that there is a strong relationship between the time when people turn on air conditioning after entering the office and the temperature. In addition to time, we add the influence of temperature, an independent variable on the survival probability. It is necessary to include temperature in the survival model to describe the air conditioning switching behavior.

3.2. Establishment of the Survival Model

As shown in 2.4, parametric methods usually assume that the survival time data follows a specific distribution, and then fit the parameters related to survival time according to the

known distribution. It is assumed that the distribution of survival time data follows distribution of Weibull. In addition to time, we add the influence of temperature, an independent variable or covariable, on survival probability, and use the ordinary least squares method to fit the functional relationship among survival time, average temperature within survival time and survival probability. Finally, the function is substituted into our action model to calculate the accuracy of the survival model.

3.2.1. Determination of the Initial Event

Generally speaking, in summer, when the environment is larger than a certain temperature, people will start to feel hot, and then they will turn on the air conditioning, which we call the tolerance temperature here. In the establishment of survival model, we will take the tolerance temperature as the start event, and the action of turning on the air conditioning as the failure event. Due to the limitations of this measurement data, we can't simply determine the temperature tolerance of occupants in the office by using the comfort level of indoor thermal environment. Here we begin by exploring the starting event used as a survival model to determine what temperature is appropriate for office workers to tolerate indoors. Figure 4 shows the functional relationship between outdoor temperature and air conditioning. It can be seen from the figure that when the outdoor temperature is less than 25°C, air conditioning is basically not turned on in the office, so the tolerance temperature T_0 is set as 25°C here.

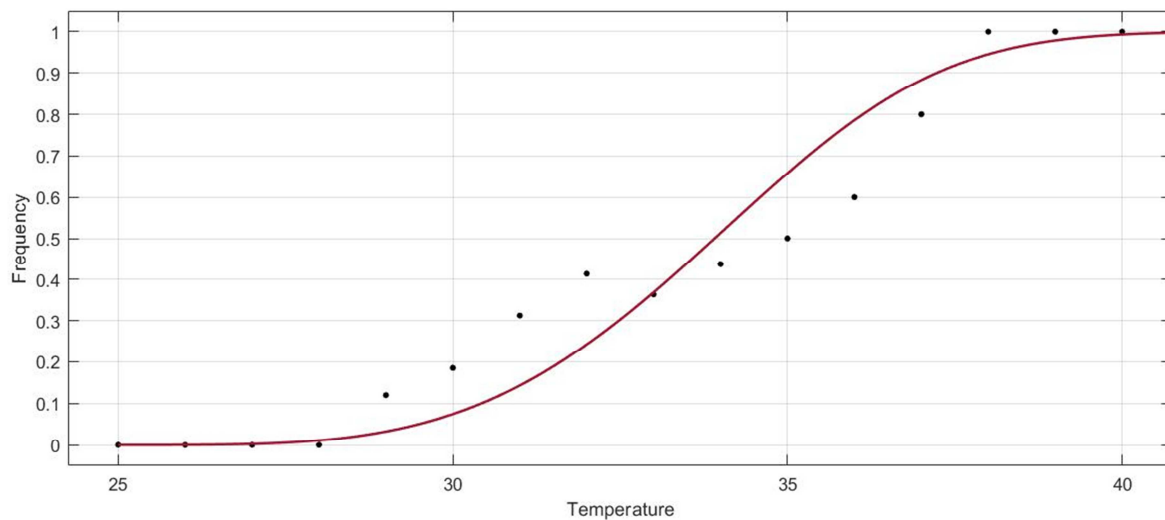


Figure 4. Outdoor temperature fitting curve of air conditioning operating behavior in summer.

