



Air infiltration rate distribution across Chinese five climate zones: A modelling study for rural residences



Ye Wang^a, Shanshan Shi^b, Zhengxu Zhou^c, Song Guo^c, Bin Zhao^{a,*}

^a Department of Building Science, School of Architecture, Tsinghua University, Beijing, 100084, PR China

^b School of Architecture and Urban Planning, Nanjing University, Nanjing, Jiangsu Province, 210093, PR China

^c Department of Urban Planning, School of Architecture, Tsinghua University, Beijing, 100084, PR China

ARTICLE INFO

Keywords:

Air infiltration rate
Rural
Indoor air
Simulation
CONTAM

ABSTRACT

Air infiltration rate is one of the fundamental parameters in the design and analysis of building environment, which is the fundamental mode of air exchange between indoor and outdoor in non-mechanically ventilated residences when doors and windows are closed. Most Chinese residential buildings have no mechanical ventilation equipment installed. Previous researches on Chinese housing have primarily focused on urban residences. However, the number of both urban households and rural households accounts for almost half of Chinese households. Considering the significant urban-rural differences in Chinese housing, there is an urgent need for researching the air infiltration rate in rural area. This study employed a multi-zone network airflow model approach to simulate the air infiltration rate of 111 rural residential samples in 13 Chinese provinces, obtaining the hour-by-hour results for a whole typical weather year of each sample. Thus, the distributions of air infiltration rate for each rural region were finally summarized. Additionally, we performed field measurements in 23 samples to validate the simulation results. According to the results, the mean value of air infiltration rate is larger than that of urban residences, and its distribution in each climatic zone generally follows a log-normal distribution. The results are expected to be applied for indoor air quality studies, building energy consumption analysis, and other frontiers related to architectural environmental design and analysis.

1. Introduction

During the course of a human lifespan, the majority of time is spent indoors [1,2], underscoring the paramount importance of constructing a conducive indoor environment. In 2019, Household Air Pollution (HAP) ranked 10th in the leading risk ranking for attributable disability-adjusted life-years (DALYs) [3]. Furthermore, in many countries in Africa and Asia, the burden percentage attributed to air pollution (including environmental particulate matter, household air pollution, and environmental ozone pollution) ranged from 10% to 15%, with HAP being the dominant risk in the African region [3]. Air change rate (ACR) is a crucial indicator in significantly influencing the indoor air pollution level due to equivalent dilution (when outdoor concentration less than indoor) or deterioration (when outdoor concentration worse than indoor) effect. It is defined as the volume of air entering or leaving the space per unit of time divided by volume of the enclosed space. Higher ACR leads to an increased import of outdoor-originated air pollution into the indoor environment. Conversely, if there are indoor pollution

sources such as building materials or cooking oil fumes, a lower ACR will hinder the removal of pollutants, resulting in an elevated indoor air pollution exposure of residents [4–8]. The variation in ACR among different regions can partly explain the disparities in health risks associated with ozone and particulate matter between these areas [9,10]. Furthermore, it affects the energy consumption and thermal environment of the building [11,12].

In the absence of mechanical ventilation, the modes of indoor-outdoor air exchange can be primarily categorized into two forms: natural ventilation through openings doors, windows, grilles and all other intentional envelope penetrations; and uncontrolled air leakage through cracks or the normal use of exterior doors, commonly referred to as air infiltration [11]. The natural ventilation rate typically ranges from 1 to 10 h^{-1} [13], constituting the primary method adopted by Chinese residents to meet their fresh air requirements. Under this condition, indoor pollutant concentrations tend to approach those of the outdoor environment. However, as the air infiltration rate varies (typically in the range of 0.1–1 h^{-1}), it should be emphasized that the air

* Corresponding author.

E-mail address: binzhao@tsinghua.edu.cn (B. Zhao).

infiltration rate in this study refers to the normalized air infiltration rate after the air volume flow rate divided by the building volume, of which the unit is h^{-1}), the indoor/outdoor ratio (I/O ratio) of pollutant concentration, such as fine particulate matter with aerodynamic diameter less than $2.5 \mu\text{m}$ ($\text{PM}_{2.5}$), fluctuates between 40% and 80% [14]. Notably, air infiltration is prevalent during heating and cooling seasons and serves as the primary air exchange mode for bedrooms in Chinese households during nighttime sleep [15]. Therefore, the air infiltration rate of Chinese residences will be the focus of this study.

In past decades, researchers have conducted measurements of air infiltration rate in multiple countries, including the United States, the United Kingdom, Spain, and others [16–27]. Moreover, previous studies have also covered numerous cities in China, such as Beijing, Tianjin, Guangzhou, and other central cities [15,28–34]. Considering that conducting actual measurements requires significant human and financial resources, some studies attempted to construct estimation models for air infiltration rate. Chan et al. organized and analyzed large amount of measured data from the U.S. air leakage database and obtained a fitting formula applicable to U.S. residences using regression methods [35–37]. Subsequently, many studies have employed this empirical formula. Breen et al. verified and further refined this formula through actual measurements [38]. Similarly, Montoya et al. derived a leakage area empirical formula applicable to the Catalonia region using air leakage data from European countries [39]. Unfortunately, only a small number of studies were based on probability samples due to the difficulty of collecting data. For instance, Persily et al. [14] conducted probabilistic sampling for U.S. residential homes and calculated air infiltration rate using CONTAM [40], developed by The U.S. National Institute of Standards and Technology. In another study, Shi et al. performed probabilistic sampling for 9024 apartment units across 180 buildings in Beijing and simulated the air infiltration rate of the sampled apartments [34]. The simulated air infiltration rates ranged from 0.02 to 0.82 h^{-1} , with a mean of 0.16 h^{-1} , and the simulated values aligned well with the measured values. The results of these studies are important input parameters used in the analysis of the relationship between indoor energy consumption, pollutant levels and air exchange: such as the Incremental Ventilation Energy (IVE) model established by Logue et al. [41]; the probability distribution functions of stock infiltration rates and total heat loss in the UK provided by Jones et al. [42]; the Chilean Housing Archetypes AiR quality Model (CHAARM) and stochastic framework presented by Molina et al. [43,44].

It should be noted that previous research on air infiltration rate in Chinese residential buildings mainly focused on urban high-rise buildings. However, there are significant differences in building types and construction quality between urban and rural residences. Urban residences are mostly multi-story buildings, while rural residences are primarily low-rise detached buildings. Reinforced concrete structures account for over 70% in urban areas but less than 30% in rural areas, where mixed and brick-wood structures are more common [45]. Additionally, rural housing represents 47% of China's total residential buildings [45], making it inadequate to rely solely on the results of urban residences to characterize the nationwide distribution of air infiltration rate. Moreover, previous studies often concentrated on specific regions, but China's climates and building characteristics vary across different areas. Therefore, using results from a single region to generalize findings for the entire country would be unreasonable.

This study gathered rural building information from 13 provinces encompassing all five major building thermal zones in China (classified by the Ministry of Housing and Urban-Rural Development of China) [46], taking into account building characteristics and nationwide meteorological features. The multi-zone network airflow model software CONTAM 3.4 [40] was employed for simulation to determine the air infiltration rate distribution in rural residential buildings in China. Furthermore, 23 rural residential samples were measured using CO_2 decay method to verify the simulation reliability.

