Journal of Building Engineering 96 (2024) 110422

Contents lists available at ScienceDirect

Journal of Building Engineering

journal homepage: www.elsevier.com/locate/jobe

Building retrofit optimization considering future climate and decision-making under various mindsets

Pengyuan Shen [a,b,*]

[a] Shenzhen International Graduate School, Tsinghua University, Shenzhen, 518055, China
[b] Harbin Institute of Technology, Shenzhen, 518055, China

A R T I C L E I N F O

Keywords:
Building retrofit
Multi-objective optimization
Hierarchical clustering
Climate change
Sensitivity analysis
Decision making

A B S T R A C T

Building retrofit is effective in reducing building energy use and improving comfort levels for existing buildings. However, conducting multi-objective optimization for individual buildings can be challenging due to the laborious computational cost of using white box models and the difficulty in visualizing and understanding the decision-making process. Additionally, the impact of climate change has not been fully considered for the post-retrofit lifecycle. This research proposes a pragmatic automated scheme that integrates a feature selection method based on marginal abatement cost analysis and variance-based sensitivity analysis, multi-objective optimization supported by non-dominated sorting differential evolution (NSDE) algorithm, tailor-made decision-making support under different mindsets, and tree-based retrospection scheme of decision-making pathways. The simulation engine used in this study is a low-order white box modeling tool developed by the research team. The proposed scheme was applied to two educational buildings with different thermal characteristics, and the results showed that a certain number of sampling sizes were needed to achieve reliable feature selection results. The hierarchical clustering based decision-making support scheme has demonstrated robustness in visualizing and supporting decision-making for Pareto front. Two retrofit mindsets - aggressive and balanced - were assumed in the decision-making process, and the proposed method produced distinct final solutions accoridng to the two mindsets. This framework can support informed decision-making, helping stakeholders implement sustainable practices and transition to a low-carbon built environment.

Nomenclature

ECM	Energy Conservation Measure
NPV	Net Present Value
NSDE	Non-dominated Sorting Differential Evolution
GJ	Gigajoule
TMY	Typical Meteorological Year
PMV	Predicted Mean Vote
RC	Resistor-Capacitor
NSGA-II	Non-dominated Sorting Genetic Algorithm II
MOGA	Multi-objective Genetic Algorithm

(*continued on next page*)

* Shenzhen International Graduate School, Tsinghua University, Shenzhen, 518055, China.
E-mail address: Pengyuan_pub@163.com.

https://doi.org/10.1016/j.jobe.2024.110422
Received 18 March 2024; Received in revised form 30 July 2024; Accepted 9 August 2024
Available online 12 August 2024
2352-7102/© 2024 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

(*continued*)

MOPSO	Multi-objective Particle Swarm Optimization
MAC	Marginal Abatement Cost
HVAC	Heating, Ventilation, and Air Conditioning
PV	Photovoltaic
DE	Differential Evolution
RMSE	Root Mean Squared Error
RCP	Representative Concentration Pathway
GCM	Global Climate Model
MATLAB	Matrix Laboratory
TRNSYS	Transient System Simulation Tool
EnergyPlus	Building Energy Simulation Software
PSO	Particle Swarm Optimization
SHGC	Solar Heat Gain Coefficient
JMIM	Joint Mutual Information Maximization
SimBldPy	Simulation Building Python
NV	Natural Ventilation
SWH	Solar Water Heating
$E_{post,k}$	Energy use post-retrofit in year k
$E_{pre,k}$	Energy use pre-retrofit in year k
C_{post}	Energy cost post-retrofit
C_{pre}	Energy cost pre-retrofit
S_i	First-order Sobol index for ith ECM
$S_{i,j}$	Second-order Sobol index for ECM i and j
I_{total}	Total investment cost
I_{pv}	Solar irradiation striking the panel surface
n_{pv}	Number of photovoltaic panels
S_{pv}	Array area of photovoltaic panels
μ_{pv}	Conversion efficiency of the solar cell
t_a	Outside air temperature

1. Introduction

The global building sector is one of the largest consumers of energy and a significant source of greenhouse gas emissions [1]. As urbanization continues and the existing building stock ages, the need for energy-efficient retrofits becomes increasingly critical. Retrofitting existing buildings for energy efficiency is not only essential for reducing energy consumption and carbon emissions but also for improving indoor comfort and the quality of human settlements. This necessity is amplified by the urgent challenges posed by climate change, which introduces new variables and uncertainties into the energy performance of buildings. Energy-saving renovation is a crucial means to enhance building energy efficiency, reduce carbon emissions, and improve the quality of human settlements in existing buildings [2]. However, as climate change becomes more severe and frequent [3], studies have indicated that it will result in greater uncertainty and impact on the future levels of building energy consumption [4]. The most impacted meteorological factors are outdoor temperature and humidity, which will undergo significant future climate changes [5]. As a result, existing buildings in the urban area would need to meet new requirements and challenges to adapt to the future climate [6]. Therefore, ensuring and reflecting the future climate adaptability of energy-saving renovation in existing buildings is a significant issue that requires thorough discussion.

The optimization problem of building energy-saving renovation involves many design variables and a wide range; the optimization goals can be diverse, such as economic benefits, initial investment, carbon reduction benefits, etc., making the existing building renovation optimization problem complex and multi-objective, and leading to difficulties in postprocess such as decision-making. Design variables, meteorological conditions, economic conditions, and a series of uncertainties and constraints brought about by the participation of owners in decision-making, etc., add more complexity to the issue. Moreover, in most of the previous studies, the energy-saving evaluation methods adopted by these studies are carried out through complex dynamic white models (such as EnergyPlus, etc.) [7], which leads to a huge dependence on computing power in the optimization process. The huge computational cost is not conducive to the application and promotion of this type of method in actual energy-saving renovation projects, as well as the deduction and implementation of the final retrofit plan, even with simplified prototype building models [8]. Therefore, this research is dedicated to addressing the growing need for sustainable building practices by proposing an efficient, scalable, and adaptable optimization method for energy-saving renovations. By integrating future climate scenarios into the optimization process, this study not only enhances the reliability of retrofit solutions but also promotes resilience against climate change. The outcomes of this research have significant implications for policymakers, building owners, and practitioners by providing a robust framework that balances energy efficiency, economic feasibility, and occupant comfort, which contributes to the broader goal of achieving sustainable urban development and mitigating the environmental impacts of the built environment.

2. Literature review

In building energy-saving retrofits, a wide range of energy-saving measures and corresponding variables are involved. For such a complex system as a building, each measure's design variables can impact the final outcomes of building energy consumption, greenhouse gas emissions, investment costs, and returns. Moreover, there are interactive effects among the variables that can either

Table 1
Summary of related research and analysis on building energy-saving renovation optimization methods.

Literature	Optimization Objective	Decision Making	Simulation Method	Optimization Algorithm	Building Type	If Climate Change is Considered
[9]	Total energy savings, internal rate of return	Weighted multi-objective	Model predictive control	Differential evolution	1960s office building	×
[10]	Energy savings, life cycle NPV, discounted payback period	Weighted multi-objective	Manual estimation	Differential evolution	1960s office building	×
[11]	Life cycle total cost	Single objective	eQuest simulation & static model	Genetic algorithm	1964 residential	×
[12]	Primary energy consumption, global energy cost, discomfort hours	Multi-objective Pareto front	EnergyPlus & Matlab simulation	Genetic algorithm	Typical 5-story Italian residential building	×
[13]	Initial investment, energy consumption, global warming potential	Multi-objective Pareto front	DIN V 18599 assessment method	NSGA-II	1900 office building	×
[14]	Retrofit cost, energy savings, indoor comfort	Weighted Tchebycheff	TRNSYS simulation	Tchebycheff procedure	1945 residential	×
[15]	Greenhouse gas emission reduction	Single objective	TRNSYS & Matlab simulation	Branch and bound	1960s office building; late 19th century school building	×
[16]	Indoor comfort, annual energy consumption	Weighted multi-objective	EnergyPlus	Multiple sampling & feature reduction	1910s office & sports building	×
[17]	Energy consumption, indoor comfort, global cost	Multi-stage analysis method	EnergyPlus & Matlab simulation	Feature reduction & branch and bound	1920 to 1070 building cluster	×
[18]	Payback period	Single objective	EnergyPlus typical building simulation	Nonlinear regression	Typical #1 - pre-1950, Typical #2–1950 to 1975, Typical #3 - post-1975	×
[19]	Retrofit cost, energy savings	Weighted multi-objective	ISO 13790 monthly RC model	Tchebycheff procedure	1945 building	×
[20]	Electricity use, gas use	Weighted multi-objective	DOE 2.2 simulation	Genetic algorithm	Existing single thermal zone building	×
[21]	Energy consumption, indoor comfort, historic building conservation adaptability	Multi-objective Pareto front	EnergyPlus	NSGA-II	Pre-1780 building	×
[22]	Global cost, primary energy consumption	Multi-objective Pareto front	EnergyPlus	Branch and bound	Historic building	×
[23]	Energy consumption, CO2 emissions, retrofit cost, indoor thermal comfort	Multi-objective Pareto front	EnergyPlus	NSGA-III	Existing public school	×
[24]	Annualized cost, life cycle greenhouse gas emissions	Single objective & multi-objective Pareto front	EnergyPlus	Epsilon-constraint method	Existing residential building	×
[25]	Energy consumption, retrofit cost, indoor thermal comfort	Single objective & multi-objective Pareto front	TRNSYS simulation	Latin hypercube sampling, artificial neural network, MOGA	1983 school building	×
[26]	Marginal abatement cost vs. greenhouse gas reduction; Discounted payback period vs. investment cost	Main objective & sub-objective	TRNSYS & Matlab simulation	Branch and bound	Pre-1960s office building	×
[27]	Energy consumption, initial investment, indoor thermal comfort	Multi-objective Pareto front	TRNSYS simulation	NSGA-II	Existing office building	×
[28]	Life cycle economic benefit, energy savings	Multi-objective Pareto front	EnergyPlus	Branch and bound	1960s residential cluster	×
[29]	Total heating & cooling load, global cost	Multi-objective Pareto front	EnergyPlus & Matlab simulation	NSGA II	Pre-1900 public building	✓
[30]	Primary energy consumption, global cost	Multi-objective Pareto front	EnergyPlus &Matlab simulation	Orthogonal exhaustive method	An industrial building in southern Italy	×
[31]	NPV, annual total energy savings, greenhouse gas emissions	Multi-objective Pareto front	Not mentioned	Particle swarm optimization (PSO) & genetic algorithm	27 non-governmental organization buildings in Delaware, USA	×

(*continued on next page*)

Table 1 (*continued*)

Literature	Optimization Objective	Decision Making	Simulation Method	Optimization Algorithm	Building Type	If Climate Change is Considered
[7]	Total energy savings	Single objective	EnergyPlus	Feature selection based on JMIM algorithm & exhaustive method	Typical residential & office buildings in Philadelphia & San Francisco	✓
[32]	CO2-eq emissions, household water usage, life cycle costs, the percentage of thermal discomfort hours	TOPSIS multi-criteria decision-making	EnergyPlus, Matlab	NSGA-III	Residential building	✓
[33]	Life cycle cost, life cycle carbon emissions, operational energy use	Pareto frontier	EnergyPlus, Honeybee	NSGA-II	high-rise public rental housing buildings	✓
[34]	Thermal discomfort hours, carbon emissions, energy consumption	Multi-objective optimization	TRNSYS	NSGA-II	Office building	✓
[35]	Lifecycle cost	Hierarchical analytical process	TRNSYS	Parabolic interpolation algorithm	Educational buildings	✓

enhance or counteract each other's energy-saving effectiveness. On the other hand, the effectiveness of building energy-saving retrofits is constrained by various limitations, such as indoor comfort requirements, owner preferences, technical constraints, the adequacy of energy-saving management, and socio-cultural customs. These factors make energy-saving retrofits a complex multi-objective optimization problem, with the final decision-making process being open, multi-dimensional, and subject to multiple constraints. Currently, the methods for solving the optimization problem of single-building energy-saving retrofits are realized by combining building simulation techniques with optimization algorithms. Table 1 lists the main simulation techniques, optimization algorithms, and optimization objectives used in the current research field of building energy-saving retrofit optimization.

At present, the method of solving the optimization problem of the energy-saving renovation of a single building is mainly realized by a combination of building simulation methods and optimization algorithms. In terms of optimization goals, the sub-objective equations commonly used in the previous mainly include: 1) Energy consumption, including total energy consumption [12,13,16,17,21–23,25,27,33,34], energy saving [7,9,10,14,19,36], cooling and heating utility [20,29,37], embodied energy [38], etc.; 2) Economic aspects, including initial investment [13,14,19,25,27], Life cycle cost [11,12,17,22,29,32,33,35,39], payback period [18,26], etc.; 3) In terms of thermal comfort, indoor thermal comfort [14,16,17,21,27,32,37], annual thermal discomfort hours [23,34,36], etc.; 4) In terms of emission reduction effects [26], carbon dioxide emissions [23,32–34], greenhouse gas emissions [15,24,39], global warming potential [13], etc. Except for [7,36], few studies have considered the impact of future climate change on building energy consumption into the process of energy-saving renovation and optimization. The future meteorological uncertainty brought about by climate change scenarios has not been sufficiently considered in the optimization objective equations of relevant studies, so the overall future performance of energy-saving retrofits under climate change conditions cannot be guaranteed—that is, the future climate adaptation of energy-saving retrofits.

In the domain of building energy-saving renovation optimization, the main method used to calculate and evaluate the energy consumption level in the target equation after building renovation is building energy consumption simulation, among which EnergyPlus [7,12,17,18,21–24,28–30,32,33,37–41] is the most frequently used building energy consumption simulation. Tools, followed by TRNSYS [14,15,25–27,34,35,42], eQuest [11,43], DOE 2 [20], etc. As dynamic simulation tools, EnergyPlus and TRNSYS are often used for building energy efficiency diagnosis and scenario simulation in various studies due to their high modeling and simulation accuracy. Among the optimization problems involved in this research, the calculation and simulation of building energy consumption are the inevitable core problems when evaluating the target equation. The use of "white box" simulation tools based on dynamic equations such as EnergyPlus can ensure the accuracy of energy consumption simulation calculation results, but its modeling process requires a large amount of parameter input and professional knowledge, labor costs, and the solution of dynamic equations.

In terms of optimization algorithm and optimization result decision-making process, literature review shows that adopted methods include single-objective optimization [7,11,18], weighted multi-objective optimization [9,10,16,20], and multi-objective optimization (Pareto frontier non-dominated solution) are mainly used at present. For multi-optimization algorithm, a series of algorithms mainly involved include NSGA-II (non-inferior ranking genetic algorithm) [13,21,23,27,29,33,34,39], NSGA-III [32], MOGA (multi-objective genetic algorithm) [25], multi-objective particle swarm optimization (MOPSO) [31], and etc. However, a major issue with the weighted multi-objective method is that decision-makers' needs may change during the decision-making process, making the obtained optimization results meaningless if weight coefficients are altered. Furthermore, the method may lack intuitiveness and comparability due to varying dimensions among multiple targets, which can be problematic for decision-makers. Based on the literature review, it is evident that the current research on building energy-saving renovation and optimization is constrained by certain factors. These limitations include:

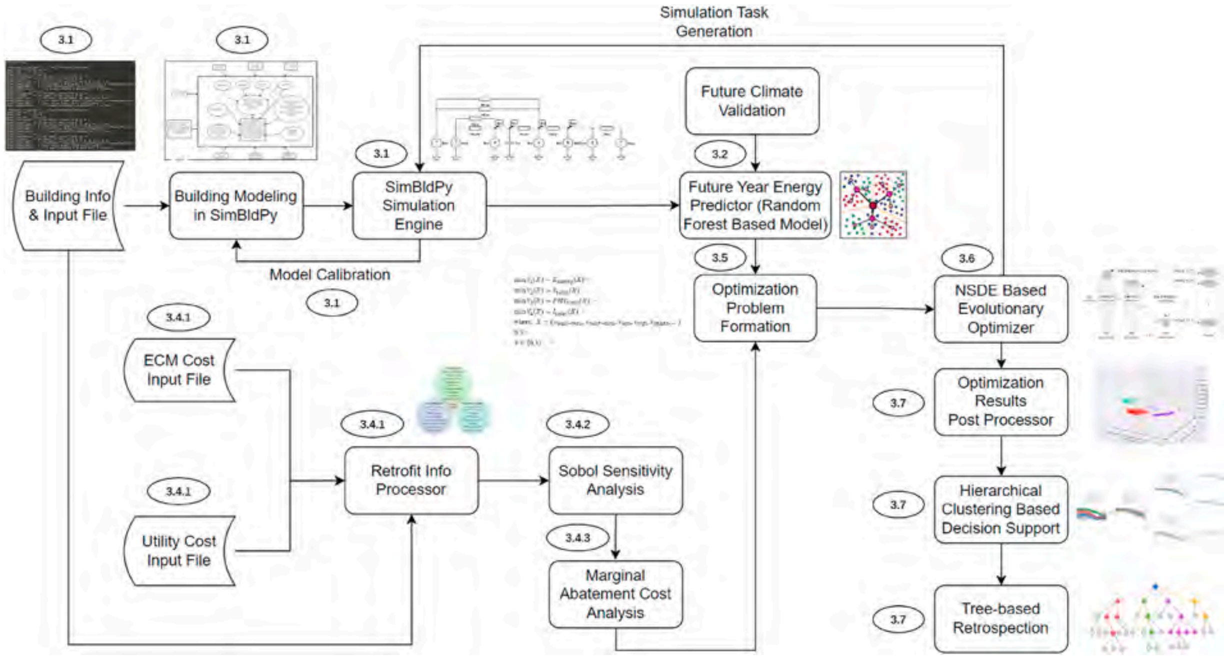


Fig. 1. Workflow and adopted methodology of the research.

