



Article

Energy Prediction and Optimization Based on Sequential Global Sensitivity Analysis: The Case Study of Courtyard-Style Dwellings in Cold Regions of China

Juanli Guo ¹, Meiling Li ¹, Yongyun Jin ^{2,*} , Chundi Shi ² and Zhoupeng Wang ² 

¹ School of Architecture, Tianjin University, Tianjin 300072, China; guojuanli@tju.edu.cn (J.G.); meilingli_418@tju.edu.cn (M.L.)

² Tianjin International Engineering Institute, Tianjin University, Tianjin 300072, China; chundi1011@tju.edu.cn (C.S.); wangzhoupeng@tju.edu.cn (Z.W.)

* Correspondence: jinyongyun@tju.edu.cn

Abstract: A great abundance of rural houses lacking design guidance exists in the cold regions of China, often accompanied by huge energy loss. Particularly, a courtyard-style dwelling (CSD) has more complex and diverse building elements than a common house, rendering the design optimization extremely costly. Sensitivity analysis (SA) can screen the significant parameters of energy consumption for prediction and optimization. In this paper, (1) the design variables related to CSDs and their data details were extracted; (2) a ranking of parameters sensitive to energy demand was formulated; (3) an energy prediction model was trained and (4) dual-objective optimization was carried out. Using the survey data from 150 units in nine villages, 25 control variables were extracted for sequential global sensitivity analysis (GSA). Thus, the ranking of sensitivity parameters was formulated with the two-stage-and-three-sort GSA method. Furthermore, an energy prediction model was then trained with Gaussian Process Regression (GPR) and compared with the other four high-precision models. Based on the obtained prediction model, optimization was then carried out on energy and economic concerns. Consequently, a GSA-based workflow for CSD optimization was proposed to help architectural designers figure out the most efficient energy-saving parameter strategy.

Keywords: global sensitivity analysis; courtyard-style dwelling; energy demand; prediction model; optimization



Citation: Guo, J.; Li, M.; Jin, Y.; Shi, C.; Wang, Z. Energy Prediction and Optimization Based on Sequential Global Sensitivity Analysis: The Case Study of Courtyard-Style Dwellings in Cold Regions of China. *Buildings* **2022**, *12*, 1132. <https://doi.org/10.3390/buildings12081132>

Academic Editors: Bo Hong, Dayi Lai, Zhi Gao, Yongxin Xie and Kuixing Liu

Received: 19 June 2022

Accepted: 27 July 2022

Published: 31 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Building performance analysis (BPA) is a powerful tool that integrates different design and operation factors into a comprehensive assessment of buildings [1]. Although complete coverage of various input parameters can improve the accuracy and robustness, it is unacceptable for complex models. Sensitivity analysis (SA) is a feasible solution to this problem since it is a science that focuses on how to allocate the uncertainty of outputs to each of the inputs [2].

SA methods can be classified into two main categories: local sensitivity analysis (LSA) and global sensitivity analysis (GSA). LSA is a one-parameter-at-a-time (OAT) method, which usually entails evaluating the significance of input parameters by calculating the derivatives at specific points [3] or averaging the derivatives at several points [4]. LSA is not appropriate for dependent input parameters and nonlinear or non-additive models [5]. Moreover, its results rely highly on the central values of the parameters and cannot be used to estimate the uncertainty of the model output [6].

Distinct from LSA, GSA allows for a more comprehensive exploration into the effects of input parameters on output results, and therefore, an increasing preference for GSA is emerging to identify the sensitive variables of the building energy model (BEM) [7,8]. GSA

methods can be divided into four primary categories: screening-based [9–11], regression-based [12–14], variance-based [15,16], and metamodel-based [17–19].

As an advanced version of LSA, the screening-based method filters out factors from a vast quantity of parameters based on no reduction in output variance. Distinct from the OAT concept of LSA, the global screening method directly alters another input argument x_2 after the first argument x_1 changes, that is, all the inputs are changed in a single iteration, which is also known as the Morris method. This method has been broadly used for BPA due to its high efficiency. Heiselberg et al. [20] used the Morris method to present the dominant design parameters for heating energy consumption, taking an office building in Denmark as an example. Maučec et al. [21] employed the Morris method to derive the overall U -value of windows, solar heat gain coefficient (SHGC), and heating setpoint temperature as the most sensitive parameters affecting the energy loads, which provides design guidelines for wood-frame buildings in various climatic conditions.

Most commonly used in BPA are the regression-based methods, including Standard Regression Coefficients (SRC), Partial Correlation Coefficients (PCC), Standardized Rank Regression Coefficients (SRRC), and Partial Rank Correlation Coefficients (PRCC). SRC calculates the regression equations using normalized data. PCC is the correlation coefficient between two variables when the effects of other variables are eliminated. SRRC and PRCC are the rank transformations of SRC and PCC, respectively, which can be employed in nonlinear monotonic functions. The various regression-based methods differ in the correlation requirements of the inputs. Tian and de Wilde [22] adopted SRC and ACOSSO to explore the uncertainty and sensitivity in predicting the building thermal performance under climate change, by taking a school building in the UK as an example. Yildiz and Arsan [23] performed the GSA and uncertainty analysis (UA) of an existing apartment building in Izmir, Turkey, using the SRRC with the Latin Hypercube Sampling technique.

To quantify the sensitivity more precisely, a variance-based approach is proposed, which decomposes the output variance into the effects of individual and combined parameters. Thus, it can quantify the independent input effects on the output as well as the joint effects between the inputs. The most popular variance-based methods are Sobol and FAST. The Sobol method assesses the parameter sensitivity by evaluating the contribution of single and multiple parameters to the output variance. It is considered a non-model approach and applies to various linear or non-linear models, provided that the inputs are mutually independent. FAST focuses on calculating the contribution of individual arguments to the output variance, which is widely applicable but cannot take into account the interactions between inputs. The variance-based method is extensively applied in BPA due to its quantifiability and reliability, even though it demands a massive computing process. Spitz et al. [24] used the Sobol method with 6669 simulations to identify influential parameters in building energy performance and to determine the influence of parameter uncertainty on the building performance based on an experimental house in France. Shen and Tzempelikos [25] applied an expanded FAST method to a private office space in Philadelphia, USA, taking seven parameters including window-to-floor ratio, shading transmittance, shading front and back reflectance, spatial aspect ratio, thermal insulation, and glass type to find the significant factors affecting daily illumination, lighting power consumption, as well as heating and cooling demand of the building.

Further, the metamodel-based method can be regarded as developing to address the drawbacks of the variance-based method with high computational expense. The basic idea is to calculate sensitivity measures with nonparametric regression models before performing variance-based SA, thus providing more efficient sensitivity indices. The frequently used metamodels include Random Forest, Neural Networks, Treed Gaussian Process, Multivariate Adaptive Regression Splines, Support Vector Regression, etc. Although such method can effectively reduce the execution time of the model, it also results in the validity domain and the applicability depending heavily on the training data [26]. Owing to the advances in computing technology, this method maintains its full vitality and potential in the field of BPA. Pang and O'Neill [27] conducted an uncertainty and sensitivity analysis on

the usage behavior of hotel domestic hot water as well as its key influencing factors using a Bayesian Network-based Sobol method. Østergård et al. [28] compared the accuracy, efficiency, ease-of-use, robustness, and interpretability of the six prevalent metamodelling techniques based on five metrics of building performance and eight test problems to find suitable metamodelling techniques for different outputs of BPA.

A comparative review of the features of the dominant GSA methods is presented in Figure 1, where the metamodel-based method is excluded since the variance-based method remains in use, bringing them in line with each other except for the lower computational expense and more efficient sensitivity indices [29]. Yet, although the GSA method has been well established in the BPA field, it is rare to find its application in the huge stock of courtyard-style dwellings (CSDs) in China. These differ significantly from urban buildings in various aspects such as building form, layout, occupancy patterns, thermal requirements, etc. Tabadkani et al. [30] employed a brute-force approach to conducting a parametric analysis of courtyard design variants in residential buildings for different climates to assess the occupant thermal comfort, energy loads, and costs in air-conditioned residential buildings. Furthermore, highly accurate deep learning models were constructed to provide superior forecasting skills for courtyard design thermal comfort and utility costs. Soflaei et al. [31] performed a library and field research study to compare the socio-environmental sustainability of traditional courtyards in Iran and China. They derived a series of socio-environmental design concepts which they suggested for application to all scenarios with similar climatic circumstances. However, the research on the huge energy consumption of traditional rural dwellings in China is still dominated by orthogonal experiments and LSA [32,33]. Worse still, the adoption of a single SA method tends to deviate the results, while a multi-category and multi-stage SA method can strengthen the robustness of the conclusions. Therefore, in this paper, a two-stage with three-method sequential SA was carried out for Chinese traditional CSDs to identify the globally significant factors regarding building energy demand. More specifically, a balance between efficiency and accuracy was achieved by first pre-screening a wide range of design parameters with the highly efficient Morris and the SRRC, followed by the second stage of SA on the previous screened results using the even more detailed variance-based Sobol method. This is the first time within our limited horizon that sequential GSA embedded in three methods has been used to obtain robust conclusions for CSDs and applied to design optimization studies. The key procedures of this research are as follows:

- Extract 25 design variables from CSDs in cold regions of China and their data details by field research
- Conduct a sequential SA to recognize and inspect energy-influential design variables
- Set up a reliable prediction model on the energy demand of CSDs
- Obtain and inspect the energy-economy-optimal solutions of CSD key parameters
- Propose a set of GSA-based workflow for CSD design optimization.

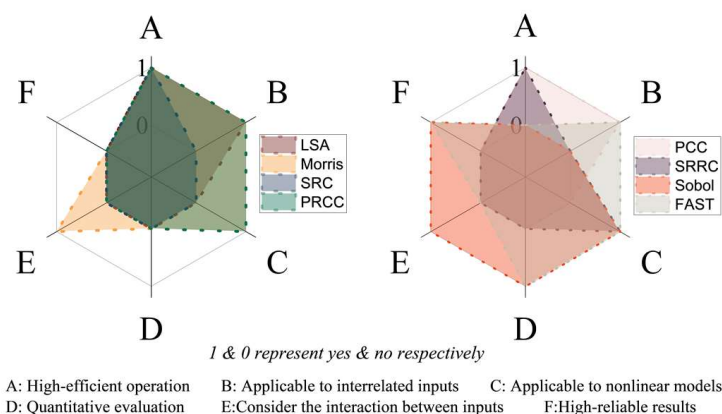


Figure 1. Performance differences and characteristics of each SA method in six aspects.