3.2.2. Establishment of Survival Model

After the tolerance temperature was set to 25°C, we sorted out the data of the office in summer, and obtained the three elements required by the survival model, namely, the time of not turning on the air conditioner after the tolerance temperature exceeded (survival time), the average temperature during this period (covariate value), and the corresponding probability of not turning on the air conditioner (survival probability). We use the ordinary least squares method to fit

Weibull function in MATLAB as shown in formula (7), and the obtained $R^2=0.5885$. It can be concluded that the relationship between survival time and temperature of other influencing factors can be expressed by Weibull function. The probability P , outdoor temperature T and the time T exceeding the tolerance temperature at the same time are shown in formula (8). We put the function P into the program in Figure 1 as $P(T,t)$ to judge the state of air conditioning.

$$S(t)_{\text{survival}} = \exp(-\exp(0.1521 * T - 11.61) t^{1.242}) \quad (7)$$

$$P_{\text{survival}} = 1 - \exp(-\exp(0.1521 * T - 11.61) t^{1.242}) \quad (8)$$

3.3. Validation of Survival Model

We defined the opening status of air conditioning as 1 and the closed status as 0. The predicted results and actual results are shown in *Table 1*. Among them, there are 4983 data points in the off state of actual results, and 1350 data points in the on state predicted by the survival model, with an accuracy rate of 73%. The actual results showed that there were 2492 data points in the open state, among which 592 data points in the closed state were predicted by the survival model, with an accuracy rate of 77%. In conclusion, the overall accuracy of survival model obtained by Weibull function parameter method reached 74%.

Table 1. Confusion matrix of survival model.

Measured result	Simulated result	
	0	1
0	3633	1350
1	592	1900
Total accuracy	0.74	

When we further screen out the data of the occupancy in the room and analyze the switching curve of the personnel in the room air conditioner, the switching curve of the five predicted results and the actual results is shown in *Figure 5*, we found that in the predicted situation, the opening state is indeed more than the opening state of the actual situation. This is mainly due to the following two reasons: The first reason is that during the seven days of August 26th-August 28th, September 4th, and September 6th, 9th, 11th, the air conditioning was not turned on for a whole day. However, by investigating the average temperature of these seven days, it can be found that the average temperature between August 26

and August 28th and September 4th was as high as 30 degrees, and the air conditioner was not turned on. This was mainly due to the different habits of switching on air conditioning caused by different people in the room, or other objective factors, which were not included in our analysis. The second reason is that we find that air conditioning is predicted to turn on earlier than it actually is every day. This indicates that the actual condition of the air conditioning tolerance temperature threshold is larger.

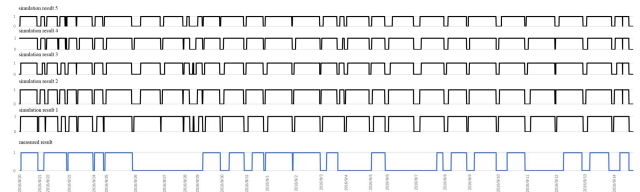


Figure 5. Switch curve (measured result and simulated results).

The survival model is mainly used to solve the specific time of turning on air conditioning when people are in the room. Therefore, we compared the actual time of people staying in the room before turning on air conditioning (survival time) with the simulated time of people in the room after 100 times before turning on air conditioning. As shown in *Figure 6*, the box diagram represents a discrete case of 100 simulation results. We found that the simulated results were basically consistent with the actual results within 0-60 min. However, in the simulation results, the time for turning on the air conditioner is mainly concentrated in 60-150 min, and there are few conditions that the air conditioner is not turned on for a long time. However, in actual situations, the time of not opening air conditioning mainly exists in 0-60 minutes and 240-300 minutes.