1) The computational complexity of the fitness function based on dynamic energy consumption simulation is high and calls for significant computing power. This is due to the high computational complexity of the energy consumption simulation software used in most research institutions.
2) The optimization process does not fully consider the regional meteorological uncertainty that may arise from climate change. The current energy-saving simulation calculation method through the Typical Meteorological Year (TMY) generates non-life-cycle optimal solutions under present climate conditions.
3) There is a lack of decision-making support and visualization tools for multi-objective optimization solution sets, which can facilitate various decision-making mindsets and trackback. To address this gap, this study proposes the development of decision-making support and visualization methods and tools for multi-objective optimization solutions. Such methods will help stakeholders in energy-saving renovation to obtain a more intuitive and in-depth understanding of the optimization solution set, thus facilitating and promoting multi-participation in building energy-saving renovation projects.

To resolve the above challenges, this study proposes an automated scheme that integrates multiple methods to address the complexities of optimizing energy performance in building renovation projects. While recent studies [7,29,32–35] have begun to consider future climate impacts in the optimization frameworks, there remains a need for more comprehensive approaches that balance multiple objectives and adapt to the varying decision-making mindsets. This research addresses this gap by employing a feature selection method based on marginal abatement cost analysis and variance-based sensitivity analysis to identify the most influential factors affecting energy consumption. A non-dominated sorting differential evolution (NSDE) algorithm is implemented to solve the multi-objective optimization problem, offering robust solutions across diverse criteria. The simulation engine used for building energy performance is a resistor-capacitor-based low-order white box building simulator developed by the research team. This tool provides a lower computational cost compared with traditional white box modeling tools like EnergyPlus, enhancing the feasibility of large-scale applications. Additionally, the integration of data-driven models enables the consideration of uncertainties in future energy consumption due to climate change, utilizing validated and selected global climate models. The proposed scheme also leverages hierarchical clustering, decision tree algorithms, and advanced data visualization techniques to support decision-making and facilitate result backtracking for high-dimensional multi-objective optimization outcomes (Pareto front). The innovations introduced in this study—including the integration of future climate scenarios, the use of a low-order white box modeling tool, and the comprehensive decision support framework—address the existing gaps in the literature. These contributions not only advance the theoretical understanding of building retrofit optimization but also offer practical solutions for real-world applications, making the proposed scheme highly relevant for rapid optimization and implementation in energy-saving renovation projects.

3. Methodology

This study presents a systematic methodology to optimize building energy retrofits considering future climate conditions. The whole framework is programmed and implemented in Python 3.7 environment. The overall workflow, illustrated in Fig. 1, includes the following key phases:

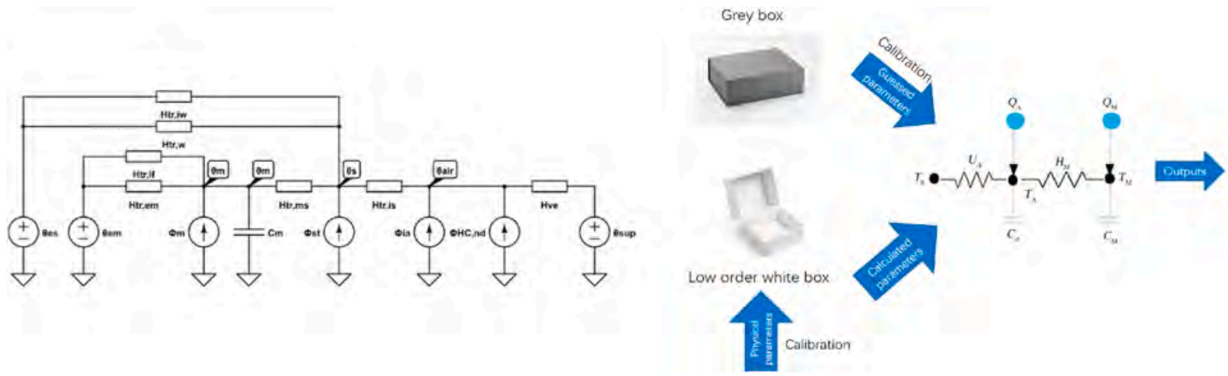


Fig. 2. The RC circuit diagram of SimBldPy [44], and the difference between lower order white box modeling and grey box modeling.

Table 2
Building parameters included in calibration and hyperparameters of DE.

Building parameter for calibration	Unit	Range	
Building Heat Capacity	J/K m^2	10000	800000
Effective Mass Area	m^2	1	5
External Wall Material U-value	W/m^2K	0.5	5
External Wall Material Absorptivity		0.3	0.95
Internal Wall Material U-value	W/m^2K	0.5	5
Window Material U-value	W/m^2K	0.5	6
Window Material Emissivity		0.6	0.95
Window Material SHGC		0.6	0.95
Roof Material U-value	W/m^2K	0.5	6
Roof Material Absorptivity		0.3	0.95
External Floor Material U-value	W/m^2K	0.5	6
Internal Floor Material U-value	W/m^2K	0.5	6
Air Infiltration Rate	h^{-1}	0.1	4
Lighting Load	W/m^2	1	20
Plug Load	W/m^2	1	30
Heating Supply Air Temperature	°C	20	40
Cooling Supply Air Temperature	°C	15	26
HVAC Distribution Loss Coefficient		0	0.3
Heating Temperature Setpoint	°C	16	24
Cooling Temperature Setpoint	°C	20	28
Hyperparameter of differential evolution	**Unit**		**Value**
Mutation Rate	/		0.1
Crossover Rate	/		0.8
Population Multiplier	/		20
Maximum Iteration	/		80
Convergence Tolerance	/		0.01

1) Data Collection: Gather building information, energy consumption data, and onsite historical climate data for the preparation of accurate building and climate modeling.
2) Building Energy Modeling: Use SimBldPy, a low-order white box modeling engine, to simulate building energy performance efficiently with low computational cost.
3) Climate Model Selection and Validation: Select and validate global climate models (GCMs) to obtain reliable future climate projections.
4) Feature Selection: Implement Sobol sensitivity analysis and the marginal abatement cost (MAC) analysis to identify influential factors and cost-effective energy conservation measures (ECMs).
5) Multi-Objective Optimization: Apply the non-dominated sorting differential evolution (NSDE) algorithm to address trade-offs between objectives like energy savings, cost, and indoor comfort.
6) Decision-making Support Scheme: Utilize hierarchical clustering and decision tree algorithms to aid decision-making and visualize decision-making pathways.
7) Implementation: Formulate and implement energy-saving retrofit measures based on the optimized solutions to two case study buildings.

Table 3
The available RCPs of the eight GCMs.

	RCP2.6	RCP4.5	RCP6.0	RCP8.5
MIROC5	✓	✓	✓	✓
GISS-E2-R	✓	✓	✓	✓
GFDL-CM3	✓	✓	✓	✓
CanESM2	✓	✓		✓
CSIRO-Mk3.6	✓	✓	✓	✓
MRI_CGCM3	✓	✓	✓	✓
INM-CM4		✓		✓
IPSL-CM5B-LR		✓		✓

3.1. Building energy modeling and calibration

The literature review reveals that previous research has predominantly utilized white box modeling tools such as EnergyPlus, TRNSYS, ESP-r, among others. However, this paper introduces a new building energy modeling and simulation method called SimBldPy [44], which is a low-order white box modeling tool developed by the authors. SimBldPy is essentially a simulation framework that is built on the resistor-capacitor (RC) analogy. It emulates the thermal transfer process using an electric circuit where thermal resistance is treated as resistors and thermal inertia is treated as capacitors. One key feature of SimBldPy is that it uses only one capacitor to represent the dynamic characteristic of the thermal transfer process, which in turn reduces the complexity in the solution of the core transfer function. SimBldPy is classified as a low order "white box" model because users are required to input model inputs that have physical significance. Unlike the traditional grey-box RC models that use lumped resistance values as a proxy input during the modeling process, SimBldPy requires fewer building-related parameters. This feature ensures that the modeling procedure is more efficient, and computation is more cost-effective, thus distinguishing SimBldPy from the traditional grey-box RC models (see Fig. 2).

To guarantee the validity of the SimBldPy building model, we employed the heuristic search method DE (differential evolution) to carry out the calibration of crucial input parameters that include building envelope, load intensity, thermal inertia, and additional elements. Table 2 exhibits the detailed parameters utilized in the calibration process and their corresponding parameter ranges.

To calibrate the model, monthly energy use bills for a full fiscal year can be collected and utilized as data. The validity of the calibrated model can be evaluated through the computation of the root mean squared error (RMSE) between the model's predictions and the metered data. The RMSE serves as the fitness function in the DE driven calibration and can be calculated in the following manner:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \widehat{x}_i)^2}{n}} \tag{1}$$

where, x_i and $\widehat{x}_i$ is the true and model predicted value.

To further corroborate the use of the developed SimBldPy model, the performance of SimBldPy in applying various ECMs across different climate zones has been compared and validated with EnergyPlus in a previously published study [44]. This validation demonstrated that SimBldPy produces results comparable to EnergyPlus while significantly accelerating the building simulation process by about ten times. SimBldPy, with its RC modeling approach coupling the heat transfer among thermal zones, can effectively reduce the computational load associated with building energy simulations while introduce controllable calculation bias, making it particularly suitable for parametric studies and optimization processes involving various ECMs in building retrofit projects. Furthermore, in this study, the developed SimBldPy model will be calibrated and validated against actual energy use data, as detailed in Section 4.1. Though evolutionary algorithms nowadays like NSGA-II can significantly release computational burdens and speed up the optimization process, it is important to note that the computational cost reduction achieved by optimization algorithms like NSGA-II is distinct from the computational burden alleviated by the building simulation engine itself.

3.2. Validated future climate model

Global climate change (GCC) is a significant environmental issue that affects many areas of the world [3]. Previous research has demonstrated that GCC can impact the effectiveness of energy conservation measures (ECMs) during building retrofit lifecycles [7]. Therefore, the optimal solution or Pareto front of multi-objective optimization may be affected by the impact of GCC if it is not considered. While there has been some work that incorporates the impact of GCC in building retrofit analysis and optimization [29,36], the selection and validation of GCC models and their downscaled results have not been well established. Zhai's research [36] has shown that various general circulation models (GCMs) produce a wide range of predictions for future weather conditions in a given location. As a result, there can be significant uncertainty and variation in the impact of GCC on building energy predictions. Our research has also found that different GCMs can produce vastly different downscaled results, and caution should be exercised when selecting which GCM to use for a particular city or district. To ensure the accuracy of model projections, historical weather data should be used to validate the selected GCM's fit with the real conditions of the target location.

This study collected and downscaled simulation results from various GCMs, which were derived from a total of 27 different emission scenarios. Specifically, the results from the GCMs were obtained from the Coupled Model Intercomparison Project Phase 5 (CMIP5) of the IPCC Fifth Assessment Report (AR5) [45], and included eight different models. The results were collected from the specific grid point that corresponds to the selected site. This research utilizes GCMs from CMIP5 with representative concentration

Table 4
Window replacement properties and cost [14,19,51].

window (SHGC, U-value (W/m2 °C))	$/m2	descriptions
(0.0, 0.0)	0.00	N/A
(0.80, 5.6)	47.0	Single glazing
(0.75, 2.8)	53.2	2bl glazing Without thermal break
(0.62, 1.6)	75.2	2bl glazing low-e window
(0.44, 1.6)	92.9	2bl glazing Window air-filled metallic frame
(0.288, 1.05)	79.2	SGSILVER
(0.585, 0.52)	98.1	SGCLIMATOP
(0.28, 0.33)	113.4	3050 SH 1.11 cm glass low-e
(0.63, 0.48)	131.7	3050 SH 1.11 cm glass
(0.25, 0.26)	183.0	3050 DH 3–7/16 insulated glass low-e krypton filled triple pane

pathway (RCP) scenarios for downscaling future weather data. Despite the availability of newer CMIP6 models incorporating SSP scenarios, we chose CMIP5 models due to their demonstrated alignment with observed warming trends. According to a recent study by Carvalho et al. [46], CMIP5 projections are more consistent with observed temperature increases over land areas from 1980 to 2020 compared to both CMIP3 and CMIP6 models. The study concluded that CMIP5 tends to slightly underestimate warming but remains closer to actual observations than CMIP6. Therefore, using CMIP5 models ensures reliable and validated projections for our study's context, providing a robust basis for our building energy retrofit optimization under future climate scenarios.

This study collected data on various weather variables, namely air temperature, relative humidity, and solar radiation. The specified CO_2 concentrations for each GCM were obtained from the Representative Concentration Pathways (RCPs) described in IPCC AR5. Specifically, the concentrations were 421 ppm (RCP2.6), 538 ppm (RCP4.5), 670 ppm (RCP6.0), and 936 ppm (RCP8.5) in 2100 as simulated in Ref. [47]. The CMIP5 database contains a total of 27 concentration pathways, which are listed in Table 3. To fit the GCM output to local meteorological characteristics, downscaling of the future projections generated by the GCM must be performed after the appropriate grid point is selected. To achieve this, an easy-to-use downscaling method proposed by Belcher et al. [48] was employed. This method has been widely used in the downscaling of future weather data [49,50]. The 27 downscaling scenarios were then compared with measured weather data obtained between 2012 and 2016. The results of the comparison will be presented and discussed in subsequent sections.

3.3. Future hourly energy prediction during retrofit lifecycle

The process for obtaining energy use savings during the retrofit lifecycle involves generating hourly weather data for future years and then subtracting the energy use of the baseline building model (i.e. the original model of the existing building) from that of the retrofitted model. However, predicting future hourly energy use using simulation methods can be overwhelming due to the high computational cost involved.

Previous research conducted by the authors has shown that the RF (random forest) algorithm is an effective approach for predicting future hourly energy use. This is due to its high performance in producing predictive models trained on hourly energy use data, specifically data simulated using future extreme-weather year data [7]. The authors found that the most reliable data-driven model is produced when features including outdoor air temperature, relative humidity, solar radiation, wind speed, occupancy schedule, lighting schedule, equipment schedule, infiltration level, cooling setpoint temperature, and heating setpoint temperature are included in the training data. In addition, the authors have demonstrated that using future extreme-weather year data, rather than TMY data, significantly improves the accuracy of energy use prediction. This is because TMY weather data does not allow the model to "learn" from experience with future, more extreme weather conditions, such as higher temperatures and humidity in summer, and lower temperatures in winter. Extreme year weather data can be created by combining half of the future weather data from the year with the most severe winter (having the lowest monthly mean temperature in the chosen future period) and the year with the most severe summer (having the highest monthly mean temperature in the chosen future period) [7]. This data can subsequently be utilized as input for building simulations and processed through the RF algorithm to develop a data-driven model. In this study, the assumed lifecycle for retrofitting is 20 years.

3.4. Pre-optimization decision variable (ECM) selection

A two-step feature selection scheme was implemented in this research to solve the issue of non-influential ECMs and computational complexity in the multi-objective optimization problem. The primary concern of the decision makers was the economic performance of the project. Hence, the research assumes that the net present value (NPV) is a robust and comprehensive criterion for evaluating the feasibility of the project. To select the pre-optimization decision variables, NPV was utilized as the objective function.

3.4.1. ECMs and parameters

A total of 14 ECMs were considered, with the majority comprising passive measures aimed at conserving energy. These measures include enhancing the thermal properties of the building envelope, optimizing air tightness, promoting natural ventilation, and regulating heating and cooling setpoints during unoccupied hours. The only non-passive measure included was the optimization of lighting efficiency. The costs associated with window replacement were detailed in Table 4. ECMs requiring monetary investment were catalogued in Table 5, while those without costs were listed in Table 6, respectively.

This study solely focused on the installation of PV (photovoltaic) panels as a renewable energy measure. The electricity generated

Table 5
ECM parameters and costs [10,13,52–54].

wall insulation (m^2 °C/W)	\$/wall m^2	roof insulation (m^2 °C/W)	\$/roof m^2	window shading	\$/window m^2	air infiltration (h^{-1})	\$/m^2	lighting efficiency improvement	\$/m^2	daylight control	\$/m^2
N/A	0	N/A	0	N/A	0	N/A	0	N/A	0	N/A	0
1.25	11.4	1.52	12.5	1	28.7	0.3	25.5	30 %	3	Applied	3
1.61	12.5	1.97	16.4	2	37.2	0.5	20.2	40 %	1.9		
1.97	13.5	2.42	20.1			0.7	14.4				
2.33	14.6	2.87	22.9			0.9	9.3				
2.69	15.7	3.32	26.8								
3.05	16.7	3.77	30.3								
3.41	18.5										
3.77	20.5										

Table 6
ECM parameters without cost.

cooling supply air temperature	natural ventilated window ratio	cooling setpoint	heating setpoint	unoccupied hour setback
N/A	N/A	N/A	N/A	N/A
17	10 %	22	18	Applied
18	20 %	23	19	
19	30 %	24	20	
20	40 %	25	21	
	50 %	26	22	
	60 %	27	23	
	70 %		24	
	80 %		25	
	90 %			
	100 %			

Table 7
ECM parameters related with PV system.

PV panel area in proportion to roof area (%)	0	20 %	40 %	60 %	80 %	100 %
PV panel angle (degree)	0	15	30	45	60	75

by the panels will be directly inverted to AC and supplied to the building. It is important to note that no battery is installed in the PV system, meaning any surplus energy generation will not be utilized by the building. However, considering the usual high electricity load of the building, it is likely that nearly all the generated power will be utilized. To calculate the power output of the PV system, the method described in Ref. [55] was used.

$$P_{pv} = n_{pv}\mu_{pv}S_{pv}\, I_{pv}(1 - 0.005(t_a - 25)) \;\; [W]$$

where, n_{pv} represents the number of panels, S_{pv} corresponds to the array area (m^2), and μ_{pv} is the conversion efficiency of the solar cell used in the array. A poly-crystalline silicon (p-Si) solar cell, with an efficiency of 14 %, is utilized in this instance [56]. Additional parameters include I_{pv}, which is the solar irradiation striking the panel surface (W/m2), and t_a, the outside air temperature. The PV panel cost amounts to \$274.7/m^2. For the installation of both PV and solar water heating (SWH) systems at a non-zero inclination angle, a frame support installation fee of \$50/m^2 will be charged. The PV system parameters consist of two factors: the installed area and the inclination degree. The installed area parameter is a percentage value of the total roof area of the building. The parameters and their range for the PV system are shown in Table 7.

