Energy & Buildings 332 (2025) 115440

Contents lists available at ScienceDirect

Energy & Buildings

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

Invited Review Article

AI-driven design optimization for sustainable buildings: A systematic review

Piragash Manmatharasan [a], Girma Bitsuamlak [b], Katarina Grolinger [a,*]

[a] Department of Electrical and Computer Engineering, Western University, London, N6A 5B9, Ontario, Canada
[b] Department of Civil and Environmental Engineering, Western University, London, N6A 5B9, Ontario, Canada

ARTICLE INFO

Keywords:
Machine learning
Building sustainability
Design optimization
Surrogate model

ABSTRACT

Buildings are major contributors to global carbon emissions, accounting for a substantial portion of energy consumption and environmental impact. This situation presents a critical opportunity for energy conservation through strategic interventions in both building design and operational phases. Artificial Intelligence (AI) has emerged as a transformative approach in this context, enhancing the efficiency and precision of energy management efforts. In the operational phase, AI is extensively utilized as smart controllers for Heating, Ventilation, and Air Conditioning (HVAC) systems and passive energy gains, as well as for fault detection. In the design phase, AI is pivotal as a surrogate model, enabling rapid and accurate evaluation of design options and allowing designers to optimize building performance with minimal computational resources. As the early-stage optimization is more cost-effective than post-construction modifications, design phase optimization has a great potential. Consequently, this paper examines recent advancements in surrogate-assisted design optimization for sustainable buildings, providing a comprehensive overview of the entire optimization process, from data preparation and surrogate model training to final optimization. The review categorizes studies based on experimental approaches and methodologies, identifying trends, gaps, and opportunities in the field. Notably, it highlights how modern AI techniques can incorporate previously unexplored dimensions into surrogate-assisted optimization, broadening the scope and potential of surrogate models. Therefore, this study provides guidance for future research and practical applications of AI-driven strategies in sustainable building practices.

Contents

* Corresponding author.
E-mail address: kgroling@uwo.ca (K. Grolinger).

https://doi.org/10.1016/j.enbuild.2025.115440
Received 29 November 2024; Received in revised form 22 January 2025; Accepted 6 February 2025
Available online 10 February 2025

1. Introduction

Buildings consume a significant amount of energy and contribute substantially to CO2 emissions. According to the United Nations Environment Programme's (UNEP's) 2023 report [1], 37% of global greenhouse emissions come from buildings. These emissions are primarily due to the materials used in construction and the energy required for building operation, highlighting the need to reduce operational and embodied carbon through sustainability measures.

Norbert Lechner proposed a 3-tier approach presented in Fig. 1 for integrating sustainable measures into buildings [2]. Tier 1 focuses on fundamental aspects of building design, such as orientation, insulation, compactness, and other early design decisions. Tier 2 involves strategies that leverage natural energies, including daylight, natural ventilation, and passive solar heating. Finally, Tier 3 addresses the selection and optimization of mechanical and electrical systems. As shown in Fig. 1, according to Lechner, strategic design decisions in Tier 1 have the potential to reduce energy consumption by 60%, with further reductions achieved through passive systems in Tier 2 (20%) and optimized equipment in Tier 3 (5%). Together, these measures can achieve up to an 85% reduction in energy consumption.

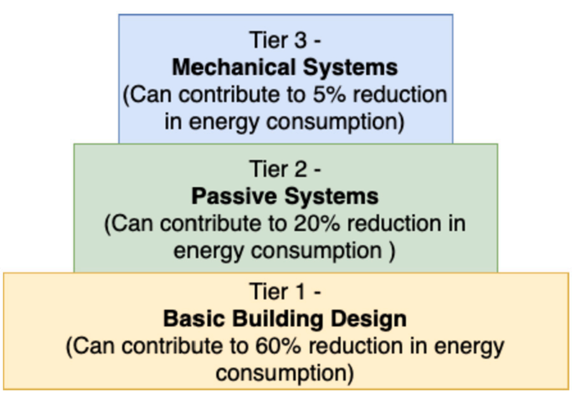

Fig. 1. Norbert Lechner's 3-tier approach for building sustainability outlining three levels of sustainable building measures [2]. According to Lechner, integrating measures from all three tiers can cumulatively reduce energy consumption by 85%. Notably, decisions made during the fundamental design stage alone can achieve up to a 60% reduction in energy use.

Design stage optimizations involve exploring the potential solution space early in the building lifecycle to identify the most suitable options, balancing performance, cost, and other relevant constraints. These early-stage optimizations are crucial because they are more cost-effective than post-construction modifications and have the greatest potential impact on the building's performance [3]. Therefore, integrating Tier 1 elements, in conjunction with strategies from Tiers 2 and 3, into the design stage optimization process is essential for achieving optimal sustainability outcomes.

Given the complexity inherent in building design and the need for comprehensive optimization exploration during the design stage, AI plays a crucial role in these processes. Machine Learning (ML) models automate complex simulations—such as energy modeling and daylight analysis—and enable the exploration of extensive design spaces, allowing for the identification of the most sustainable building designs within a constrained timeframe. ML models that replicate the outcomes of physics-based simulations are commonly known as *surrogate models*, as they approximate the performance of these traditional simulation methods.

Recognizing AI's role, specifically surrogate models, in sustainable building practices, our review focuses on building design optimization using surrogate models, with an emphasis on sustainability. While Westermann et al. [4] provided a comprehensive review of surrogate modeling approaches, their work encompassed a broader range of applications, including sensitivity analysis, uncertainty analysis, design optimization, and conceptual design, while we focus solely on design optimization.

More recently, Elwy et al. [5] and Cruz et al. [6] conducted reviews to identify trends in building performance optimization using surrogate models, providing an overview of applications and methodological advancements. We also review performance optimization focusing on sustainability; however, unlike the studies by Elwy et al. and Cruz et al. [5,6], we define the surrogate modeling process by breaking it into components and reviewing each component individually. By doing this, we provide researchers and practitioners with a way to select each component according to their specific use case scenario. Additionally, our study emphasizes advanced ML and deep learning (DL) techniques, providing the ML perspective as well as incorporates a guidance framework for practitioners throughout the modeling process while offering researchers directions for future advancements.

Consequently, this paper provides a systematic examination of recent studies that have advanced surrogate-assisted design optimization methods in the context of sustainability. This offers a structured synthesis, emphasizing experimental approaches and key methodological strategies from a computer engineering and ML perspective. The review spans the entire optimization workflow, encompassing data preparation, surrogate model selection, training and validation, and final design optimization, while identifying and discussing methodologies and techniques at each stage. By analyzing nuances in these studies, this review not only highlights the current research landscape but also identifies gaps and opportunities for further advancement in the field.

The remainder of the paper is organized as follows: Section 2 explores AI applications in building sustainability and details the systematic review process. Section 3 introduces surrogate models in design optimization, while Section 4 reviews the modeling process. Section 5 examines key focus areas, and Section 6 identifies gaps and opportunities in the field. Finally, Section 7 concludes with a summary of the findings.

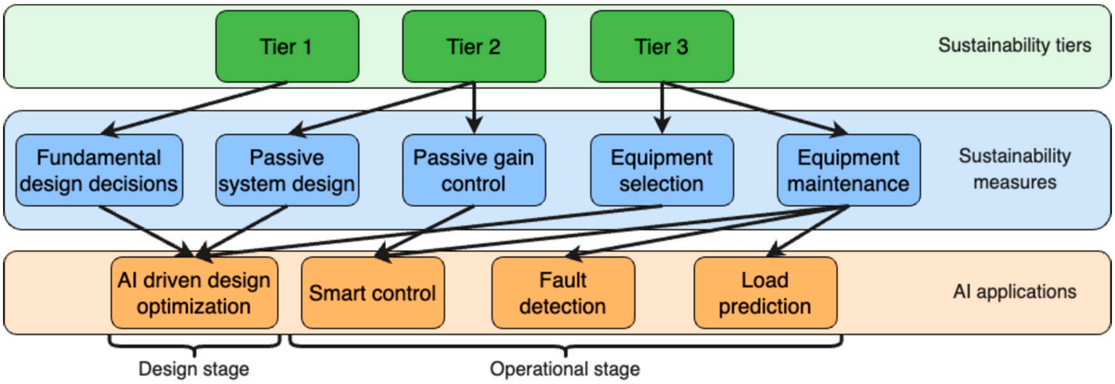

Fig. 2. AI integration in building sustainability. AI is applied during both the design and operational stages. The design stage focuses on optimizing decisions from the three tiers, while the operational stage includes smart control, fault detection, and load prediction.