The paper is organized in the following way: Section 2 presents the sequential SA method with BPA and the route of the whole study. Section 3 describes the data details of the extracted important variables and the established parametric BEM. In Section 4, the findings of the two-stage GSA are compared and discussed in detail. In Section 5, based on the reliable energy prediction model, dual-objective optimization solutions are successively obtained. Finally, the principal research findings, limitations, and recommendations for future work are summarized.

2. Methodology

2.1. SA Method with BPA

The benefit of the SA method with BPA lies in that it is based on the Monte Carlo method to generate pseudo-random values and sample point sets from joint probability distributions, which can substantially conserve the resource consumption of the simulations while ensuring that the outcomes do not have significant bias occurring. To achieve this goal, this paper establishes a link between Simlab and Grasshopper platform, through Honeybee & Ladybug and TT Toolbox to enable the traversal calculation of samples, with all inputs and outputs automatically integrated into Excel and then read by Simlab to complete the GSA. The whole process of the research is shown in Figure 2.

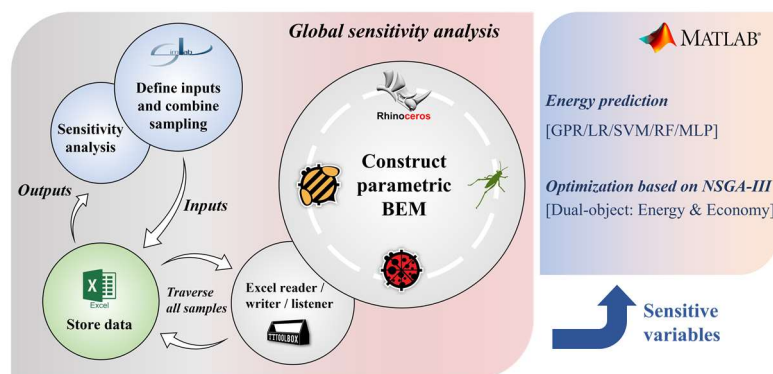


Figure 2. Flow chart of sensitivity analysis and subsequent research.

2.1.1. Step 1: Define Inputs and Combine Sampling

The first step, always the most critical one, tends to have the greatest impact on the final results. To begin with, the details of the design variables, derived from the literature review and field research, were specified in Simlab as in Section 3.1 (containing the data type, interval length, upper and lower thresholds, and distribution type). Then, the type and size of sampling were determined by various SA methods as shown in Table 1, thus completing the sampling of the input arguments. Eventually, the sampled N sets of data were stored in Excel.

Table 1. The sampling type and scale specified by different SA methods in different stages.

SA Method	Sampling Type	Sampling Size	Calculation Method (K as Custom Factor)
SRRC	Latin Hypercube Sampling	250 ($K = 10$)	$K \cdot N$, $K \geq 10$ [34]
Morris	Morris Sampling	216 ($K = 8$)	$K \cdot (N + 1)$, $K = 4, 6, 8 \dots$
Sobol	Sobol Sampling	3328 for cooling (>2600 when $K = 100$) 3328 for heating (>3000 when $K = 100$)	$K \cdot (2N + 2)$, $K = 100, 200, 500 \dots$ [35]

2.1.2. Step 2: Construct Parametric BEM

To conduct an SA, a tunable parametric model is essential. Rhino & Grasshopper were employed to create a prototype CSD following the field research (as Section 3.1). On this

basis, Ladybug & Honeybee were called to set up an energy simulation program to convert the geometric model into a BEM that has thermal properties.

2.1.3. Step 3: Read and Store Data Automatically

GSA requires a substantial volume of samples that would be virtually infeasible to manipulate manually. To automate the data reading and storage, TT Toolbox was called to access Excel, which could read the N sets of the sample data stored in Step 1 and could assign the variables of each set of the sample data to the BEM. The greatest feature of TT Toolbox is the ability to automatically traverse N sets of input arguments, enabling the BEM to perform N calculations. The TT Toolbox then wrote the N sets of inputs and outputs to Excel.

2.1.4. Step 4: Run SA Program

The data stored in Excel were imported into Simlab to identify the sensitivity of cooling and heating demands to each design variable. The option of SA methods is also critical, a wide range of which is provided in Simlab, such as SRC, PCC, SRRC, PRCC, Morris, Sobol, FAST, etc.

2.2. Sequential SA Method

Since different SA methods have different characteristics, adopting multi-sort and multi-stage SA methods provides the benefits of (1) mutual validation to prevent single-method unreliability, (2) a comprehensive conclusion drawn by complementary methods, and (3) balanced accuracy and efficiency. The hybrid SA method requires that the proper methods be specified in a reasonable order. Thus, a two-stage SA was performed to provide more robust conclusions.

2.2.1. Stage 1: Morris and SRRC

In the initial stage, Morris and SRRC were used to preselect the relatively significant ones from a vast set of variables, thereby fixing or eliminating the minor ones. These two methods both have their own strengths in SA. Morris offers high operational efficiency and the ability to account for the interactions between input variables; SRRC is similarly computationally small and applicable to the nonlinear monotonic functions between inputs and outputs. However, neither can exactly quantify the variance of the output for various input elements, and the aim of the pairing is to synthesize and identify which input variables have a major impact on the output uncertainty of the high-dimensional models. Morris has two evaluation indicators: μ is used to assess the primary effect of the input on the output, while σ can be employed to determine the interaction of the parameters with nonlinear response. The primary computational process of Morris can be expressed as [5,20,34]:

$$EE_i = \frac{y(x + e_i \Delta_i) - y(x)}{\Delta_i}, \quad (1)$$

$$\mu_i = \frac{1}{r} \sum_{i=1}^r EE_i, \quad (2)$$

$$\sigma_i = \sqrt{\frac{1}{(r-1)} \sum_{i=1}^r (EE_i - \mu_i)^2}, \quad (3)$$

where EE_i represents the elementary effect to determine the effect of input variations on outputs; $y(x)$ and $y(x + e_i \Delta_i)$ represent the model outputs before and after input variations, respectively; e_i is zero vector; Δ_i is variation value of x_i ; μ_i and σ_i are mean value and standard deviation, respectively.

Regression-based methods have numerous types, all of which establish the correlation between inputs and outputs by modeling the mathematical expressions, but with different application conditions. Therefore, the correlation of the independent variables and the

monotonic nonlinearity of the model need to be determined well in advance before it starts formally working. SRRC was employed in this paper, with its pre-processing discussed in more detail in Section 4.1.

2.2.2. Stage 2: Sobol Indices

In the second stage, the effect of the inputs on the outputs was quantitatively evaluated using the even more robust Sobol, which not only quantifies the output variances due to each of the inputs, but also takes into account the interactions between the variables. Nevertheless, it requires mutually independent inputs and a high sample size to ensure model convergence, dramatically increasing the computational effort. That is the key benefit of setting Stage 1. The Sobol results consist of the first-order effect and total effect indices of the variables. The variance-based SA are premised on decomposing a square integrable mathematical function into [35,36]:

$$Y(x) = Y_0 + \sum_i^k Y_i(x_i) + \sum_{i<j}^k Y_{ij}(x_i, x_j) + \dots + Y_{12\dots k}(x_1, x_2 \dots x_k), \quad (4)$$

where Y_0 is a constant; x_i represents an input.

The total model variance $V(Y)$ can be decomposed into the variance generated by the independent action and the joint effect of parameters [37]:

$$V(Y) = \sum_i V_i + \sum_{i<j} V_{ij} + \dots + V_{12\dots k}, \quad (5)$$

where V_i denotes the effect of the i th input to the output; V_{ij} and $V_{12\dots k}$ are the effect of the interactions between two inputs and k inputs on the output.

Thus, the first-order effect that reflects the contribution of x_i to the total variance of Y can be expressed as S_i , which is an estimate of the expected variance fraction in the model output [37]:

$$S_i = \frac{V_i}{V(Y)}, \quad (6)$$

For both sides of Equation (5) divide simultaneously by $V(Y)$ to be able to find $\sum_i S_i \leq 1$, where $\sum_i S_i = 1$ for linear and additive models:

$$1 = \sum_i S_i + \sum_{i<j} S_{ij} + \dots + S_{12\dots k}, \quad (7)$$

Further, the total effect indices S_{Ti} can be calculated as [38]:

$$S_{Ti} = S_i + S_{ij} + S_{12\dots k} = 1 - \frac{V_{\sim i}}{V(Y)}, \quad (8)$$

3. Case Study

3.1. Control Variables and Parameters

Rural and urban housing share distinctly different patterns in terms of both design and use. Consequently, the design lessons of urban buildings cannot be transferred to rural ones. From nine villages in the cold IIB region of China, 150 units were studied in the field to bring the findings as close to practice as possible. Details on this work can be found in Appendix A. By field research and literature review [31,39–41], 25 variables were identified for SA, where No. 1–13 refer to the spatial parameters while No. 14–25 refer to the envelope parameters. All variables were set as uniformly distributed continuous data. The abbreviated names, intervals, and reference values of each variable follow in Table 2. The reference value comes close to the mean of the actual measured parameters of CSDs.

Table 2. Extracted design variables and their details.

#	Variable	Abbr.	Interval	Ref. Value	Unit
1	EastWingSpacing	EWS	[1.00, 4.00]	2.50	m
2	WestWingSpacing	WWS	[1.00, 4.00]	2.50	m
3	EastWingWide	EWV	[0.00, 4.50]	3.60	m
4	WestWingWide	WVW	[0.00, 4.50]	3.60	m
5	EastWingHeight	EWH	[2.80, 4.00]	3.00	m
6	WestWingHeight	WVH	[2.80, 4.00]	3.00	m
7	CourtyardWallHeight	CWH	[2.80, 4.00]	3.00	m
8	Orientation	ORT	[-45.00, 45.00]	0.00	deg
9	FloorHeight	FLH	[2.80, 4.00]	3.80	m
10	ShadingRatio	SDR	[0.20, 1.20]	0.80	/
11	Width	WDH	[11.00, 17.00]	12.60	m
12	Length	LGH	[4.50, 7.50]	7.00	m
13	RoofSlope	RFS	[15.00, 45.00]	30.00	deg
14	EastWindowWallRatio	EWWR	[0.00, 0.30]	0.15	/
15	SouthWindowWallRatio	SWWR	[0.20, 0.50]	0.40	/
16	WestWindowWallRatio	WVWR	[0.00, 0.30]	0.15	/
17	NorthWindowWallRatio	NVWR	[0.00, 0.30]	0.10	/
18	SolarHeatGainCoefficient	SHGC	[0.20, 0.50]	0.35	/
19	U-valueofWindows	UWD	[1.20, 5.00]	4.00	W/m ² K
20	U-valueofExteriorWall	UEW	[0.20, 2.00]	1.50	W/m ² K
21	U-valueofRoof	URF	[0.20, 1.70]	1.50	W/m ² K
22	U-valueofGround	UGD	[0.40, 3.40]	3.00	W/m ² K
23	AirTightness	ATN	[0.17, 1.00]	0.50	h ⁻¹
24	RoofSolarAbsorptionRate	RSAR	[0.10, 0.90]	0.48	/
25	ExteriorWallSolarAbsorptionRate	WSAR	[0.10, 0.90]	0.48	/

3.2. Parametric BEM

Single-story CSDs of L-shaped and U-shaped layouts were designated for the research, which could be interchanged by adapting the dimensions of the wings. The BEM built with Rhino & Grasshopper was used to estimate the annual heating and cooling demands, with heating and cooling setpoints at 15 °C and 26 °C, respectively. The occupant density was 0.1 person/m², and the occupancy rate as indicated in Figure 3. The lighting power density was set to 4.0 W/m² which remained on between 6:00–8:00 and 18:00–22:00. Other less-used devices dissipating heat were neglected. The calculated U-values of the interior walls and ceilings were 1.72 W/m²K and 1.62 W/m²K, respectively. The U-values of the interior walls and ceilings, roughly calculated by layers and material properties, were 1.72 W/m²K and 1.62 W/m²K, respectively.