3.4.2. *Variance-based sensitivity analysis*

Sobol's method has gained popularity in recent literature as a variance-based method due to its ease of implementation [57]. In cases where non-linearity characteristics are present, such as building retrofit analysis, Sobol's method can provide insight into both the individual impact of ECMs on the system as a whole and the interactive effects among ECMs. The first order effect, also known as the main effect, measures the impact of a single input parameter on the model output variance. This effect demonstrates the specific impact of an ECM on the objective function, which in this study is the NPV. The calculation for the first order effect is as follows:

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

where, Y is the univariate model output – NPV in our case; S_i is the first-order Sobol index for ith ECM; $V_i(Y)$ is the variance brought by ith ECM; and $V(Y)$ is the total variance of the output variable.

The second order Sobol index of ECM i and j (S_{ij}) can be calculated by:

$$S_{ij} = \frac{V_{ij}(Y)}{V(Y)}$$

where, $V_{ij}(Y)$ is the variance of output Y under the synergetic impacts from ith and jth ECM. The total effects, which means the combined first and second order effect on the model output can be then calculated as:

$$\sum_{i=1}^{d} S_i + \sum_{i<j}^{d} S_{ij} + \cdots + S_{12\ldots d} = 1$$

where d is the dimension of the design space.

The present research employs Saltelli's extension of the Sobol method for sample generation within the design space [58]. Moreover, we have conducted an analysis of the sampling size to provide a comparison for the two case study buildings. Our aim is to determine if the returned Sobol' indices can converge over an increasing number of samples.

3.4.3. Marginal abatement cost analysis

Prior to conducting the sensitivity analysis for feature selection, a precedent measure should be adopted in advance to prevent input variables that may have a negative impact on building performance from being included in the optimization process. In this study, the Marginal Abatement Cost (MAC) analysis method has been employed to determine whether a particular ECM will have a positive effect on the chosen objective function.

The MAC curve method, which has been extensively used to estimate the balance between environmental benefits and costs, is worth mentioning [59–62]. By evaluating different abatement measures individually, MAC curves can be generated. Incremental costs of diverse abatement measures are calculated based on the initial energy system situation, and the costs of CO2 emissions abatement for individual measures are ranked to derive the MAC curve [63].

3.5. The optimization problem

In this research, we have identified four objectives for the multi-optimization approach: energy saving in gigajoule (GJ) E_s, energy saving in dollar ($) S, investment cost ($) I, and the aggregation of absolute PMV (predicted mean vote) values over the lifetime of the retrofit. It should be noted that the "PMV sum" in this study refers to the total number of absolute PMV values across all thermal zones in the building throughout its lifecycle. This metric is calculated by summing the absolute PMV values for each hour across all thermal zones and then multiplying by the number of occupied hours per day, days per year, and the number of years in the lifecycle. As thermal neutrality corresponds to a PMV of 0, a lower PMV sum indicates a more comfortable indoor environment throughout the building's lifecycle. Therefore, the PMV sum provides a comprehensive measure of indoor thermal comfort, with smaller values signifying better overall comfort for the building's occupants. While the single objective of NPV is used in the sensitivity analysis, it can be represented through the previous three sub-objectives of the optimization problem. However, it is necessary to list all three sub-objectives in the optimization because decision makers may not be willing to invest in high initial retrofit costs, even if the NPV suggests that the retrofit is viable. Additionally, the inclusion of the thermal comfort index is important to prevent Pareto solutions that compromise indoor thermal comfort. The optimization problem is described as follows:

$$\min Y_1(X) = E_s$$

$$\min Y_2(X) = S_{total}$$

$$\min Y_3(X) = P$$

$$\min Y_4(X) = I_{total}$$

s.t.

$$X \in [0, 1)$$

In order to accommodate the requirements of the NSDE algorithm, a set of ECM parameters X has been normalized for each ECM in a continuous space. This can be achieved by pre-processing each input variable into a continuous space subjected to uniform distribution. Consequently, the energy savings in GJ have been transformed into primary energy use with transforming factors F to enable comparability between different end-use utilities. The calculation of energy savings is provided below:

$$E_s = \sum_u \left(\sum_{k \le L} \left(E^u_{post,k} - E^u_{pre,k} \right) * F_u \right)$$

where, $E_{pre,k}$ and $E_{post,k}$ is kth year's annual energy use if retrofit does not take place, which is presumed to be the baseline case, and the energy use after kth years of retrofit for a utility u, in GJ. We utilize a retrofit lifecycle period length (L) of 20 years in this research to assess the long-term effectiveness of retrofit measures.

The calculation of energy savings in dollars can be conducted by comparing the utility cost of a retrofitted building with that of a baseline building. This comparison should be aggregated over the retrofit lifecycle. To determine the energy savings in dollars, the formula can be expressed as follows:

$$S_{total} = \sum_u \sum_k \frac{(1 + \tau_u)^k * \left(C^u_{post,k} - C^u_{pre,k} \right)}{(1+r)^k}$$

where, τ_u (%) is the cost increase for a certain utility type u, and $C^u_{post,k}, C^u_{pre,k}$ is the energy cost of utility type u during the retrofit life cycle for the retrofitted building and baseline building in year k.

The PMV calculation is typically carried out under standard conditions, assuming a constant metabolic rate of 1.1 met and a steady indoor air velocity of 0.5 m/s. Additionally, when determining investment costs, it is important to account for maintenance expenses occurring every five years. This maintenance cost is estimated to be 15 % of the total investment cost. As a result, investors can calculate their total investment cost by using this formula:

$$I_{total} = I_0 + \begin{cases} \sum_k \frac{(1+\tau_m)^k * I_k}{(1+r)^k}, k \,\%\, 5 = 0 \\ 0 \ , \ \ otherwise \end{cases}$$

where, τ_m (%) is the increase in maintenance fee of each year, r is the discount rate (%), and I_k is assumed to be proportional to the initial investment of each ECM. The discount rate is assumed to be 4 %. A discount rate of 4 % is assumed in this research.

3.6. Optimization algorithm

The optimization algorithm utilized in this study is the non-dominated sorting differential evolution (NSDE), which is based on the multi-objective genetic algorithm. NSDE is a highly capable variant of genetic algorithms that can effectively address problems that are continuous in domain, noisy, and subject to change over time. The algorithm works by iteratively improving a candidate solution with respect to the fitness function.

Research has been conducted to compare the performance of three algorithms and their application in various optimization problems. The results show that NSDE outperforms the others in terms of both computational time and accuracy [64–66]. In other words, NSDE provides higher quality solutions while requiring less time or cost to converge. As a result, NSDE was chosen for the optimization of building retrofit in this study. Further information on the NSDE algorithm used in this research can be found in Ref. [36]. The effectiveness of NSDE in Ref. [36] has been confirmed for optimizing building retrofit problems by validating the chosen hyperparameters. The algorithm employs a generator based on a gaussian random Latin hypercube sampling method. In the research, tournament selection was utilized as the selector due to its computational efficiency and suitability for parallel implementation, as opposed to the classic rank selection method [67]. A crossover rate of 0.85 and a uniform crossover operator were assigned to the algorithm while employing a gaussian mutation operator with a mutation rate of 0.01. The population size of each generation was set to be twenty times the total number of parameters.

It is important to mention that a customized criterion for convergence of optimization was utilized in this study. In order to determine if the optimization process converges, we introduced a novel criterion that assesses the quantity of "newcomers" to the non-dominated solution set. Typically, the commonly used method to assess whether optimization has reached convergence is by observing if the non-dominated solution count remains constant across generations. However, this approach is inadequate as it does not take into account situations where the number of new solutions entering the solution archive is nearly equivalent to those eliminated. To address this issue, the suggested criterion effectively prevents the optimization from reaching a "false" convergence.

3.7. Decision-making support and tree-based retrospection

For multi-objective optimization problems that involve multiple subobjectives, the resulting Pareto front can have a high-dimensional decision space. This often leads to a large number of non-dominated solutions, sometimes numbering in the hundreds, which can prove challenging to visualize and make decisions from. To address this issue, an unsupervised machine learning algorithm known as the hierarchical clustering method [68] has been developed to provide decision-making support. This method is widely accessible to decision-makers as it is easy to understand [69], the decision-making process can be clearly tracked, and the algorithm itself is not difficult to implement.

The decision-making process for building retrofit optimization was supported by hierarchical clustering techniques to group similar retrofit strategies based on multiple criteria, which was described in detail in Ref. [36]. The procedure involves normalizing the optimization results and applying agglomerative hierarchical clustering to classify the data into clusters. The "elbow" method is used to determine the appropriate number of clusters at each layer by identifying the point where the distance growth between clusters stabilizes. This process is visualized through dendrograms, which illustrate the merging of clusters. At each layer, parallel coordinates plots and heat maps are employed to visualize the trade-offs among different retrofit options, aiding stakeholders in making informed decisions. Decision tree algorithms are then used to retrospectively analyze the selected pathways, ensuring transparency and traceability in the decision-making process. This approach provides an intuitive and structured framework for exploring and evaluating various retrofit strategies.

After examining each subcluster of interest to the decision makers, the optimal solution for implementing the retrofit will consist of a set of ECMs or a single best combination. To further visualize the entire decision-making process, the decision tree algorithm is applied. The decision trees involves a supervised algorithm that utilizes a binary tree graph, with each node having two children, to assign a target value to each data sample [70]. The process of decision tree learning entails seeking out the optimal rules in each internal tree node based on the selected metric. Through the tagging of all non-dominated solutions with a cluster number in each layer of clustering, the decision tree algorithm is capable of retracing how decisions were made by the hierarchical clustering algorithm. In our case study, we employed the decision tree method to retrospectively examine the decision-making process utilized by both the hierarchical clustering algorithm and the decision maker. The results of our analysis will be discussed in the following chapter using two case studies.

4. Case studies of two educational buildings

4.1. Building energy modeling of the two case study buildings

In this study, we selected two on-campus educational buildings to evaluate our proposed optimization scheme. These buildings have similar functions, including classrooms, office, and lab spaces. However, they differ in architectural design and surface-to-volume

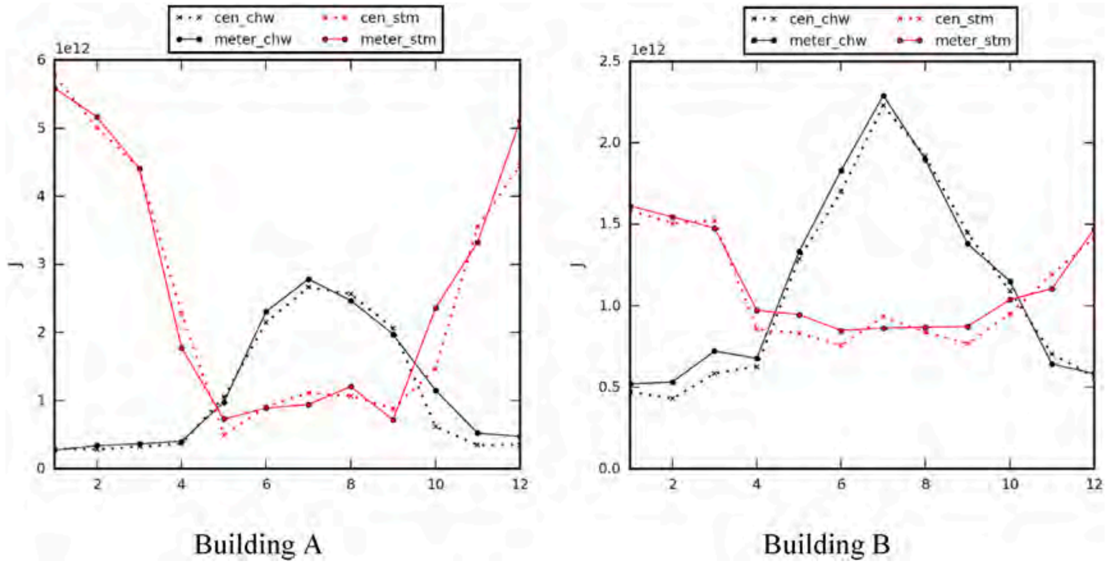


Fig. 3. The calibration results of the monthly energy use of the two buildings.

Table 8
Key modeling parameters of the building energy models of the two case study buildings.

Modeling Parameter	Building A	Building B
Building Area (m^2)	14260	7780
Envelope Heat Capacity (J/K m^2)	135679	213468
External Wall U-value ($W/m^2 \cdot K$)	1.87	1.34
Internal Wall U-value ($W/m^2 \cdot K$)	1.96	0.89
Window U-value ($W/m^2 \cdot K$)	4.16	4.31
Roof U-value ($W/m^2 \cdot K$)	0.92	1.78
HVAC System	Variable Air Volume (VAV)	Variable Air Volume (VAV)
Lighting Load (W/m^2)	4 to 10 depending on thermal zone	8 to 14 depending on thermal zone
Occupancy (m^2/person)	19	19
Metabolic Rate (W/person)	120	120
Appliance Load (W/m^2)	4 to 15 depending on thermal zone	6 to 22 depending on thermal zone
Indoor Temperature Setpoints (°C)	Varies (e.g., 18 to 27 depending on time and conditioned thermal zone)	Varies (e.g., 18 to 24 depending on time and conditioned thermal zone)
Energy Source of Heating and Cooling	Chilled water and steam provided by district cooling and heating system	Chilled water and steam provided by district cooling and heating system

ratio (S/V). Building A has a larger floor area of 13900 m^2 and a higher S/V compared to Building B, which has a floor area of 7781.4 m^2. We chose these two buildings for our case studies because they represent two types of buildings in terms of thermal behavior - climate dominated and internal load dominated. The data on monthly energy consumption has been collected annually and used to calibrate a model, the results of which are depicted in Fig. 3. The solid lines in the figure depict the metered energy usage data, while the dotted lines represent the results from the calibrated models simulated via SimBldPy. The findings demonstrate that the DE algorithm-based approach exhibits exemplary performance in matching the simulated models to monthly metered data and reliably yields reasonable building parameters. The monthly energy usage of two buildings located on the same campus and exposed to a similar microclimate exhibits significant differences. Building A uses more heating energy, while Building B uses more cooling energy annually. This variation is attributed to the differences in building design and thermal inertia properties of the structures. Building A, with higher S/V, is more influenced by climatic factors than Building B. On the other hand, Building B is more influenced by internal loads, which results in a higher cooling load during summer.

Table 8 summarizes the essential building energy modeling parameters, such as building dimensions, envelope characteristics, HVAC system details, and occupancy loads, for both buildings.

4.2. Selection of GCM and generation of future hourly weather data

Various GCM models were employed to validate the predictions of $\overline{DNT}$ (monthly mean daily minimum temperature), $\overline{DXT}$ (monthly mean daily maximum temperature) and $\overline{MMT}$ (monthly mean temperature) using different RCPs and GCMs against historical weather data in Philadelphia. To evaluate the prediction performance, the RMSEs of different emission scenarios and GCMs were computed and evaluated. As shown in Fig. 4, the performance of the IPSL-CM5B-LR model is the worst. In particular, the RMSE of $\overline{DNT}$

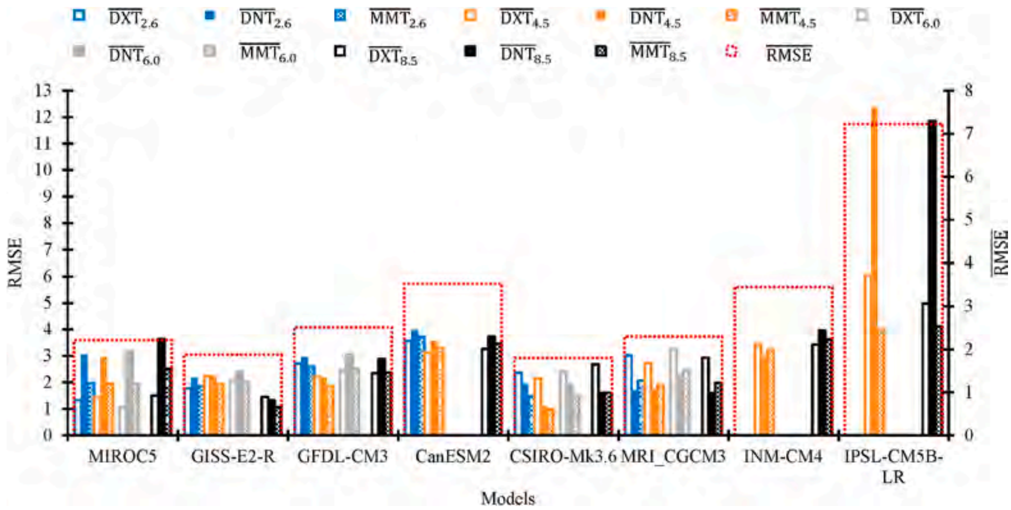


Fig. 4. The RMSE of temperature for each model and RCP.

