E3S Web of Conferences **686**, 01014 (2026) https://doi.org/10.1051/e3sconf/202668601014
ARFEE 2025

Regional Sensitivity Analysis for Multi-Performance Optimization of Rural Dwellings: A Case Study in Tianjin

Yunhui Wu[1], Yuan Gao[1,2], Ruoxi Han[3], Xiaozhu Lin[4,*]

[1]School of Architecture and Art Design, Hebei University of Technology, Tianjin 300132, China
[2]Key Laboratory of Healthy Human Settlements in Hebei Province, Tianjin 300132, China
[3]School of Architecture and Design, Beijing Jiaotong University, Beijing100044, China
[4]College of information Engineering, Beijing Institute of Petrochemical Technology, Beijing 102617, China

Abstract: With the advancement of building energy efficiency simulation technology, sensitivity analysis has become increasingly important in optimizing architectural design. This study examines typical rural dwellings in the Beijing–Tianjin–Hebei region, using Regional Sensitivity Analysis (RSA) to assess how design parameters affect energy consumption, carbon emissions, costs, and thermal comfort. Eleven parameters were analyzed through 1,000 simulations, including insulation type, envelope geometry, and photovoltaic configuration. Eave depth showed the highest sensitivity for all metrics, while roof insulation thickness and PV panel angle strongly influenced thermal performance and cost. The study also applies an RSA workflow to mixed continuous and categorical variables, with binning adjustment improving sensitivity results for discrete parameters. Based on sensitivity findings, differentiated optimization strategies were proposed: prioritizing passive shading and roof insulation for comfort optimization, while focusing on PV system scale for carbon emission and cost control. This study provides quantifiable evidence for low-carbon, efficient, and comfortable design in rural buildings.

1. Introduction

Under the “dual carbon” targets[1], the construction industry must reduce energy use and emissions while maintaining comfort and cost-effectiveness. Building Performance Simulation (BPS) supports design optimization by quantifying parameter impacts[2], yet efficiently determining the relative significance of multidimensional variables in materials, structures, and equipment remains a core challenge.

Sensitivity analysis (SA) has a long history in building performance simulation (BPS), predating energy consumption models from the 1980s and evolving into a core method for evaluating design uncertainties[3,4]. Traditional local SA relies on gradient information, offering intuitive insights but overlooking multi-parameter nonlinearity; while global SA (GSA) comprehensively assesses full-range impacts, it struggles to capture condition dependency [5,6]. Regional sensitivity analysis (RSA), as a semi-global approach, reveals how parameters influence performance across different ranges through the binning mechanism of output cumulative distribution functions [7]. This exposes the conditional dependency of parameters on varying performance levels, facilitating the identification of critical performance points and enabling differentiated optimization strategies[8].

For rural dwellings in northern China, the architectural forms are relatively uniform, primarily featuring brick-concrete structures with rudimentary envelope systems[9]. Heating and cooling systems are also simplified, typically relying on scattered coal or electric heating [10]. Performance optimization primarily relies on envelope design and passive measures, with significant feasibility impacts stemming from residents' economic conditions [11], material availability, and policy direction [12]. Therefore, applying the RSA method to this building category can not only identify conventional “key parameters” but also reveal how their relative importance shifts across different objectives—energy efficiency, low-carbon performance, comfort, and cost [13]. This facilitates the development of more targeted and actionable optimization strategies.

This study examines typical rural dwellings in the Beijing–Tianjin–Hebei region, selecting 11 design variables spanning geometry, materials, and photovoltaic configurations. Using SALib to implement the RSA method, sensitivity and pattern analyses were conducted for four metrics: annual electricity use, carbon emissions, global cost increment, and thermal discomfort hours. The main contributions are: (1) an RSA workflow for mixed discrete–continuous variables, reducing sensitivity bias in categorical data; (2) identification of variation patterns in parameter importance across performance levels; and (3) a differentiated optimization strategy providing quantitative guidance for low-carbon, efficient, and comfortable rural housing design.