2. Methodology

2.1. Sample selection

Air infiltration rate is decided by stack-ventilation and wind pressure ventilation, affected by factors including meteorological factors and building's own characteristics [38,47]. Indoor-outdoor temperature difference and building height jointly affect stack-ventilation, while outdoor wind environment and building orientation influence wind pressure ventilation. Furthermore, the leakage area and the building's spatial configuration respectively influence the airflow entering the interior and the airflow pathways inside the building. In this study, we calculated the leakage area using empirical formulas with the construction year and floor area as independent variables [35]. A larger leakage area suggests a less dense building envelope, leading to higher air infiltration rate. It is worth noting that the leakage area includes the leakage from doors, windows, and walls. As a result, factors which may influence air infiltration rate are as follows: meteorological condition, construction year, floor area, number of floors, story height, building orientation and building spatial configuration. We did not perform probability sampling for all rural households in China, considering the difficulty in obtaining information of rural residences and the lack of comprehensive survey data available online. 3705 questionnaires from 11 provinces (Jiangsu, Zhejiang, Anhui, Guangdong, Guangxi, Fujian, Guizhou, Yunnan, Inner Mongolia, Hebei, Shaanxi) in 5 climatic zones (Severely Cold Zone (SC), Cold Zone (C), Moderate Zone (M), Hot Summer Cold Winter Zone (HSCW), and Hot Summer Warm Winter Zone (HSWW)) were collected via the on-site investigation of Department of Urban Planning at the School of Architecture at Tsinghua University and the Centre of Rural Revitalization Praxis at the School of Architecture and Urban Planning at Nanjing University (RRPCS, NJU-SAUP) instead. Considering there exist similar meteorological conditions within the same climatic zone and similar rural building characteristics within the same province, the data were comparable. The basic information, including construction year, floor area, orientation, number of floors, roof type, and presence of chimneys, was the focus of the questionnaire collection. Since the total housing stock is huge, we analyzed the representative of survey sample size following the method described by Ref. [48], as detailed in “**Representative analysis of survey samples**” in Supporting Information. The margin of error represents the error of population mean estimated based on existing sample size at acceptable confidence level (95% in this study). And we considered the error presented in Table 1 as acceptable. Among survey samples, 111 rural houses had their floor plans surveyed for simulation analysis. Table 2 presents the proportion of various types of houses in the survey and mapping samples, demonstrating the good representativeness of the mapping sample. Details are available in the “**Information of samples**” section of Supporting Information.

2.2. Simulation

The multi-zone network modeling software CONTAM 3.4 was uti-

Table 1
Representative of survey samples.

Region	Rural housing stock [45]	Number of survey samples	Margin of error (CI 95%)
SC	18943740	532	3%
C	61695547	591	2%
M	10497968	526	3%
HSCW	71754062	1337	2%
HSWW	21760668	719	2%

SC represents severely cold zone, C represents cold zone, M represents moderate zone, HSCW represents how summer and cold winter zone, and HSWW represents hot summer and warm winter zone.

CI represents the confidence interval.

Table 2

Proportion of different types of residences in research.

Characteristic	Simulated samples, number(proportion%)						Survey samples, number(proportion%)					
	SC	C	M	HSCW	HSWW	All	SC	C	M	HSCW	HSWW	All
All	22(100)	10(100)	14(100)	45(100)	20(100)	111(100)	532(100)	591(100)	526(100)	1337(100)	719(100)	3705(100)
Year built												
≥2011	2(9)	6(60)	8(57)	12(26)	8(40)	36(32)	42(7)	342(57)	234(44)	312(23)	209(29)	1139(30)
2001–2010	12(54)	4(40)	1(7)	11(24)	0(0)	28(25)	287(53)	189(31)	201(38)	357(26)	182(25)	1216(32)
1991–2000	7(31)	0(0)	2(14)	6(13)	8(40)	23(20)	156(29)	46(7)	67(12)	297(22)	163(22)	729(19)
1981–1990	1(4)	0(0)	3(21)	11(24)	0(0)	15(13)	28(5)	12(2)	24(4)	279(20)	123(17)	466(12)
≤1980	0(0)	0(0)	0(0)	5(11)	4(20)	9(8)	19(3)	2(0)	0(0)	92(6)	42(5)	155(4)
Floor area(m ²)												
≤72.2	4(18)	0(0)	5(35)	17(37)	11(55)	37(33)	95(17)	12(2)	174(33)	433(32)	373(51)	1087(29)
72.2–110.8	10(45)	1(10)	5(35)	16(35)	4(20)	36(32)	312(58)	130(21)	159(30)	447(33)	188(26)	1236(33)
≥110.8	8(36)	9(90)	4(28)	12(26)	5(25)	38(34)	125(23)	449(75)	193(36)	457(34)	158(21)	1382(37)
Number of floors												
1	22(100)	9(90)	4(28)	9(20)	2(10)	46(41)	468(87)	502(84)	152(28)	179(13)	54(7)	1355(36)
2	0(0)	1(10)	5(35)	26(57)	7(35)	39(35)	49(9)	67(11)	203(38)	653(48)	248(34)	1220(32)
3	0(0)	0(0)	5(35)	9(20)	8(40)	22(19)	13(2)	13(2)	147(27)	435(32)	291(40)	899(24)
≥4	0(0)	0(0)	0(0)	1(2)	3(15)	4(3)	2(0)	9(1)	24(4)	70(5)	126(17)	231(6)

lized to simulate residential buildings, and we finally established 111 models of which input parameters were specific values according to our survey information. The simulation process is as follows.

Step1: Setting floor plans for each level based on the actual conditions and characteristics of the building in the sketchpad of CONTAM 3.4 software.

Step2: Defining the airflow paths of the building envelope. Since Chinese rural residences typically lack mechanical ventilation systems, this study did not involve modeling air handling systems. The focused airflow paths included external wall leakage area, internal wall leakage area, external doors, internal doors, and stairwells. The external wall leakage area encompassed all intentional and unintentional openings in the building envelope, excluding doors. The size of the external wall leakage area depended on building size and construction year, and the calculation formula is as follows:

$$NL = \exp(\beta_0 + \beta_1 \times \text{year built} + \beta_2 \times \text{floor area} + \epsilon) \quad (2-1)$$

$$ELA = 0.001 \times A_f \times NL \times (H/2.5)^{-0.3} \quad (2-2)$$

This empirical formula was obtained through multiple regression by researchers using data from the American air leakage database [36]. NL represents the normalized leakage area, β_0 , β_1 and β_2 (m^{-2}) are constant coefficients, and for low-income residential buildings, they are taken as 11.1, -5.37×10^{-3} and -4.18×10^{-3} respectively. “*year built*” is the year of construction for the house, “*floor area*” (m^2) and A_f (m^2) denote the building area, H (m) represents the floor height, and ELA (m^2) stands for the total leakage area of the building envelope. To obtain the unit area leakage area, we divided ELA by the external surface area of the building. It's worth noting that this empirical formula is derived from summarizing data on American residential buildings and may not necessarily be applicable to Chinese residential buildings. However, considering that American residences are primarily low-rise detached buildings, which is somewhat similar to Chinese rural buildings, and given that this empirical formula is currently the most reliable one available to us, previous scholars have successfully used this formula to conduct valuable research on residential buildings in Beijing [34]. Therefore, we continued to use this empirical formula in this study. Detailed information of other airflow path elements is presented in “Floor plan and airflow path” in Supporting Information.

Step3: Setting the indoor temperature. The indoor-outdoor temperature difference drives stack-ventilation, so the indoor temperature settings varied for different climate zones and seasons. The specific methods of dividing seasons and temperature setting principles can

be found in “Season classification and indoor temperature setting” in Supporting Information.

Step4: Performing hour-by-hour simulations throughout the entire year on the residential buildings using weather data corresponding to each sampled location which is provided in EnergyPlus weather format from Chinese Standard Weather Data (CSWD) [49].

Step5: Exporting the time-averaged transient zone airflow for each sampled residence and saving it as an Excel worksheet for each season. The average air infiltration rate for each season was calculated by dividing the total airflow entering the residence from outdoors by the corresponding volume.

2.3. Field measurement

This study conducted measurements of the air infiltration rate for 8 rural households in Yunnan Province and Guizhou Province and 15 residences (results provided by RRPCS, NJUSAUP) in Jiangsu Province using the CO₂ decay method. The measurements were conducted in the months of July and August. The experiments were conducted under clear and calm weather conditions, with wind speeds below 3 m/s to avoid the influence of wind on the experimental results [32], because it's difficult for us to measure the wind pressure during on-site experiments. Before the experiments, GE 7001 carbon dioxide detectors (with an accuracy of ±50 ppm) and temperature and humidity data loggers (with an accuracy of ±0.5 °C and ±3%) were strategically placed in both the living room and the bedroom. Each GE 7001 carbon dioxide detector was equipped with a HOBO data logger. All instruments were positioned at a height of 1 m above the ground, placed near the centre of the rooms, away from corners, doors, and windows. Temperature and humidity data loggers were also positioned outdoors to record the outdoor weather conditions. And we took the measured weather data (mainly temperature data) as the input parameter for simulation validation. During the experiment, the experimenter walked steadily inside the room while holding a carbon dioxide fire extinguisher and a fan, releasing carbon dioxide and promoting rapid mix of the gas. This process persisted until concentrations exceeding 2000 ppm were consistently detected by all carbon dioxide detectors, with displayed values in close proximity. Upon achieving this condition, the release was terminated, and the experiment personnel were relocated to the outdoor area. To ensure accurate measurement of the air infiltration rate, which specifically reflects building infiltration and is not influenced by occupants' behavior, we maintained an unoccupied condition during the measurements. In this period, the external doors and windows were closed, while the internal doors and windows remained open.