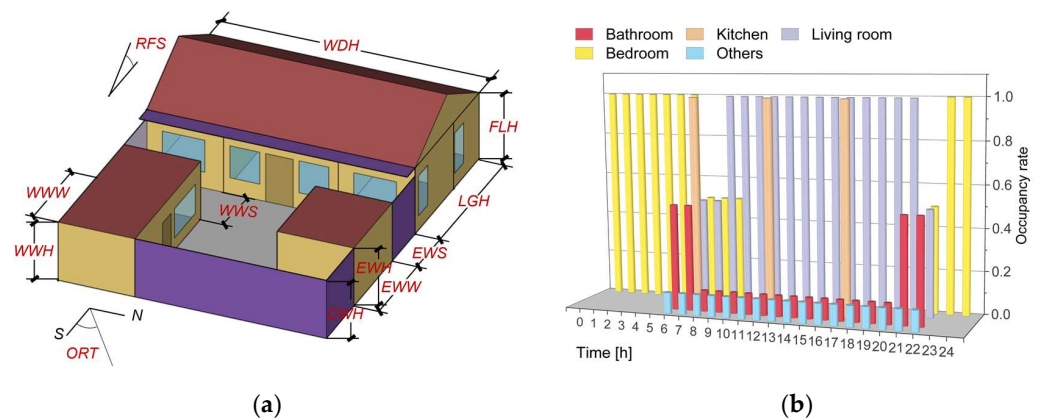


Figure 3. Parametric building energy model with geometric variables and the floor plan of the case: (a) parametric BEM; (b) occupancy rate.

4. Results and Discussion

4.1. Parameters Screening by SRRC & Morris

The parameter correlation and model nonlinearity always dictate the selection of SA methods, especially the regression-based method. Therefore, the monotonicity of the model needs to be determined first to ascertain the validity of the method. A parametric sweep of 25 variables using the OAT approach yields the fluctuation trend of energy demand as shown in Figure 4. Definitively, the outputs maintained monotonically vary to different inputs, without extreme points within the interval, which could provide the basis for the subsequent SA method selection.

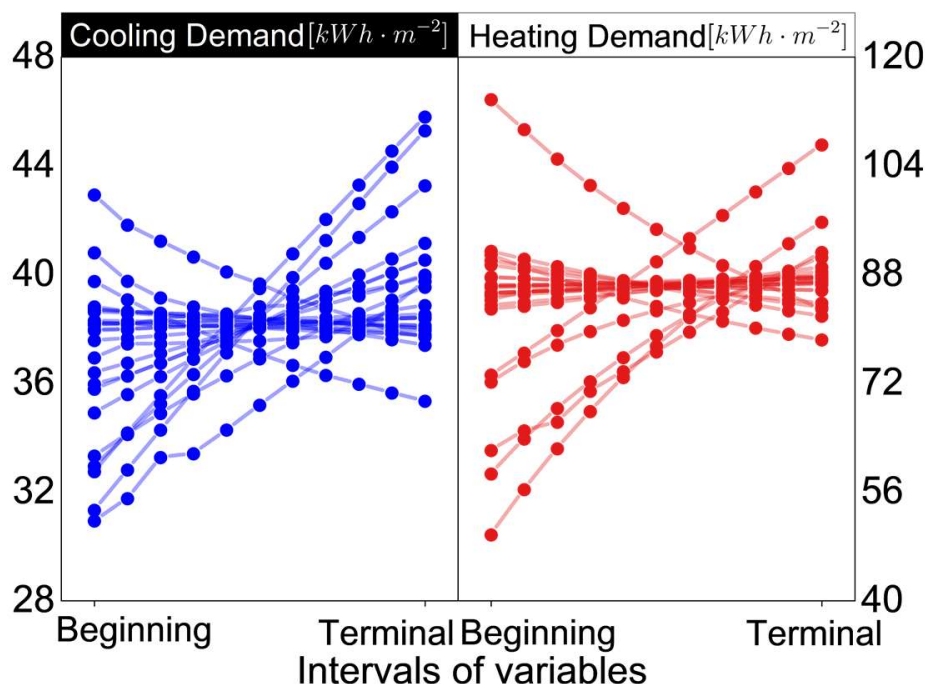


Figure 4. Fluctuation trend of energy demand with 25 variables.

The correlations between the inputs, as shown in Figure 5, were identified with the Pearson correlation coefficient. The entire graph was shaded extremely faintly, indicating that the variables had very weak mutual correlations. The elliptical direction of the upper triangle indicated the positive and negative correlations between the variables, and the bottom triangle gave the precise values. The overwhelming majority of the variables had correlation coefficients with absolute values less than 0.1, while a small minority were also below 0.2. It suggested that the variables were so weakly correlated that they could be regarded as uncorrelated. Thus, for the regression-based method, SRC and SRRC could be initially identified as the applicable SA methods for this study. For further examination of the approach validity, whether the model was linear or not was then discussed, and a linear regression study of 25 variables versus annual cooling and heating demands was conducted. It was evident that the linear correlation of annual energy demand with variables was less robust ($R^2 = 0.51$ for cooling demand; $R^2 = 0.62$ for heating demand), that is, only 51% of the cooling demand data and 62% of the heating demand data could be linearly explained with the inputs. Furthermore, the p -values of cooling and heating demands were 4.89×10^{-25} and 2.19×10^{-44} , respectively, which were both far below 0.001. Therefore, a significant nonlinear correlation could be concluded between the inputs and outputs. To sum up, SRRC and Morris were used as the mutually complementary SA methods at the first stage to screen the significance parameters rationally.

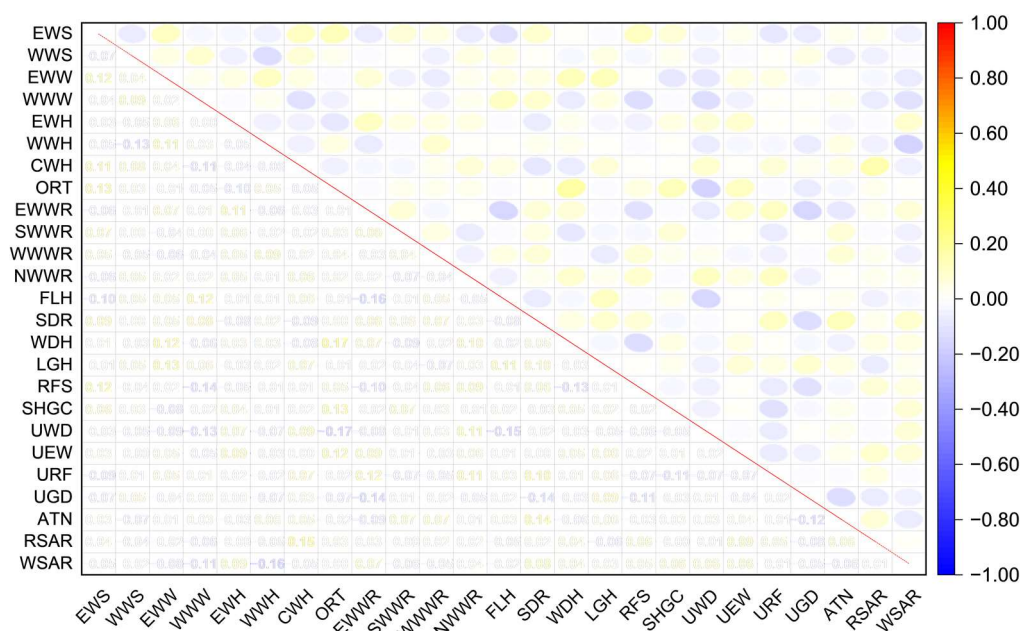


Figure 5. The correlation between 25 variables expressed by Pearson correlation coefficient.

With regard to cooling demand, Figure 6 shows the sensitivity rankings of the variables presented by Morris and SRRC. Surprisingly, the two SA methods exhibited fairly consistent results on this issue. From the 25 parameters, the top 12 highly sensitive parameters were sorted out and used as the inputs for the subsequent SA rounds, namely RSAR, ATN, FLH, WSAR, SHGC, UEW, LGH, NWWW, WWWW, EWWR, SDR, WDH. Despite that, the rankings of the 12 parameters still differed slightly between the two methods. First, the shading ratio was the most sensitive variable in terms of the spatial parameters, except for the width, the depth, and the floor height. It was once speculated that the presence of wings and fences in courtyard houses might interfere with the solar absorption of the principal building and thus the energy demand. However, apparently, the dimensional parameters concerning the courtyard fences and the wings, inclusive of the spacing between the wings and the principal building, were identified as secondary parameters ranked in the bottom ten. The phenomenon most likely arose because these variables even when taken to the boundary values, still could not shade the principal building: in fact, the heights of the wings and fences were less than the floor heights of the principal building and the wings had a distance of 1 m closest to the principal building. Then it also suggested that in certain higher latitudes, the sensitivity rankings of these parameters might shift forward. The impact of south-side window-to-wall ratio was less significant than that of the north-side one due to the sun-shading effect, while the highest influential factor was found on the east–west side. Second, in terms of envelope parameters, the U -value ranked high as expected, but only for those of exterior walls, while those of other parts such as exterior windows, roofs, and floors showed only limited significance. Compared to the U -value, the radiation properties of the envelope had a more significant effect on cooling demand. The solar absorptivity of roofs and façades as well as SHGC ranked in the top five by both methods. The impact of envelope parameters on cooling demand tended to be more significant than that of spatial ones, owing to the pronounced role of solar heat gain on building cooling.