2. AI in building sustainability

This section first explores the applications of AI in building sustainability, highlighting its multifaceted roles in this domain, before focusing on the specific areas selected for this review. Next, the systematic review process used to identify the relevant studies is detailed.

2.1. AI applications

AI is increasingly being used across various domains to improve efficiency and precision, particularly in the realm of sustainability. Fig. 2 illustrates our categorization of AI applications in building sustainability, encompassing the three tiers across both design and operational phases.

Tier 3 which involves the selection and optimization of mechanical and electrical systems, has seen substantial AI applications including equipment selection during design optimization [7–9], smart control of HVAC systems [10–21], fault detection of electronic devices [22–30], and load prediction [31,32]. AI's ability to optimize these systems ensures that mechanical and electrical components are not only efficiently selected, but also dynamically tuned during operation for peak performance.

Tier 2 focuses on optimizing passive design strategies to efficiently use natural energy sources, such as daylight, ventilation, and solar heating. AI augments these strategies by providing precise optimization of passive elements during the design process [33–36,9]. Additionally, AI serves as a smart controller for systems that influence passive gains—such as shades, mechanical ventilation systems, and operable windows—ensuring these strategies are dynamically optimized to maintain sustainability [37–39].

Tier 1, which covers fundamental design considerations including orientation and insulation, is integrated into the broader design optimization processes [4,5,40–44]. Here, AI facilitates the exploration of optimal design configurations that maximize energy efficiency.

AI-driven design stage optimization leveraging surrogate models, which integrates elements from all three tiers, plays a pivotal role in achieving comprehensive sustainability in building design. By enabling the incorporation of strategic sustainable measures during the design stage, AI enhances the sustainability of modern buildings early in the building life cycle avoiding expensive post-construction retrofits. Accordingly, our review focuses on surrogate-assisted design optimization for sustainable buildings.

2.2. Review process

Our review process employs the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) [45] methodology to systematically identify relevant papers for this review, ensuring a comprehensive and unbiased selection process. Fig. 3 illustrates the step-by-step approach for identifying, screening, and selecting studies for analysis.

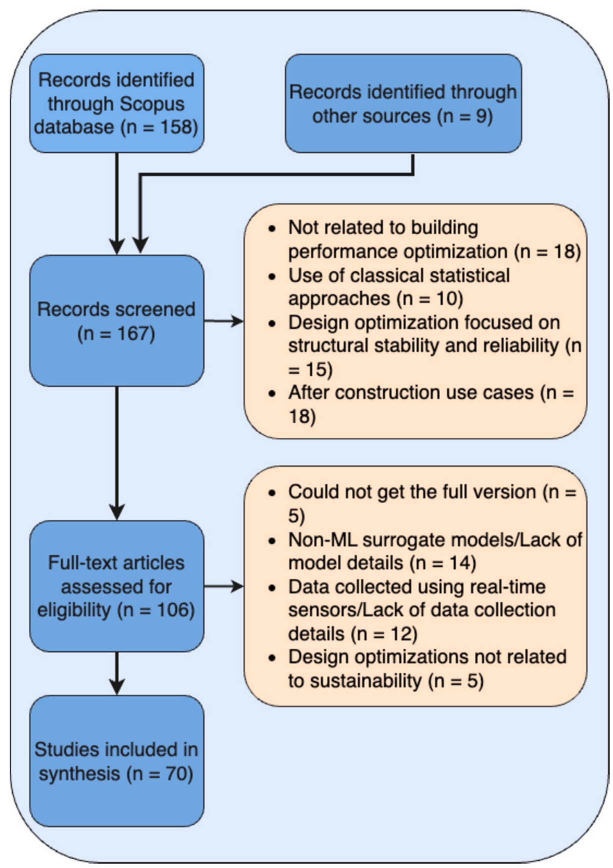

Fig. 3. Literature review process. The initial search through Scopus yielded 158 publications, with 9 additional records identified through other sources, totaling 167. After screening, 61 publications were excluded. Out of 106 full-text articles assessed for eligibility, 36 were excluded, resulting in 70 included papers.

The literature search began with a search of publications indexed in the Scopus database, focusing on the domain of building design optimization with surrogate models. The primary search terms included *"building design optimization"*, *"building performance and design optimization"*, and *"building energy and design optimization"*. To further refine the search and specifically target studies employing surrogate ML models, additional keywords were added: *"machine learning"*, *"artificial intelligence"*, *"surrogate model"*, *"meta model"*, *"deep learning"*, and *"neural network"*. The final search query applied was as follows:

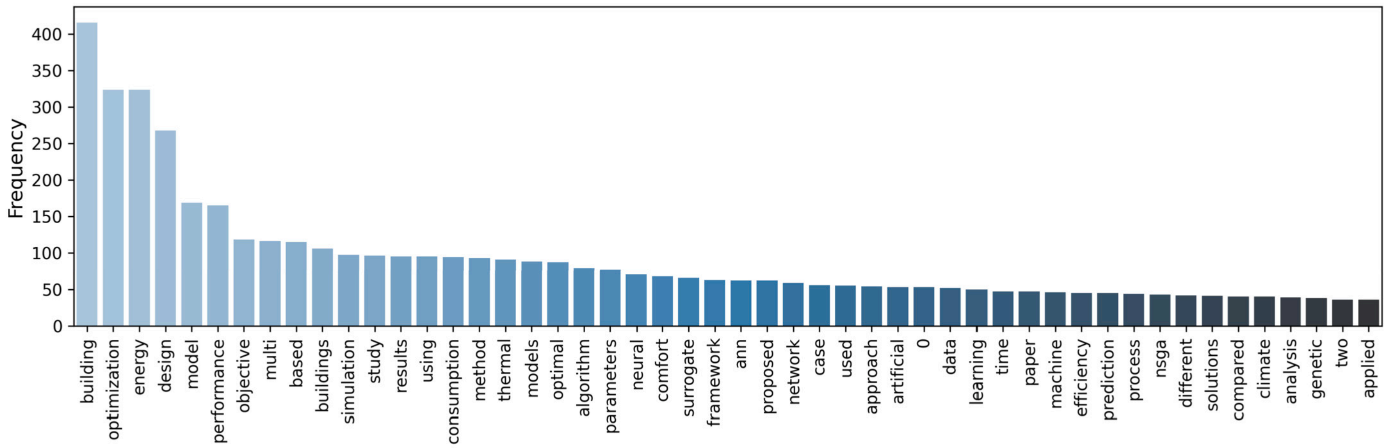

Fig. 4. Keyword frequency analysis for the titles and abstracts of the reviewed studies. The chart showcases key research themes and trends within the surrogate modeling domain.

(("building design" AND "optimization") OR
("building performance" AND
"design optimization") OR
"design optimization")) AND
("neural network" OR "artificial intelligence" OR
"machine learning" OR "surrogate model" OR
"meta model" OR "deep learning")

Only peer-reviewed journal articles published in English within the last fifteen years were included, ensuring the currency of the methodologies and technologies discussed. This resulted in 158 publications. Nine more publications were identified through a manual search of references and author's knowledge, bringing the total to 167 publications for screening. There are numerous studies on surrogate models in other domains such as reliable design [46,47] but those are not included as they do not consider sustainable building design.

Each paper was evaluated based on its abstract and keywords to determine its relevance to the review's focus on the surrogate model in building design optimization. Papers not related to design optimization, such as those on load prediction, fault detection, and solar panel energy generation prediction, were removed. Papers focusing on optimization for structural stability and reliability, pedestrian wind flow around buildings, or operational stage optimizations—such as HVAC control and energy consumption—were also not considered, as well as works based on statistical methods and other non-AI approaches. A total of 61 papers were excluded, leaving 106 papers selected for full-text review.