The accumulation of indoor carbon dioxide in a well-mixed single zone can be represented by equations (2) and (3):

$$dC_{in}/dt = \alpha(C_{out} - C_{in}) \quad (2-3)$$

C_{in} (ppm) stands for the concentration of indoor carbon dioxide, and C_{out} (ppm) represents the concentration of outdoor carbon dioxide. α (h^{-1}) is air infiltration rate. Assuming that the air infiltration rate remains constant during the experiment, equations (2) and (3) can be derived to yield equations (2)–(4):

$$\ln(C_{in} - C_{out}) = -\alpha t + C \quad (2-4)$$

In which C_{out} was set as 400 ppm considering outdoor CO₂ concentration often varies from 350 to 450 ppm in reality [15]. Also, it is consistent with the average value of ambient CO₂ concentrations which we measured during August 2022, in the studied regions. According to error analysis, the impact of fluctuation of outdoor CO₂ concentration in air infiltration rate is less than 10%, which is acceptable [15]. C (ppm) is a constant, and the air infiltration rate α (h^{-1}) represents the slope of the fitted line. Since the prerequisite for the above equation to hold is the complete mixing within a single zone, we only selected data intervals where the difference of concentrations between different detectors was less than 15%, as demonstrated in gray section in Fig. 1.

3. Results

3.1. Comparisons between field measurements with simulations

As shown in Fig. 2(a), the relative error between the simulated and the measured air infiltration rate of 6 rural residences in Yunnan and Guizhou regions is smaller than 15%. This relative error is within an acceptable range, and the simulation results are reliable. In addition, Fig. 2(b) shows a comparison between 15 data points obtained from rural areas in Jiangsu and simulated values for HSCW zone. Since there was no mapping performed for the measured samples, a one-to-one comparison method was not applicable. It can be seen that the mean air infiltration rate obtained from simulation is 0.44 h^{-1} , while the mean measured result is 0.51 h^{-1} . The medians are 0.40 h^{-1} and 0.31 h^{-1} , respectively. The results are close, but it can be observed from the graph that the standard deviation of the measured results is larger and more dispersed. This can be attributed to the simulation method's limitation in accounting for unique variations between individual samples,

resulting in a smaller standard deviation in the results. In contrast, measurements can reflect specific variations such as openings and superior sealing of newly built houses, leading to more dispersed results, with most values being small and a few values being large, resulting in a lower median. From the perspective of providing reference values, the average value of the simulation results is still accurate. In summary, the reliability of the simulation results has been verified.

3.2. Distribution of the modeled air infiltration rate

The annual and seasonal average air infiltration rate of rural residences in various climatic zones are statistically summarized in Table 3. In the SC region, the annual average air infiltration rates range from 0.53 to 1.39 h^{-1} , with a geometric mean (GM) of 1.03 h^{-1} and a median of 1.12 h^{-1} . In the C region, the annual average air infiltration rates range from 0.21 to 0.69 h^{-1} , with a GM of 0.49 h^{-1} and a median of 0.58 h^{-1} . In the M region, the annual average air infiltration rates range from 0.18 to 0.40 h^{-1} , with a GM of 0.25 h^{-1} and a median of 0.24 h^{-1} . In the HSCW region, the annual average air infiltration rates range from 0.19 to 1.17 h^{-1} , with a GM of 0.42 h^{-1} and a median of 0.38 h^{-1} . In the HSWW region, the annual average air infiltration rates range from 0.35 to 1.27 h^{-1} , with a GM of 0.61 h^{-1} and a median of 0.60 h^{-1} . GM instead of arithmetic mean was chosen to represent the distribution, because each distribution follows log-normal distribution according to our following test. The arithmetic mean of a true log-normal distribution is often susceptible to the large tail number, much larger than the median value, while GM is equal to the median value, which is more representative of the statistical results of data [50,51].

Fig. 3 displays the histogram of the annual average results for 5 plotted climate zones, which visually shows a strong right-skewed tendency. Therefore, a two-parameter log-normal distribution was conducted fitting to the results, represented as follows:

$$\ln(\alpha) \sim N[\mu, \sigma^2] \quad (3-1)$$

α (h^{-1}) represents the air infiltration rate, μ is the mean value of the distribution, and σ is the standard deviation. The fitting results are shown in Fig. 4, where the vertical axis of the scattered points represents the logarithm of the simulated values, and the horizontal axis of the

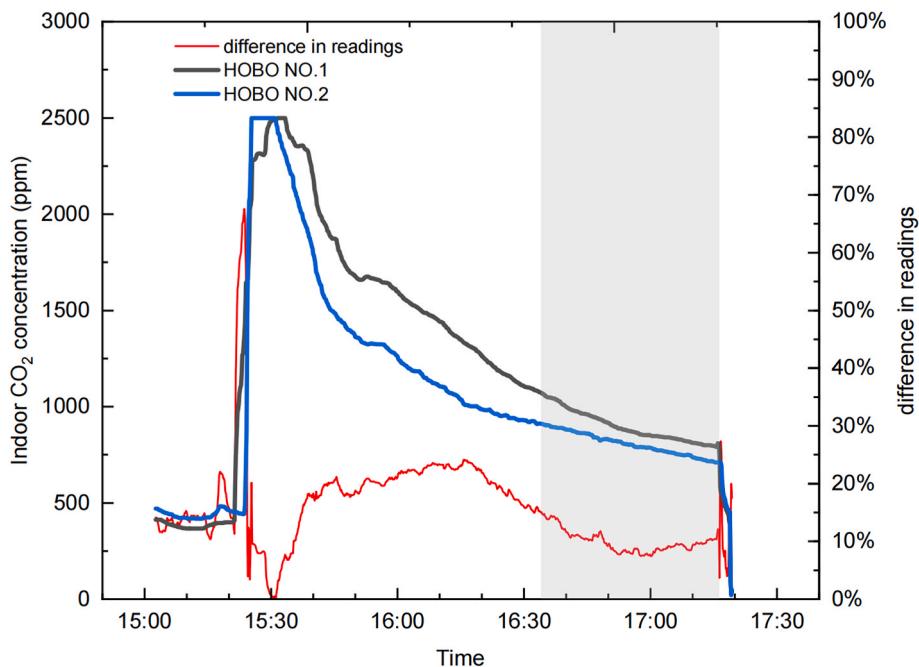
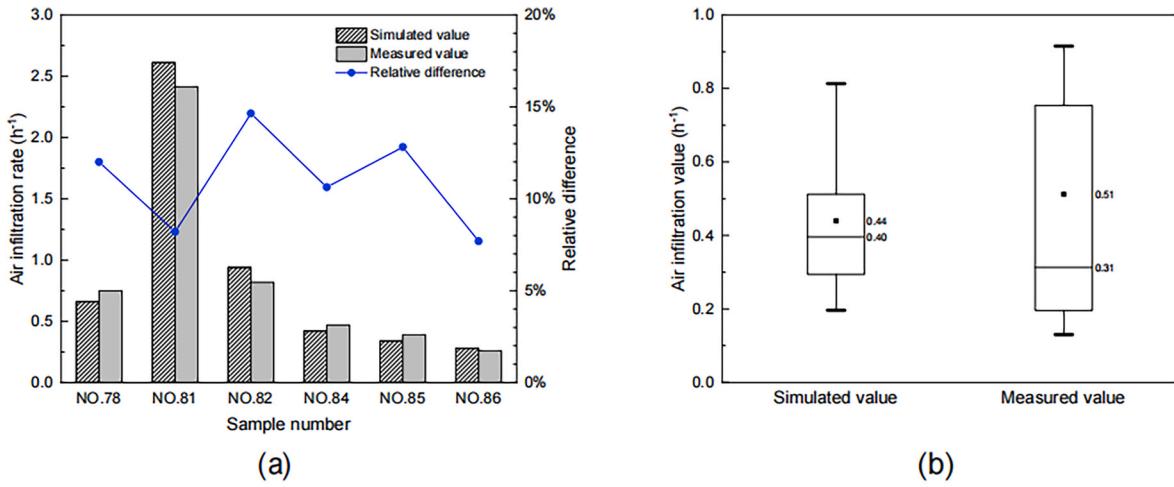


Fig. 1. Measured CO₂ concentration from different detectors.