Several of the previous patterns were retained for the heating demand, but there were a few differences. Figure 7 shows the sensitivity ranking of the variables for the heating energy consumption presented by Morris and SRRC. Both methods consistently identified 14 highly sensitive parameters, which, as expected, still had slightly different rankings. The inputs screened for the second SA round consisted of UEW, UWD, ATN, FLH, LGH, URF, SHGC, WDH, WSAR, SDR, NWWW, SWWR, RSAR, RFS. Roof slope was the spatial parameter that had a significant effect on the heating demand in addition to the

four aforementioned sensitive variables. The previous situation was reversed in terms of envelope parameters: the heat transfer coefficient became the primary significant parameter. At this point, the heat transfer coefficients of either exterior walls, exterior windows, or roofs exerted a strong impact on the heating demand. The reason for such a result might be that radiation was relatively weak in winter, while the indoor–outdoor temperature difference was much larger than that in summer, so the heat transfer driven by the temperature difference would be even greater if the thermal resistance of the large area of the envelope was poor. In addition, the airtightness of the building had a remarkable impact on heating demand. Similar to the results for cooling demand, the spatial parameters had a minor impact on energy consumption compared to the envelope ones.

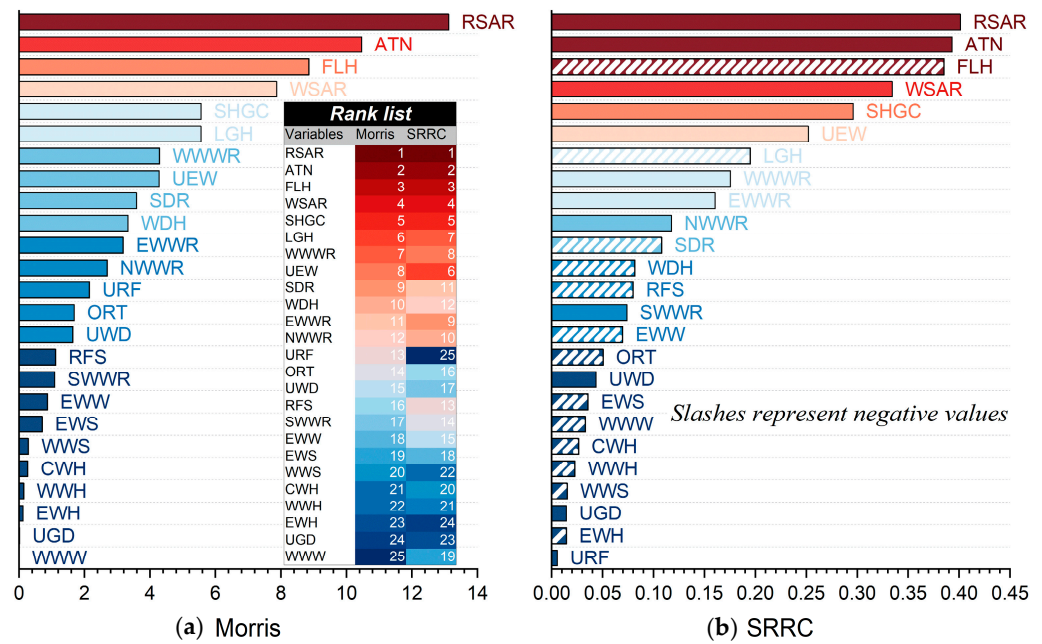


Figure 6. Ranking of sensitivity parameters to cooling demand obtained by the two methods in the first-stage SA: (a) Morris (left); (b) SRRC (right).

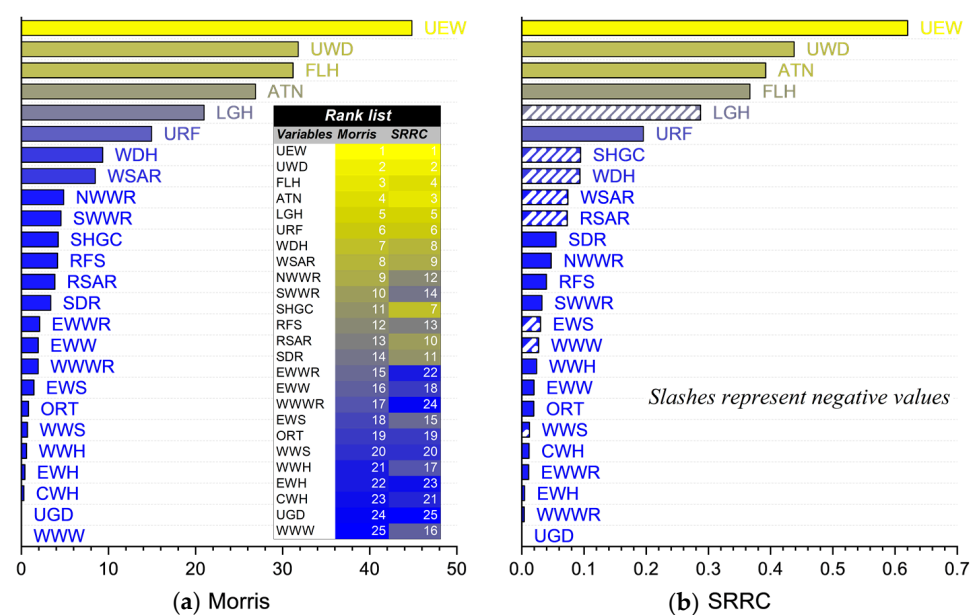


Figure 7. Ranking of sensitivity parameters to heating demand obtained by the two methods in the first-stage SA: (a) SRRC; (b) Morris.

4.2. Quantitative Analysis by Sobol

From the initial SA screening, 12 and 14 significant parameters remained on cooling and heating demands, respectively. Although both Morris and SRRC gave superior sensitivity rankings for these variables in their SA results, the rankings within these variables differed. To rectify this deficiency and to quantify the sensitivities more precisely, Sobol, a variance-based method, was used in the second-stage SA.

Figure 8 shows the distributions of the cooling and heating demands corresponding to the 3328 sample sets, which were attempted to be depicted as lognormal vs. normal. The cooling demand was positively skewed, and its values were concentrated in the interval 27.4, 45.3. At this point, the geometric mean μ_g was 34.9 ± 0.2 kWh/m² ranging from 21.7 kWh/m² to 64.6 kWh/m², which indicated the significant effect of these 12 parameters on cooling demand. Likewise, the heating demand also showed a positive skewed distribution whose values concentrated at the interval 30.3, 87.1. At this point, the geometric mean μ_g was 53.1 ± 1.2 kWh/m² ranging from 12.2 kWh/m² to 135.9 kWh/m², suggesting that the effect of these 14 parameters on the heating demand was considerable. Then, the normal Q-Q plot revealed that the cooling and heating demands could meet the normal distribution. The mean values of cooling and heating demands μ were 37.1 kWh/m² and 62.5 kWh/m² respectively at that time, with the standard deviation σ of 6.9 and 21.1, respectively.

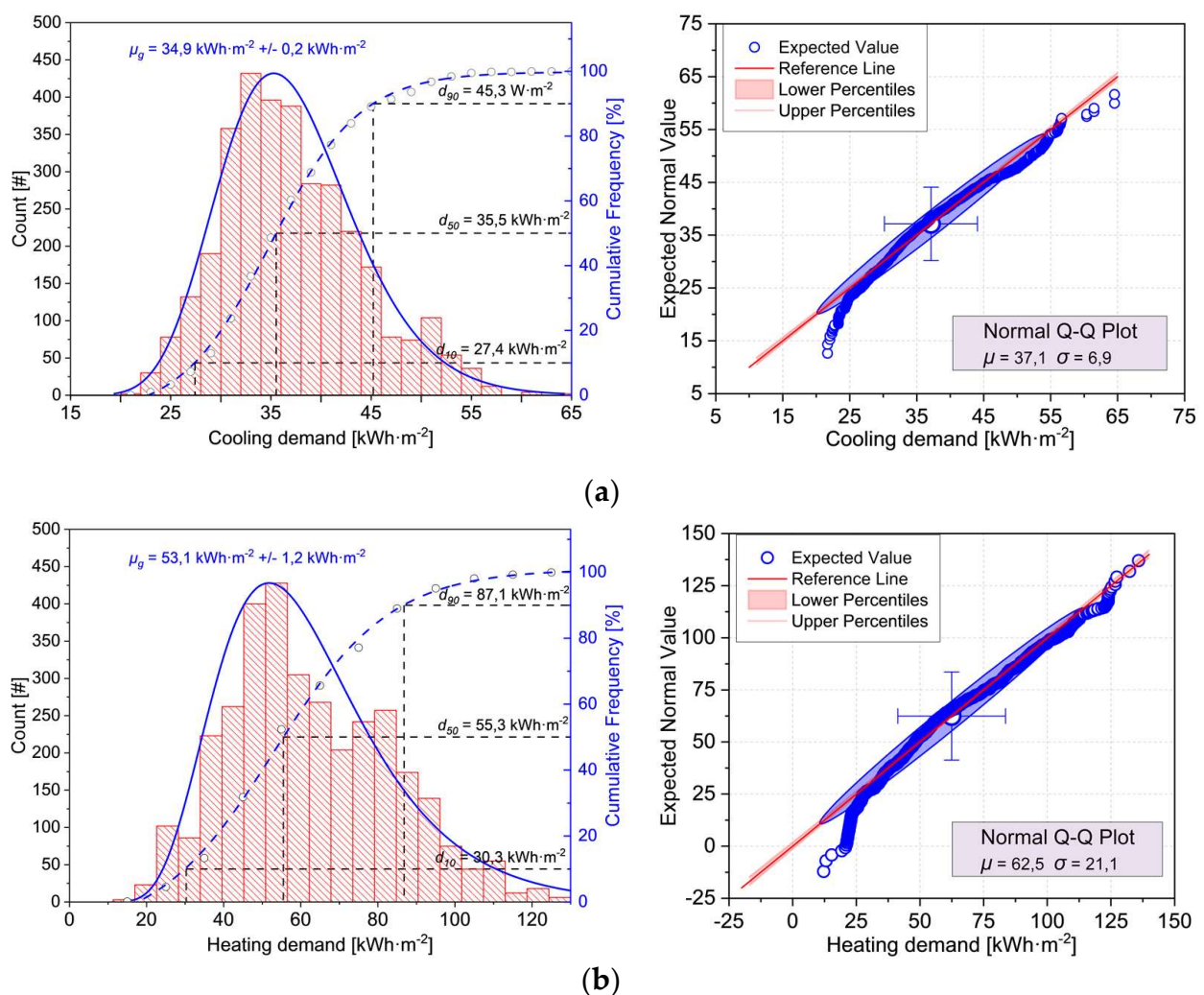


Figure 8. Frequency distribution of cooling and heating demands and normal Q-Q plot: (a) cooling demand; (b) heating demand.