From the full-text article review, another 36 papers were removed due to the following reasons: the lack of model details, the use of non-ML meta-heuristic algorithms, insufficient information on data collection methods, and data collected from existing buildings using sensors and other measures. Studies that utilized data collected from existing buildings were excluded, as these data are gathered during the operational stage and are not fully available during the design stage, and our focus is on the design stage. Some papers that initially appeared to have a potential for sustainability considerations were found not to include any sustainability measures in their optimization. Additionally, five papers could not be fully accessed. After this screening, 70 papers were retained for review and analysis.

To highlight the emerging trends in these studies, a keyword frequency analysis is presented in Fig. 4. As expected, building, optimization, energy, and design, dominate, followed by performance and objectives.

3. Surrogate models in sustainable design optimization

We analyzed the selected papers to identify the methodologies, techniques, and algorithms employed in building design optimization and understand how these approaches contribute to achieving sustainable designs. The insights gained from this analysis not only shed light on current trends but also highlight the evolving role of advanced computational methods in enhancing sustainability in building practices.

Fig. 5 illustrates the core process of building design optimization using surrogate models, which can be divided into two main stages: surrogate model development and optimization using the surrogate model. In the first stage, a surrogate model is developed to replicate the task of a physics-based model, which reduces the computational burden of physics-based models. Once validated, the model is employed in the second stage, where it facilitates the design optimization process to derive optimal parameters.

The core steps in surrogate model development include the identification of objectives, design variable selection and parameter combination generation, design and simulation of building performance, and, finally, surrogate ML model selection, training and validation. Optionally, the model training step may include performance tuning techniques, such as hyperparameter tuning or adaptive sampling, to enhance model accuracy. Once the model is built, it is utilized in an optimization process with a suitable optimization algorithm to find the optimal values for the selected design parameters.

Based on the part of the surrogate model-driven optimization process they address, the studies can be divided into four categories as seen in Table 1: model development only, model development with performance tuning, model development and design optimization, and design optimization with the tuned model. Within each of these categories, the works are further grouped based on their primary focus. While the studies share a common surrogate modeling process, they emphasize different aspects of this process.

Model development with performance tuning: This category includes studies that focused on both developing a surrogate model and implementing performance-tuning techniques to improve model accuracy. This includes an optimization of the ML model's hyperparameters which aims to improve the ML model accuracy. Note that this is the improvement of the ML model itself and does not involve building design optimization.

These cases involved optimizing the model's hyperparameters through various techniques to achieve higher accuracy. Techniques such as genetic algorithms (GAs), Bayesian optimization, and grid search were commonly employed for this [48–50].

Model development and design optimization: Falling into this category are works that developed a surrogate model and then used it

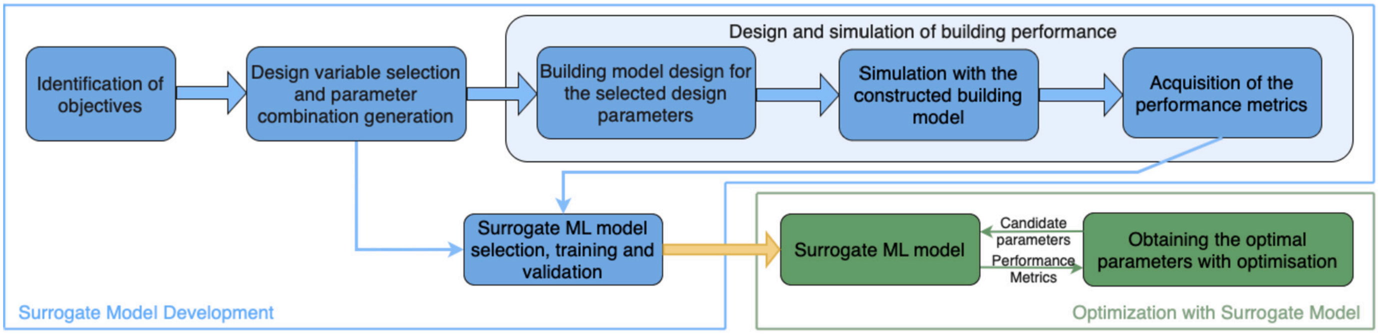

Fig. 5. Design optimization using surrogate model consists of surrogate model development and optimization with the surrogate model. First, objectives are identified and variables are determined. This follows by building design and simulation which create data for surrogate model training. This trained model is subsequently utilized in the building optimization process.

to derive optimal parameters for a selected building. These studies offer a complete process for design optimization.

These papers are further classified based on the specific factors they focused on: climate change [51,35], ensemble models [52,53,22], meta-heuristic optimization [54–56], systematic sampling [57], Python framework [58], renovation [59,41], feature selection (e.g., [60,61]), passive energy elements (e.g., [34,35]), active energy elements (e.g., [9,8]), and case studies (e.g., [36,62]).

Model Development Only: This group encompasses works that focused exclusively on surrogate model development. These studies addressed all the necessary steps involved in building and evaluating a surrogate model but did not include a case study demonstrating the optimization process related to design variable selection.

The studies in this category are further grouped based on the key aspects they emphasize: DL-aided models [63,40,43,42], data generation [44], feature selection [64,65], occupant behavior [66], explainable AI [67], shape descriptors [68], ventilation [69], and optimized FFNN [70].

Design optimization with tuned model: This group contains the publications that developed a ML model, tuned its performance, and then used it for deriving sustainable designs. The key difference from the previous category is that these studies involved a two-stage optimization process, where the ML model itself is optimized before being used for parameter derivation.

The works in this category are further categorized based on the distinct elements they used in the two-stage optimization process. This includes Reinforcement Learning (RL) optimization [92] in the second stage, and parameter optimization (e.g., [99,98,97]) and adaptive sampling (e.g., [100,101]) in the first stage.

4. Review of the surrogate modeling process

This section provides a comprehensive review of the core methodologies essential to the surrogate-assisted design optimization process. It covers key stages, including data preparation, surrogate model development, and optimization techniques, highlighting the methods and tools used at each step.

4.1. Identification of objectives

Through the review of the selected literature, we identified three primary sustainability objectives commonly addressed in optimization studies: reducing operational energy demand, enhancing daylight performance levels, and reducing carbon emissions. As seen in Table 1, each study included at least one of these objectives in the optimization, with some addressing multiple sustainability objectives simultaneously as a part of a multi-objective optimization approach.

Operational Energy Demand (OED): Buildings consume a substantial amount of energy during operation for heating, cooling, lighting, and equipment operation. Therefore, minimizing this energy consumption is a common objective in the sustainability domain and is widely addressed in the literature on sustainable buildings. In our review, 56 studies identified the reduction of operational energy as one of the primary objectives.

Daylight Performance Level (DPL): Another common objective in design optimization is to increase DPLs within buildings. This approach reduces the need for artificial lighting, lowering energy consumption and enhancing indoor environmental quality. In the analyzed literature, 12 studies identified DPLs as one of their objectives. Among these, except for three studies [49,42,87], all others (e.g., [51,76,77]) had multiple sustainability objectives.

Carbon Emissions (CE): Several studies aim to minimize total carbon emissions by addressing both embedded carbon and OED. Embedded carbon refers to the carbon content in the materials used for constructing the building, including emissions from material extraction, manufacturing, and transportation. In our review, 11 studies specifically targeted the reduction of both embedded and operational carbon [63,44,53,74].

4.2. Design variable selection and parameter combination generation

Once the objectives are identified, the next step is to determine the building design variables that should be optimized to achieve these objectives. This is followed by a parameter combination generation process, which creates input datasets for the surrogate ML model.

4.2.1. Design variable selection

Certain design variables have a greater influence on the optimization process than others, and their importance can vary depending on the specific objectives selected. The selection of these influential variables can be informed by domain expertise or determined through techniques such as sensitivity analysis. This subsection examines the most frequently employed variables in the current literature on surrogate models for building design optimization. Fig. 6 provides a summary of commonly used variables, indicating the percentage of reviewed studies that utilized each feature.