Samples in figure(a) come from the HSWW region, and the simulated values correspond one-to-one with the measured values. Samples in figure(b) come from the HSCW region. In Figure(b), the top and bottom of the box represent the 75th percentile and the 25th percentile respectively, the whiskers correspond to the 95th percentile and the 5th percentile, the black dot represents the median, and the horizontal line represents the mean.

Fig. 2. Comparison of simulations and measurements.

Table 3
Statistic results of annual and seasonal average air infiltration rate.

Region	Season	Number of samples	Minimum	P25	P50	P75	Maximum	Geometric mean	Geometric standard deviation
SC	Spring	22	0.51	0.86	1.16	1.35	1.51	1.04	1.38
	Summer	22	0.35	0.59	0.77	0.90	1.17	0.71	1.37
	Autumn	22	0.42	0.71	0.99	1.13	1.26	0.87	1.39
	Winter	22	0.85	1.34	1.55	1.77	1.91	1.48	1.24
	Annually	22	0.53	0.89	1.12	1.26	1.39	1.03	1.31
C	Spring	10	0.22	0.46	0.61	0.67	0.76	0.52	1.46
	Summer	10	0.20	0.35	0.40	0.44	0.50	0.37	1.29
	Autumn	10	0.18	0.32	0.45	0.52	0.58	0.39	1.43
	Winter	10	0.26	0.57	0.81	0.91	0.99	0.67	1.52
	Annually	10	0.21	0.43	0.58	0.63	0.69	0.49	1.43
M	Spring	14	0.19	0.23	0.28	0.38	0.56	0.30	1.38
	Summer	14	0.18	0.20	0.23	0.29	0.35	0.24	1.25
	Autumn	14	0.18	0.20	0.23	0.30	0.37	0.24	1.25
	Winter	14	0.18	0.19	0.22	0.26	0.33	0.23	1.21
	Annually	14	0.18	0.21	0.24	0.32	0.40	0.25	1.27
HSCW	Spring	45	0.22	0.27	0.42	0.61	1.41	0.47	1.62
	Summer	45	0.20	0.24	0.36	0.50	1.20	0.40	1.56
	Autumn	45	0.20	0.25	0.39	0.53	1.25	0.43	1.59
	Winter	45	0.16	0.26	0.31	0.44	1.00	0.37	1.58
	Annually	45	0.19	0.26	0.38	0.53	1.17	0.42	1.58
HSWW	Spring	20	0.38	0.48	0.61	0.76	1.27	0.61	1.39
	Summer	20	0.33	0.47	0.58	0.70	1.10	0.58	1.39
	Autumn	20	0.30	0.48	0.57	0.77	1.25	0.59	1.43
	Winter	20	0.34	0.47	0.62	0.77	1.47	0.63	1.48
	Annually	20	0.35	0.49	0.60	0.78	1.27	0.61	1.40

SC represents severely cold zone, C represents cold zone, M represents moderate zone, HSCW represents how summer and cold winter zone, and HSWW represents hot summer and warm winter zone. Geometric mean (GM) = $\sqrt[n]{\prod_{i=1}^n x_i}$, Geometric standard deviation (GSD) = $\exp(\sqrt{\sum_{i=1}^n (\ln(x_i/GM))^2/n})$.

scattered points corresponds to the z-scores of the simulated results arranged in ascending order. The solid line represents the best-fitting line for the series of scattered points, with its slope representing μ and intercept representing σ . R^2 aims to assess the goodness of fit, and the upper ($\exp(\mu + \sigma)$) and lower ($\exp(\mu - \sigma)$) bounds of the 68% confidence intervals are shown in middle parentheses. It can be observed that as the sample size for each region increases, R^2 gradually improves. The region with the largest sample size, the HSCW region, has the best fitting result with an R^2 of 0.983. For region with fewer samples, R^2 can reach 0.851,

indicating that the research results align well with the log-normal distribution, especially for the HSCW and HSWW climate zones. However, it's worth noting that for SC and C region, the log-normal distribution may overestimate extremely low and high air infiltration rate while underestimating intermediate value. Conversely, for M region, the results exhibit the opposite behavior.

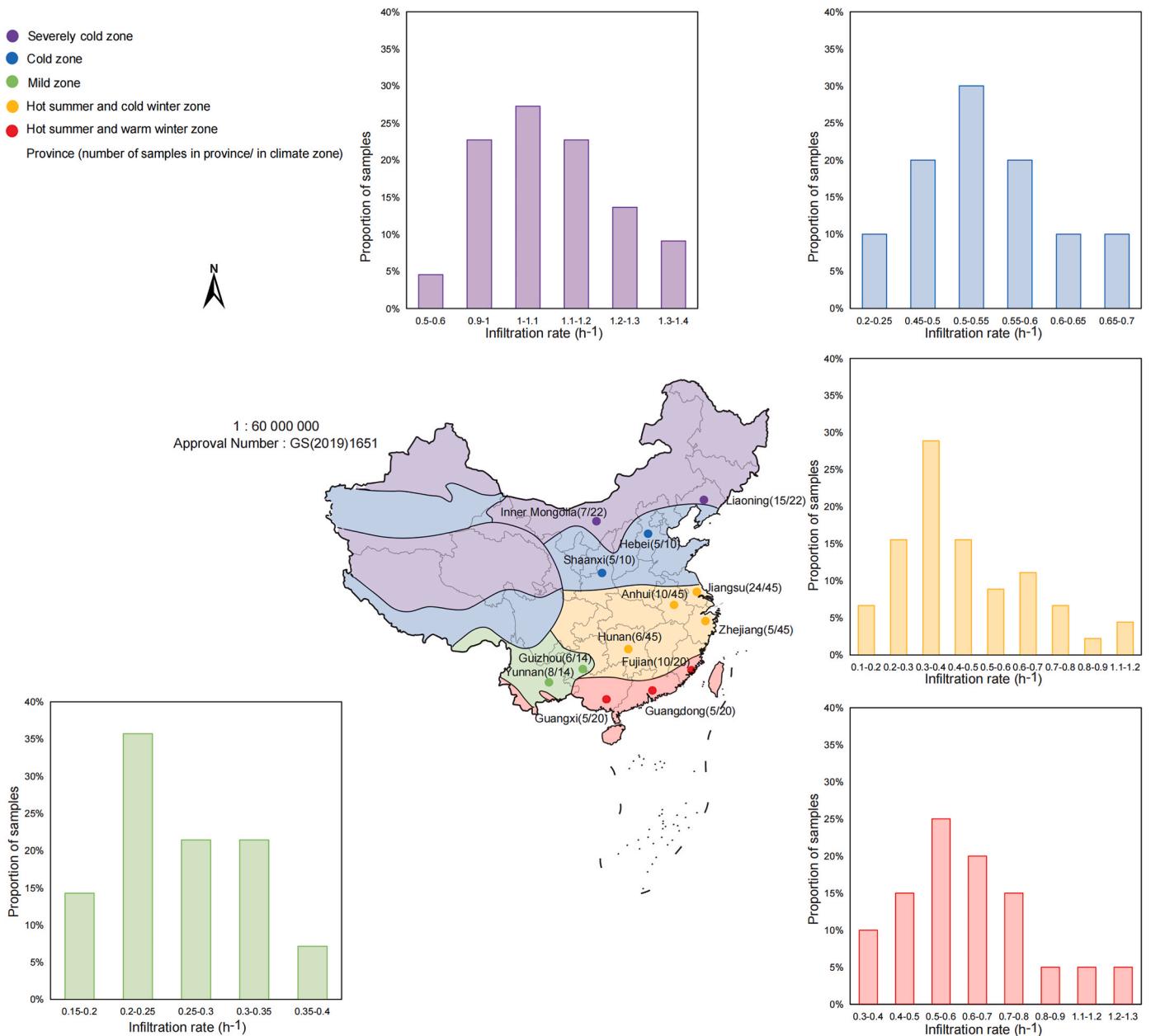


Fig. 3. Distribution of sample's climate zones and probability histogram of annual average air infiltration rate.