Figure 9 indicates the first-order effect indices (S_1) and the total effect indices (S_t) obtained by Sobol. Among the parameters that had a significant effect on the annual cooling demand, those concerning the solar absorptivity of the envelope occupied three of the top five. The solar absorptivity of roofs and facades as well as SHGC could induce variations in cooling demand of 23.2%, 17.4%, and 11.7% (52.3% in total), respectively. Meanwhile, the airtightness, the floor height, and the heat transfer coefficient of the exterior wall were also parameters not to be ignored, which were capable of causing 17.4%, 16.7%, and 8.1% (42.2% in total) variations in cooling energy consumption independently. The sensitivity of the cooling demand to the solar absorptivity of the facade increased by 4% when the interaction of other parameters was considered. In addition, the floor height, the U -value of exterior walls, and the airtightness were also significantly affected by the interaction, while the other parameters showed little difference. Out of the significant parameters affecting the annual heating demand, the U -value and airtightness of the envelope were the most prominent. The U -value of exterior walls and windows as well as the airtightness could induce 32.5%, 21.1%, and 18.1% (71.7% in total) variations in heating demand separately. Similar to the cooling demand, the floor height had considerable sensitivity to heating demand, which acted alone to induce a 13.3% variation in heat demand. Regarding the parameter interaction, except for the heat transfer coefficient of the exterior windows and the floor height, the S_1 and S_t of the remaining parameters were quite similar, that is, the variability of the parameter sensitivity due to the interaction was not considerable.

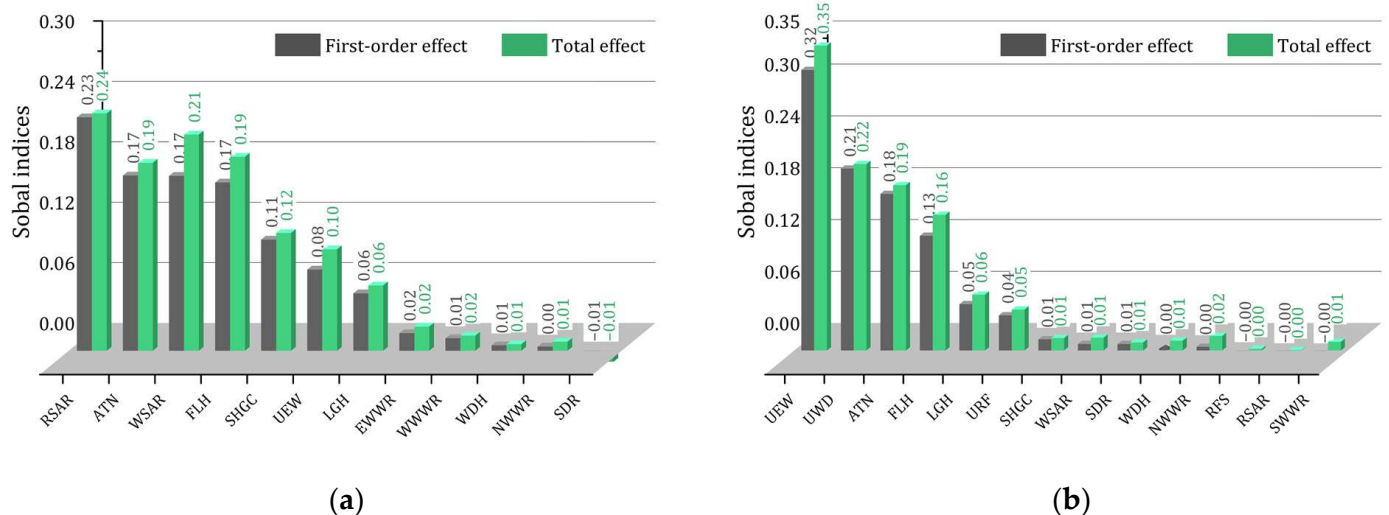


Figure 9. First-order effect indices (S_1) and total effect indices (S_t) of sensitivity parameters to energy demand obtained by the Sobol method in the second-stage SA: (a) cooling demand; (b) heating demand.

The Sobol method had been proven to be the most accurate of the three methods in previous studies [38,42]. In this study, it was also easy to find the conclusion differences between the three methods given, as shown in Table 3. Out of the 12 significant variables for cooling demand, the three methods consistently provided the top five parameters, even if their rankings differed somewhat. The sensitivity ranking of the parameters was essentially the same for SRRC compared to S_1 , where their maximum ranking difference did not exceed two. Since the interaction between these 12 parameters was not very prominent, similar findings could be found in the comparison of SRRC vs. S_t , while Morris yielded a three-position difference in rankings compared to Sobol. As can be seen, although Morris could effectively screen most of the significant parameters, its ranking of parameter sensitivity was not yet quite robust, for which one potential solution was to expand the sample size. Among the 14 significant variables for heating demand, the three methods were in general agreement in determining the top six parameters in the sensitivity ranking. For these six parameters, SRRC agreed exactly with the ranking given by Sobol, while Morris presented the different one for ATN and FLH. For the latter eight parameters in the

sensitivity ranking, SRRC was relatively close to the conclusion reached by S_1 , while Morris fully underestimated the significant effect of SHGC on heating demand in the process. In general, the ranking discrepancy between SRRC, Morris, S_1 , and S_t could be attributed to two main aspects: (1) the mis-ranking of the not significant parameters might be because their sensitivity indices were so small that the absolute differences were negligible; (2) the ranking discrepancy between S_t and the others might be due to the fact that the sensitivity indices themselves were small enough that although the interaction was not obvious, it was sufficient to shift the sensitivity ranking of those parameters.

Table 3. The difference comparison of parameter ranking of energy demand given by various methods in the sequential SA process.

Var.	Cooling Demand				Heating Demand				
	Morris	SRRC	S_1	S_t	Var.	Morris	SRRC	S_1	S_t
RSAR	1	1	1	1	UEW	1	1	1	1
WSAR	4	4	3	2	UWD	2	2	2	2
FLH	3	3	4	3	ATN	4	3	3	3
ATN	2	2	2	4	FLH	3	4	4	4
SHGC	5	5	5	5	LGH	5	5	5	5
UEW	8	6	6	6	URF	6	6	6	6
LGH	6	7	7	7	NWWR	9	12	11	7
EWWR	11	9	8	8	WSAR	8	9	8	8
WWWR	7	8	9	9	SHGC	11	7	7	9
NWWR	12	10	11	10	WDH	7	8	10	10
WDH	10	12	10	11	SWWR	10	14	14	11
SDR	9	11	12	12	SDR	14	11	9	12
		/			RFS	12	13	12	13
					RSAR	13	10	13	14

5. Energy Prediction and Optimization Study

5.1. Prediction Model Description

Building energy simulations tend to require a vast number of complex inputs, requiring substantial effort to access and difficult to control uncertainty. SA helps the designer to exclude a large amount of secondary information and thus obtain a highly accurate output by the saliency parameters alone, which supports the design aspects that need to be adjusted iteratively to optimize the building performance. Here in this section, the prediction model of cooling and heating demands was developed based on the significant parameters identified in the previous study. Several assumptions should be stated in advance to ensure the model validity: (1) the experimental building form could be approximately described by the established model as in Figure 3; (2) the parameter ranges of the experimental building could be contained in or approximated to those of the established model; (3) the fixed parameters in the model building, such as heating and cooling setpoints, user schedules, etc., could describe closely the actual situation; (4) the meteorological conditions should be consistent with or similar to those of the established model.

Gaussian Process Regression (GPR) was used to train prediction models, which were typically appropriate for addressing nonlinearities. Since the GPR achieved prediction by the very process of probabilistic inference, it was capable of quantifying the prediction uncertainty in a principled way. However, the drawback of GPR was that as a non-parametric model, its computational complexity surged with the volume of data; specifically, for N samples, it would have a complexity of $O(N^3)$, exactly the problem solved by the previous SA. Likewise, Support Vector Machine (SVM), Random Forest (RF), and Multi-layered Perceptron (MLP) were all recognized as excellent prediction algorithms; thus they were used to verify the training precision. In addition, Linear Regression (LR) was employed to compare the results with the machine learning algorithms.

5.2. Model Performance Evaluation

During the model training, the significant parameters with S_t larger than 0.04 were retained, namely RSAR, WSAR, FLH, ATN, SHGC, UEW, LGH, UWD, and URF. The GPR model was trained and tested with 10 cross-validation folds based on a set of 7122 samples. Several factors might affect the accuracy of GPR models, such as basis function, kernel function, kernel scale, sigma, etc., which gave difficulty in specifying artificially the optimal or near-optimal values. Thus, Bayesian optimization was employed to identify the specific values of the hyperparameters so as to obtain a highly accurate GPR model. At this point, the acquisition function was set as expected improvement per second plus and the signal standard deviations of the energy demand models were predefined as 4.91 and 14.95, respectively. Table 4 shows the optimized hyperparameters of the GPR model as well as their search range. Eventually, the optimization process for both the cooling demand and heating demand models went through 30 iterations. Figure 10 shows the optimizable GPR models and their prediction results for energy. From the results, it could be seen that the accuracy of the model was significantly improved by optimization, with good consistency between the predicted and actual values, and the residuals could be kept within an acceptable range.