*Corresponding author: linxiaozhu@bipt.edu.cn

2. Methodology

2.1. Overview of the Principles of Regional Sensitivity Analysis (RSA)

Regional Sensitivity Analysis (RSA) is a sensitivity analysis method based on output distribution characteristics. Its core approach involves binning the cumulative distribution function (CDF) of model outputs to compare differences in the conditional distributions of input variables across distinct performance intervals, thereby characterizing how variable importance shifts under varying performance levels. This method can be regarded as a semi-global, distribution-based sensitivity analysis capable of revealing nonlinearities, threshold effects, and conditional dependencies in input-output relationships [14,15].

The model is formalized as:

$$\mathbf{X} = (X_1, X_2, \ldots, X_k), \quad k \in N^+ \tag{1}$$

$$Y = f(\mathbf{X}) \tag{2}$$

2.2. Partitioning Mechanism (Based on Cumulative Distribution Function CDF)

In RSA, the “regional” aspect is defined by partitioning the model output space into bins. A cumulative distribution function (CDF) of all outputs is first computed:

$$F_Y(t) = \mathrm{P}(Y \leq t) \tag{3}$$

Quantile points are then obtained as:

$$q_j = F_Y^{-1}\left(\frac{j}{m}\right), \quad j = 0,1,\ldots,\mathrm{m} \tag{4}$$

With (m = 20) in this study—this value is the default in the SALib library and is widely used for RSA sensitivity analysis—producing intervals:

$$\mathrm{B_j} = \{ y^{(i)} \mid q_{j-1} < y^{(i)} \leq q_j \} \tag{5}$$

This ensures each bin contains outputs of similar performance levels and adequate sample numbers, which is more robust than direct slicing by (Y) value. For categorical (discrete) inputs, whose values change in jumps rather than smoothly, sensitivity estimates may be biased by category codes within different CDF regions. To mitigate this, discrete variables are one-hot encoded before RSA, preserving distribution trends without numeric ranking effects.

2.3. Conditional Probability Interpretation Framework

For an input variable X_u, its conditional distribution within the jth interval can be described by the conditional mean and variance:

$$\mu_{u|B_j} = \frac{1}{|B_j|}\Sigma_{y^{(i)}\in B_j} x_u^{(i)} \tag{6}$$

$$\sigma_{u|B_j}^2 = \frac{1}{|B_j|}\sum_{y^{(i)}\in B_j}(x_u^{(i)} - \mu_{u|B_j})^2 \tag{7}$$

RSA compares the front interval (low percentile, e.g., B_{low}) with the rear interval (high percentile, e.g., B_{high}):

Low-value interval (front of array): Represents excellent performance (e.g., low energy consumption, low carbon emissions, high comfort)

High-value range (rear of array): Represents degraded performance (e.g., high energy consumption, high carbon footprint, low comfort)

The sensitivity index can be defined as:

$$S_u = \left|\mu_{u|B_{\mathrm{low}}} - \mu_{u|B_{\mathrm{high}}}\right| \tag{8}$$

2.4. Analytical Tools and Sample Size Determination

To balance analytical precision and computational efficiency, this study employs the SALib open-source Python library to implement the RSA method. SALib's built-in RSA module efficiently performs:

- Construction and binning of the CDF for output data
- Calculation of conditional distribution parameters
- Evaluation and ranking of sensitivity indices S_u

For sample selection, Latin Hypercube Sampling (LHS) was employed to generate (N = 1000) sample points covering the full range of input variables. This sample size ensures numerical stability in sensitivity assessment while aligning with computational resource constraints in engineering calculations[16].

3. Case Study

3.1. Research Subjects and Reference Model

This study analyzes energy-saving retrofits for typical rural residences. Consequently, such projects are generally cost-sensitive, employ straightforward and easily implementable retrofit technologies. They also ensure the benchmark model used in this study possesses realistic representativeness and engineering feasibility. These characteristics make rural housing retrofits an ideal testing ground for validating sensitivity analysis methods and exploring differentiated optimization strategies.