3.3. Influencing factors analysis

Fig. 5 displays the simulation results for the four seasons and the annual average of five climate regions. The annual averages, in descending order, are for SC, HSWW region, C region, HSCW region, and M region. Specifically, the figure shows that the differences in air infiltration rate among seasons in the north (i.e., SC and C regions) are most pronounced, with significantly higher rates in winter and spring compared to summer and autumn. This is because during the heating season, the study set the indoor temperature in the north not lower than 14°C , while the outdoor temperature can drop to -20°C , resulting in a large indoor-outdoor temperature difference. As a result, the significant temperature difference between indoor and outdoor environments in the north during winter dominates the air infiltration rate through thermal pressure-driven ventilation [15]. At the same time, outdoor wind speed is higher in spring and winter than in summer and autumn. Under the combined effects of thermal pressure and wind pressure, the results for the heating season in the north are significantly higher than

those in other regions, and the annual average is consequently higher than in the south. In the south, especially in M region, there are no significant seasonal differences in the results. As for the relative relationships among the results of M region, HSCW region, and HSWW region, further analysis will be provided in the following paragraph.

Fig. 6 depicts the air infiltration rate for residential buildings of different floor areas, construction years, and number of stories. The values 72.2 m^2 and 110.8 m^2 represent the 33rd and 66th percentiles, respectively, for the 111 sampled building floor areas. Based on these percentiles, all samples were divided into three floor area intervals. It can be observed from the figure that as the floor area increases, the air infiltration rate decreases. Specifically, for rural residences with floor areas less than 72.2 m^2 compared to those greater than or equal to 110.8 m^2 , the median air infiltration rate decreases from 0.60 h^{-1} to 0.44 h^{-1} . Regarding the age of buildings, as buildings grow older, there is an increase in air infiltration rate. This can be attributed to the leakage area calculation equation proposed by Chan et al., which suggests that older buildings often exhibit higher leakage rates [36]. This aligns with the

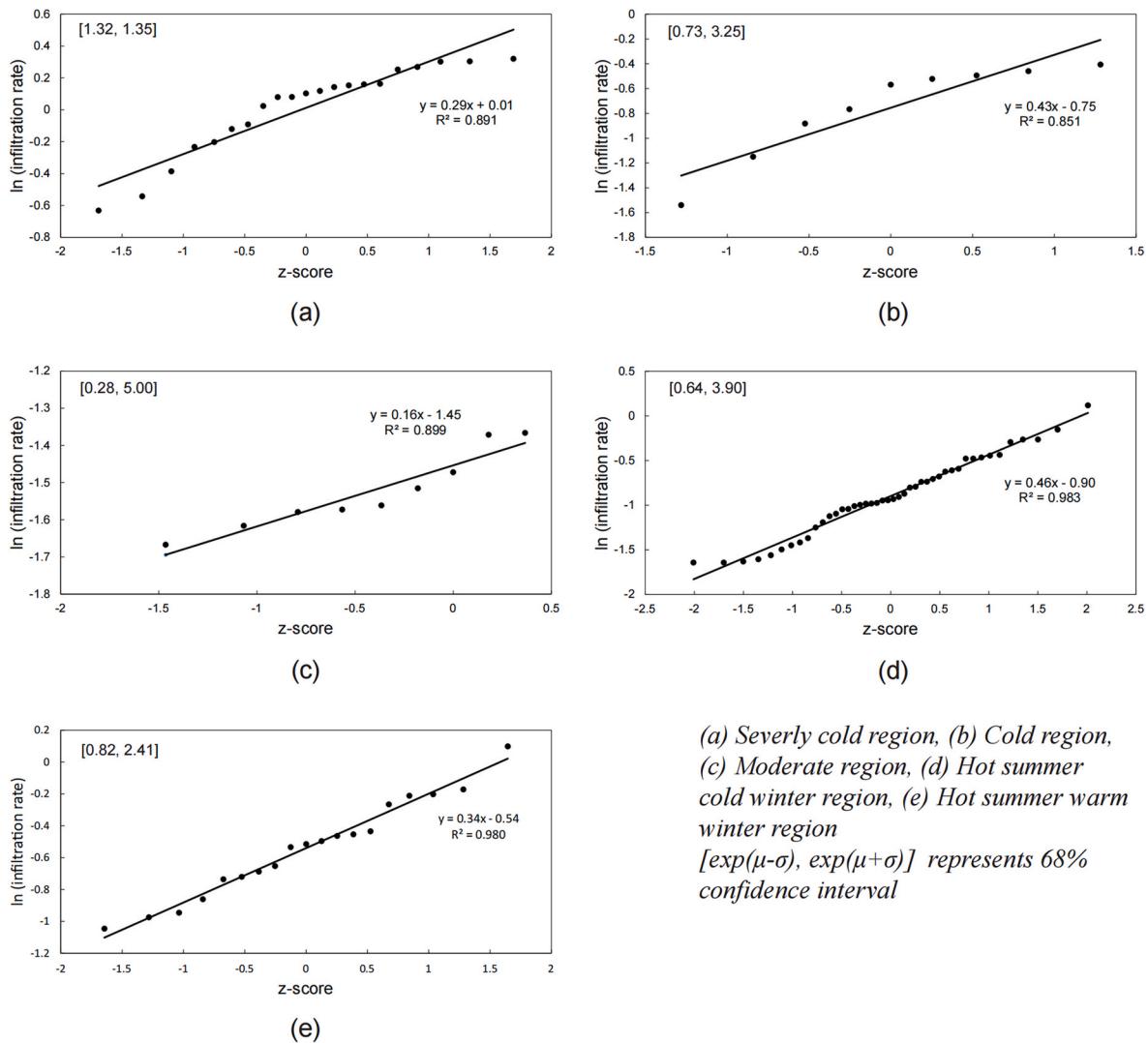
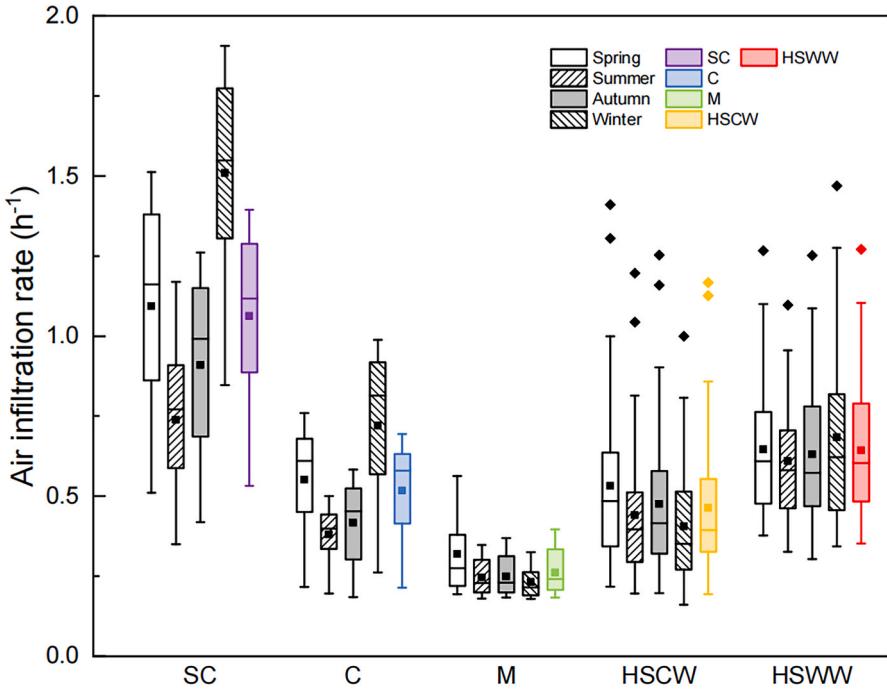