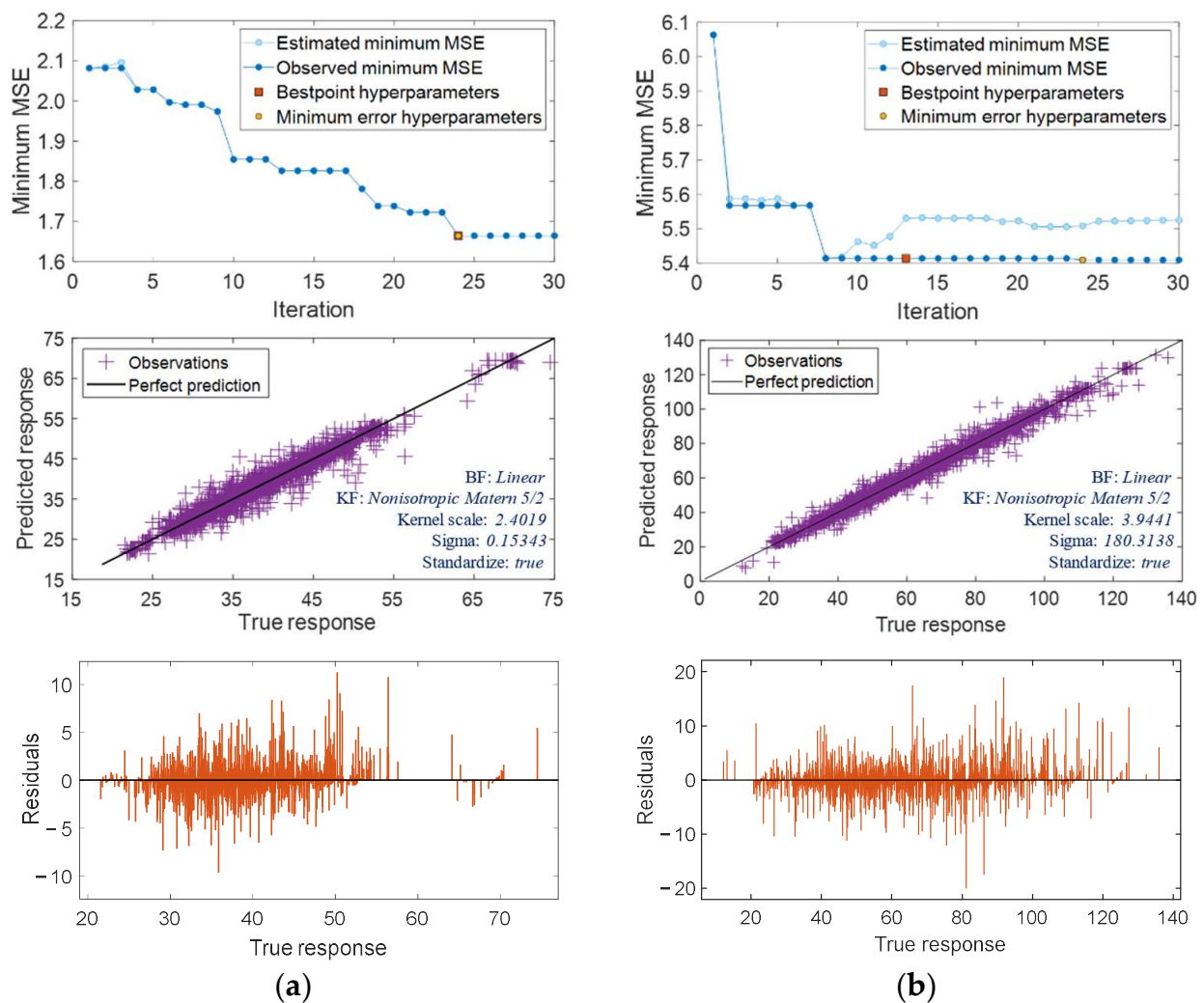


Figure 10. Energy prediction model trained by optimizable GPR with 10 cross-validation folds through 30 iterations: (a) cooling demand; (b) heating demand.

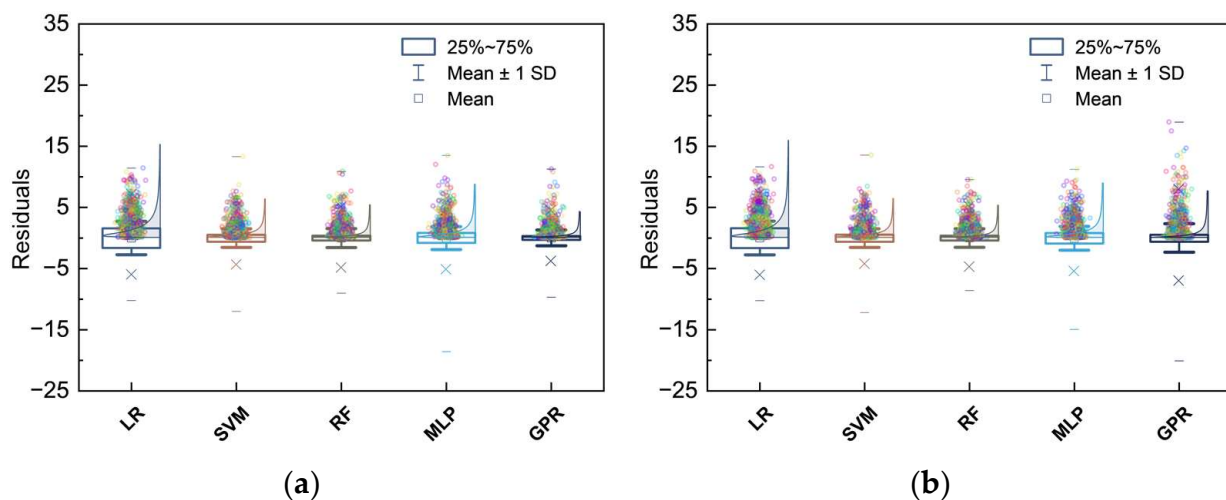
Table 4. Hyperparameter search range of optimizable GPR model training.

#	Hyperparameter	Search Range
1	Sigma	0.0001~69.38 for cooling demand, 0.0001~211.45 for heating demand
2	Basis function	Constant, Zero, Linear
3	Kernel function	Nonisotropic Exponential, Nonisotropic Matern 3/2, Nonisotropic Matern 5/2, Nonisotropic Rational Quadratic, Nonisotropic Squared Exponential, Isotropic Exponential, Isotropic Matern 3/2, Isotropic Matern 5/2, Isotropic Rational Quadratic, Isotropic Squared Exponential
4	Kernel scale	0.0029062~2.9062 for cooling, 0.00465~4.65 for heating
5	Standardize	true, false

A comparative experiment was conducted using linear regression and machine learning algorithms (including SVM, RF, and MLP) to further evaluate the prediction accuracy of the GPR model. The specific configurations of these algorithms are shown in Table 5, and the performance of all these algorithms was compared with 10-fold cross-validation. Figure 11 is a box plot comparing the prediction residuals of the models derived from GPR vs. LR and three machine learning algorithms. As we can see, the residuals of all five models were well controlled, with the absolute value of the residuals distributed below three for the vast majority of the data. However, whether for predicting cooling or heating demand, the GPR model remained the best performance among these models. For cooling demand prediction, the RMSEs of each model were ranked as LR (2.7564) > MLP (1.9426) > SVM (1.5231) > RF (1.5109) > GPR (1.2901). Similarly, in heating demand, the RMSE ranking of the models was LR (4.8936) > MLP (3.5804) > RF (3.3442) > SVM (2.8163) > GPR (2.3259).

Table 5. Configurations of the regression models concerned.

#	Model Type	Configurations
1	Linear regression	Terms = 'Interactions, Robust option = 'On' Kernel function = 'Gaussian', Kernel scale = 2.6, Box constraint = 6.601, Epsilon = 0.6601, Standardize data = 'On' (for cooling)
2	Support vector machines	Kernel function = 'Gaussian', Kernel scale = 2.4, Box constraint = 22.72, Epsilon = 22.72, Standardize data = 'On' (for heating)
3	Random forest regression	Minimum leaf size = 8, No. of learners = 30 No. of fully connected layers = 3, Each layer size = 10,
4	Multi-layered perceptron	Activation = 'ReLU', Iteration limit = 1000, Regularization strength = 0; Standardize data = 'Yes'

**Figure 11.** The comparison on the accuracy of 10-fold cross validation using linear regression and machine learning: (a) cooling demand; (b) heating demand.

5.3. Optimization

In this part, optimal design was used as an instance of the advanced usage of prediction models. Naturally, the main purpose of this paper would be to provide decision support for the design of CSDs, thus the optimization was centered on minimizing the total annual energy demand with the above screened sensitive parameters. On this basis, construction cost was introduced as the second objective, which in turn led to a dual-objective optimization for the nine parameters. Here, the construction cost was assumed to conform to a linear function, fitted by the collected data during field research. The optimization process was performed based on the Nondominated Sorting Genetic Algorithm III (NSGA-III). The crossover fraction was set to 0.8, and the population size and stall generation limit were set to 200 and 500, respectively. Figure 12 illustrates the final optimization result. The random data shown in Figure 12 were the output solutions corresponding to the random matrix generated with the rand function following the interval of nine variables for verifying the reliability of the Pareto front.

After the dual-objective optimization, 80 sets of optimal combinations regarding the nine variables were obtained. To comprehensively evaluate and compare their technical economy (TE), the definition of R_B was introduced as:

$$R_B = \begin{cases} \frac{E_{avg} - E_{opt}}{C_{opt} - C_{avg}} & , E_{avg} > E_{opt} \\ 0 & , E_{avg} < E_{opt} \end{cases} \tag{9}$$

where E_{avg} and E_{opt} represent the mean energy demand of the random matrix output and the energy demand of the Pareto front, respectively; Similarly, C_{avg} and C_{opt} are the mean cost of the random matrix output and the cost of the Pareto front, respectively.

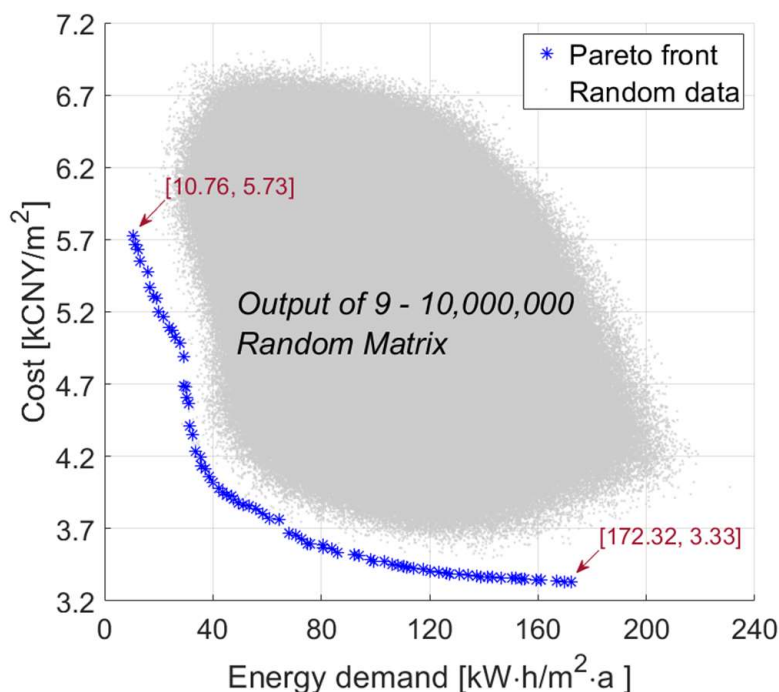


Figure 12. Pareto front generated by NSGA-III dual-objective optimization and target distribution generated by 9–10,000,000 random matrix.

When using R_B for evaluation, the larger its value indicates that the solution would yield better energy investment benefits, and solutions without energy conservation effect were excluded as zero. Figure 13 gives a comparative evaluation of the 80 optimal parameter sets. From the selected optimal TE solutions, it could be found that most of the variables were not completely biased towards the interval boundaries. The bias of the optimal variables

towards the interval boundary might be explained in two reasonable ways: (1) the additional investment in the variable exceeded the improvement in energy efficiency by a large margin; (2) the same energy objective could be achieved more efficiently by adding investments to other variables. Therefore, it could be believed that the specific state of the optimal solution depended heavily on the mutual response between the two objective functions, which indicated that the uncertainty of the objective function would significantly cut the generalizability of the optimal solution. Nevertheless, the resulting Pareto front could still preserve its learning value for design decisions.