For a comprehensive list of the parameters used in each reviewed study, refer to the Appendix. Table A.1 provides an exhaustive list of studies and the specific wall and roof parameters they utilized, while Table A.2 lists the remaining parameters employed in the studies. Some parameters that were used in specific studies but are not widely adopted are not included in these tables. The main categories of the design parameters are as follows:

Windows: Windows serve as a conduit for sunlight and airflow while providing connectivity to the outdoors. Increasing the glazing area can enhance solar gains and natural lighting, but it may also reduce insulation effectiveness. As a result, window parameters are among the most frequently optimized variables in sustainable design studies. The

Table 1
Classification of studies based on the part of the surrogate model optimization process they address and the primary focus. The table also highlights the objectives, ML models, and optimization algorithms employed.

Category	Primary Focus	Author	Year	Ref	Objectives			ML Models			Optimization
					OED	DPL	CE	FFNN	Tree	Other	
Model development only	DL aided models	Mateusz	2020	[63]			✗			CNN & FFNN	-
		He	2021	[42]		✗				CNN & GAN	-
		Westermann	2022	[40]	✗					TCN & FFNN	-
		Yue	2023	[43]	✗					RDPG	-
	Data generation	Venkatraj	2023	[44]			✗	✗			-
	Feature selection	Didwania	2023	[64]	✗			✗			-
		Seyedzadeh	2019	[65]	✗				✗		-
	Occupant behavior	Li	2020	[66]	✗					SVR	-
	Explainable AI	Barbaresi	2022	[67]	✗				✗		-
	Shape descriptors	Storcz	2023	[68]	✗			✗			-
	Ventilation	Alghamdi	2024	[69]	✗			✗			-
	Optimized FFNN	Himmetoglu	2021	[70]	✗			✗			-
Model development with performance tuning	Parameter Optimization	Garcia	2020	[48]	✗			✗			-
		Han	2021	[49]		✗		✗			-
		Cai	2023	[50]	✗					SVR	-
Model development and design optimization	Climate change	Zou	2021	[51]	✗	✗		✗			GA
		Li	2023	[35]	✗				✗		NSGA-ii
	Ensemble models	Chen	2023	[53]			✗			Ensemble	NSGA-iii
		Yang	2023	[71]	✗					Ensemble	NSGA-iii
		Shen	2024	[52]	✗	✗				Ensemble	NSGA-iii
	Metaheuristic optimization	Yu	2015	[54]	✗			✗			NSGA-ii
		Chegari	2021	[55]	✗			✗			MOPSO
		Chegari	2022	[72]	✗			✗			MOPSO
	Systematic sampling	Zheng	2024	[57]	✗			✗			GA
	Python framework	Hocine	2023	[58]	✗			✗			GA
	Renovation	Arjomandnia	2023	[59]	✗			✗			PSO
		Asadi	2014	[41]	✗			✗			GA
	Feature selection	Chen	2018	[37]	✗					SVR	NSGA-ii
		Li	2019	[73]	✗			✗			NSGA-ii
		Serbouti	2021	[74]			✗	✗			NSGA-ii
		Wang	2021	[75]	✗			✗			NSGA-ii
		Chen	2022	[61]			✗	✗			NSGA-iii
		Razmi	2022	[76]	✗	✗		✗			NSGA-iii
		Zhan	2024	[60]		✗	✗	✗			NSGA-iii
	Passive energy elements	Gou	2018	[33]	✗			✗			NSGA-ii
		Lin	2021	[77]	✗	✗		✗			Antlion
		Li	2023	[78]	✗					SVR	PSO
		Alsharif	2023	[34]	✗				✗		Manta-Ray
	Active energy elements	Magnier	2010	[79]	✗	✗		✗			NSGA-ii
		Xu	2021	[80]	✗	✗		✗			NSGA-ii
		Amini	2022	[7]	✗					GPR	NSGA-ii
		Li	2024	[8]	✗				✗		DE
		Zong	2024	[9]			✗		✗		NSGA-iii
	Case studies	Gossard	2013	[81]	✗			✗			NSGA-ii
		Li	2017	[82]	✗			✗			MODE
		Prada	2018	[83]	✗			✗		MARS	NSGA-ii
		Xu	2018	[84]	✗			✗			Iterative
		Si	2019	[85]	✗			✗			NSGA-ii
		Zhao	2021	[86]	✗	✗		✗			NSGA-ii
		Wang	2021	[87]		✗		✗			NSGA-ii
		Xue	2022	[88]			✗	✗			NSGA-ii
		Wu	2022	[89]			✗		✗		NSGA-iii
		Saryazdi	2022	[90]	✗			✗			GA
		Elbeltagi	2022	[91]	✗			✗			GA
		Kubwimana	2023	[62]	✗			✗			GA
		Ji	2024	[36]			✗	✗			NSGA-iii
Design optimization with tuned model	RL optimization	Pan	2024	[92]	✗			✗			DDPG
	Parameter optimization	Sun	2020	[93]	✗	✗		✗			SPEA2
		Garcia	2022	[94]	✗			✗			GA
		Liu	2023	[95]			✗		✗		NSGA-ii
		Shen	2023	[96]	✗				✗		AGE-MOEA
		Wu	2023	[97]	✗				✗		NSGA-iii
		Khan	2024	[98]	✗				✗		AGE-MOEA
		Si	2024	[99]	✗				✗		GA
	Adaptive sampling	Bre	2020	[100]	✗			✗			NSGA-ii
		Bamdad	2020	[101]	✗			✗			Ant Colony
		Yue	2021	[102]	✗			✗			NSGA-ii
		Batres	2023	[103]	✗			✗			PSO
		You	2023	[104]	✗			✗			MCOA
		Lahmar	2024	[105]	✗			✗			GA

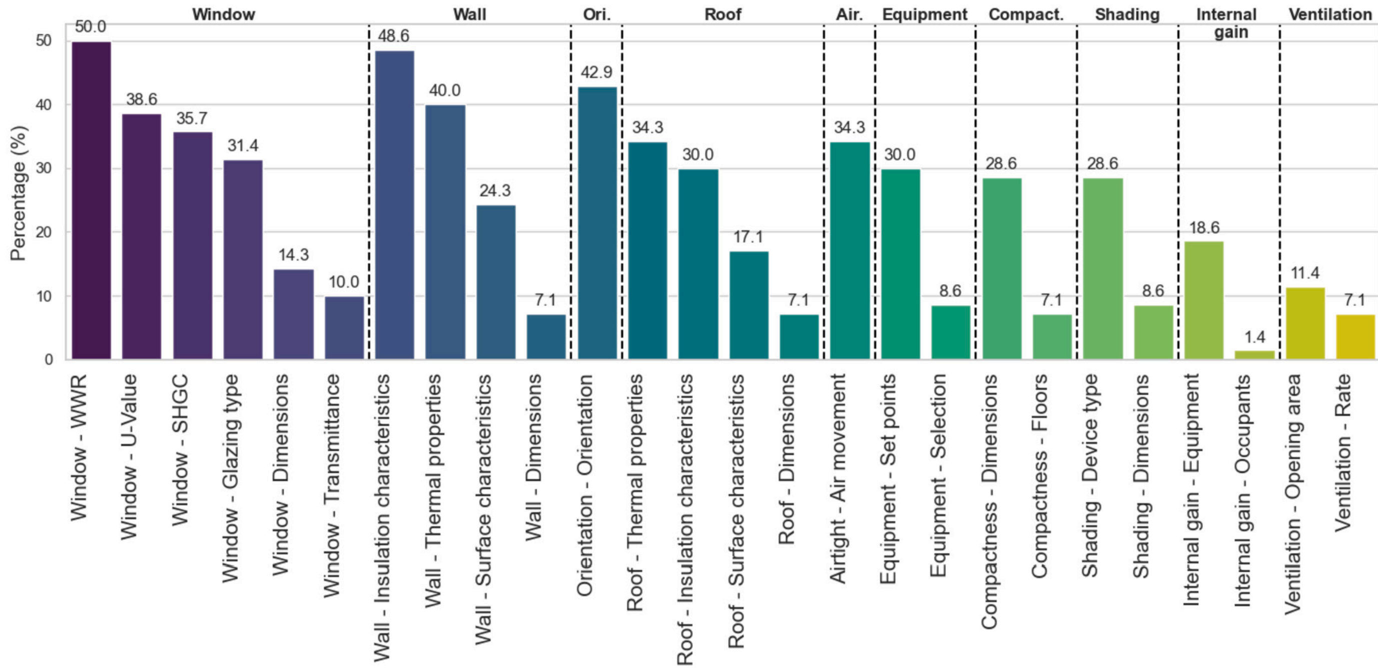

Fig. 6. Parameters commonly optimized using surrogate models for sustainability. Optimization is performed to obtain the optimal values for selected design variables, which are chosen based on their impact on the defined objectives.