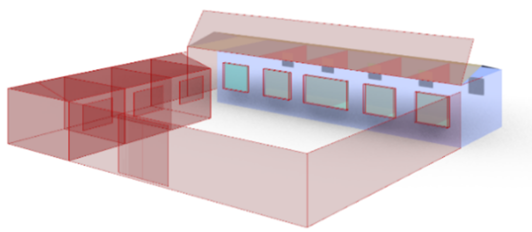

Fig. 1. Baseline Model Example

A sample survey revealed [10] that the main house lengths ranged from 10.50 m to 18.10 m, with a median of 13.85 m and an average of 13.91 m. Based on preservation status and structural integrity, Farmhouse A was selected as the reference model (see Figure 1). Constructed using the Rhino plugins Ladybug+ Honeybee, it supports dynamic thermal-humidity coupled simulation. Model geometric parameters were derived from field measurements: wall thickness 0.41 m (clay brick + mortar); The HVAC system includes air conditioning + electric heating (COP=3.2), with loads based on TMY meteorological data (Tianjin station). To validate the model, calibration was performed using monthly bills and field monitoring data, achieving RMSE<10%. Operating conditions reflect the impact of rural residential behaviors and clothing thermal resistance on building loads, providing a foundation for subsequent thermal comfort index calculations [17].

3.2. Parameter Definitions and Coding

3.2.1 Definition of Independent and Dependent Variables

To investigate the mechanisms by which different design parameters influence building performance, this study defines 11 independent variables (X1–X11), detailed in Table 1. Independent variables encompass insulation material types, geometric dimension parameters, and photovoltaic system configurations; dependent variables include: Y1: Electricity consumption (kWh); Y2: Carbon emissions (tCO_2e); Y3: Incremental global cost (CNY); Y4: Thermal discomfort hours (h), reflecting occupant comfort.

The study's independent variables include both continuous parameters (e.g., insulation thickness, eave depth, PV panel angle) and discrete parameters (e.g., insulation material type, window type, exterior wall/roof cladding material). This mixed-type input poses unique challenges for sensitivity analysis methodologies.

Table 1. Summary of Design Variables and Ranges

Code	Variable	Type / Range
X1	Insulation material	Categorical (3)
X2	Wall insulation thickness	0.14–0.29 m
X3	Roof insulation thickness	0.20–0.30 m
X4	Floor insulation thickness	0.07–0.20 m
X5	Window type	Categorical (2)
X6	Eave depth	0.6–1.5 m
X7	Window trim length	Categorical (2)
X8	Exterior wall finish	Categorical (3)
X9	Roof finish	Categorical (3)
X10	PV module quantity	1–24
X11	PV tilt	3–15°

3.2.2 Discrete Variable Handling and One-Hot Encoding

In regional sensitivity analysis (RSA), when binning cumulative frequencies, samples may concentrate in a limited number of bins if discrete variables have few categories. This leads to systematic amplification of sensitivity indices, compromising the accuracy of variable importance rankings.

To mitigate this bias, this study employs one-hot encoding to transform each categorical variable into multiple binary features (taking values 0 or 1 to indicate membership in the category) [2]. This allows them to be treated as "continuous variables" during binning. This approach ensures distribution uniformity while avoiding spurious ordinal relationships between categories. Following encoding, these features can be input alongside continuous parameters into simulation models and RSA analysis.

4. Results and Discussion

4.1. Comprehensive Processing of Results Data

In the RSA results, multiple NaN values appeared in the sensitivity indices for Y1 (electricity consumption). Cumulative distribution plots (Figure 2) show Y1's data distribution starts at a 50% cumulative frequency, causing NaNs at lower percentiles. Consequently, sensitivity calculations at lower percentiles produce NaN values. This does not indicate analysis failure but rather signifies insufficient variation in certain regions of the data distribution to compute sensitivity.

When a large number of samples cluster around the same value, the binning mechanism of the RSA method fails to effectively distinguish between different samples. This leads to division by zero or undefined conditions during sensitivity index calculations, resulting in NaN values. This phenomenon is not unusual in building energy consumption analysis; rather, it reflects the physical reality that buildings can achieve their theoretical minimum electricity consumption under specific parameter combinations.