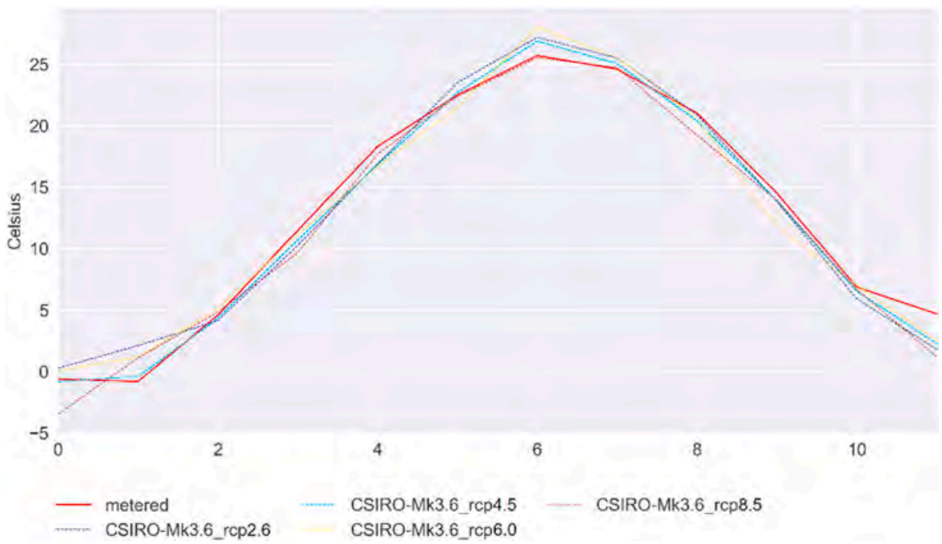


Fig. 5. The metered vs. the predicted $\overline{MMT}$ by CSIRO-Mk3.6 model.

can reach up to 12.37, and the RMSE of $\overline{DXT}$ and $\overline{MMT}$ is also higher than other models. The overall accuracy of the CSIRO-Mk3.6 model is the highest, with $\overline{RMSE}$ of 1.81. This is because that the model can simulate $\overline{DNT}$ and $\overline{MMT}$ well. However, the RMSE of $\overline{DXT}$ is relatively high for CSIRO-Mk3.6 model. In fact, MIROC5 model has the best performance when $\overline{DXT}$ is simulated, but which shows lower accuracy when $\overline{DNT}$ is predicted. And the RMSE of MRI_CGCM3 model for $\overline{DNT}$ prediction is low. In general, even if the serious error of the IPSL-CM5B-LR model is not considered, the whole error of $\overline{DNT}$ is also the largest, and the error of $\overline{MMT}$ is the smallest. The climate predictions under different RCP scenarios did not show obvious differences. It can also be said that the model selection has a greater impact on the simulation results.

CSIRO-Mk3.6 model has the highest accuracy when $\overline{MMT}$ was projected, and the comparison between the simulated $\overline{MMT}$ and the metered data is shown in Fig. 5. All RCP scenarios can simulate the $\overline{MMT}$ change trend. From December to March, the simulated values under the RCP4.5 scenario are the closest to the metered data. In general, the RMSE under the RCP4.5 scenario is the smallest of 0.97. The simulated value of $\overline{MMT}$ in November was underestimated, and the simulated value of $\overline{DXT}$ was also significantly lower than the measured value.

4.3. *Feature selection - sensitivity analysis*

4.3.1. *MAC analysis*

During the analysis of MAC, a simulation will be conducted for each ECM and parameter independently to calculate their respective

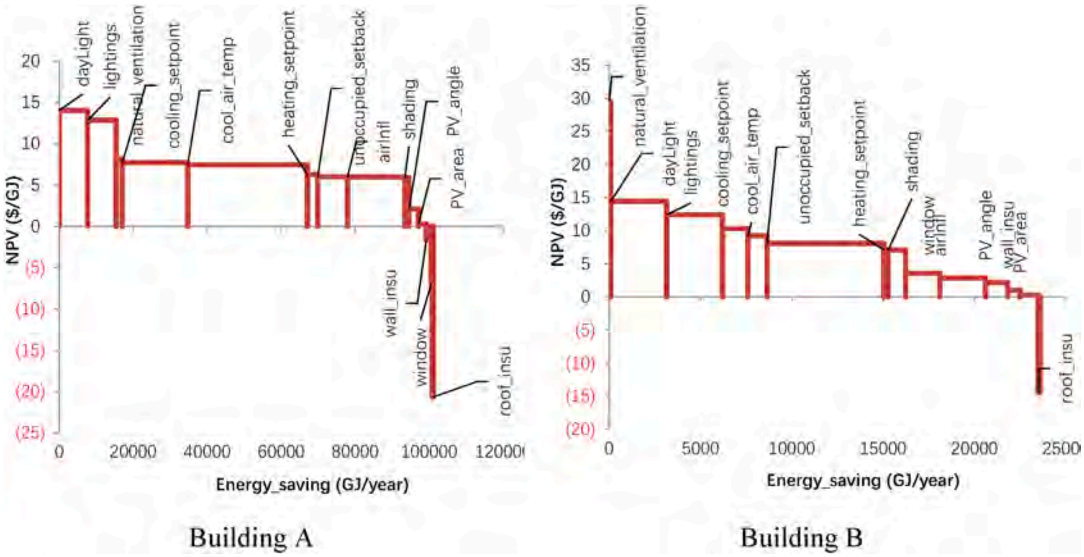


Fig. 6. Marginal abatement cost curve of the two buildings.

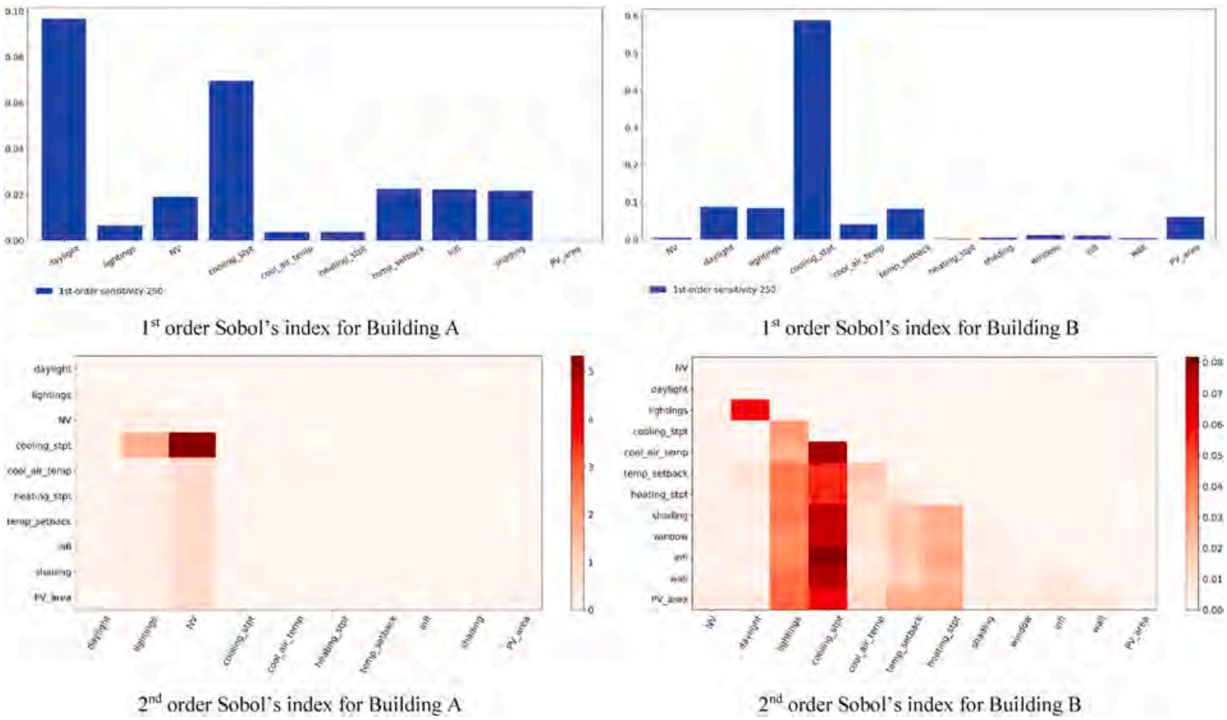


Fig. 7. Sobol sensitivity analysis results for the two buildings. stpt refers to setpoint, infl refers to air infiltration, and NV refers to natural ventilation.

cost. The total number of parameters in the analysis is 84 and includes factors such as the installation area and inclination angle of the PV system. The results of this analysis are depicted in Fig. 6 where the x-axis represents the energy savings in GJ per year, categorized as "abatement," and the y-axis represents the cost in NPV (in $ per GJ), categorized as "cost." The ECMs are then ranked according to their unit cost performance, from left to right.

Fig. 6 shows the performance of different ECMs with and without monetary investment. ECMs that do not increase future energy use have a positive MAC. The ranking of the ECMs in terms of NPV per gigajoule (GJ) varies between the two buildings, indicating different levels of performance for each. Any negative ECM with MAC is eliminated from consideration at this stage. For Building A, the eliminated ECMs are wall insulation, window replacement, and roof insulation, while roof insulation is the only ECM eliminated for Building B. The ranking for Building A begins with day lighting and follows with lighting system upgrade, natural ventilation, cooling

Table 9
Final selections of seven ECMs for the two buildings.

Ranking	Building A	Building B
1	Daylight control	Cooling Setpoint
2	Cooling setpoint	Daylight control
3	Unoccupied hour setback	Lighting system upgrade
4	Air infiltration level	Unoccupied hour setback
5	Window shading	PV area
6	Natural ventilation	Cooling supply air temperature
7	Lighting system upgrade	Window replacement

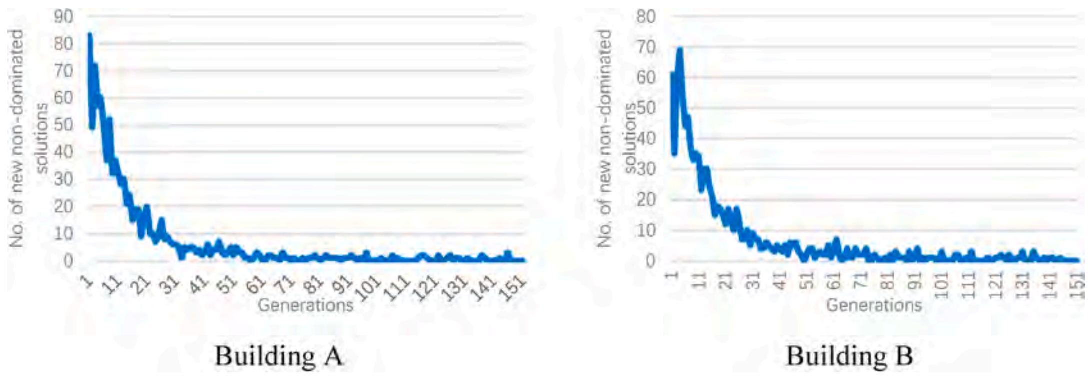


Fig. 8. Convergence of the optimization process.

setpoint, cooling supply air temperature, heating setpoint, unoccupied hour setpoint setback, air infiltration, window shadings, PV angle, and PV area. Meanwhile, for Building B, natural ventilation takes the top spot, followed by daylighting, lighting system upgrade, cooling setpoint, cooling supply air temperature, unoccupied hour setpoint setback, heating setpoint, window shadings, window replacement, air infiltration, PV angle, wall insulation, and PV area.

4.3.2. *Sobol sensitivity analysis results*

The computational time for the Sobol sensitivity analysis was 0.87 h for Building A and 1.32 h for Building B, respectively when N = 200 (the multiple of the unit population, which will be further elaborated in Section 5.1) using parallel computation. During the Sobol analysis phase, parameters related to the ECMs eliminated through MAC analysis are excluded. Additionally, PV angle parameters are also excluded since a 30-degree optimal inclination angle was determined during MAC analysis. The PV angle parameter for the PV system is considered an independent factor, and its values do not interact with other ECMs. As a result, only 9 ECMs are included in the variance-based sensitivity analysis for Building A, while Building B includes 12 ECMs, all of which were automatically generated by MAC analysis. The results of the variance-based sensitivity analysis are displayed in Fig. 7.

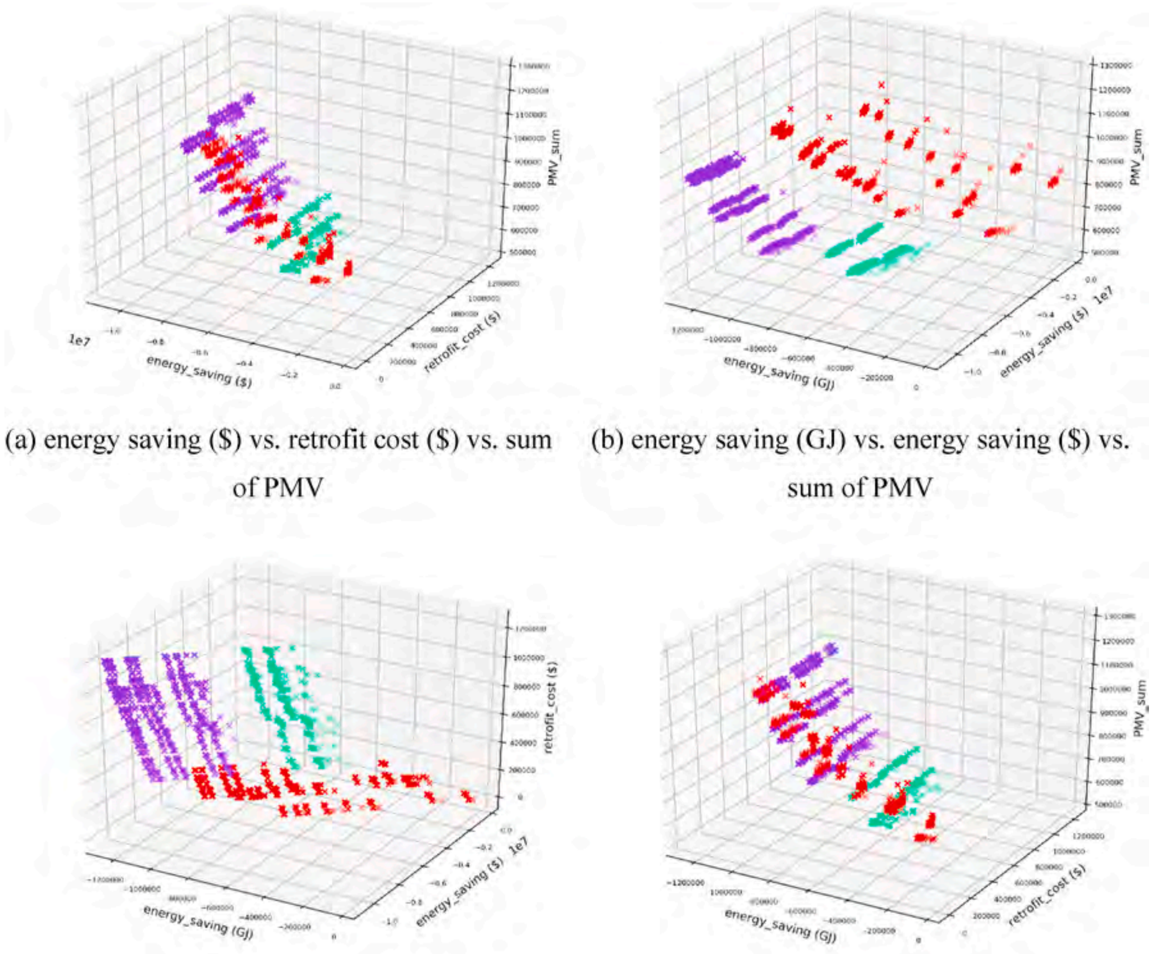
Based on the findings of global sensitivity analysis, it has been determined that the ECMs with the greatest influence on the NPV of buildings during their post-retrofit lifecycle can be identified by examining the 1st order Sobol's indices. In the case of Building A, the most impactful ECMs were identified as daylighting control and cooling setpoint, with the second tier ECMs being natural ventilation, unoccupied hour setpoint setback, air infiltration level, and window shading. In contrast, the most sensitive ECMs for Building B included cooling setpoint, daylighting control, lighting system upgrade, unoccupied hour setpoint setback, cooling supply air temperature, and PV area. These results serve as valuable insights for decision-makers hoping to optimize building performance and assess retrofit strategies.

The cooling setpoint is considered one of the most influential ECMs affecting the NPV of the two buildings under study. This can be attributed to the significant contribution of cooling energy to the total energy consumption of both buildings. However, the buildings differ in the sensitivity of their ECMs for heating energy use. For Building A, the infiltration level and heating setpoint are more effective in achieving the objective of maximizing the NPV compared to Building B. This is due to the higher proportion of heating energy to total energy use during winter in Building A, as compared to Building B in an annual perspective.

To ensure a streamlined optimization process and enable a fair comparison of results between the two buildings, we have limited the number of ECMs included in the analysis. Specifically, we have used Sobol sensitivity analysis to select seven ECMs for consideration. By focusing on 1st order Sobol's indices, we are able to capture both the individual impact of each ECM and its interactions with other ECMs. For a comprehensive overview of the chosen ECMs, please see Table 9 where we have listed the top seven selected for both buildings.

4.4. *Multi-objective optimization*

Fig. 8 displays the convergence plots for the optimization of both buildings. It is observed that the quantity of "newcomers,"



(a) energy saving ($) vs. retrofit cost ($) vs. sum of PMV

(b) energy saving (GJ) vs. energy saving ($) vs. sum of PMV

(c) energy saving (GJ) vs. energy saving ($) vs. retrofit cost ($)

(d) energy saving (GJ) vs. retrofit cost ($) vs. sum of PMV

Fig. 9. Obtained Pareto front of Building A (colors painted according to 1st layer clustering).