Window-to-Wall Ratio (WWR), which is the ratio of window area to the total wall surface area, is the most commonly optimized window design variable. This parameter was employed in 35 (50%) of the studies, where strategic window placement maximized passive solar gains and day lighting levels (e.g., [53,52]).

Beyond WWR, optimizing window specifications—such as U-value, Solar Heat Gain Coefficient (SHGC), transmittance, and glazing type—is essential for enhancing energy efficiency (e.g., [53,99]). Additionally, some studies explored the optimization of window dimensions to further improve sustainability (e.g., [104]).

Wall specifications: Wall characteristics are a common focus in the optimization studies reviewed. As illustrated in Fig. 6, 34 (48.6%) of the literature incorporated wall insulation-related features. Some studies concentrated on optimizing the type of insulation (e.g., [41,94,44]), while others focused on insulation thickness (e.g., [35,88,40]), with a subset addressing both parameters in tandem (e.g., [33,51,70]).

Thermal properties of wall materials were targeted in 28 (40%) of the literature: key metrics include thermal conductivity (e.g., [81,62, 96]), specific heat capacity (e.g., [81,37]), heat inertia [54], and heat transfer coefficient (e.g., [72,95,97]). Additionally, surface properties such as solar absorptance (e.g., [35,53]) and reflectance (e.g., [49,77]) were optimized in 17 of the papers. Wall dimensions (e.g., [50,92]) also received attention in 5 of the studies.

Orientation: The orientation of a building relative to the cardinal directions influences solar gain, natural lighting, and wind flow. Also referred to as azimuth, this design variable was optimized in 30 (42.9%) of the reviewed studies (e.g., [91,78])

Roof specifications: Roofs, together with walls, form the building's enclosure and serve as a barrier against environmental conditions. Design variables related to roofs found in the literature mirror those optimized for walls and include insulation characteristics (e.g., [99,43]), thermal properties (e.g., [81,54]), surface characteristics (e.g., [72,95]), and dimensions (e.g., [50,35]). Roof thermal properties were optimized in 24 (34.3%) of the studies reviewed, while insulation characteristics and surface characteristics were optimized in 21 and 11 studied, respectively. In 5 of the publications, roof dimensions were among the optimized design variables.

Air tightness: Air tightness and infiltration rate significantly influence internal temperature and are crucial for maintaining indoor air quality. In 24 (34.3%) of the studies reviewed, parameters related to air movement, such as air gaps and infiltration, were included in the optimization process (e.g., [94,97]).

Equipment selection and setpoints: In 6 of the reviewed studies, equipment selection was also incorporated into the optimization process. Common equipment selections identified in the literature include decisions on HVAC system types [41,94,9,83] and lighting system [102,43]. In 21 (30%) of the studies, HVAC set points were also used as parameters. This gives optimal set point configuration for the design (e.g., [89,85]).

Compactness: Compactness is typically assessed by the volume-to-surface area ratio, where a higher ratio reduces the potential for heat loss through the building's envelope. In the reviewed studies, parameters such as the number of floors and floor dimensions (e.g., [86,54]) were optimized to improve compactness.

Shading: Shading can be used for controlling passive heat gain through windows, and optimizing shading parameters can reduce excessive heat gain during summer. In 20 (28.6%) of the reviewed studies, the type of shading elements was optimized (e.g., [51,77]). Additionally, 6 of the reviewed studies focused on optimizing shading element dimensions (e.g., [34,100]) to further manage passive solar gains.

Internal gains: Internal energy gains can affect the internal temperature and contribute to the total energy consumption of a building. Internal gains from equipment were used as a variable in 13 (18.6%) of the studies reviewed (e.g., [90,53]). Additionally, Shen et al. [96] included occupancy density as one of the parameters.

Ventilation: Ventilation is an essential passive cooling element in certain weather conditions, and the optimization of ventilation openings was addressed during the design stage in 8 (11.4%) of the studies (e.g., [78,43]), while the ventilation rate was specifically optimized in 5 (7.1%) of the studies (e.g., [90,61]).

In addition to the discussed commonly used variables, some studies have explored other variables. Two of these important yet underutilized variables, which are particularly relevant for enhancing the reusability of models across different sites are weather-related features and urban area effect. Weather-related features are rarely used as inputs for surrogate models; however, incorporating these features is essential for ensuring that surrogate models perform accurately under different climate conditions. Westermann et al. [40] utilized a deep Temporal Convolutional Network (TCN) to process hourly weather data, automatically extracting temporal patterns and key climatic features. These extracted features, combined with building design parameters, were used to predict energy demands for different locations. Similarly, Zheng et al. [57] incorporated weather-related data into their feature engineering process by deriving aggregated metrics such as annual solar irradiance and average ambient temperature. These engineered features, integrated with building-specific parameters, allowed the surrogate model to generalize across diverse climate zones without requiring retraining. Both approaches demonstrated the critical role of weather variables in enhancing the accuracy and generalizability of surrogate models.

Similarly micro-climate effects caused by adjacent structures are seldom considered, although they play a major role in energy consumption. Mateusz et al. [63] proposed Convolutional Neural Networks (CNNs) to identify urban effects from image data, and incorporated this information into performance predictions. Similarly, Wang et al. [87] demonstrated the importance of optimizing the layout of residential buildings to enhance daylight levels, taking into account the urban effects from adjacent structures.

4.2.2. Parameter combination generation

After selecting the parameters to be optimized, the next step is to establish a range of values for each parameter, which will be used in creating the training dataset for the surrogate model. These ranges vary based on the nature of each parameter. Some parameters are numerical and may have continuous or discrete values, while others are categorical. To adequately train the surrogate model, a comprehensive training set with combinations of these parameters must be generated. Various sampling techniques are employed to select parameter values, with commonly used sampling techniques including:

- **Latin Hypercube Sampling (LHS):** This technique ensures that samples are evenly distributed across the entire range of each parameter, providing better coverage of the input space compared to random sampling [106]. LHS reduces redundant sampling and guarantees representation across all areas of the design space, making it especially suitable for high-dimensional problems. These characteristics made LHS the most commonly used sampling approach in the reviewed papers.
- **Monte Carlo Sampling:** This technique is simple to implement and useful for approximating complex distributions. Four of the studies reviewed used Monte Carlo methods for sampling [35,53,37,60].
- **Sobol Sequences:** The Sobol technique provides low-discrepancy sequences that cover the parameter space uniformly. In the reviewed literature, five of the studies employed this technique for the sampling process [74,43,64,61,83].
- **Orthogonal Sampling:** The Orthogonal method ensures that the sampled points are orthogonal in the parameter space, leading to wider coverage and reliable results. Only four studies employed this method in their surrogate modeling process [95,97,98,92].
- **Random Sampling:** In this approach, samples are selected randomly from the parameter space, which can be effective for a wide variety of problems but may require a large number of samples to achieve good coverage. It was preferred by three studies for the parameter generation [63,78,8].
- **Adaptive Sampling:** Adaptive methods dynamically adjust the sampling strategy based on model performance, improving efficiency. In five of the very recent studies, adaptive strategies were employed [100,101,103–105].

Since the goal of a surrogate model is to reduce the number of physics-based simulations, it is essential to employ a representative but limited training set. These techniques aim to create the training dataset which is both representative and robust, allowing the surrogate model to accurately learn and predict building performance across a wide range of design variables.

4.3. Design and simulation of building performance

The subsequent step in the surrogate modeling process entails generating building designs for selected design parameters using advanced design tools. These designs are then imported into simulation software to evaluate the performance metrics for each parameter combination. This process yields a comprehensive assessment of how different design variables influence the building's overall performance.