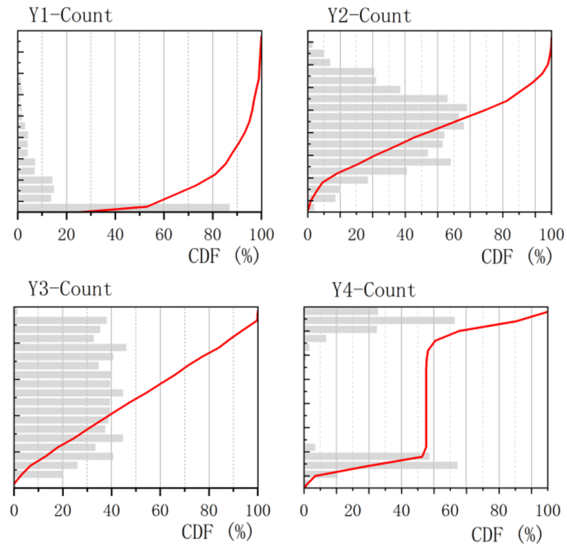

Fig. 2. Dependent variable statistical relationship, with the vertical axis representing Y values and the bars indicating technology.

Data processing for RSA sensitivity analysis was conducted as follows. For Y1 (electricity consumption), two data subsets exhibiting anomalous spikes were identified and excluded; the arithmetic mean of the RSA sensitivity index was then calculated using the remaining valid data. For Y2, Y3, and Y4, data distributions were stable, allowing direct calculation of the arithmetic mean.

The final RSA sensitivity analysis results shows that discrete variables consistently exhibit higher sensitivity values than continuous variables, including insulation material type (X1), window type (X5), window sill length (X7), exterior wall finish material (X8), and roof finish material (X9), with consistent rankings across all four indicators. This reflects both the foundational influence of these construction types and the systematic amplification effect of the RSA partitioning mechanism on categorical variables.

4.2. Trends and Physical Interpretation of Regional Sensitivity Analysis

In the sensitivity array, the front section reflects lower Y values and the rear higher Y values. For Y1 (electricity consumption), low values imply more energy-efficient configurations; for Y4 (thermal discomfort hours), high values indicate less comfort. This conditional-probability framework reveals how parameter importance shifts across performance levels. Data with unexplained fluctuations—e.g., separation variables Y3, X5 and X7—are omitted.

4.2.1 Trend of RSA Results for Separated Variables

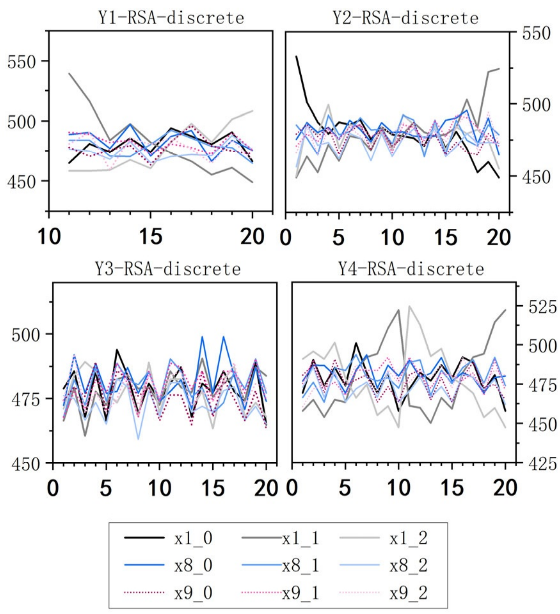

Fig. 3. Trend of RSA Results for Separated Variables

Fig. 3 reveals distinct response patterns among separated variables. In Y2 (carbon emissions), the intersection of X1_0 (EPS) and X1_1 (XPS) at the CDF midpoint defines a critical performance threshold. While the overall sensitivity of cladding (X8, X9) is stable, specific material traits shape their peak responses: cement tile (high absorption) responds more at low emissions, light linoleum (low conductivity) at medium, and steel exhibits marked fluctuations.