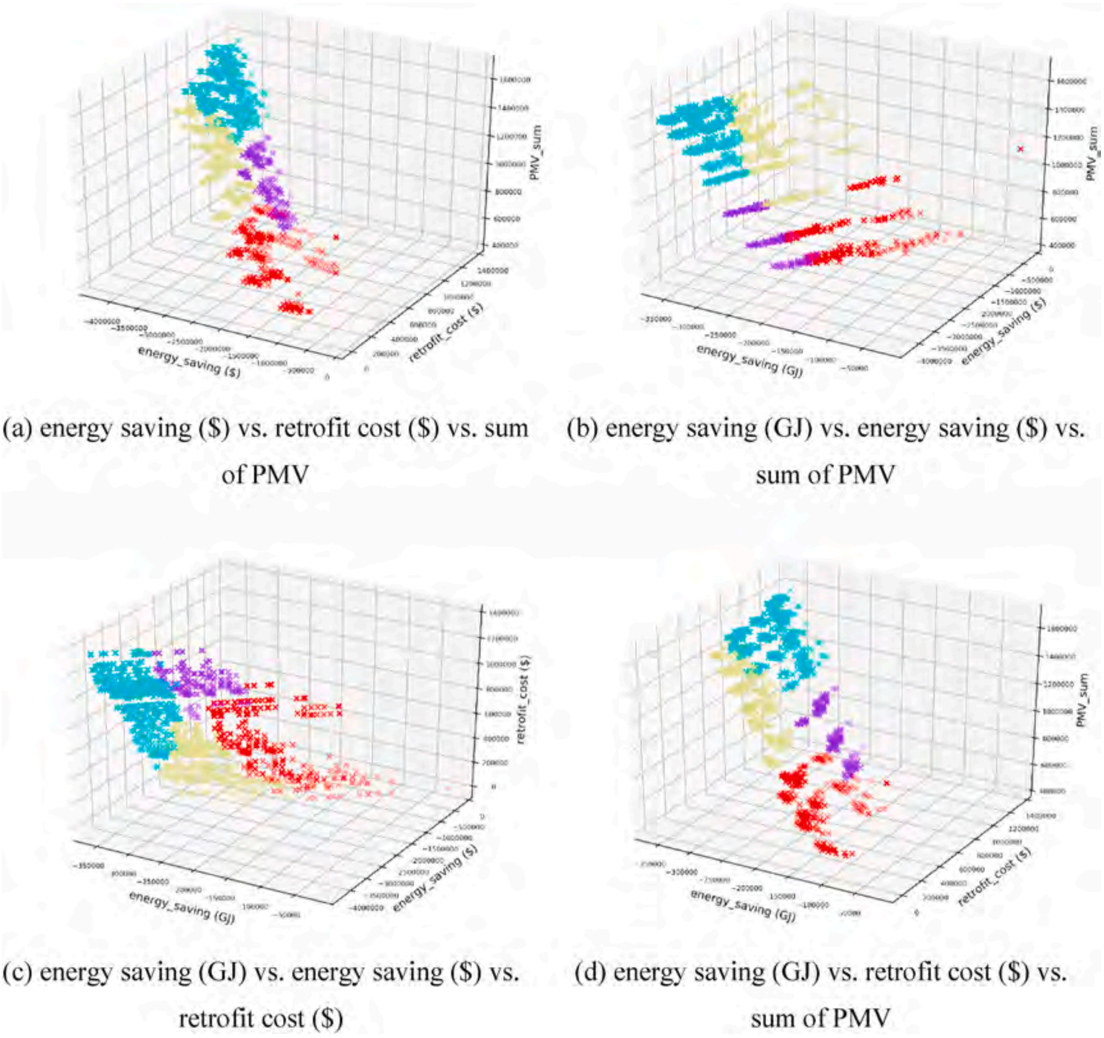
referring to the newfound non-dominated solutions in each generation, decreases rapidly after approximately 50 generations of iterations. The multi-objective optimization process took 1.51 h for Building A and 1.39 h for Building B using parallel computation. This demonstrates that the application of SimBldPy can potentially speed up the building simulation process, making it feasible to conduct extensive parametric studies and optimization tasks efficiently in 2 h.

Following optimization, the 3D Pareto front for each building is displayed in Figs. 9 and 10. As there are four subobjectives, four separate 3D subfigures are necessary to portray the complete picture of the resulting Pareto front. The colors utilized across the figures represent the 1st layer hierarchical clustering, which is intended to facilitate decision-making. The total number of non-dominated solutions obtained for Building A and Building B are 1008 and 1029, respectively.

The differences in the Pareto front between Building A and B are noteworthy. Building A's Pareto front is more scattered than that of Building B, particularly in regard to energy savings measured in GJ and dollars. This suggests that the energy savings in Building A may come from various utilities to similar degrees, while in Building B a particular utility may dominate the energy savings. Nevertheless, it is challenging to make practical decisions based on the plotted Pareto front alone, even with the utilization of clustering techniques. Consequently, we are exploring alternative visualization methods to aid decision making in the initial clustering layer.

4.5. Decision-making based on clustering technique

In this study, we employed hierarchical clustering techniques for retrofit decision-making support and decision tree algorithms for the retrospection of selected decision-making pathways. The rationale behind using hierarchical clustering lies in its ability to group similar retrofit strategies based on multiple criteria, providing a clear visualization of the trade-offs between different retrofit options. This approach allows stakeholders to easily identify and select optimal retrofit strategies that align with their specific priorities, such as



(a) energy saving ($) vs. retrofit cost ($) vs. sum of PMV

(b) energy saving (GJ) vs. energy saving ($) vs. sum of PMV

(c) energy saving (GJ) vs. energy saving ($) vs. retrofit cost ($)

(d) energy saving (GJ) vs. retrofit cost ($) vs. sum of PMV

Fig. 10. Obtained Pareto front of Building B (colors painted according to 1st layer clustering).

cost-effectiveness or energy efficiency. While traditional Multi-Criteria Decision-Making (MCDM) techniques are powerful and widely used, they often require extensive input information and can be less intuitive for stakeholders to understand. Hierarchical clustering offers a more flexible and interpretable method for visualizing and comparing retrofit options. Additionally, decision tree algorithms are utilized for the retrospection of selected decision-making pathways, allowing stakeholders to trace back the decision process and understand the rationale behind each choice. This combination of hierarchical clustering for decision-making support and decision trees for retrospection enhances the transparency and accessibility of the decision-making process, facilitating better communication and understanding among stakeholders.

Once the Pareto front has been obtained, we create a parallel coordinates plot to enhance the front's visual representation and aid decision-making. The parallel coordinates plot presents the selected ECMs' parameters and the four subobjectives in parallel, allowing us to analyze the Pareto front's distribution. We terminate clustering when the number of solutions in a subcluster falls below 30. In this study, we can accomplish this entire process using a three-layered clustering approach for both buildings.

For the two case study buildings, we present two distinct retrofitting mindsets that can be used to simulate the decision-making process - aggressive and balanced. We then share the results of the decision-making process and provide an explanation of the outcomes. The aggressive mindset entails the highest level of energy savings, despite the high investment cost. However, this focus may come at the expense of indoor thermal comfort, as it is not necessarily a priority in this approach. On the other hand, the balanced mindset prioritizes a constrained fiscal policy while still pursuing high cost-effective retrofitting options that also improve indoor thermal comfort. This approach seeks to strike a balance between achieving energy savings and maintaining a comfortable indoor environment. By tailoring retrofit strategies to different stakeholder priorities and mindsets, this approach can support informed and customized decision-making in real-world scenarios.

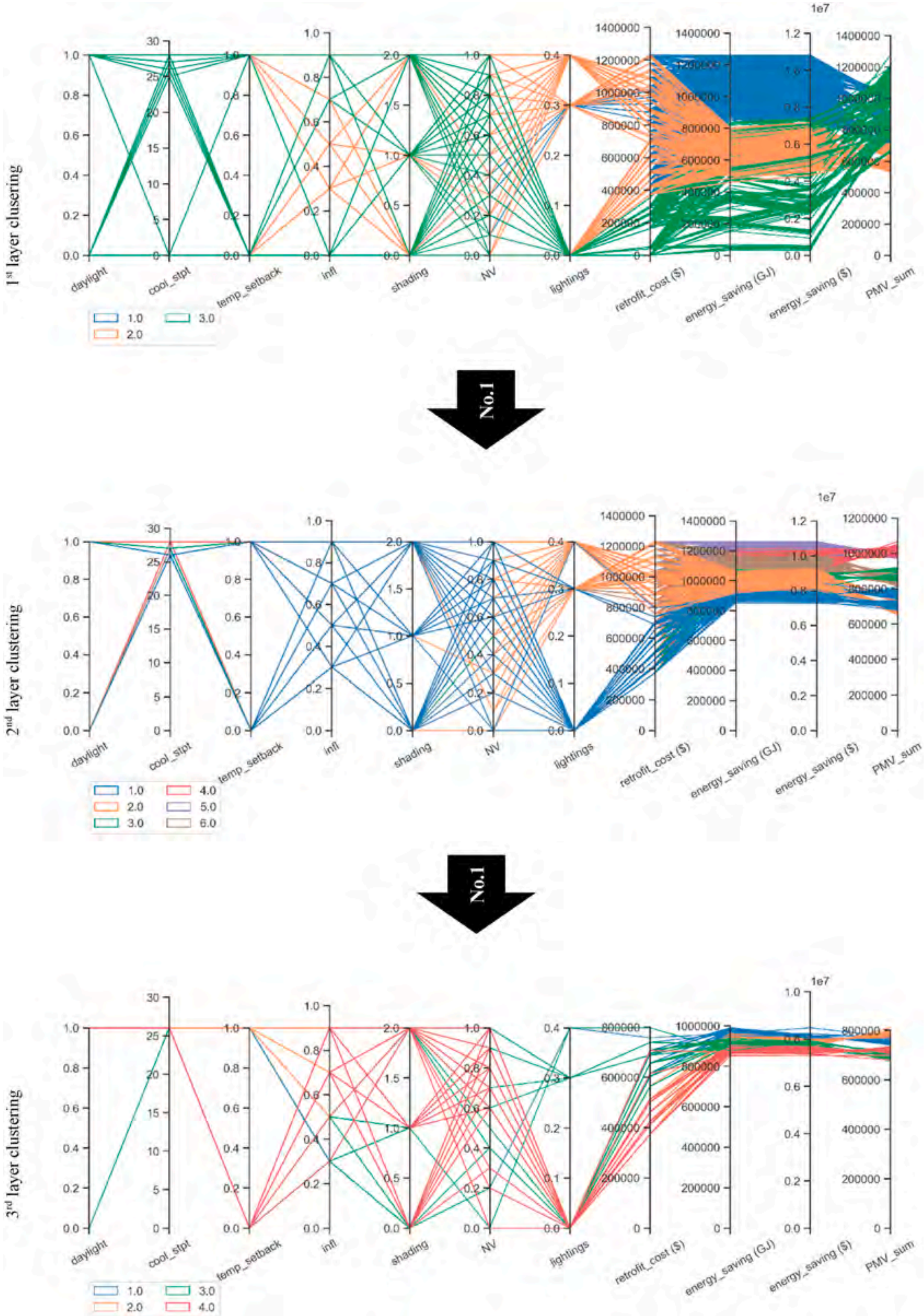


Fig. 11. Decision making based on hierarchical clustering for Building A (aggressive mindset). stpt refers to setpoint, infl refers to air infiltration, and NV refers to natural ventilation.

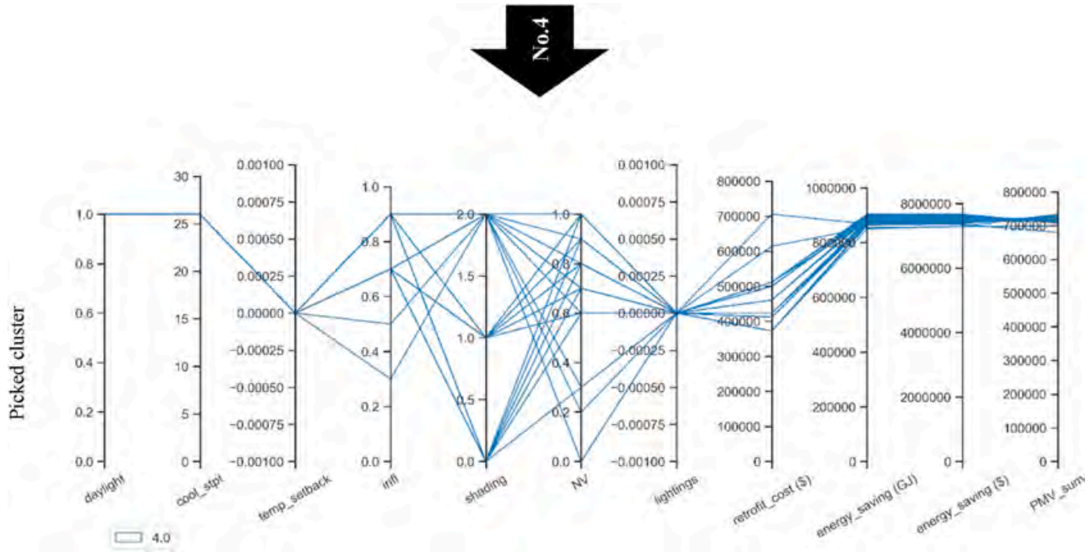


Fig. 11. (*continued*).

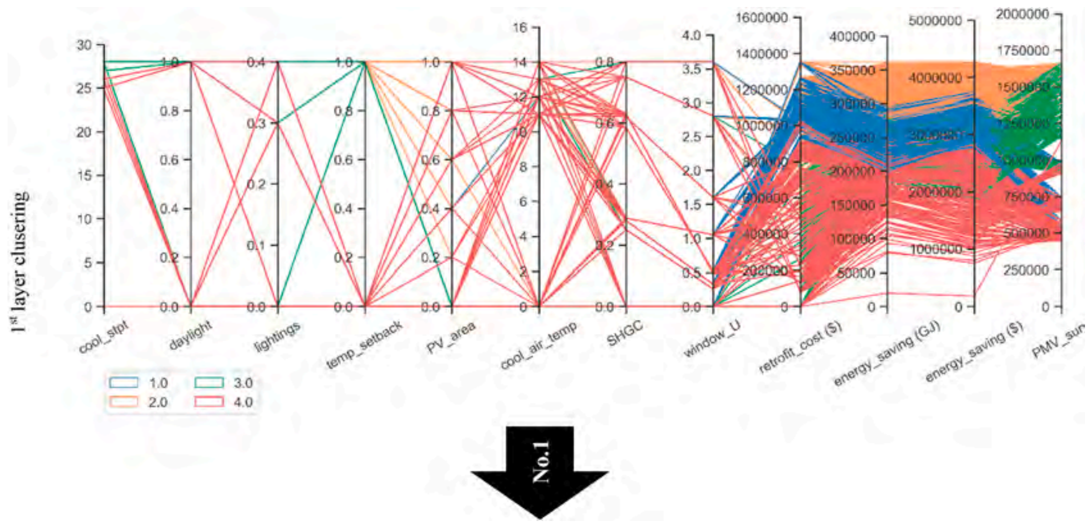


Fig. 12. Decision making based on hierarchical clustering for Building B (aggressive mindset). stpt refers to setpoint.

4.5.1. Decision-making results under an aggressive retrofitting mindset

Figs. 11 and 12 depict the decision-making pathways for two buildings that underwent an aggressive retrofitting approach. Cluster 1 is selected as the first layer of clustering for both buildings due to its comparatively higher economic return in energy saving during the lifecycle, as well as its high investment. Although Cluster 2 in Building B offers a higher economic return than Cluster 1, it is not chosen due to its less-than-ideal indoor thermal conditions when compared with Cluster 1.

In the second and third layers of clustering and decision making, it was found that Cluster 1 was the optimal choice for both buildings and layers. This is because Cluster 1 provided a good economic return with a relatively lower investment, and most importantly, it ensured the best indoor thermal comfort among all the clusters at this layer. It is interesting to note that both buildings followed the same decision-making pathway under the aggressive retrofitting mindset. However, it was later discovered that the decision-making pathway differed between the two buildings when a balanced mindset was implemented.

4.5.2. Decision-making results under a balanced retrofitting mindset

Figs. 13 and 14 illustrate the decision-making pathways for two buildings that are being retrofitted in a balanced manner. The outcomes of the decision-making process differ considerably for Building A and Building B under this approach.

Cluster 2 was chosen over Cluster 1 for retrofitting Building A due to its lower cost. However, Cluster 2 includes non-cost-effective

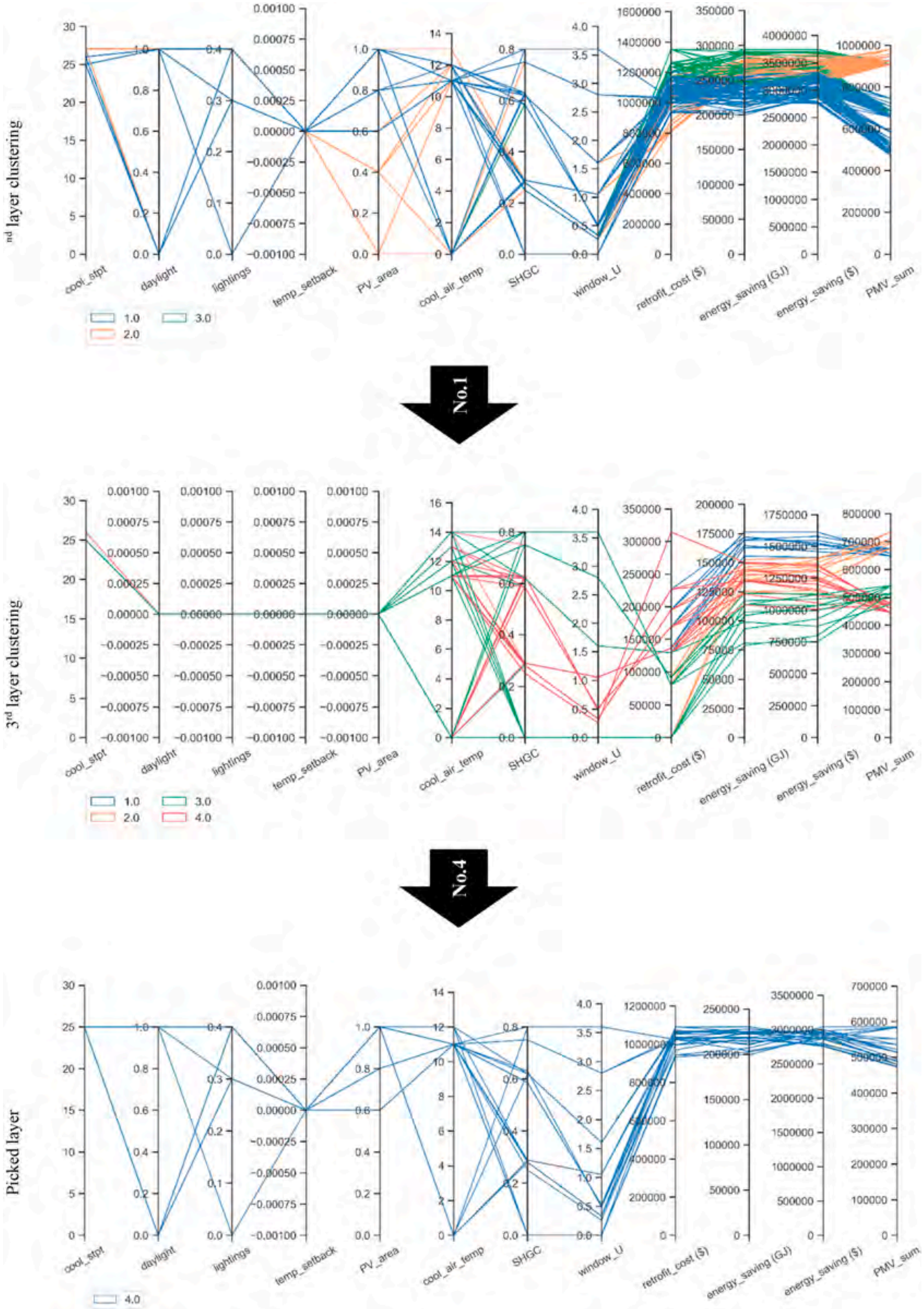


Fig. 12. (*continued*).