In the literature, tools such as SketchUp, Revit, Rhino, Grasshopper, and Dynamo are frequently utilized for generating building designs due to their ability to create detailed and precise models. These tools output a design file—typically in formats such as IFC, gbXML, or DXF—which represents the building model that can be modified for parameterized simulations. To streamline this process, these design files are modified parametrically, allowing for efficient adjustments to key variables. This parametric modification enables the systematic alteration of specific design aspects, such as insulation levels or window properties, within predefined ranges.

After the design phase, the next step is to simulate the building to obtain performance metrics. Fig. 7 shows the range of simulation tools frequently used across different studies to assess various aspects of building performance, along with the percentage of studies that employed each tool. These tools include:

EnergyPlus: EnergyPlus utilizes weather data and building attributes to perform simulations, providing assessments of heating, cooling, ventilation, lighting, water use, and other energy flows within buildings. Its large user community and robust features make EnergyPlus a reliable choice for researchers and practitioners. As shown in Fig. 7, 53 (75.7%) of the reviewed studies employed EnergyPlus, either directly or in combination with other tools, to evaluate the energy performance of building designs.

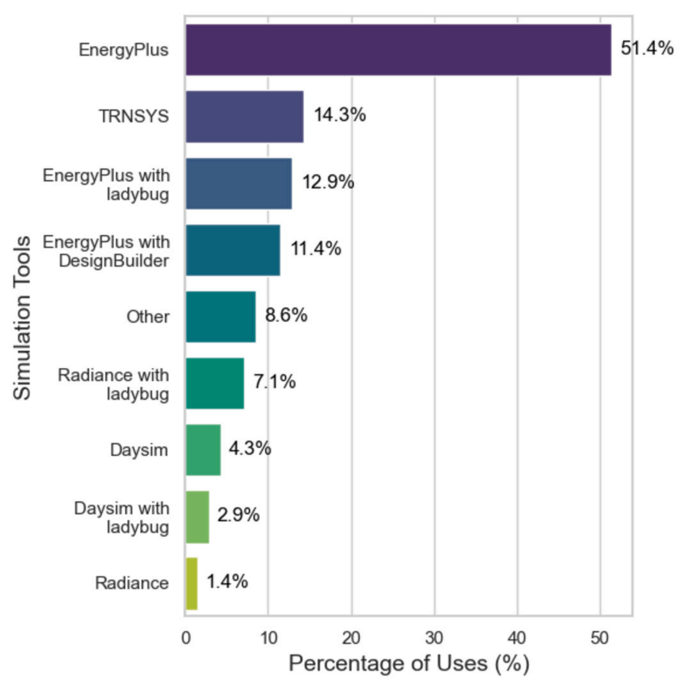

Fig. 7. Distribution of simulation tool usage in literature. EnergyPlus is the most frequently used tool, either alone or in combination with other tools such as DesignBuilder and Ladybug.

Radiance: Radiance is used to perform daylight simulations that determine how light interacts with building surfaces [107]. It offers high

Table 2
Comparison of commonly used ML algorithms in surrogate modeling, focusing on their data requirements, training complexity, learning methodologies, and main use cases.

Algorithm	Data Need	Training Complexity	Learning Methodology	Main Use Cases
FFNN	High	High	Backpropagation with gradient descent	Complex nonlinear relationships
SVR	Moderate	Moderate	Quadratic programming (optimization)	Small to medium datasets, simple nonlinear problems
CNN	Very High	Very High	Backpropagation with convolutional filters	Image data, spatiotemporal patterns
Tree-Based Models	Low	Low	Recursive partitioning and ensemble methods (e.g., boosting, bagging)	Explainability, tabular data, handling categorical data

precision in predicting visual comfort and lighting levels within buildings: it is used for optimizing artificial lighting and window openings. Radiance was utilized directly or through plugins in six of the reviewed studies to assess lighting levels.

Daysim: Daysim utilizes advanced ray tracing algorithms from Radiance, as it is built on the Radiance engine [108]. Compared to Radiance, Daysim offers a more user-friendly platform, making it easier for users to perform detailed daylight analyses. It was employed in five of the reviewed studies to assess the daylight levels in selected building designs.

Transient System Simulation Tool (TRNSYS): TRNSYS is another commonly used simulation tool for assessing the energy performance of buildings. It offers high flexibility, allowing users to create custom components and adapt simulations to meet specific requirements [109]. In 10 of the reviewed studies, TRNSYS was used as the performance evaluator.

Ladybug and Honeybee: Ladybug and its extension Honeybee serve as a bridge between powerful parametric modeling performed with Rhino/Grasshopper and robust simulation engines like EnergyPlus, Radiance, and similar [110]. Fig. 7 includes the ladybug plugin in conjunction with the corresponding simulation engines. In 16 of the reviewed studies, these plugins were used to run simulations on EnergyPlus, Radiance, or Daysim engines.

DesignBuilder: DesignBuilder is a comprehensive simulation tool widely used for building design optimization. It integrates EnergyPlus for energy simulation and offers a user-friendly interface for creating detailed building models. In the reviewed literature, eight (11.4%) studies employed DesignBuilder to evaluate the energy performance of their designs, as illustrated in Fig. 7.

Other simulation tools: In addition to these commonly used tools, a few studies also employed Ecotect and eQuest for performance analysis. Although Ecotect is not as widely used as other tools, it is valued for its ability to provide rapid feedback on design decisions, which aids in the iterative design process. On the other hand, eQuest is designed to facilitate detailed energy modeling with minimal input effort, making it a useful tool for quick energy performance assessments. These tools were preferred by 7.5% of the studies.

Carbon assessment tools are utilized when the objective is to reduce the embedded carbon in buildings, as calculating embedded carbon typically does not require simulations. As this is not a typical simulation tool, it has been excluded from Fig. 7. Serbouti et al. [74] incorporated the carbon cost of building materials by using data from the Koordinationskonferenz der Bau- und Liegenschaftsorgane der öffentlichen Bauherren (KBOB) database [111], a Swiss repository containing life cycle assessment data. Similarly, Zong et al. [9] utilized a knowledge-based database to obtain carbon details of building materials. Zhan et al. [60] calculated the embedded carbon using carbon emission factors for materials and transport, sourced from various databases. Chen et al. [53] and Ji et al. [36] also used carbon emission factors in their life cycle carbon emission assessments while Xue et al. [88] employed Carbon Emission Factors (CEFs) to determine the embedded carbon, sourcing these CEFs from the China Life Cycle Database.

Simulation time depends on factors such as granularity, duration, design complexity, and hardware. Most studies used hourly granularity for a one-year period, though some adopted finer resolutions, such as 30-minute intervals [98,96] and 10-minute steps [93,33]. Zhan et al. [60] simulated an 8,276 m^2, seven-story senior apartment building with EnergyPlus and Radiance on a standard laptop, requiring 2.4 minutes per design. Yue et al. [43] simulated four gymnasium buildings across different Chinese climate zones using EnergyPlus, completing hourly annual data simulations in 24 hours with a quad-core laptop, averaging 4.32 seconds per sample. These examples illustrate how simulation time is influenced by hardware, building complexity, and simulation tools. High-quality simulations can ensure accurate optimization results, as they provide reliable data for training surrogate models. Studies should address simulation quality to ensure the overall reliability of the optimization process.

4.4. Surrogate model selection, training and validation

After generating performance metrics through simulations for selected designs, the next step in the surrogate modeling process is model selection, training, and validation to develop an ML model that approximates the simulation task. This process ensures the chosen model provides accurate predictions and is suitable for optimization tasks.

4.4.1. Model selection

ML models approximate relationships between inputs and outputs, enabling predictive capabilities. However, each ML model offers distinct advantages and varying performance characteristics. The most suitable model for a task depends on specific requirements and constraints. Therefore, model selection is a critical step in the surrogate modeling process. Table 2 provides a summary of the key comparisons between commonly used algorithms in surrogate modeling.

Table 1 lists the ML models employed as surrogates in the reviewed studies, while Fig. 8 provides a summarized visualization of their usage. These models include:

Feed Forward Neural Networks (FFNN): FFNNs are the most common surrogate models in this domain due to their strong ability to recognize patterns. In an FFNN, neurons—the fundamental units of the network—are organized into interconnected layers: an input layer, one or more hidden layers, and an output layer, with information flowing in one direction from input to output. The nonlinear output functions of neurons enable the formation of complex nonlinear relationships between inputs and outputs [112]. FFNNs are typically trained using the backpropagation algorithm, where the gradients of the loss function are used to update the network weights [113]. However, training methods are not limited to backpropagation; the weights of a neural network can also be optimized using various optimization algorithms. In 46 out of the 70 studies reviewed, FFNN was the preferred surrogate model.