Fig. 4. Logarithm probability of air infiltration rate of each climate zone.

fact that newer buildings tend to be constructed with energy-efficient designs. The median air infiltration rate for rural residences constructed before 1980 is 0.81 h^{-1} , while for new constructions after 2011, it is only 0.40 h^{-1} . However, there is no significant difference in results between the years 1981 and 2010, and in fact, the mean value slightly increases during this period. In terms of the number of stories, there is an initial decrease followed by an increase in the air infiltration rate as the number of stories increases. Compared to single-story houses and buildings with four or more stories, rural residences with 2–3 stories have the lowest air infiltration rate. In Fig. 7, we conducted a correlation analysis between rural residential building floor area, number of stories, and shape coefficient of building (the ratio of the exterior surface area of a building exposed to outdoor air to its enclosed volume). The box plots on the left indicate that buildings with more stories tend to have smaller floor areas, and the scatter plot shows that as floor area increases, the shape coefficient of building decreases. In other words, the ratio of building surface area to volume decreases with larger floor areas. Combining these two graphs reveals that as the number of stories increases, rural residence morphology transitions from 'short and stout' to 'tall and slender', and the air infiltration rate exhibits a corresponding trend of decreasing and then increasing. The residences with the most 'average' shape coefficient of building, typically having 2–3 stories and moderate floor areas, have the lowest air infiltration rate. We speculate that this can account for the abnormal result in Fig. 6(b). Samples from

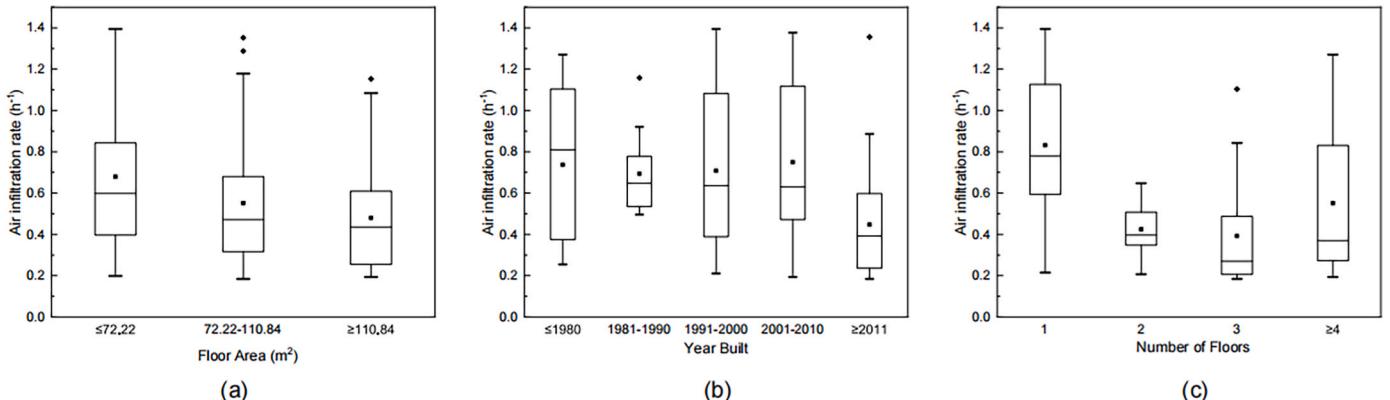
1991 to 2000 are characterized by small area and high floor, while from 2001 to 2010, the average sample area is large with low floor and large shape coefficient, thus the morphological factor replaces the year factor.

The aforementioned conclusion can explain the relative relationship of the results between the HSCW and HSWW regions. Table 2 reveals that in terms of building type, the proportion of samples with floor areas less than 72.2 m^2 in HSWW region is 55%, significantly higher than the proportions in M and HSCW regions at 35% and 37%, respectively. Additionally, buildings with three or more stories account for 55% of the sample in this region, also higher than the other two regions at 35% and 22%. This indicates that single-story, small-area rural buildings with multiple stories are more prevalent in HSWW region, which is consistent with the statistical trends of the Chinese population census [45], consequently leading to higher air infiltration rate in this region. In contrast, rural residences in M and HSCW regions align more closely with the 'average' type mentioned earlier, with minimal differences between the two regions. However, the samples in M region include a higher proportion (57%) of new constructions built after 2011, likely due to the recent rural poverty alleviation and relocation efforts in this area, which may have influenced the sampling results. As a result, the proportion of new constructions in M region is much higher than in HSCW region, where samples are evenly distributed across various years, leading to the situation where the results for M region are lower than those for HSCW region.



The colored boxplots represent the annual average simulated results for each climatic zone. The top and bottom of the box represent the 75th percentile and the 25th percentile respectively, the whiskers correspond to the 95th percentile and the 5th percentile, the dot represents the median, the diamond-shaped dot represents an outlier and the horizontal line represents the mean.

Fig. 5. Box plot of annual and seasonal average air infiltration rate.



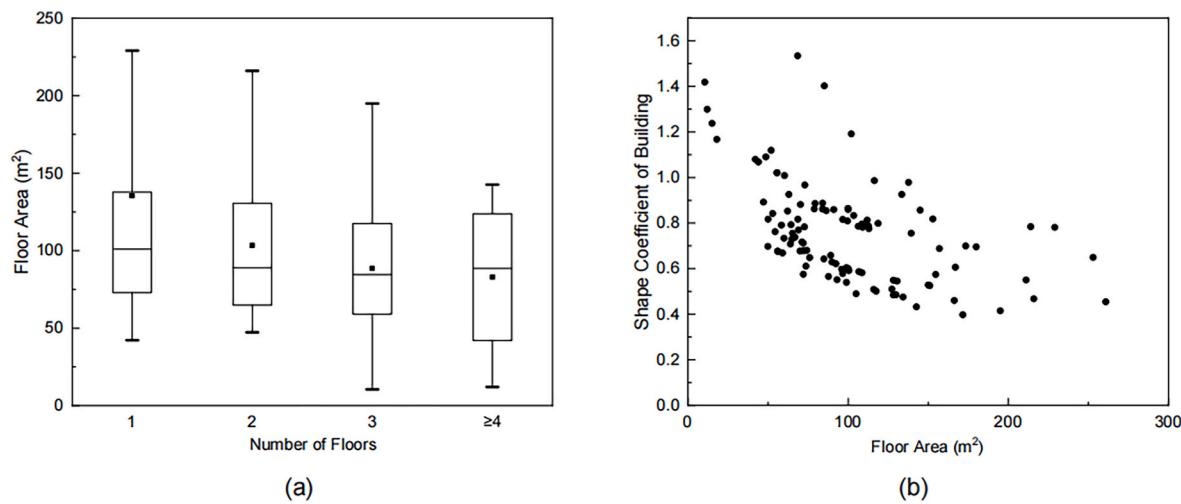
The top and bottom of the box represent the 75th percentile and the 25th percentile respectively, the whiskers correspond to the 95th percentile and the 5th percentile, the dot represents the median, the diamond-shaped dot represents an outlier and the horizontal line represents the mean.

Fig. 6. Air infiltration rate for residences with different (a) Floor area (b) Year built (c) Number of floors.

3.4. Comparisons between urban and rural areas

Table 4 presents a comparison of the air infiltration rate between rural area and urban area. Urban data were collected from studies by Cheng at Tsinghua University [52] and Hou et al. [15] at Tianjin University. Considering the varying geographical classification methods employed by different researchers, for the purpose of facilitating comparison, this study reorganized the results according to the geographical division methods used in the two urban studies. The rural research outcomes, except for the southwest temperate region, are consistently

higher than those of the corresponding urban areas. This disparity can be attributed to the fact that rural residences, influenced by economic factors, tend to be older and less well-constructed compared to urban dwellings, making these results readily understandable. For the severely cold and cold regions, geographically designated as northern China, the simulated values for these areas are significantly higher than the urban results. Numerically, the median values for cities are as low as 0.24 h^{-1} and 0.30 h^{-1} , while the rural medians are substantially higher at 1.17 h^{-1} and 1.12 h^{-1} . The differences between the two reach 389% and 273%, respectively. One possible explanation is that the significant



(a): The top and bottom of the box represent the 75th percentile and the 25th percentile respectively, the whiskers correspond to the 95th percentile and the 5th percentile, the dot represents the median, the diamond-shaped dot represents an outlier and the horizontal line represents the mean.