In Y4 (thermal discomfort hours), material responses diverge contextually: XPS surges under high discomfort, whereas RW1 maintains elevated sensitivity in low-discomfort ranges before declining, and EPS remains stable. For X8 materials, lime mortar and cement mortar peak at moderate and high discomfort levels respectively, influenced by solar absorption, while tile's low sensitivity relates to its high thermal mass. An intersection near CDF group 15 suggests a thermal-comfort inflection point.

Overall, these trends confirm diminishing marginal benefits: high-performance materials (e.g., XPS) are crucial under stringent or extreme conditions, while compensatory options (e.g., RW1) become significant in lower-efficiency scenarios.

4.2.2 Trends in RSA Results for Continuous Variables

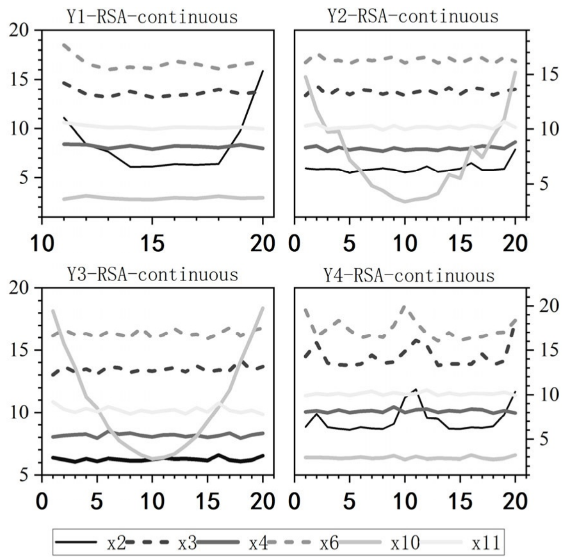

Fig. 4. Trends in RSA Results for Continuous Variables

Fig. 4 reveals distinct response patterns among continuous variables. X10 shows a U-shaped curve in Y2 and Y3, critical under both low and high carbon-cost extremes but with differing primary mechanisms (carbon reduction vs. cost recovery). In contrast, its sensitivity remains low in Y1 and Y4. X3 is highly volatile in Y1 and Y4 but stable in Y2/Y3. X6 is most sensitive in Y4, while X11 shows moderate sensitivity (~10-13).

Although Y2 and Y3 share similar overall sensitivity distributions, X10's unique nonlinear pattern (decreasing then increasing in Y2, fluctuating then rising in Y3) differentiates them. Overall, shading and roof insulation act as foundational drivers, PV specializes in carbon/cost optimization, and exterior walls serve as conditional modulators. These trends reveal indicator dependencies for continuous variables: shading and roof insulation serve as foundational drivers, PV specializes in carbon/cost optimization, while exterior walls function as conditional modulators.

5. Conclusion

5.1. Key Findings

This study applied RSA to assess the conditional sensitivity of design variables across key performance metrics. The results confirm that variable impacts are highly nonlinear and contingent on building performance levels. Key findings include: (1) For discrete variables, the intersecting sensitivity curves of insulation (X1) and finish materials (X8/X9) define critical performance thresholds, dictating material selection (e.g., XPS in high-efficiency vs. RW1 in low-efficiency scenarios). Windows (X5) are a consistently high-sensitivity element. (2) For continuous variables, eave depth (x10) is predominant, photovoltaic quantity (x11) shows a U-shaped carbon/cost sensitivity, and roof insulation (x6) is central to thermal comfort. These patterns arise from the dynamic interactions of thermophysical properties across performance levels.

5.2. Method Applicability, Engineering Implications, and Future Research Directions

The findings translate into tiered design strategies: prioritize eave depth (x10) and roof insulation (x6) for energy/comfort optimization; tailor PV capacity (x11) to avoid low-sensitivity troughs for carbon/cost control; and follow a roof > walls > floor insulation sequence, selecting materials (e.g., XPS or RW1) based on target performance thresholds. High-performance windows (X5) should be standardized.