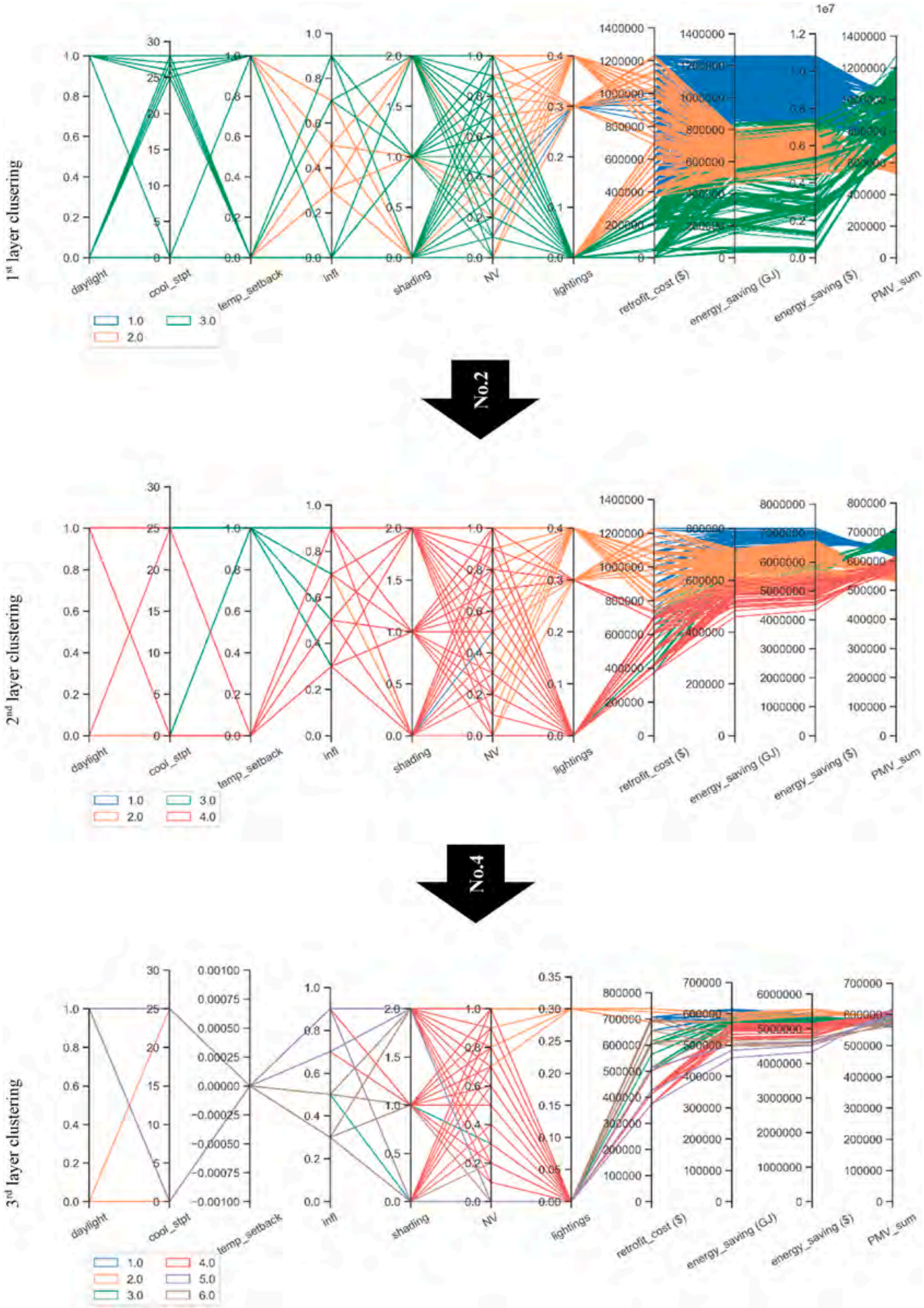


Fig. 13. Decision making based on hierarchical clustering for Building A (balanced mindset). stpt refers to setpoint, infl refers to air infiltration, and NV refers to natural ventilation.

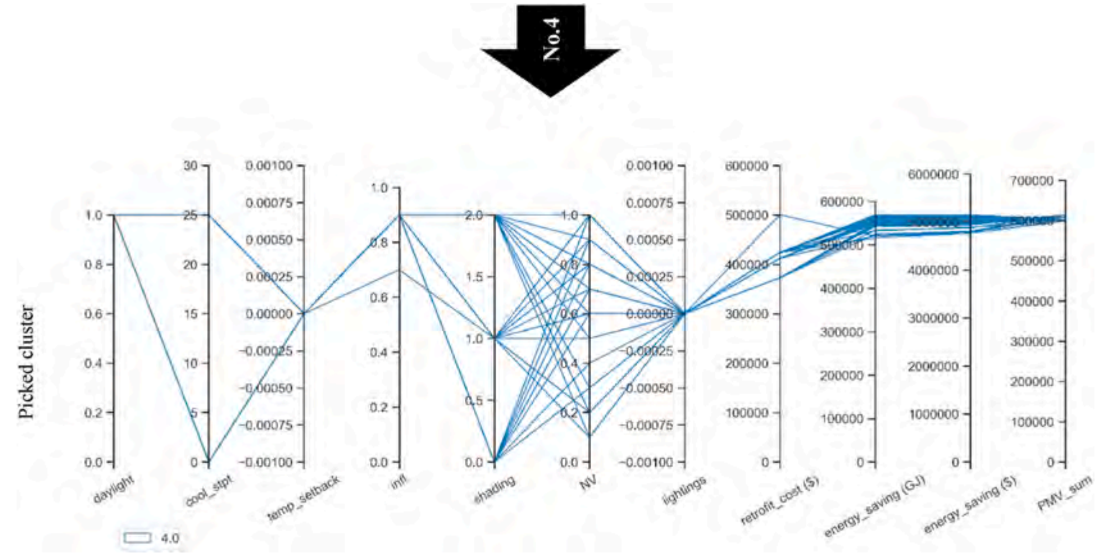


Fig. 13. (*continued*).

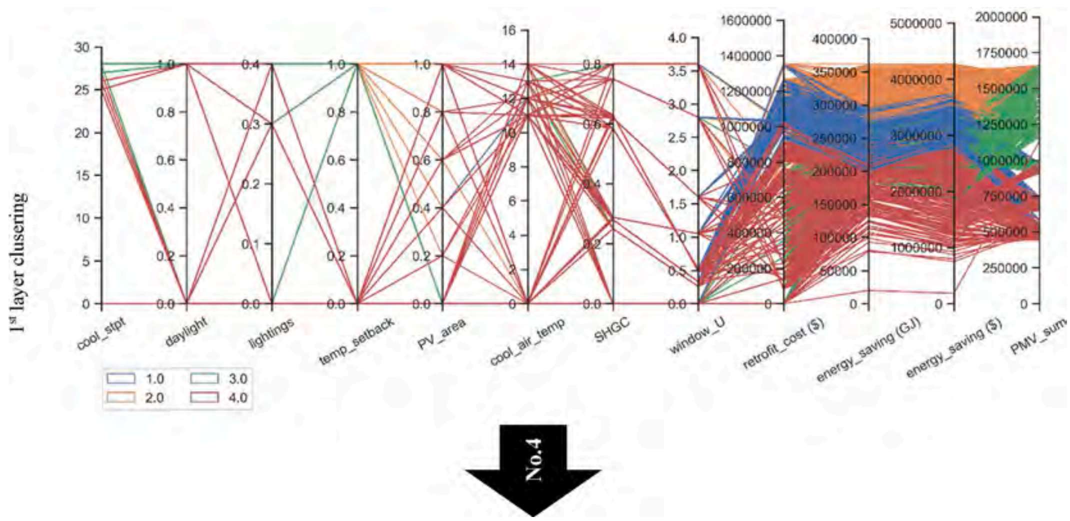


Fig. 14. Decision making based on hierarchical clustering for Building B (balanced mindset). stpt refers to setpoint.

solutions as well, as indicated by the negative slope from retrofit cost to energy saving. Despite this, Cluster 2 was still preferred over Cluster 3 because it offered better thermal comfort solutions and had high cost-effective retrofit options with a positive slope from retrofit cost to energy saving. Subsequently, in the second layer and third-layer clustering, Cluster 4 (in different sublayers) was selected for its superior cost-effective performance compared to other clusters. It should be noted that these high cost-effective retrofitting options were previously identified in Cluster 2 during the first-layer clustering process.

Cluster 4 is chosen as the initial choice for Building B due to its excellent indoor thermal comfort performance and low economic investment. Despite Cluster 3 having a low investment cost compared to the other clusters at this level, its thermal comfort performance is not optimal, and therefore it is not selected. In the second level of clustering, Cluster 2 is selected for its superior thermal comfort performance at a lower cost. This same rationale is applied again in the third-level clustering, selecting Cluster 3. The decision-making process for Building B reflects a conservative retrofitting approach, as retrofit bundles with no investment are included in the final result.

4.5.3. Final decisions based on two different mindsets

As demonstrated in the previous section regarding the decision-making process, it is apparent that trade-offs inevitably arise during this process. Ultimately, the responsibility of making subjective decisions based on varying propensities or mindsets falls upon the

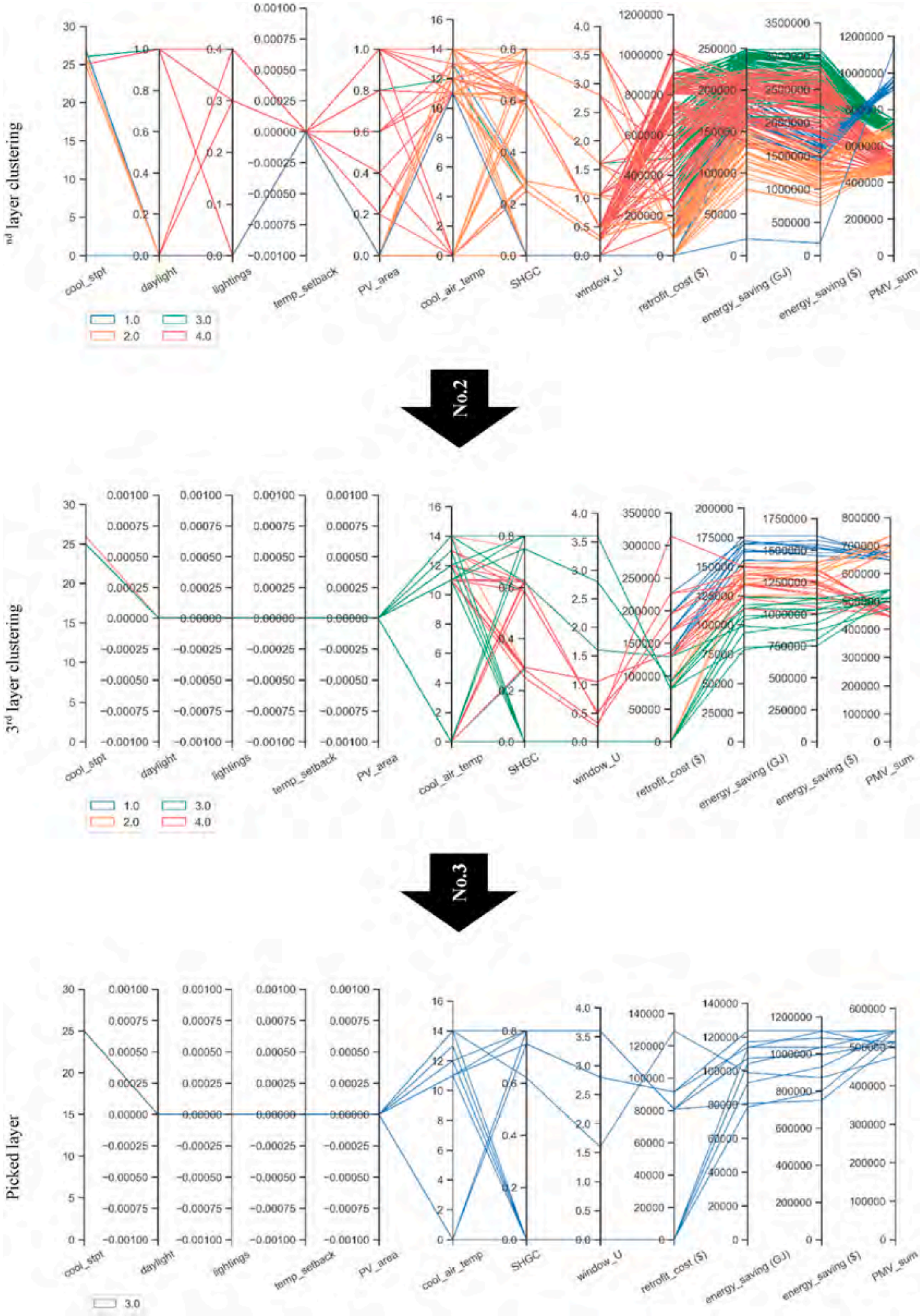


Fig. 14. (*continued*).

Table 10
Top five selected ECM combinations under two different mindsets for Building A.

	daylight	cooling setpoint	temperature setback	air infiltration	shading	natural ventilation	lightings	retrofit cost ($)	energy saving (GJ)	energy saving ($)	PMV sum
aggressive	1	26	0	0.7	2	1	0	511932	−903820	−7671121	706769
	1	26	0	0.7	2	0.8	0	511932	−901954	−7657479	707537
	1	26	0	0.7	2	0.7	0	511932	−900427	−7645867	707808
	1	26	0	0.7	2	0.6	0	511932	−897875	−7624925	708160
	1	26	0	0.7	1	1	0	500208	−896456	−7610690	713179
balanced	1	0	0	0.9	2	1	0	423795	−569203	−5148311	601649
	1	25	0	0.9	2	1	0	423795	−568930	−5146114	601651
	1	0	0	0.9	2	0.9	0	423795	−567723	−5135360	601294
	1	0	0	0.9	2	0.8	0	423795	−565980	−5130907	600915
	1	25	0	0.9	2	0.8	0	423795	−565698	−5128495	600948

Table 11
Top five selected ECM combinations under two different mindsets for Building B.

	cooling setpoint	daylight	lightings	temperature setback	PV area	cooling air temperature	Window SHGC	window U-value	retrofit cost ($)	energy saving (GJ)	energy saving ($)	PMV sum
aggressive	25	1	0.4	0	0.6	17	0.585	0.52	1088788	−230209	−2891855	486931
	25	1	0.4	0	0.6	17	0.288	1.05	1056512	−227692	−2945308	521726
	25	0	0.3	0	1	17	0.585	0.52	1064131	−227647	−2915738	479829
	25	0	0.3	0	1	17	0.288	1.05	1031855	−225254	−2970578	514116
	25	1	0.4	0	0.8	17	0.62	1.6	1060983	−224986	−2956162	512909
balanced	25	0	0	0	0	17	0.75	2.8	91398	−123621	−1123064	497866
	25	0	0	0	0	17	0.8	3.6	80734	−117087	−1084629	513535
	25	0	0	0	0	0	0.75	2.8	91398	−113901	−1033638	497866
	25	0	0	0	0	17	0	0	0	−113621	−1120479	542065
	25	0	0	0	0	18	0.8	3.6	80734	−107359	−995020	513535