Support Vector Regression (SVR): SVR, a variant of Support Vector Machines (SVM) designed specifically for regression problems, has also been used as a surrogate model. SVR finds a function that deviates from the actual observed values by a small margin, effectively capturing the underlying relationship between input and output variables [114]. The training process involves solving a convex optimization problem. SVR was employed as a surrogate model in four reviewed studies.

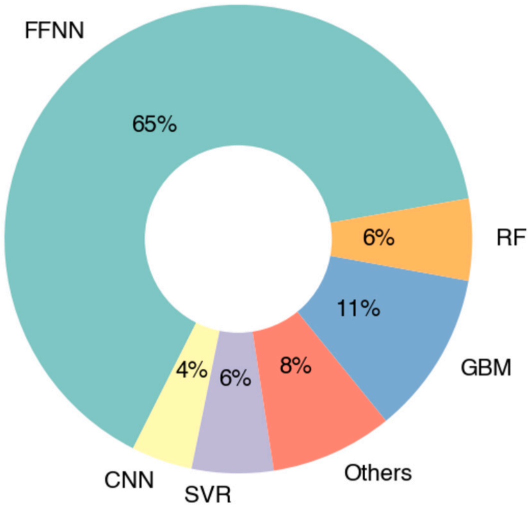

Fig. 8. ML models applied in design optimization across the reviewed studies. Among these, FFNN are the most common technique, with a growing integration of other DL models in the most recent studies.

Tree based models: Tree-based models, including decision trees, Random Forests (RFs), and Gradient Boosting Machines (GBMs), have become increasingly popular in recent studies involving surrogate models. A decision tree consists of multiple nodes, where the data are split at each node based on a feature that minimizes the error. RFs combine multiple decision trees to enhance accuracy and robustness [115]. GBMs are ensemble models in which trees are sequentially cascaded, with each subsequent model attempting to correct the errors made by its predecessor [116]. Tree-based models are favored for their explainability compared to most other ML models and were utilized in 12 recent studies.

CNNs: CNNs are a type of DL model particularly effective at capturing information from image data. They consist of convolutional layers that extract local patterns such as edges, textures, and shapes [117]. While CNNs primarily use 2D filters to extract features from images; 1D CNNs, which utilize 1D filters, are adept at capturing local patterns, such as trends or periodicity, in time series data [118]. For example, Mateusz et al. [63] utilized a CNN model to capture urban effects from image data, while Westermann et al. [40] applied a 1D CNN to extract information from weather data.

Other surrogate models: In addition to the models discussed, several other approaches have been employed as surrogates in the reviewed studies. Yue et al. [43] demonstrated that RL models can also serve as surrogates. Gaussian Process Regression (GPR) was used as a meta-model by Amini et al. [7], while Prada et al. [83] utilized Multivariate Adaptive Regression Splines (MARS) for the same purpose. Ensemble techniques were employed by Chen et al. [53], Yang et al. [71], and Shen et al. [52] to enhance model accuracy and robustness.

4.4.2. Model training and validation

Training is the process by which ML models learn the relationship between inputs and outputs. As highlighted in Table 2, FFNNs and CNNs are most suitable for addressing complex nonlinear relationships and spatiotemporal patterns, though they require significant data and computational resources. In comparison, SVR and tree-based models utilize simpler techniques with lower resource demands, making them more appropriate for smaller datasets and applications where model explainability is important. Despite these differences, all training methodologies aim to minimize the discrepancy between actual and predicted values. This discrepancy is typically evaluated using metrics such as Root Mean Square Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), which capture different aspects of error magnitude. The Coefficient of Determination (R^2) which measures the variance explained by the model, is also a commonly used metric. Other metrics, such as Relative Error (RE), Normalized Mean Bias Error (nMBE), and Huber Loss, are also used in some cases.

ML models generally perform well on training data but can become overly dependent on them, leading to poor generalization on unseen data. To evaluate generalization, all the reviewed studies used a separate test dataset from that used for training to evaluate model performance on unseen data. In most cases, the test data size ranged between 10% and 30% of the total dataset. Additionally, some studies included a validation set, which allows for performance evaluation during training and aids in fine-tuning hyperparameters or preventing overfitting (e.g., [69,55,87,51,75]). A few studies also employed k-fold cross-validation, where the data are divided into k subsets, with the model trained on k-1 subsets and tested on the remaining subset (e.g., [99,105,52,34,97]). This process is repeated k times, ensuring that each subset is used as test data exactly once.

ML models offer a significantly faster alternative to traditional simulation tools for evaluating building designs. Ji et al. [36] demonstrated that a surrogate model reduced evaluation time from approximately 1.2 minutes per design using the physics-based EnergyPlus model to nearly instantaneous, achieving over a 99% reduction in computational time. Similarly, Venkatraj et al. [44] reported that ML models provided results in under 1 second, compared to 2–3 minutes per run for EnergyPlus simulations. However, training these models and generating the training data can be computationally expensive, especially for DL models. Westermann et al. [40] noted that training a model with feature learning took approximately 8 hours using a Tesla K80 GPU but only 4 minutes with manually engineered features. Mateusz et al. [63] reported that training a CNN-based model required about 1 hour for 300 epochs with a batch size of 10 on a dataset of 3,000 samples. The efficiency of surrogate-assisted optimization also depends on training time, making training requirements a key factor in model selection.

4.5. Optimization

The final step in the surrogate model design process is optimization, where the goal is to identify the best design solutions based on predefined objectives. The optimization process typically leverages advanced algorithms to explore a vast design space, ensuring that the best possible outcomes are identified within the project's constraints.

Table 1 lists the optimization techniques identified in the reviewed studies, while Fig. 9 offers a summarized visualization of their usage. The remainder of this section examines the various optimization techniques employed in the reviewed studies, highlighting their applications and effectiveness in achieving sustainability goals.

Genetic Algorithm (GA): GA is inspired by the natural process of biological evolution, specifically the Darwinian theory of survival of the fittest. It operates through the processes of selection, crossover, and mutation, which are applied to a population of candidate solutions, referred to as chromosomes, to evolve improved solutions over successive generations [119]. In 10 out of the 55 reviewed studies that involved optimization, GA was employed as the optimization technique.

Non-Sorted Genetic Algorithm - II (NSGA-II): Traditional GAs often struggle with issues such as clustered solutions or poor approximation of the Pareto-optimal front. NSGA-II addresses these limitations by combining effective convergence strategies with robust diversity maintenance, making it superior to traditional GAs for multi-objective optimization [120]. NSGA-II is the most frequently used optimization algorithm in the literature, employed in 19 of the reviewed studies. However, more recent works have shown a preference for NSGA-III over NSGA-II due to its enhanced capabilities in handling many-objectives optimization problems.

Non-Sorted Genetic Algorithm - III (NSGA-III): NSGA-III is an extension of NSGA-II, designed to introduce a more robust mechanism for addressing many-objectives optimization problems [121]. This feature

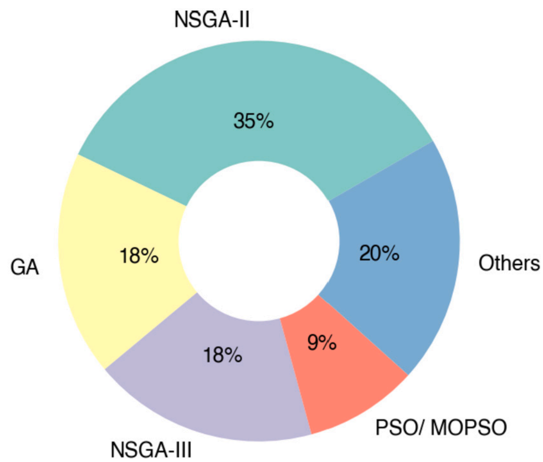

Fig. 9. Optimization algorithms applied in design optimization across the reviewed studies. NSGA-II emerges as the preferred optimization algorithm, while NSGA-III is increasingly being adopted in more recent studies.

is particularly crucial given the complexity and multitude of objectives typically encountered in building design optimization. Although NSGA-III was found in 10 of the reviewed studies, it has become the most frequently used algorithm in recent years, surpassing others in popularity.