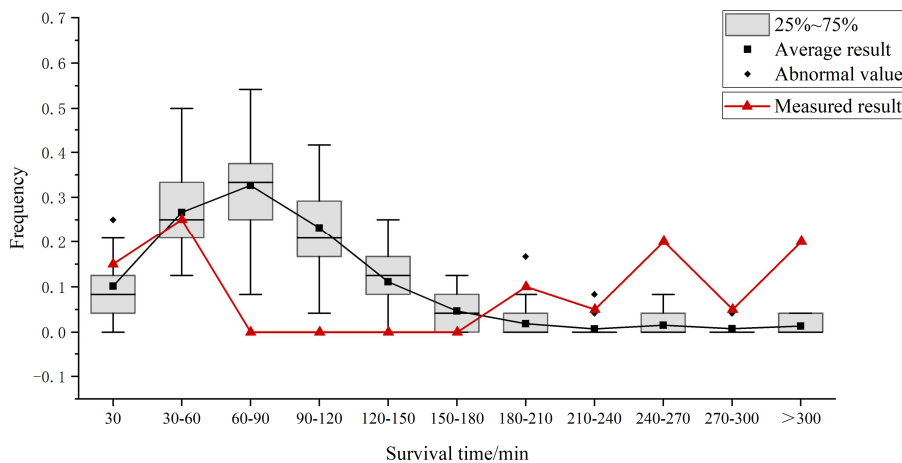


Figure 6. The simulated probability of not turning air conditioning and survival time.

3.4. Model Comparison Between Survival Model and Traditional Three-parameter Weibull Model

It is necessary to compare Weibull model with other existing models to discuss its practicability. The traditional three-parameter Weibull distribution is used to describe the

influence of outdoor temperature in previous window-opening behavior studies. The fitting result of $R^2=0.952$. Therefore, this study compared the survival model based on Weibull with the traditional three-parameter Weibull model. The mathematical expression of traditional three-parameter Weibull model is shown in Equation (9).

$$P = \begin{cases} 1 - e^{-5\left(\frac{T-25}{15}\right)^{3.795}} & (T > 25) \\ 0 & (T \leq 25) \end{cases} \quad (9)$$

Substitute the above P into the model of air conditioner switch judgment in 2.4 to judge the air conditioner switch status. In the following analysis, the status of the air conditioner is 1 and 0 respectively. Table 2 shows the accuracy of survival model Weibull model and common Weibull model in predicting air conditioning on or off behavior respectively. The overall accuracy in predicting window opening and closing behavior was 74 %and 76 %. In general, both survival model Weibull model and ordinary Weibull model can predict window opening and closing behavior. However, the overall prediction accuracy of general Weibull model is slightly lower than that of survival model.

Table 2. Confusion matrix of survival model and weibull.

Measured result	Simulated result			
model	0	1	0	1
model	survival		weibull	
0	3633	1350	3702	1281
1	592	1900	468	2024
Total accuracy	0.74		0.76	

4. Discussion

In the survival model, the definition and final outcome of the death event can be greatly influenced. For an office, the temperature of tolerance is a topic worth discussing. We built a model for the tolerance temperature of office staff at 25-35 degrees Celsius each time, and then we used the above method to analyze and verify the accuracy of each model. We found that due to the previous discussion, the threshold of office personnel turning on air conditioning is too large. As the temperature increases, the model accuracy increases gradually. We found that the model has the accuracy of 88% at outdoor temperatures up to 34 °C. In our subsequent research, we will increase the tolerance temperature of the model. As for the specific degree, we may need to discuss further.

Table 3. Model accuracy at different tolerance temperatures.

Number of model	Tolerance temperature	Accuracy
1	25	0.74
2	26	0.76
3	27	0.78
4	28	0.80
5	29	0.81
6	30	0.80
7	31	0.83
8	32	0.83
9	33	0.85
10	34	0.88
11	35	0.85

5. Conclusion

In previous studies, no matter the air conditioning on or off action, it only studies the mathematical function relationship

between it and one or more environmental parameters. When we arrive at the office and feel uncomfortable, we don't open the window immediately. Instead, we put up with it for a while. In view of the above shortcomings, this study proposed an event - and temperature-based survival model to predict air-conditioning opening behavior. The specific conclusions of this paper are as follows:

1. This paper established a survival model based on two parameters of outdoor temperature and tolerance time, and took Weibull model as the main function to predict the air-conditioning regulation behavior of open offices in summer. Through the verification and analysis of the model, we found that the accuracy of the survival model to predict the condition of the air conditioner was up to 74%, and the prediction behavior of the air conditioner was very accurate within 0-60 min.
2. Although the accuracy of the common Weibull model is slightly higher than that of the common Weibull model by 2%, through the discussion of the tolerance temperature of the survival model, we find that the Weibull accuracy based on the survival model can be as high as 88% with the increase of the tolerance temperature. This shows that the survival model is more applicable.