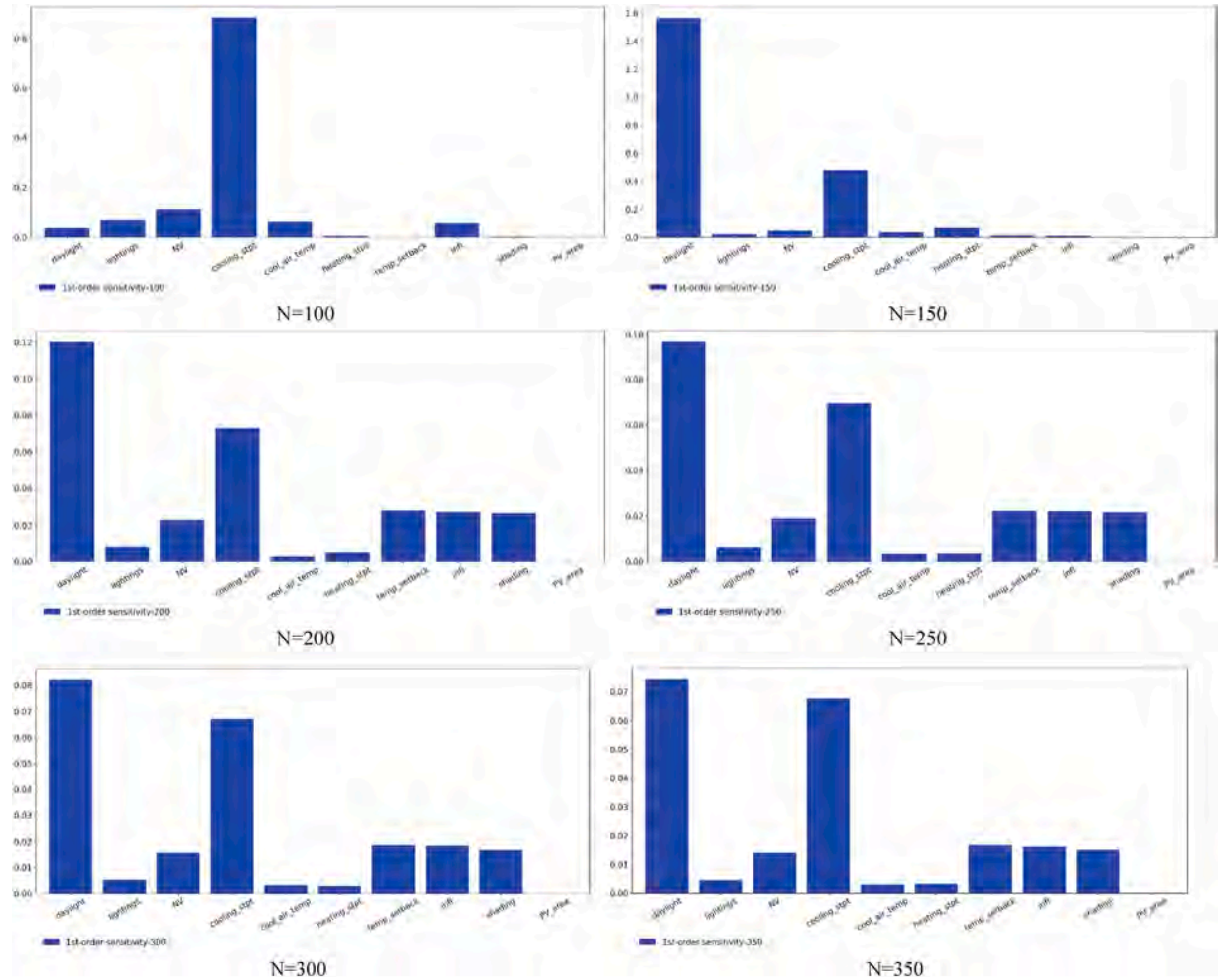


Fig. 15. Sampling size independence study of Sobol sensitivity analysis for Building A. stpt refers to setpoint, infl refers to air infiltration, and NV refers to natural ventilation.

decision makers. The purpose of our experiment, which investigates two different mindsets, aims to showcase that differing mindsets result in the generation of distinct non-dominated solutions during the final stage.

Tables 10 and 11 present the top five energy-saving solutions for each building and mindset. However, in the case of Building A, it was observed that unoccupied hour setpoint setback and lighting system upgrade were not considered for both aggressive and balanced mindsets. This could be attributed to the fact that unoccupied hour setback may result in poorer thermal comfort when compared with retrofit bundles that do not use this measure. Additionally, the lighting system upgrade may not be as cost-effective as other ECMs for this particular building. The level of air infiltration plays a significant role in the division of two sets of solutions into different mindsets. Therefore, the difference in retrofit costs is mainly due to the variations in air tightness levels. Building A, which we previously identified as a climate-dominated building, can achieve greater energy savings with better airtightness. This is especially important given the chilly winters in Philadelphia, which necessitates having a well-insulated building envelope.

Building B commonly uses specific parameters for selected ECMs, such as cooling supply air temperature of 17 °C, a cooling setpoint of 25 °C, and no unoccupied hour setback. However, when it comes to other ECMs, two decision-making mindsets lead to dramatically different options. For those with an aggressive mindset, daylighting and lighting system upgrades are crucial, and increasing the PV area is also essential to maximize energy savings. Unfortunately, these measures, along with window replacements of lower SHGC and U-values, come with a significantly high initial investment. These three ECMs are the primary reason the ten ECMs may fall into two mindset categories. Furthermore, a marginal diminishing effect of pursuing energy-saving returns is evident by comparing the two different sets of solutions. While the cost for retrofit bundles in the aggressive mindset group is ten times higher than that of the balanced mindset group, the resulting energy savings only double or triple. Interestingly, the indoor thermal comfort level is similar in both groups. This outcome implies that the balanced mindset group is more adept in weighing the tradeoffs between retrofit costs and energy savings.

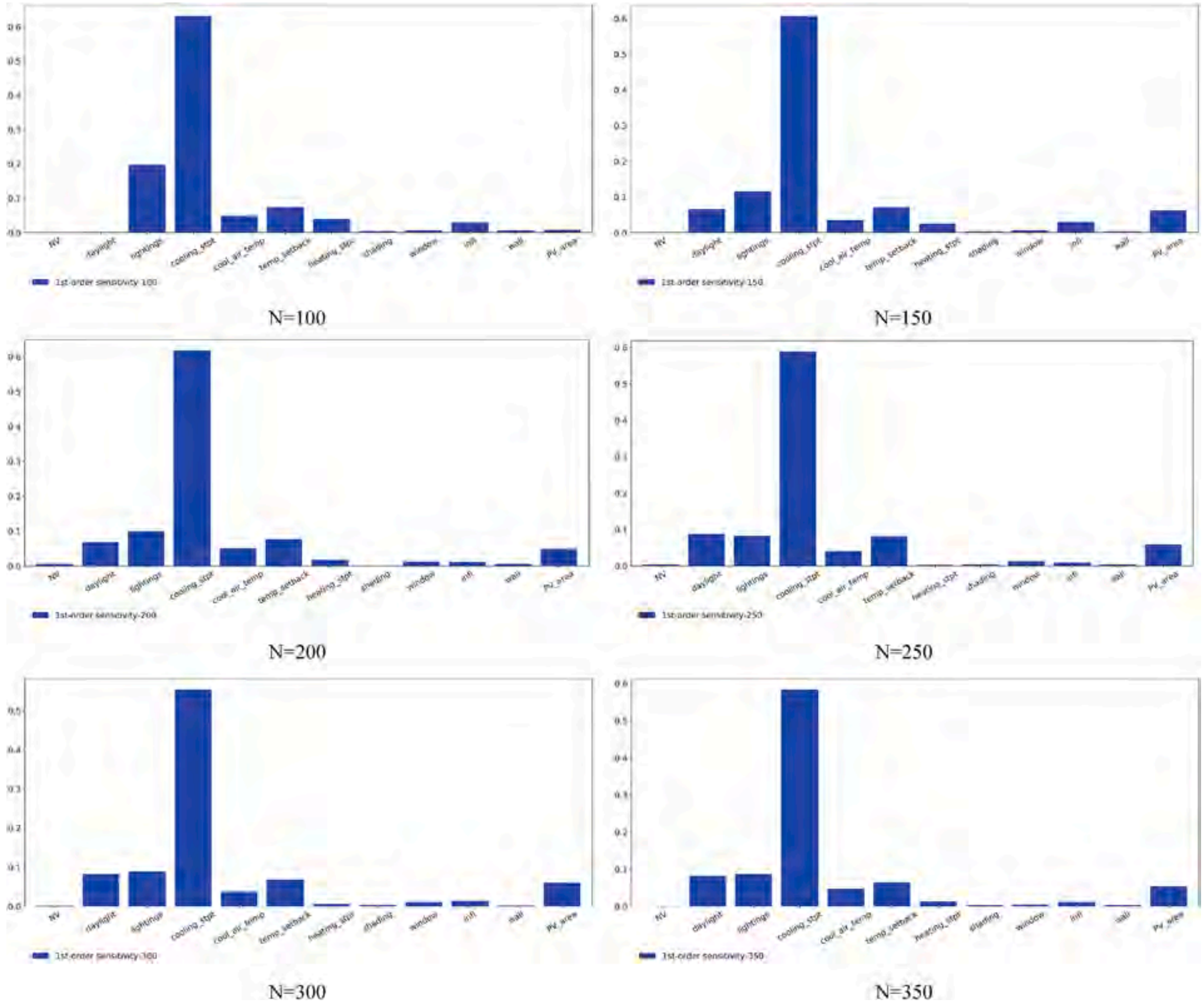


Fig. 16. Sampling size independence study of Sobol sensitivity analysis for Building B. stpt refers to setpoint.

5. Discussions

5.1. Application of sobol sensitivity analysis

During Sobol sensitivity analysis, it was found that Building A involved 9 ECMs in the analysis whereas Building B involved 12. The term "unit population" refers to the total number of parameters of all ECMs. At this stage, it was observed that the unit population of Building B is larger than that of Building A.

The precision of the final rankings of each input variable in variance-based sensitivity analysis depends greatly on the sampling size. If the sampling size is too small, Sobol analysis will not capture enough information to determine the variances that reflect the true relationships among the variables. Conversely, if the sampling size is too large, the computation becomes difficult to undertake. Therefore, it is essential to determine an optimal sampling size that balances the need for obtaining enough variance information with computational efficiency for this study.

As is commonly acknowledged, performing a grid independence study is a crucial element of conducting a computational fluid dynamic (CFD) simulation [71]. This approach serves as a means of verifying that a given solution is not reliant upon the size of the geometry grid. The primary objective of this type of study is to demonstrate that the mesh used is sufficiently small such that further reductions in size would not result in a divergent solution. In the present investigation, a similar framework is utilized to identify an appropriate sample size that would ensure sensitivity analysis results were independent. The distinguishing factor of the current study is that the sample size is increased based on the unit population specified in the previous paragraph.

In this study, we explored the impact of varying N, the multiple of the unit population, on the results of sensitivity analysis for each case study building. Specifically, we tested six different values of N: 100, 150, 200, 250, 300, and 350. Our results, as shown in Figs. 15 and 16, indicate that a sampling size smaller than 200 times the unit population can lead to skewed sensitivity analysis results. However, when N is larger than 250, the Sobol sensitivity results can be considered to have reached certain independence. This

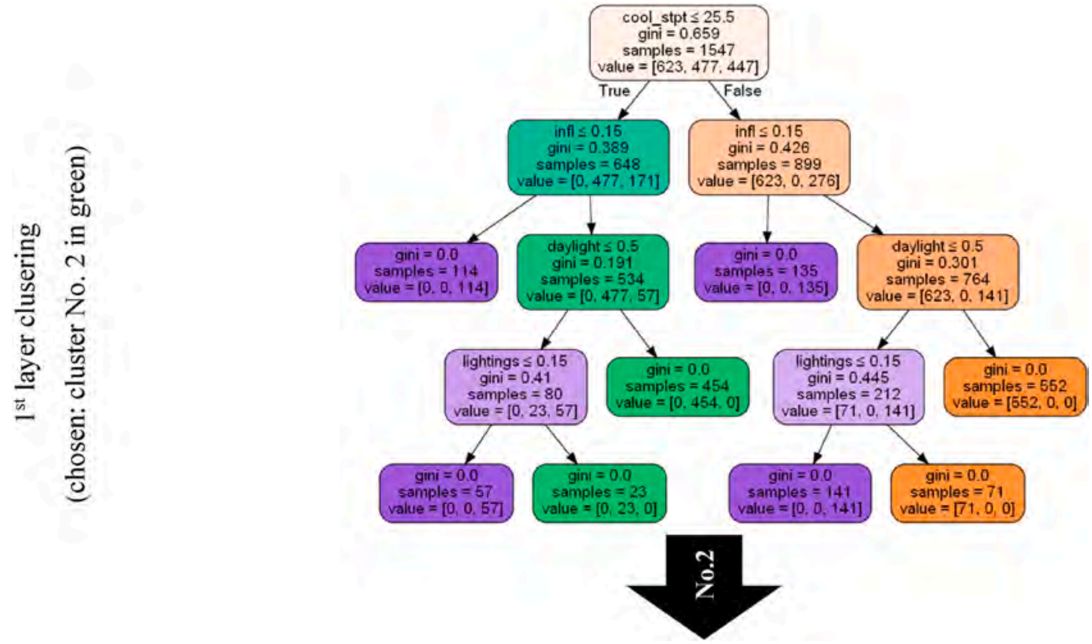


Fig. 17. Decision making retrospection based on decision tree for Building An under balanced mindset. stpt refers to setpoint, infl refers to air infiltration, and NV refers to natural ventilation.

suggests that a larger sampling size than 250-unit populations will have limited marginal benefit. Based on this finding, we determined that N should be set at 250 for the sensitivity analysis in this study.

The sensitivity results for the two buildings suggest that local sensitivity analysis, such as the use of MAC for building retrofitting problems, is insufficient. Due to the complexity and non-linearity of buildings, it is necessary to analyze 2nd order Sobol indices in order to determine how the energy savings brought about by one ECM may be affected by the introduction of another. Therefore, global sensitivity analysis is essential for building retrofit. Additionally, analysis of the results shown in Fig. 7 reveals that each individual building is unique. The Sobol sensitivity analysis conducted in our case studies indicates that the two selected buildings have differing priorities for ECM selection and exhibit distinct feedback when presented with multiple ECMs (as evidenced by variations in 2nd order Sobol analysis). It is thus imperative to approach each existing building as a unique, independent case in retrofitting analysis.

5.2. Decision making retrospection

A proposed method for identifying the decision-making process is the implementation of a decision-tree based backtracking method after obtaining the decision-making pathway. This method involves clustering to determine how the decision is made. Hierarchical clustering algorithms are used in combination with decision trees to analyze the logic behind the clustering of each cluster with respect to ECM parameters. To illustrate this approach, we have provided an example of the decision-making results for Building An under a balanced mindset. Fig. 17 displays a tree plot of the retrospection results.

The decision-making process for selected Cluster 2 (in green leaves) can be easily understood when the clustering results are visualized with a decision tree model, as shown in Fig. 16. The first-layer clustering traceback indicates that the process is based on ECM parameters: [cooling setpoint $\leq$ 25 °C & air infiltration level $>$0.15 & daylighting control not implemented & lighting system upgrade implemented], or [cooling setpoint $\leq$ 25 °C & air infiltration level $>$0.15 & daylighting control implemented]. The visualization of the clustering results with the decision tree model makes the decision-making process more intuitive.

During the later stages of the second-layer clustering process, a decision was made regarding ECM parameters for Cluster 4. The decision was either to leave the unoccupied hour setback and lighting system upgrades unimplemented or to implement the lighting system upgrade while leaving the unoccupied hour setback unimplemented. Additionally, if the lighting system upgrade was implemented, the decision was to also leave daylighting control unimplemented and maintain an infiltration level greater than 0.6.

The third layer for Cluster 3 is characterized by an infiltration level of 0.9, with no implementation of a lighting system upgrade but natural ventilation is in place. To draw conclusions for decision-making, the conditions for each layer's model with overlapping nodes must be merged into the final selected solution set. As such, the selected solution set consists of the following conditions: cooling setpoint of 25 °C or below, implementation of daylighting control, no implementation of lighting system upgrade, no implementation of unoccupied hour setback, infiltration level of 0.9, and implementation of natural ventilation. The use of a decision tree-based retrospection method can provide clarity on the decision-making pathways taken and serve as a useful tool for gaining a better intuitive understanding of how each cluster is formed by the uninterpretable machine learning method of hierarchical clustering.

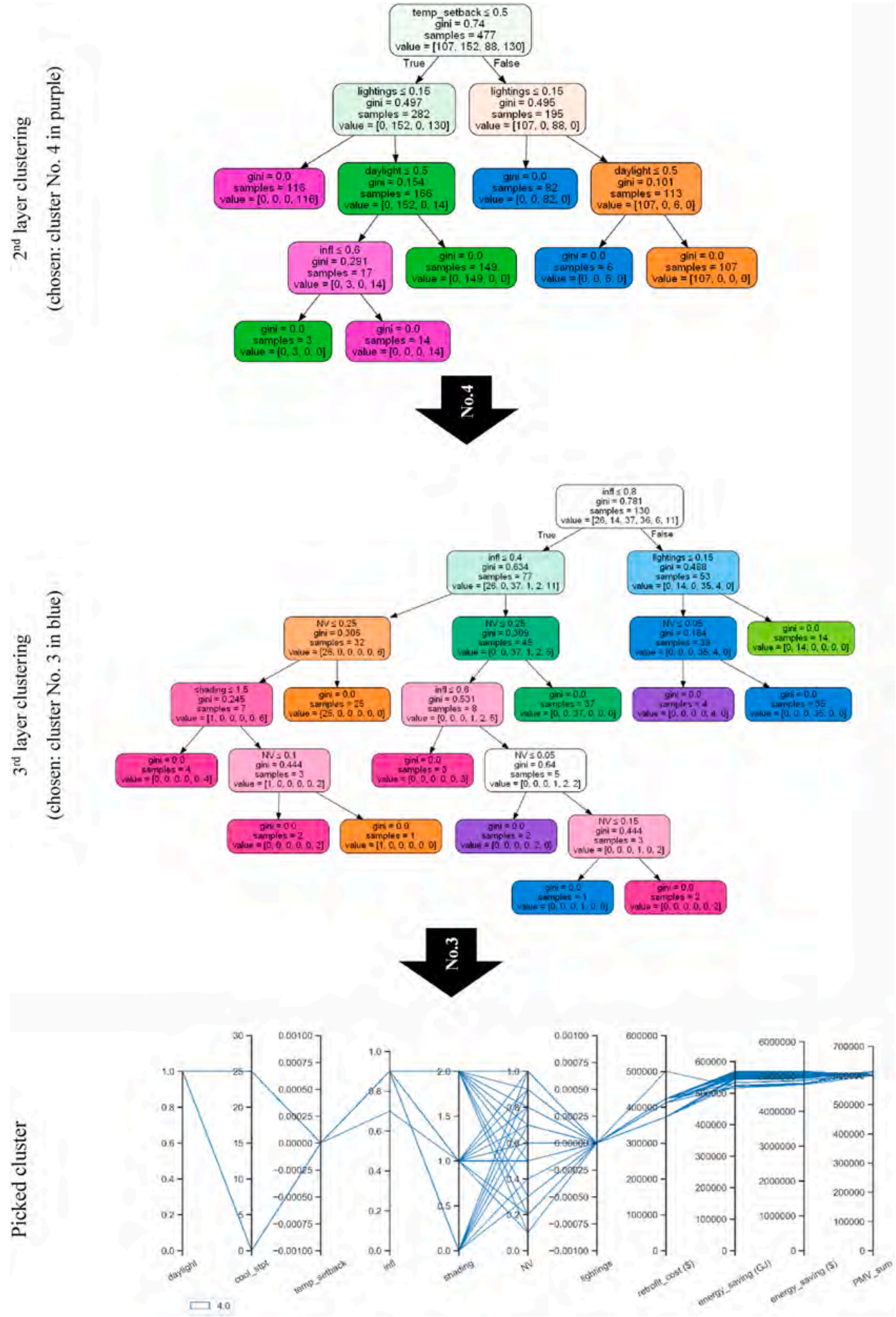


Fig. 17. (*continued*).