Fig. 7. Correlation analysis of floor area, number of floors, and shape coefficient of building.

Table 4
Comparison of air infiltration rate in urban and rural area.

Reference	Urban	Rural	Relative difference of median value
Cheng (2018) [52].	Median (standard deviation) h^{-1}		
North-east 0.24 (0.67)	North-east 1.17 (0.17)	389%	
North& North-west 0.28(0.88)	North& North-west 0.63(0.25)	124%	
East& Centre 0.31 (0.79)	East& Centre 0.45 (0.26)	46%	
South-east& South-west 0.38 (0.86)	South-east& South-west 0.37 (0.15)	-3%	
Hou et al. (2019) [15].	Median h^{-1}		
SC region 0.30	SC region 1.12	273%	
C region 0.31	C region 0.58	87%	
M region 0.38	M region 0.24	-37%	
HSCW region 0.41	HSCW region 0.38	-7%	
HSWW region 0.34	HSWW region 0.60	76%	

$$\text{relative difference} = (\text{urban} - \text{rural})/\text{rural}.$$

indoor-outdoor temperature difference during winter and spring seasons in the northern regions magnified the impact of the ELA difference between rural and urban residences, which means even a small increment in ELA can lead to a significant increase in air infiltration rate in conditions of significant indoor-outdoor temperature. Interventions such as building envelope modifications and energy-efficient retrofits, which can enhance the air-tightness of rural homes, may not have been adequately represented in the simulation process. Furthermore, this study utilized an empirical formula for residential air leakage area primarily based on data from northern U.S. states. The use of this formula may have led to an underestimation of the building envelope air leakage area for southern residences, which in turn could have resulted in the simulated values for rural homes in the southwestern region being lower than those for urban areas.

4. Discussion

The innovative significance of this study lies in exploring a research path regarding the air infiltration rate distribution and filling the gap of raw data in Chinese rural residences. However, it is worth noting that the subjects selected for this study are not probabilistic samples. Although we have provided representative descriptions, our sample size remains relatively small compared to the vast housing stock. In the future, it will be necessary to continue collecting rural housing information to expand the sample size and enhance the reliability of the results. At this phase, the results obtained in this study also hold practical value. For instance, they can be used to adjust research findings related to outdoor air pollutant exposure among Chinese residents and energy consumption calculation. Previously, Hu et al. utilized Monte Carlo approach to simulate the air infiltration rate and personal exposure to major environmental pollutants ($\text{PM}_{2.5}$, PM_{10} , O_3 , NO_2 , and SO_2) across 31 provinces in China by season, gender, and age [53]. The air infiltration rate in residential buildings was considered as a crucial input parameter in the calculation process, relying on the results from Cheng's 2018 study [52]. However, there was no differentiation between urban and rural areas in previous study. Therefore, by updating the air infiltration rate for rural residences while keeping the input parameters for urban areas unchanged, a more precise understanding of individual exposure to major environmental pollutants in China can be achieved. Since the air infiltration rate of rural residences obtained in this study is larger than urban in most areas, it means that rural residents suffer a higher risk of air pollution exposure originated outdoor, which will prompt policymakers to pay more attention to the equality between urban and rural areas. Also, the analysis of the connection between multiple factors and air infiltration rate in this study can provide guidance for the improvement initiatives.

5. Conclusion

- 1) The distribution of residential air infiltration rate in the same climatic zone generally follows a log-normal distribution. Particularly, for hot summers and cold winters as well as hot summers and warm winters climatic zones, the log-normal distribution may overestimate very low and very high air infiltration rate while underestimating intermediate value. In contrast, for moderate climatic zones, the log-

- normal distribution may underestimate very low and very high air infiltration rate while overestimating intermediate value.
- 2) The geometric mean of annual average simulated air infiltration rates, from highest to lowest, are as follows: 1.03 h^{-1} for SC region, 0.61 h^{-1} for HSWW region, 0.49 h^{-1} for C region, 0.42 h^{-1} for HSCW region, and 0.25 h^{-1} for M region.
 - 3) Significant seasonal variations of air infiltration rate are observed in SC and C regions, with simulated values in winter and spring significantly higher than those in summer and autumn.
 - 4) The simulated air infiltration rates are influenced by the housing morphology, decreasing as the building area increases. The simulated results are also related to the year of construction, with air infiltration rate decreasing as the construction year approaches.

CRediT authorship contribution statement

Ye Wang: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Shanshan Shi:** Validation, Resources, Investigation, Data curation. **Zhengxu Zhou:** Validation, Resources, Data curation. **Song Guo:** Resources, Data curation. **Bin Zhao:** Writing – original draft, Validation, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author used ChatGPT in order to polish English writing. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We express our gratitude to Department of Urban Planning at the School of Architecture at Tsinghua University and the Centre of Rural Revitalization Praxis at the School of Architecture and Urban Planning at Nanjing University for providing data for this study. At the same time, specifically thanking the reviewers for their valuable comments and suggestions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.buildenv.2024.111284>.