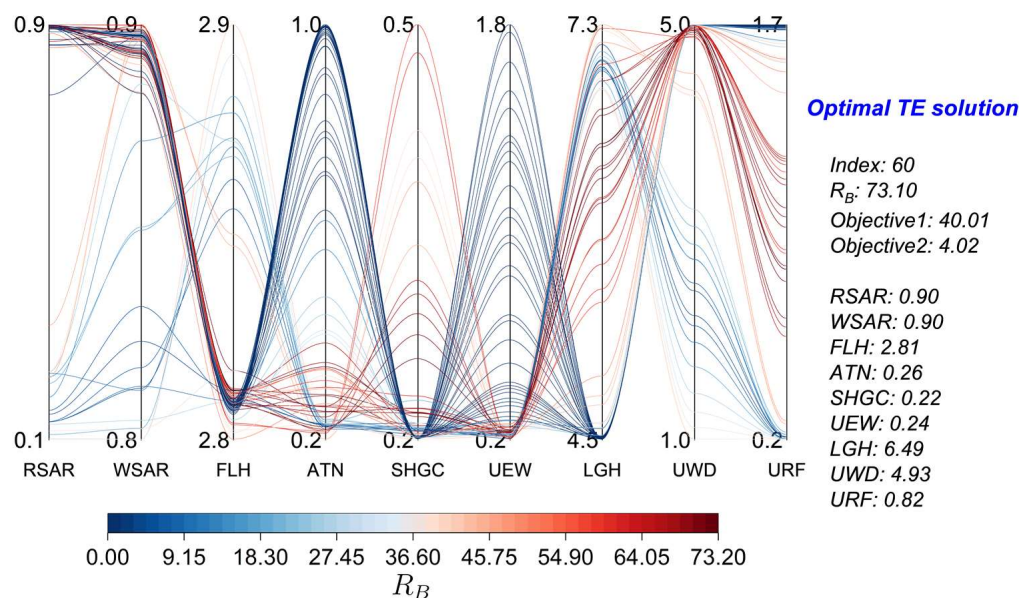


Figure 13. Evaluation and comparison of the technique-economy of 80 optimal parametric combinations by using R_B .

6. Conclusions

The present research set out to organize a GSA-based workflow for the optimal design of CSDs in cold regions of China. This workflow comprised several fragmentation processes such as UA, GSA, energy prediction model regression, and multi-objective optimization. First, the principles and characteristics of the four prevailing SA methods were reviewed and identified. Second, the uncertainty of energy demand to parameters and the sensitivity ranking of parameters were acquired by the sequential SA with Morris, SRRC, and Sobol. Third, a GPR model capable of predicting the annual energy demand was trained and outperformed in comparison with the accuracy of LR, SVM, RF, and MLP. Fourth, based on the obtained prediction models, an optimal design of the case building with a dual objective of energy and economy was accomplished by using NSGA-III. These efforts provide a relatively complete and robust strategy to support the CSD design, screening nine significant parameters on energy demand to aid qualitative optimization, as well as a set of 80 techno-economically sound Pareto front solutions for quantitative strategies. Further, the GSA-based optimization workflow ties the above fragmented studies into a logical sequence, offering a framework for further depth and richness in each area of research.

One limitation of this study lies in the insufficient sample size, which may make these findings less generalizable to CSDs in each region. For practical analytical reasons, it is quite challenging to retrieve the perfect parameter intervals and to fit the full range of cost data, which will lead to a certain bias of the final Pareto front. To remedy this deficiency, the research samples were persistently expanded to cover as much information as possible, even though this entails a substantial cost, which in fact will still take quite a bit of time for the vast size of China. Despite its limitations, the study certainly adds to our understanding of the problems of energy-saving design strategy support of CSDs,

which would be a fruitful area for further work. Thus, greater efforts might be needed to expand the scope of application as follows: (a) Developing a plug-in for real-time parameter adjustment—a python-based framework is under development to address the dependency of multiple fragmented studies of this workflow on their respective software platforms, and in the near future, the workflow could be easily available to architects as a plug-in of grasshopper; (b) Refining more efficiency–accuracy balanced sequential SA methods—the hybrid SA approach can effectively integrate the merits of various SA methods, of which the sequential SA method with Morris/SRRC and Sobol proposed in this paper is just one, and numerous other possible rational hybrid methods deserve to be explored; (c) Establishing a set of constantly improving rural housing databases—our team is persisting with extensive field research and statistics on rural building data in cold regions of China, covering the construction historical information, building prototype clustering, occupant behavior data, construction cost statistics, etc. This work serves as a data basis for the replication of more generalized findings.

Author Contributions: Conceptualization, J.G. and Y.J.; methodology, Y.J.; software, C.S. and Y.J.; formal analysis, Y.J. and M.L.; data curation, C.S. and Z.W.; writing—original draft preparation, Y.J.; writing—review and editing, J.G. and Y.J.; supervision, J.G.; funding acquisition, J.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Key R & D Program of China, Grant Number 2019YFD1101004.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

BPA	Building performance analysis
SA	Sensitivity analysis
LSA	Local sensitivity analysis
GSA	Global sensitivity analysis
OAT	One-parameter-at-a-time
BEM	Building energy model
SHGC	Solar heat gain coefficient
SRC	Standard Regression Coefficients
PCC	Partial Correlation Coefficients
SRRC	Standardized Rank Regression Coefficients
PRCC	Partial Rank Correlation Coefficients
UA	Uncertainty analysis
CSD	Courtyard-style dwelling
S_1	First-order effect indices
S_t	Total effect indices
GPR	Gaussian Process Regression
SVM	Support Vector Machine
RF	Random Forest
MLP	Multi-layered Perceptron
LR	Linear Regression
NSGA-III	Nondominated Sorting Genetic Algorithm III
TE	Technical economy

Appendix A

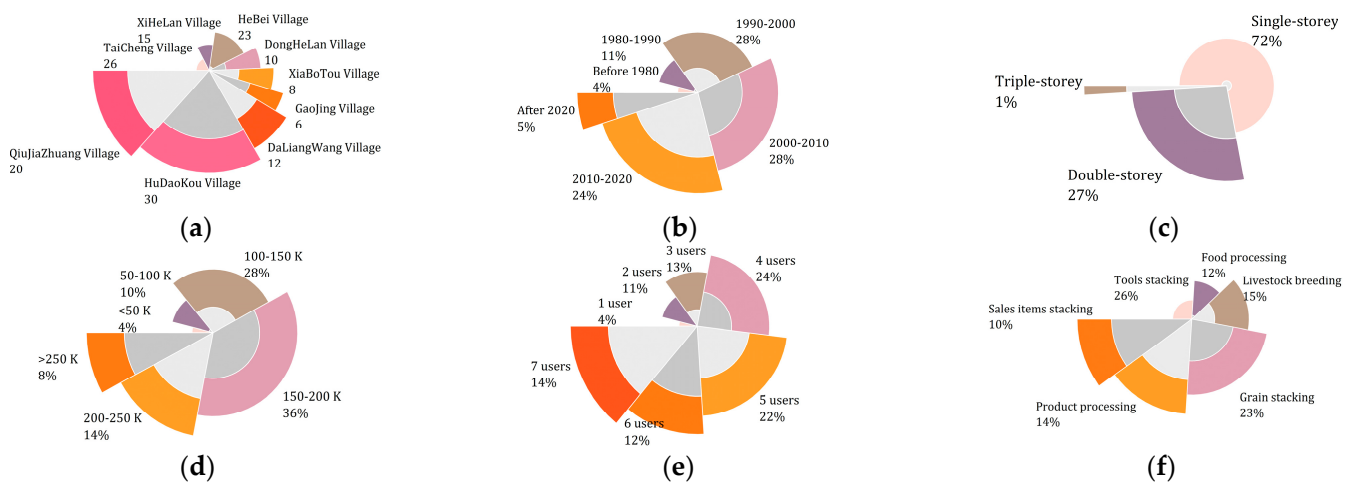


Figure A1. Basic characteristics of field research samples: (a) building location; (b) construction age; (c) floor count; (d) total cost; (e) user count; (f) production mode.

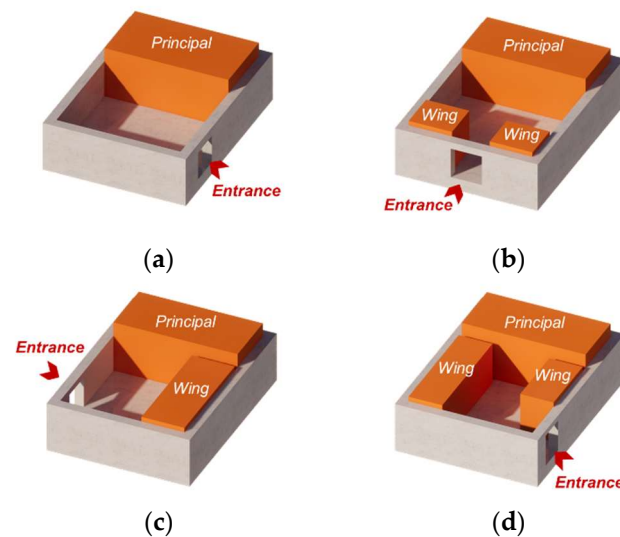


Figure A2. Common courtyard layout of CSDs in the survey: (a) I-shaped; (b) II-shaped; (c) L-shaped; (d) U-shaped.

Table A1. Common enclosure structure form of rural houses in the survey.