6. Conclusions

This study proposes an automated multi-objective optimization scheme for building retrofit, considering future climate change scenarios. The scheme involves three major stages: pre-optimization feature selection, multi-objective optimization, and post-optimization Pareto front and decision making. To facilitate the process of building simulation, a lightweight simulation tool was developed by the research team, based on resistor-capacitor thermal modeling method [36]. Two educational buildings with different thermal characteristics (one is climate dominated, and the other is internal load dominated) were selected as case studies. The application of SimBldPy can significantly speed up the building simulation process. The proposed hierarchical clustering method allows users to identify targeted solution sets for specific niches within the front, while maintaining different mindsets. Additionally, the decision-tree based retrospection method can reconstruct the process of how each cluster was conducted, using detailed ECM parameters. While subjectivity may be present in the decision-making process, this method offers visualization, preservation, and restoration of the entire decision-making pathway. As a result, the decision-making process for multi-objective solutions is more transparent and intuitive. In summary, the main findings of this research can be concluded as following points:

- This study successfully integrated future climate projections into the building energy retrofit optimization process, enhancing the adaptability and resilience of retrofit strategies.
- The use of SimBldPy, a simplified resistor-capacitor (RC) modeling tool, demonstrated significant reductions in computational costs while maintaining comparable accuracy to more complex white-box simulation tools like EnergyPlus.
- The implementation of the non-dominated sorting differential evolution (NSDE) algorithm effectively addressed the multi-objective optimization problem, balancing trade-offs between energy savings, cost, and indoor thermal comfort.
- The hierarchical clustering technique and decision tree algorithms provided robust decision-making support, allowing stakeholders to visualize and understand the trade-offs among different retrofit options.
- The application of the methodology to the two case study buildings demonstrated the practical applicability and effectiveness in optimizing building energy retrofits under future climate conditions.

Despite the promising application potential of the proposed framework for existing building retrofit in the future, this study has several limitations. Firstly, the use of a simplified resistor-capacitor (RC) modeling approach in SimBldPy, while computationally efficient, may not capture all the complexities of building energy dynamics compared to more detailed simulation tools like EnergyPlus. Secondly, the study relies on specific global climate models (GCMs) for future climate projections, which may introduce uncertainties due to model variability and ongoing emission progress. Future research should focus on integrating more detailed simulation models to enhance accuracy, exploring the use of multiple GCMs to address uncertainties, and evaluating retrofit strategies under a wider range of climate scenarios. Furthermore, expanding the methodology to include more diverse building types and climate zones would enhance the generalizability of the findings. Incorporating real-time data and adaptive control strategies in the future could also improve the robustness and responsiveness of the retrofit solutions.

CRediT authorship contribution statement

Pengyuan Shen: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

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.

Acknowledgement

The work presented in this paper was partly funded by National Natural Science Foundation of China (52008132).

References

[1] P. Dechamps, The IEA World Energy Outlook 2022–a brief analysis and implications, European Energy & Climate Journal 11 (3) (2023) 100–103.
[2] X. Zhang, Z. Wu, Y. Feng, P. Xu, "Turning green into gold": a framework for energy performance contracting (EPC) in China's real estate industry, J. Clean. Prod. 109 (2015) 166–173.
[3] IPCC, IPCC Fifth assessment synthesis report-climate change 2014 synthesis Report, in: IPCC Fifth Assessment Synthesis Report-Climate Change 2014 Synthesis Report, 2014.
[4] Z.J. Zhai, J.M. Helman, Implications of climate changes to building energy and design, Sustain. Cities Soc. 44 (2018).
[5] P. Shen, M. Wang, J. Liu, Y. Ji, Hourly air temperature projection in future urban area by coupling climate change and urban heat island effect, Energy Build. 279 (2023) 112676.
[6] Y. Chen, Z. Ren, Z. Peng, J. Yang, Z. Chen, Z. Deng, Impacts of climate change and building energy efficiency improvement on city-scale building energy consumption, J. Build. Eng. 78 (2023) 107646.
[7] P. Shen, W. Braham, Y. Yi, The feasibility and importance of considering climate change impacts in building retrofit analysis, Appl. Energy 233–234 (2019) 254–270.

[8] P. Shen, H. Wang, Archetype building energy modeling approaches and applications: a review, Renew. Sustain. Energy Rev. 199 (2024) 114478.
[9] B. Wang, X. Xia, Optimal maintenance planning for building energy efficiency retrofitting from optimization and control system perspectives, Energy Build. 96 (2015) 299–308.
[10] B. Wang, X. Xia, J. Zhang, A multi-objective optimization model for the life-cycle cost analysis and retrofitting planning of buildings, Energy Build. 77 (2014) 227–235.
[11] A. Jafari, V. Valentin, An optimization framework for building energy retrofits decision-making, Build. Environ. 115 (2017) 118–129.
[12] F. Ascione, N. Bianco, G. Maria Mauro, D.F. Napolitano, Building envelope design: multi-objective optimization to minimize energy consumption, global cost and thermal discomfort. Application to different Italian climatic zones, Energy 174 (2019) 359–374.
[13] Y. Shao, P. Geyer, W. Lang, Integrating requirement analysis and multi-objective optimization for office building energy retrofit strategies, Energy Build. 82 (2014) 356–368.
[14] E. Asadi, M.G. da Silva, C.H. Antunes, L. Dias, A multi-objective optimization model for building retrofit strategies using TRNSYS simulations, GenOpt and MATLAB, Build. Environ. 56 (2012) 370–378.
[15] A.M. Rysanek, R. Choudhary, A decoupled whole-building simulation engine for rapid exhaustive search of low-carbon and low-energy building refurbishment options, Build. Environ. 50 (2012) 21–33.
[16] B. Eisenhower, Z. O'Neill, S. Narayanan, V.A. Fonoberov, I. Mezić, A methodology for meta-model based optimization in building energy models, Energy Build. 47 (2012) 292–301.
[17] G.M. Mauro, M. Hamdy, G.P. Vanoli, N. Bianco, J.L.M. Hensen, A new methodology for investigating the cost-optimality of energy retrofitting a building category, Energy Build. 107 (2015) 456–478.
[18] S.E. Chidiac, E.J.C. Catania, E. Morofsky, S. Foo, A screening methodology for implementing cost effective energy retrofit measures in Canadian office buildings, Energy Build. 43 (2) (2011) 614–620.
[19] E. Asadi, M.G. da Silva, C.H. Antunes, L. Dias, Multi-objective optimization for building retrofit strategies: a model and an application, Energy Build. 44 (0) (2012) 81–87.
[20] V. Siddharth, P.V. Ramakrishna, T. Geetha, A. Sivasubramaniam, Automatic generation of energy conservation measures in buildings using genetic algorithms, Energy Build. 43 (10) (2011) 2718–2726.
[21] F. Roberti, U.F. Oberegger, E. Lucchi, A. Troi, Energy retrofit and conservation of a historic building using multi-objective optimization and an analytic hierarchy process, Energy Build. 138 (2017) 1–10.
[22] S. Tadeu, C. Rodrigues, A. Tadeu, F. Freire, N. Simões, Energy retrofit of historic buildings: environmental assessment of cost-optimal solutions, J. Build. Eng. 4 (2015) 167–176.
[23] H. Son, C. Kim, Evolutionary multi-objective optimization in building retrofit planning problem, Procedia Eng. 145 (2016) 565–570.
[24] R. Wu, G. Mavromatidis, K. Orehounig, J. Carmeliet, Multiobjective optimisation of energy systems and building envelope retrofit in a residential community, Appl. Energy 190 (2017) 634–649.
[25] E. Asadi, M.G.d. Silva, C.H. Antunes, L. Dias, L. Glicksman, Multi-objective optimization for building retrofit: a model using genetic algorithm and artificial neural network and an application, Energy Build. 81 (2014) 444–456.
[26] A.M. Rysanek, R. Choudhary, Optimum building energy retrofits under technical and economic uncertainty, Energy Build. 57 (2013) 324–337.
[27] F.P. Chantrelle, H. Lahmidi, W. Keilholz, M.E. Mankibi, P. Michel, Development of a multicriteria tool for optimizing the renovation of buildings, Appl. Energy 88 (4) (2011) 1386–1394.
[28] O. Pombo, K. Allacker, B. Rivela, J. Neila, Sustainability assessment of energy saving measures: a multi-criteria approach for residential buildings retrofitting—a case study of the Spanish housing stock, Energy Build. 116 (2016) 384–394.
[29] F. Ascione, N. Bianco, R.F. De Masi, G.M. Mauro, G.P. Vanoli, Resilience of robust cost-optimal energy retrofit of buildings to global warming: a multi-stage, multi-objective approach, Energy Build. 153 (2017) 150–167.
[30] F. Ascione, N. Bianco, T. Iovane, G.M. Mauro, D.F. Napolitano, A. Ruggiano, L. Viscido, A real industrial building: modeling, calibration and Pareto optimization of energy retrofit, J. Build. Eng. 29 (2020) 101186.
[31] Y. He, N. Liao, J. Bi, L. Guo, Investment decision-making optimization of energy efficiency retrofit measures in multiple buildings under financing budgetary restraint, J. Clean. Prod. 215 (2019) 1078–1094.
[32] F. Mostafazadeh, S.J. Eirdmousa, M. Tavakolan, Energy, economic and comfort optimization of building retrofits considering climate change: a simulation-based NSGA-III approach, Energy Build. 280 (2023) 112721.
[33] S. Liu, Y. Wang, X. Liu, L. Yang, Y. Zhang, J. He, How does future climatic uncertainty affect multi-objective building energy retrofit decisions? Evidence from residential buildings in subtropical Hong Kong, Sustain. Cities Soc. 92 (2023) 104482.
[34] B. Gao, X. Zhu, J. Ren, J. Ran, M.K. Kim, J. Liu, Multi-objective optimization of energy-saving measures and operation parameters for a newly retrofitted building in future climate conditions: a case study of an office building in Chengdu, Energy Rep. 9 (2023) 2269–2285.
[35] C. Zhuang, R. Choudhary, A. Mavrogianni, Uncertainty-based optimal energy retrofit methodology for building heat electrification with enhanced energy flexibility and climate adaptability, Appl. Energy 341 (2023) 121111.
[36] P. Shen, W. Braham, Y. Yi, E. Eaton, Rapid multi-objective optimization with multi-year future weather condition and decision-making support for building retrofit, Energy 172 (2019) 892–912.
[37] Q. Li, L. Zhang, L. Zhang, X. Wu, Optimizing energy efficiency and thermal comfort in building green retrofit, Energy 237 (2021) 121509.
[38] F. Shadram, S. Bhattacharjee, S. Lidelöw, J. Mukkavaara, T. Olofsson, Exploring the trade-off in life cycle energy of building retrofit through optimization, Appl. Energy 269 (2020) 115083.
[39] F. Rosso, V. Ciancio, J. Dell'Olmo, F. Salata, Multi-objective optimization of building retrofit in the Mediterranean climate by means of genetic algorithm application, Energy Build. 216 (2020) 109945.
[40] D.B. Crawley, L.K. Lawrie, F.C. Winkelmann, W.F. Buhl, Y.J. Huang, C.O. Pedersen, R.K. Strand, R.J. Liesen, D.E. Fisher, M.J. Witte, J. Glazer, EnergyPlus: creating a new-generation building energy simulation program, Energy Build. 33 (4) (2001) 319–331.
[41] S. Li, M. Wang, P. Shen, X. Cui, L. Bu, R. Wei, L. Zhang, C. Wu, Energy saving and thermal comfort performance of passive retrofitting measures for traditional rammed earth house in lingnan, China, Buildings 12 (10) (2022) 1716.
[42] S.A. Klein, e. al, Trnsys 17: a transient system simulation program, in: Solar Energy Laboratory, University of Wisconsin, Madison, USA, 2010.
[43] J.J. Hirsch, in: L.B.N. Laboratory (Ed.), eQuest, U.S., 2016.
[44] P. Shen, W. Braham, Y. Yi, Development of a lightweight building simulation tool using simplified zone thermal coupling for fast parametric study, Appl. Energy 223 (2018) 188–214.
[45] K. Oleson, Contrasts between Urban and rural climate in CCSM4 CMIP5 climate change scenarios, J. Clim. 25 (5) (2012) 1390–1412.
[46] D. Carvalho, S. Rafael, A. Monteiro, V. Rodrigues, M. Lopes, A. Rocha, How well have CMIP3, CMIP5 and CMIP6 future climate projections portrayed the recently observed warming, Sci. Rep. 12 (1) (2022) 11983.
[47] IPCC, in: T.F. Stocker, D. Qin (Eds.), Climate Change 2013– the Physical Science Basis, Fifth Assessment Report of the IPCC, 2013. New York.
[48] S.E. Belcher, J.N. Hacker, D.S. Powell, Constructing design weather data for future climates, BUILDING SERV ENG RES TECHNOL 26 (1) (2005) 49–61.
[49] P. Shen, Impacts of climate change on U.S. building energy use by using downscaled hourly future weather data, Energy Build. 134 (2017) 61–70.
[50] A.L.S. Chan, Developing future hourly weather files for studying the impact of climate change on building energy performance in Hong Kong, Energy Build. 43 (10) (2011) 2860–2868.
[51] DOE, Building Component Cost Community, DOE, 2015.
[52] N.R.E. Laboratory, in: N.R.E. Laboratory (Ed.), National Residential Efficiency Measures Database, 2013. Golden, CO.
[53] H. Monteiro, J.E. Fernández, F. Freire, Comparative life-cycle energy analysis of a new and an existing house: the significance of occupant's habits, building systems and embodied energy, Sustain. Cities Soc. 26 (2016) 507–518.

[54] G. Kokogiannakis, J. Clarke, P. Strachan, Impact of using different models in practice-a case study with the simplified methods of ISO 13790 standard and detailed modelling programs, International Building Performance Simulation Association (IBPSA) (2007) 39–46.
[55] D.G. Erbs, S.A. Klein, J.A. Duffie, Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation, Sol. Energy 28 (4) (1982) 293–302.
[56] B. Agrawal, G.N. Tiwari, Life cycle cost assessment of building integrated photovoltaic thermal (BIPVT) systems, Energy Build. 42 (2010) 1472–1481.
[57] I.M. Sobol, Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Math. Comput. Simulat. 55 (1) (2001) 271–280.
[58] A. Saltelli, Making best use of model evaluations to compute sensitivity indices, Comput. Phys. Commun. 145 (2) (2002) 280–297.
[59] M.G. Prina, V. Casalicchio, C. Kaldemeyer, G. Manzolini, D. Moser, A. Wanitschke, W. Sparber, Multi-objective investment optimization for energy system models in high temporal and spatial resolution, Appl. Energy 264 (2020) 114728.
[60] M.G. Prina, M. Lionetti, G. Manzolini, W. Sparber, D. Moser, Transition pathways optimization methodology through EnergyPLAN software for long-term energy planning, Appl. Energy 235 (2019) 356–368.
[61] S.K. Huang, L. Kuo, K.-L. Chou, The applicability of marginal abatement cost approach: a comprehensive review, J. Clean. Prod. 127 (2016) 59–71.
[62] F. Kesicki, N. Strachan, Marginal abatement cost (MAC) curves: confronting theory and practice, Environ. Sci. Pol. 14 (8) (2011) 1195–1204.
[63] M.G. Prina, F.C. Fornaroli, D. Moser, G. Manzolini, W. Sparber, Optimisation method to obtain marginal abatement cost-curve through EnergyPLAN software, Smart Energy 1 (2021) 100002.
[64] H. Monsef, M. Naghashzadegan, A. Jamali, R. Farmani, Comparison of evolutionary multi objective optimization algorithms in optimum design of water distribution network, Ain Shams Eng. J. 10 (1) (2019) 103–111.
[65] A. Hojjati, M. Monadi, A. Faridhosseini, M. Mohammadi, Application and comparison of NSGA-II and MOPSO in multi-objective optimization of water resources systems, J. Hydrol. Hydromechanics 66 (3) (2018) 323–329.
[66] M. Song, D. Chen, A comparison of three heuristic optimization algorithms for solving the multi-objective land allocation (MOLA) problem, Spatial Sci. 24 (1) (2018) 19–31.
[67] M. Mitchell, An Introduction to Genetic Algorithms, 1998.
[68] R.K. Blashfield, Mixture model tests of cluster analysis: accuracy of four agglomerative hierarchical methods, Psychol. Bull. 83 (3) (1976) 377.
[69] Y. Li, L. Li, P. Shen, Probability-based visual comfort assessment and optimization in national fitness halls under sports behavior uncertainty, Build. Environ. (2023) 110596.
[70] J.R. Quinlan, Induction of decision trees, Mach. Learn. 1 (1) (1986) 81–106.
[71] H.K. Esfeh, A. Azarafza, M. Hamid, On the computational fluid dynamics of PEM fuel cells (PEMFCs): an investigation on mesh independence analysis, RSC Adv. 7 (52) (2017) 32893–32902.