References

- [1] N.E. Klepeis, W.C. Nelson, W.R. Ott, J.P. Robinson, A.M. Tsang, P. Switzer, et al., The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants, *J. Expo. Sci. Environ. Epidemiol.* 11 (2001) 231–252.
- [2] J.A. Leech, W.C. Nelson, R.T. Burnett, S. Aaron, M.E. Raizenne, It's about time: a comparison of Canadian and American time-activity patterns, *J. Expo. Sci. Environ. Epidemiol.* 12 (2002) 427–432.
- [3] G.R.F. Collaborators, Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019, *Lancet* 396 (10258) (2020) 1223–1249.
- [4] N. Yamamoto, D.G. Shendell, A.M. Winer, J. Zhang, Residential air exchange rates in three major US metropolitan areas: results from the Relationship Among Indoor, Outdoor, and Personal Air Study 1999–2001, *Indoor Air* 20 (1) (2010) 85–90.
- [5] H. Guo, L. Morawska, C. He, D. Gilbert, Impact of ventilation scenario on air exchange rates and on indoor particle number concentrations in an air-conditioned classroom, *Atmos. Environ.* 42 (4) (2008) 757–768.
- [6] P. Blondeau, V. Iordache, O. Poupart, D. Genin, F. Allard, Relationship between outdoor and indoor air quality in eight French schools, *Indoor Air* 15 (1) (2005) 2–12.
- [7] N. Canha, C. Mandin, O. Ramalho, G. Wyart, J. Riberon, C. Dassonneville, et al., Assessment of ventilation and indoor air pollutants in nursery and elementary schools in France, *Indoor Air* 26 (3) (2016) 350–365.
- [8] C.Y.H. Chao, T.C. Tung, An empirical model for outdoor contaminant transmission into residential buildings and experimental verification, *Atmos. Environ.* 35 (2001) 1585–1596.
- [9] C. Chen, B. Zhao, C.J. Weschler, Indoor exposure to “outdoor PM10” assessing its influence on the relationship between PM10 and short-term mortality in US cities, *Epidemiology* 23 (2012) 870–878.
- [10] C. Chen, B. Zhao, C.J. Weschler, Assessing the influence of indoor exposure to “Outdoor ozone” on the relationship between ozone and short-term mortality in US communities, *Environ. Health Perspect.* 120 (2012) 235–240.
- [11] ASHRAE, *Handbook of Fundamentals*, Atlanta, 2021.
- [12] C. Younes, C.A. Shdid, G. Bitsuamakl, Air infiltration through building envelopes: a review, *J. Build. Phys.* 35 (3) (2012) 267–302.
- [13] W.J. Riley, T.E. McKone, A.C. Lai, W.W. Nazaroff, Indoor particulate matter of outdoor origin: importance of size-dependent removal mechanisms, *Environ. Sci. Technol.* 36 (2) (2002) 200–207.
- [14] A. Persily, A. Musser, S.J. Emmerich, Modeled infiltration rate distributions for U.S. housing, *Indoor Air* 20 (6) (2010) 473–485.
- [15] J. Hou, Y. Sun, Q. Chen, R. Cheng, J. Liu, X. Shen, et al., Air change rates in urban Chinese bedrooms, *Indoor Air* 29 (5) (2019) 828–839.
- [16] K. Isaacs, J. Burke, L. Smith, R. Williams, Identifying housing and meteorological conditions influencing residential air exchange rates in the DEARS and RIOPA studies: development of distributions for human exposure modeling, *J. Expo. Sci. Environ. Epidemiol.* 23 (3) (2013) 248–258.
- [17] M.I. Montoya, E. Pastor, E. Planas, Air infiltration in Catalan dwellings and sealed rooms: an experimental study, *Build. Environ.* 46 (10) (2011) 2003–2011.
- [18] G. Hong, B.S. Kim, Field measurements of infiltration rate in high rise residential buildings using the constant concentration method, *Build. Environ.* 97 (2016) 48–54.
- [19] FdA. Gonzalo, M. Griffin, J. Laskosky, P. Yost, R.A. Gonzalez-Lezcano, Assessment of indoor air quality in residential buildings of new england through actual data, *Sustainability* 14 (2) (2022).
- [20] F.-M. Jesus, P.-C. Irene, G.-L. Roberto Alonso, P. Cristina, E. Victor, A. de larriva Rafael, et al., Methodology for the study of the envelope airtightness of residential buildings in Spain: a case study, *Energies* 11 (4) (2018).
- [21] J. Feijo-Munoz, C. Pardal, V. Echarri, J. Fernandez-Aguera, R. Assiego de Larriba, M. Montesdeoca Calderin, et al., Energy impact of the air infiltration in residential buildings in the Mediterranean area of Spain and the Canary islands, *Energy Build.* 188 (2019) 226–238.
- [22] J. Feijo-Munoz, R. Alonso Gonzalez-Lezcano, I. Poza-Casado, M. Angel Padilla-Marcos, A. Meiss, Airtightness of residential buildings in the Continental area of Spain, *Build. Environ.* 148 (2019) 299–308.
- [23] G. Hong, C. Kim, Experimental analysis of airtightness performance in high-rise residential buildings for improved code-compliant simulations, *Energy Build.* (2022) 261.
- [24] U. Mathur, R. Damle, Impact of air infiltration rate on the thermal transmittance value of building envelope, *J. Build. Eng.* (2021) 40.
- [25] A. Salehi, I. Torres, A. Ramos, Experimental analysis of building airtightness in traditional residential Portuguese buildings, *Energy Build.* 151 (2017) 198–205.
- [26] G. Hong, D.D. Kim, Airtightness of electrical, mechanical and architectural components in South Korean apartment buildings using the fan pressurization and tracer gas method, *Build. Environ.* 132 (2018) 21–29.
- [27] A.V. Pasos, X. Zheng, L. Smith, C. Wood, Estimation of the infiltration rate of UK homes with the divide-by-20 rule and its comparison with site measurements, *Build. Environ.* (2020) 185.
- [28] H.K. Dai, C. Chen, Air infiltration rates in residential units of a public housing estate in Hong Kong, *Build. Environ.* (2022) 219.
- [29] X. Sui, Z. Tian, H. Liu, H. Chen, D. Wang, Field measurements on indoor air quality of a residential building in Xi'an under different ventilation modes in winter, *J. Build. Eng.* (2021) 42.
- [30] M. Yao, B. Zhao, Distribution of air change rates in residential buildings in Beijing, China, in: al ZWe (Ed.), *Proceedings of the 11th International Symposium on Heating, Ventilation and Air Conditioning, ISHVAC 2019*, 2019, pp. 1149–1156.
- [31] H. Yin, C. Liu, L. Zhang, A. Li, Z. Ma, Measurement and evaluation of indoor air quality in naturally ventilated residential buildings, *Indoor Built Environ.* 28 (10) (2019) 1307–1323.
- [32] P.L. Cheng, X. Li, Air infiltration rates in the bedrooms of 202 residences and estimated parametric infiltration rate distribution in Guangzhou, China, *Energy Build.* 164 (2018) 219–225.
- [33] Y. Ji, L. Duanmu, X. Li, Field test on the air tightness of A multi-storey apartment in dalian, *Build. Sci.* 33 (4) (2017) 57–63, 109.
- [34] S. Shi, C. Chen, B. Zhao, Air infiltration rate distributions of residences in Beijing, *Build. Environ.* 92 (2015) 528–537.

- [35] W.R. Chan, W.W. Nazaroff, P.N. Price, M.D. Sohn, A.J. Gadgil, Analyzing a database of residential air leakage in the United States, *Atmos. Environ.* 39 (19) (2005) 3445–3455.
- [36] W.R. Chan, J. Joh, M.H. Sherman, Analysis of air leakage measurements of US houses, *Energy Build.* 66 (2003) 616–625.
- [37] W.R. Chan, J. Joh, M.H. Sherman, Analysis of air leakage measurements of US houses, *Energy Build.* 66 (2013) 616–625.
- [38] M.S. Breen, B.D. Schultz, M.D. Sohn, T. Long, J. Langstaff, R. Williams, et al., A review of air exchange rate models for air pollution exposure assessments, *J. Expo. Sci. Environ. Epidemiol.* 24 (2014) 555–563.
- [39] M.I. Montoya, E. Pastor, F.R. Carrie, G. Guyot, E. Planas, Air leakage in Catalan dwellings: developing an airtightness model and leakage airflow predictions, *Build. Environ.* 45 (6) (2010) 1458–1469.
- [40] W. Dols, B. Polidoro, CONTAM User Guide and Program Documentation Version 3.2, Technical Note (NIST TN), National Institute of Standards and Technology, Gaithersburg, MD, 2015, <https://doi.org/10.6028/NIST.TN.1887>.
- [41] J.M. Logue, W.J.N. Turner, I.S. Walker, B.C. Singer, A simplified model for estimating population-scale energy impacts of building envelope air tightening and mechanical ventilation retrofits, *Journal of Building Performance Simulation* 9 (2016) 1–16.
- [42] B. Jones, P. Das, Z. Chalabi, M. Davies, I. Hamilton, R. Lowe, et al., Assessing uncertainty in housing stock infiltration rates and associated heat loss: English and UK case studies. *Building and environment, Build. Environ.* 92 (2015) 645–656.
- [43] C. Molina, B. Jones, I.P. Hall, M.H. Sherman, CHAARM: a model to predict uncertainties in indoor pollutant concentrations, ventilation and infiltration rates, and associated energy demand in Chilean houses, *Energy Build.* 230 (2021) 110539.
- [44] C. Molina, M. Kent, I. Hall, B. Jones, A data analysis of the Chilean housing stock and the development of modelling archetypes, *Energy Build.* (2020) 206.
- [45] P. Wang, H. Cui, L. Ye, in: P. Wang, H. Cui, L. Ye (Eds.), *China Population Census Almanac-2020*, China Statistics Press, Beijing, 2020.
- [46] P.R.C. tMoHaU-Rdot, Code for the Thermal Design of Civil Building, 2017.
- [47] P.F. Linden, The fluid mechanics of natural ventilation, *Annu. Rev. Fluid Mech.* 31 (1999) 201–238.
- [48] H. Du, Z. Lian, L. Lan, D. Lai, Application of statistical analysis of sample size: how many occupant responses are required for an indoor environmental quality (IEQ) field study, *Build. Simulat.* 16 (2023) 577–588.
- [49] US Department of Energy's web site for Energy Plus [Available from: <https://energyplus.net/weather>].
- [50] T.B.L. Kirkwood, Geometric means and measures of dispersion, *Biometrics* 35 (4) (1979) 908–909.
- [51] E. Limpert, W.A. Stahel, M. Abbt, Log-normal distributions across the sciences, *Bioscience* 51 (5) (2001) 341–352.
- [52] P.L. Cheng, Natural Ventilation Rate Distribution in Dwellings in China's 4 Major Cities, Tsinghua University, Beijing, 2018.
- [53] Y. Hu, M. Yao, Y. Liu, B. Zhao, Personal exposure to ambient PM_{2.5}, PM₁₀, O₃, NO₂, and SO₂ for different populations in 31 Chinese provinces, *Environ. Int.* 144 (2020) 106018.