Acknowledgements

This work is supported by National Natural Science Foundation of China (52078117), China Railway Corporation Technology Research and Development Project (P2018G049).

References

- [1] Y. Zhang, Da. Yan, S. Hu, S. Guo, Modelling of energy consumption and carbon emission from the building construction sector in China, a process-based LCA approach, Energy Policy 134 (2019) 110949, <https://doi.org/10.1016/j.enpol.2019.110949>.
- [2] S. Hu, Da. Yan, S. Guo, Y. Cui, B. Dong, A survey on energy consumption and energy usage behavior of households and residential building in urban China, Energy Build. 148 (2017) 366–378.
- [3] H. Sha, P. Xu, C. Hu, et al., A simplified HVAC energy prediction method based on degree-day, Sustain. Cities Soc. 51 (2019) 101698. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp. 68–73.
- [4] X Ren, Da. Yan, C Wang, Air-conditioning usage conditional probability model for residential buildings, Building and Environment, 81 (2014) 172-182.
- [5] Y. Peng, Z. Nagy, A. Schluter, Temperature-preference learning with neural networks for occupant-centric building indoor climate controls, Build. Environ. 154 (MAY) (2019) 296–308.
- [6] J. Tanimoto, A. Hagishima, State transition probability for the Markov Model dealing with on/off cooling schedule in dwellings, Energy Build. 37 (3) (2005) 181–187.

- [7] S. H. Mun, Y. Kwak, J. H. Huh, A case-centered behavior analysis and operation prediction of ac use in residential buildings, *Energy Build.* 188 (APR.) (2019) 137–148.
- [8] G. Mark, Towards a Residential Air-Conditioner Usage Model for Australia, *Energies* 10 (9) (2017) 1256.
- [9] J. Wang, J. Zhu, Z. Ding, et al., Typical energy-related behaviors and gender difference for cooling energy consumption, *J. Cleaner Prod.* 238 (2019) 117846.
- [10] J. Tanimoto, A. Hagishima, State transition stochastic model for predicting off to on cooling schedule in dwellings as implemented using a multilayered artificial neural network, *J. Build. Perform. Simul.* 5 (1) (2012) 45–53.
- [11] M. Schweiker, M. Shukuya, Comparison of theoretical and statistical models of air-conditioning-unit usage behaviour in a residential setting under Japanese climatic conditions, *Build. Environ.* 44 (10) (2009) 2137–2149.
- [12] J. Ryu, J. Kim, W. Hong, et al., Quantifying householder tolerance of thermal discomfort before turning on air-conditioner, *Energy Build.* 211 (2020) 109797.
- [13] J. Kim, R. De Dear, T. Parkinson, et al., Understanding patterns of adaptive comfort behaviour in the Sydney mixed-mode residential context, *Energy Build.* 141 (2017) 274–283.
- [14] T. Hong, H. Sun, Y. Chen, et al., An Occupant Behavior Modeling Tool for CoSimulation, *Energy Build.* 117 (2015) 272–281.
- [15] X. Zhou, D. Yan, X. Feng, et al., Influence of household air-conditioning use modes on the energy performance of residential district cooling systems, *Build. Simul.* 9 (4) (2016) 429–441.
- [16] Y. Peng, A. Rysanek, Z. Nagy, A. Schlüter, Using machine learning techniques for occupancy-prediction-based cooling control in office buildings, *Appl. Energy* 211 (2018) 1343–1358.