Type	Main Material	Construction Layer (Inside Out)		
		Layer 1	Layer 2	Layer 3
Exterior wall	Clay brick	Gypsum/PVC	Main structure (Without insulation)	Mortar/Tile/Vacant
	Hollow block	Gypsum/PVC		Mortar/Tile/Vacant
	Adobe	Wrapped clay brick/ Choi steel/Vacant		Gypsum
Roof	Hollow slab/	Main structure (Without insulation)	Cinder lime	Tile/Vacant
	Concrete slab		Tile/Choi steel	/
Window	Single glass	Curtain/Vacant	Wood/PVC/Aluminum alloy/Bridge-cut-off aluminum alloy	Plastic cloth/Vacant
	Double glass			
Door	Wood/Stainless steel /Aluminum alloy	Main structure	Curtain	/

References

1. Hensen, J.; Lamberts, R. *Building Performance Simulation for Design and Operation*; Taylor & Francis: Abingdon, UK, 2011; ISBN 978-0-415-47414-6.
2. Saltelli, A. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*; Wiley: Hoboken, NJ, USA, 2004; ISBN 978-0-470-87093-8.
3. Iooss, B.; Lemaître, P. A Review on Global Sensitivity Analysis Methods. *arXiv* **2014**, arXiv:1404.2405.
4. Hamby, D.M. A Review of Techniques for Parameter Sensitivity Analysis of Environmental Models. *Environ. Monit. Assess.* **1994**, *32*, 135–154. [[CrossRef](#)]
5. Kristensen, M.H.; Petersen, S. Choosing the Appropriate Sensitivity Analysis Method for Building Energy Model-Based Investigations. *Energy Build.* **2016**, *130*, 166–176. [[CrossRef](#)]
6. Xu, C.; Gertner, G. Understanding and Comparisons of Different Sampling Approaches for the Fourier Amplitudes Sensitivity Test (FAST). *Comput. Stat. Data Anal.* **2011**, *55*, 184–198. [[CrossRef](#)] [[PubMed](#)]
7. Goffart, J.; Woloszyn, M. EASI RBD-FAST: An Efficient Method of Global Sensitivity Analysis for Present and Future Challenges in Building Performance Simulation. *J. Build. Eng.* **2021**, *43*, 103129. [[CrossRef](#)]
8. Neale, J.; Shamsi, M.H.; Mangina, E.; Finn, D.; O'Donnell, J. Accurate Identification of Influential Building Parameters through an Integration of Global Sensitivity and Feature Selection Techniques. *Appl. Energy* **2022**, *315*, 118956. [[CrossRef](#)]
9. Paleari, L.; Movedi, E.; Zoli, M.; Burato, A.; Cecconi, I.; Errahouly, J.; Pecollo, E.; Sorvillo, C.; Confalonieri, R. Sensitivity Analysis Using Morris: Just Screening or an Effective Ranking Method? *Ecol. Model.* **2021**, *455*, 109648. [[CrossRef](#)]
10. Wang, C.; Peng, M.; Xia, G. Sensitivity Analysis Based on Morris Method of Passive System Performance under Ocean Conditions. *Ann. Nucl. Energy* **2020**, *137*, 107067. [[CrossRef](#)]
11. King, D.M.; Perera, B.J.C. Morris Method of Sensitivity Analysis Applied to Assess the Importance of Input Variables on Urban Water Supply Yield—A Case Study. *J. Hydrol.* **2013**, *477*, 17–32. [[CrossRef](#)]
12. Zhao, J.; Zhang, J.J.; Grunewald, J.; Feng, S. A Probabilistic-Based Method to Evaluate Hygrothermal Performance of an Internally Insulated Brick Wall. *Build. Simul.* **2021**, *14*, 283–299. [[CrossRef](#)]
13. Tøndel, K.; Vik, J.O.; Martens, H.; Indahl, U.G.; Smith, N.; Omholt, S.W. Hierarchical Multivariate Regression-Based Sensitivity Analysis Reveals Complex Parameter Interaction Patterns in Dynamic Models. *Chemom. Intell. Lab. Syst.* **2013**, *120*, 25–41. [[CrossRef](#)]
14. Haahtela, T.J. Regression Sensitivity Analysis for Cash Flow Simulation Based Real Option Valuation. *Procedia-Soc. Behav. Sci.* **2010**, *2*, 7670–7671. [[CrossRef](#)]
15. Li, X.; Li, L.; Yang, Y.; Zhao, G.; He, N.; Schmidt, E. Variance-Based Sensitivity Analysis for the Influence of Residual Stress on Machining Deformation. *J. Manuf. Process.* **2021**, *68*, 1072–1085. [[CrossRef](#)]
16. Pohya, A.A.; Wicke, K.; Kilian, T. Introducing Variance-Based Global Sensitivity Analysis for Uncertainty Enabled Operational and Economic Aircraft Technology Assessment. *Aerosp. Sci. Technol.* **2022**, *122*, 107441. [[CrossRef](#)]
17. Chen, X.; Yang, H.; Sun, K. Developing a Meta-Model for Sensitivity Analyses and Prediction of Building Performance for Passively Designed High-Rise Residential Buildings. *Appl. Energy* **2017**, *194*, 422–439. [[CrossRef](#)]
18. Yun, W.; Lu, Z.; He, P.; Jiang, X.; Dai, Y. Parameter Global Reliability Sensitivity Analysis with Meta-Models: A Probability Estimation-Driven Approach. *Aerosp. Sci. Technol.* **2020**, *106*, 106040. [[CrossRef](#)]
19. Zhao, Y.; Guo, Z.; Niu, F.; Yu, Y.; Wang, S. Global Sensitivity Analysis of Passive Safety Systems of FHR by Using Meta-Modeling and Sampling Methods. *Prog. Nucl. Energy* **2019**, *115*, 30–41. [[CrossRef](#)]
20. Heiselberg, P.; Brohus, H.; Hesselholt, A.; Rasmussen, H.; Seinre, E.; Thomas, S. Application of Sensitivity Analysis in Design of Sustainable Buildings. *Renew. Energy* **2009**, *34*, 2030–2036. [[CrossRef](#)]
21. Maučec, D.; Premrov, M.; Leskovar, V.Ž. Use of Sensitivity Analysis for a Determination of Dominant Design Parameters Affecting Energy Efficiency of Timber Buildings in Different Climates. *Energy Sustain. Dev.* **2021**, *63*, 86–102. [[CrossRef](#)]
22. Tian, W.; de Wilde, P. Uncertainty and Sensitivity Analysis of Building Performance Using Probabilistic Climate Projections: A UK Case Study. *Autom. Constr.* **2011**, *20*, 1096–1109. [[CrossRef](#)]
23. Yıldız, Y.; Arsan, Z.D. Identification of the Building Parameters That Influence Heating and Cooling Energy Loads for Apartment Buildings in Hot-Humid Climates. *Energy* **2011**, *36*, 4287–4296. [[CrossRef](#)]
24. Spitz, C.; Mora, L.; Wurtz, E.; Jay, A. Practical Application of Uncertainty Analysis and Sensitivity Analysis on an Experimental House. *Energy Build.* **2012**, *55*, 459–470. [[CrossRef](#)]
25. Shen, H.; Tzempelikos, A. Sensitivity Analysis on Daylighting and Energy Performance of Perimeter Offices with Automated Shading. *Build. Environ.* **2013**, *59*, 303–314. [[CrossRef](#)]
26. Pang, Z.; O'Neill, Z.; Li, Y.; Niu, F. The Role of Sensitivity Analysis in the Building Performance Analysis: A Critical Review. *Energy Build.* **2020**, *209*, 109659. [[CrossRef](#)]
27. Pang, Z.; O'Neill, Z. Uncertainty Quantification and Sensitivity Analysis of the Domestic Hot Water Usage in Hotels. *Appl. Energy* **2018**, *232*, 424–442. [[CrossRef](#)]
28. Østergård, T.; Jensen, R.L.; Maagaard, S.E. A Comparison of Six Metamodeling Techniques Applied to Building Performance Simulations. *Appl. Energy* **2018**, *211*, 89–103. [[CrossRef](#)]
29. Sun, H.; Leng, M. Analysis on Building Energy Performance of Tibetan Traditional Dwelling in Cold Rural Area of Gannan. *Energy Build.* **2015**, *96*, 251–260. [[CrossRef](#)]

30. Tabadkani, A.; Aghasizadeh, S.; Banihashemi, S.; Hajirasouli, A. Courtyard Design Impact on Indoor Thermal Comfort and Utility Costs for Residential Households: Comparative Analysis and Deep-Learning Predictive Model. *Front. Archit. Res.* **2022**, *1*–18. [[CrossRef](#)]
31. Soflaei, F.; Shokouhian, M.; Zhu, W. Socio-Environmental Sustainability in Traditional Courtyard Houses of Iran and China. *Renew. Sustain. Energy Rev.* **2017**, *69*, 1147–1169. [[CrossRef](#)]
32. Wang, F.; Liu, Y. Thermal Environment of the Courtyard Style Cave Dwelling in Winter. *Energy Build.* **2002**, *34*, 985–1001. [[CrossRef](#)]
33. Deng, Q.; Wang, G.; Wang, Y.; Zhou, H.; Ma, L. A Quantitative Analysis of the Impact of Residential Cluster Layout on Building Heating Energy Consumption in Cold IIB Regions of China. *Energy Build.* **2021**, *253*, 111515. [[CrossRef](#)]
34. Loepky, J.L.; Sacks, J.; Welch, W.J. Choosing the Sample Size of a Computer Experiment: A Practical Guide. *Technometrics* **2009**, *51*, 366–376. [[CrossRef](#)]
35. Sobol, I.M. Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates. *Math. Comput. Simul.* **2001**, *55*, 271–280. [[CrossRef](#)]
36. Mara, T.A.; Tarantola, S. Application of Global Sensitivity Analysis of Model Output to Building Thermal Simulations. *Build. Simul.* **2008**, *1*, 290–302. [[CrossRef](#)]
37. Saltelli, A.; Annoni, P.; Azzini, I.; Campolongo, F.; Ratto, M.; Tarantola, S. Variance Based Sensitivity Analysis of Model Output. Design and Estimator for the Total Sensitivity Index. *Comput. Phys. Commun.* **2010**, *181*, 259–270. [[CrossRef](#)]
38. Homma, T.; Saltelli, A. Importance Measures in Global Sensitivity Analysis of Nonlinear Models. *Reliab. Eng. Syst. Saf.* **1996**, *52*, 1–17. [[CrossRef](#)]
39. Zheng, H.; Long, E.; Cheng, Z.; Yang, Z.; Jia, Y. Experimental Exploration on Airtightness Performance of Residential Buildings in the Hot Summer and Cold Winter Zone in China. *Build. Environ.* **2022**, *214*, 108848. [[CrossRef](#)]
40. Lu, Y.; Xiang, Y.; Chen, G.; Liu, J.; Wang, Y. On-Site Measurement and Zonal Simulation on Winter Indoor Environment and Air Infiltration in an Atrium in a Severe Cold Region. *Energy Build.* **2020**, *223*, 110160. [[CrossRef](#)]
41. Evans, M.; Yu, S.; Song, B.; Deng, Q.; Liu, J.; Delgado, A. Building Energy Efficiency in Rural China. *Energy Policy* **2014**, *64*, 243–251. [[CrossRef](#)]
42. Yang, S.; Tian, W.; Cubi, E.; Meng, Q.; Liu, Y.; Wei, L. Comparison of Sensitivity Analysis Methods in Building Energy Assessment. *Procedia Eng.* **2016**, *146*, 174–181. [[CrossRef](#)]