Future work should validate the robustness of these sensitivity rankings through multi-method comparison (e.g., Sobol or Morris methods) and expand parametric coverage under extreme conditions. Cross-climate and building-type validations are needed to test the universality of the proposed strategies.

References

1. P. Shen, Y. Li, X. Gao, Y. Zheng, P. Huang, A. Lu, W. Gu, S. Chen. Recent progress in building energy retrofit analysis under changing future climate: A review, Appl. Energy, **383**, 125441 (2025)
2. D. Koutsandreas, I. Keppo. Harnessing machine learning algorithms to unveil energy efficiency investment archetypes, Energy Rep., **12**, 3180–95 (2024)
3. Z. Pang, Z. O'Neill, Y. Li, F. Niu. The role of sensitivity analysis in the building performance analysis: A critical review, Energy Build., **209**, 109659 (2020)
4. W. Tian. Renew. Sustain. A review of sensitivity analysis methods in building energy analysis, Energy Rev., **20**, 411–9 (2013)
5. R. Roka, A. Figueiredo, A. Vieira, C. Cardoso. A Systematic Review of Sensitivity Analysis in Building Energy Modeling: Key Factors Influencing Building Thermal Energy Performance, Energies, **18**, 2375 (2025)
6. B. Wang, J. Li. J. Global sensitivity analysis based on multi-objective optimization of rural tourism building performance, Cleaner Prod., **417**, 137917 (2023)
7. F. Pianosi, K. Beven, J. Freer, J. W. Hall, J. Rougier, D. B. Stephenson, T. Wagener. Sensitivity analysis of environmental models: A systematic review with practical workflow, Environ. Model. Softw., **79**, 214–32 (2016)
8. L. Li, S. Zu, S. Miraba, S. A. H. H. Dehkordi, A. K. Mehr, M. Baghoolizadeh, Sh. Baghaie, R. Marzouki. Multi-objective optimization of building performance: Integrating sensitivity analysis for energy efficiency and comfort, Int. Commun. Heat Mass Transfer, **165**, 109009 (2025)
9. Z. Li, Y. Gao, Y. Wu, Q. Geng. Economic Evaluation and Sustainable Optimisation of Envelope Parameters of Building Energy-Efficiency Design Standards in Cold Regions: A Case Study of Shijiazhuang, China, Sustainability, **17**, 9065 (2025)
10. Y. Gao, S. Luo, J. Jiang, Y. Rong. Environmental-thermal-economic performance trade-off for rural residence retrofitting in the Beijing–Tianjin–Hebei region, Northern China: A multi-objective optimisation framework under different scenarios, Energy Build., **286**, 112910 (2023)
11. T. Han, P. Liu, C. Niu, Q. Li. Evaluation of energy-saving retrofit projects of existing rural residential envelope structures from the perspective of rural residents: the Chinese case, Environ. Dev. Sustain., **25**, 8419–46 (2023)
12. M. Shan, P. Wang, J. Li, G. Yue, X. Yang. Energy and environment in Chinese rural buildings: Situations, challenges, and intervention strategies, Build. Environ., **91**, 271–82 (2015)
13. R. Srikanthan, Q. J. Wang, Y. Zhang. Use of regional sensitivity analysis for diagnosing parsimony of models: A water model case study, Environ. Model. Softw., **195**, 106727 (2026)
14. A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola. *Global sensitivity analysis: the primer* (John Wiley & Sons, 2008)
15. Y. Fujii, K. Matsuoka, R. Koide, N. Shinojima, S. Murakami. Regional sensitivity analysis to assess critical parameters in circular economy interventions: An application to the dynamic MFA model, J. Ind. Ecol., **29**, 429–39 (2025)
16. A. Karatzetzou. Uncertainty and Latin Hypercube Sampling in Geotechnical Earthquake Engineering, Geotechnics, **4**, 1007–25 (2024)
17. T. Szul, K. Nęcka, T. G. Mathia. Neural Methods Comparison for Prediction of Heating Energy Based on Few Hundreds Enhanced Buildings in Four Season's Climate, Energies, **13**, 5453 (2020)