Particle Swarm Optimization (PSO)/ Multi-Objective Particle Swarm Optimization (MOPSO): PSO is widely used due to its simplicity and effectiveness in handling complex, non-linear optimization problems [122]. MOPSO extends PSO's capabilities to address multi-objective optimization challenges. In the studies reviewed, PSO and MOPSO were employed in five works.

Other optimization algorithms: In addition to the previously discussed algorithms, other optimization techniques have also been applied, albeit with less frequency. These include Differential Evolution (DE) [8], Antlion optimizer [77], Ant Colony Optimization (ACO) [101], Multi-Criteria Optimization Algorithm (MCOA) [104], and Multi-Objective Evolutionary Algorithm (MOEA) [96,98]. The Strength Pareto Evolutionary Algorithm 2 (SPEA2) was employed by Sun et al. [93], while Pan et al. [92] utilized RL in the optimization process to determine the optimal parameters.

5. Analysis of papers based on their primary focus

In this section, the papers are discussed according to the approaches they followed. In Section 3, the main categories are listed, while here details are provided taking advantage of components discussed in that section. Table 1 provides a detailed categorization of each study considered in the review, aligning them with their primary focus.

5.1. Model development only

This section includes studies that focus exclusively on the development of surrogate models, without addressing the design variable optimization. As seen from Table 1, the primarily focus of these studies includes:.

DL-aided models: Mateusz et al. [63] integrated traditional ML with CNNs, using urban layout images to account for overshadowing effects in surrogate modeling. Their architecture combined CNN-based image analysis with FFNN numerical inputs to predict operational and embodied carbon footprints. Similarly, He et al. [42] trained ResNet (a CNN-based architecture) on residential floor plan images to predict DPL metrics and used a pix2pix GAN to visualize illuminance distributions.

Westermann et al. [40] proposed a location-independent surrogate model using a deep TCN to extract relevant features from high-dimensional weather data. The extracted features, combined with building design parameters, trained an FFNN to address the retraining challenges of traditional surrogate models. Yue et al. [43] compared six surrogate modeling techniques, including DL approaches such as LSTM, CNN, and Recurrent Deterministic Policy Gradient (RDPG). The RDPG model, combining LSTM and RL, outperformed the others, showcasing the potential of DL to replace traditional surrogate models.

Overall, these studies introduced new dimensions to the optimization process and demonstrated the effective application of modern DL techniques in this domain. However, such models often come with high data requirements and training demands. Their true potential lies in multi-scenario use cases, where a trained model can be leveraged for multiple optimization tasks, making the investment in training more worthwhile.

Data generation: Venkatraj et al. [44] proposed a framework for generating building datasets that automate energy simulations. By leveraging a parametric approach, the framework ensures efficiency and flexibility, enabling the exploration of diverse building designs without manual intervention. However, adapting it for varied optimization tasks remains a challenge.

Feature selection: Didwania et al. [64] proposed a surrogate model paradigm combined with the Morris method for feature selection. Similarly, Seyedzadeh et al. [65] employed the Sobol method to assess the impact of each feature and eliminate less significant ones. Feature selection methods significantly reduce computational requirements by eliminating unnecessary features early in the process.

Occupant behavior: Occupant behavior significantly influences energy consumption. Li et al. [66] generated a comprehensive database incorporating various occupancy pattern scenarios, which they used as a feature in the optimization of building design alongside other design variables. Although optimizing based on occupancy patterns is beneficial, modeling occupancy, particularly for medium or large-sized buildings, remains a significant challenge.

Explainable AI: Barbaresi et al. [67] employed SHapley Additive exPlanations (SHAP) to enhance the interpretability of predictions made by the eXtreme Gradient Boosting (XGB) model. Their objective was to improve the transparency of the model's decision-making process by quantifying the impact of each input feature on the predicted energy consumption of the building. Broader perspectives on explainability methods are discussed in comprehensive surveys, such as those by Mersha et al. [123] and Abusitta et al. [124], which provide detailed comparisons of interpretability techniques and their applications across various fields.

Shape descriptors: Storcz et al. [68] demonstrated shape-focused optimization potential by analyzing shape parameter impact on energy demand and comfort. Using a 3D Coordinates descriptor with 18 spatial points, they dynamically represented building layouts. However, adapting to complex or irregular designs may require advanced techniques such as GANs or DL models for better spatial pattern capture.

Ventilation: Alghamdi et al. [69] developed a simple FFNN surrogate model to predict energy consumption, incorporating mechanical ventilation to account for its impact on the optimization process.

Optimized FFNN: Himmetoglu et al. [70] introduced the Particle Swarm and Ant Colony Optimization Neural Network (PSACONN) approach, a hybrid method combining PSO and ACO to train a neural network. Instead of traditional backpropagation, this method employs optimization algorithms to determine the optimal weights. However, the absence of a direct comparison with traditional backpropagation limits the ability to evaluate its relative performance and efficiency.

5.2. Model development with performance tuning

Similar to the previous section, this section also focuses on studies that primarily develop surrogate models. However, these works incorporate an additional performance tuning step to further optimize the models' accuracy.

Garcia et al. [48] applied a GA to optimize both the structure and hyperparameters of an FFNN, aiming to improve model performance.

Similarly, Han et al. [49] used Bayesian optimization to determine the optimal hyperparameters for an FFNN, including the number of hidden layers, neurons per layer, and batch size.

Cai et al. [50] conducted a comparative study, optimizing an SVR model for predicting heating and cooling loads. The SVR parameters were tuned using six different meta-heuristic algorithms: Artificial Ecosystem-based Optimization (AEO), Artificial Bee Colony (ABC), Slime Mold Algorithm (SMA), Arithmetic Optimization Algorithm (AOA), Sparrow Search Algorithm (SSA), and Gray Wolf Optimizer (GWO).

5.3. Model development and design optimization

In this section, we examine studies where a surrogate model is developed and subsequently utilized to derive optimal building design parameters for a specific building. This category comprises the majority of the literature reviewed. Moreover, we analyze the approaches employed in the optimization process across these studies, highlighting key methodologies and their effectiveness in improving design outcomes.

Climate change: Using past weather data in optimization may overlook future climate shifts. Several studies address this by incorporating future climate scenarios into the optimization process.

Zou et al. [51] employed General Circulation Models (GCM) and Representative Concentration Pathways (RCP) to generate high-resolution future weather data through a morphing technique, which was then used in simulations to create training datasets. Similarly, Li et al. [35] applied GCM-based approach with future weather data for Huangshan, China, spanning 2020–2050. Uncertainties in future weather predictions can affect optimization accuracy and reliability, and should be carefully considered.

Ensemble models: In ensemble models, multiple ML models are combined to reduce variance, typically through bagging or boosting. Chen et al. [53] combined RF, GBRT, and FFNN models into an ensemble to predict carbon emissions. Similarly, Shen et al. [52] used LightGBM, RF, SVR, naive Bayes, and LSTM as base models, with an XGBoost meta-model for final predictions. Yang et al. [71] integrated outputs from eight base models using a multiple linear regression meta-model. While these methods enhance predictive performance, they also increase computational costs and complexity of training.

Metaheuristic optimization: In metaheuristic optimization, the surrogate model is trained using an optimization algorithm instead of traditional backpropagation. Yu et al. [54] used a GA to optimize the weights of an FFNN surrogate model. Chegari et al. [55] adopted a hybrid approach, applying Gradient Descent (GD) for initial weight estimates followed by PSO to refine them. A similar method was applied by Chegari et al. [72] in a different study. A detailed comparison of metaheuristic optimization and traditional backpropagation is essential to highlight the advantages of metaheuristic methods.

Systematic sampling: This approach samples data systematically to cover the relevant design space. Zheng et al. [57] performed sensitivity analysis on an initial 900 samples to identify six key parameters affecting energy consumption, which were then varied to generate 2,160 additional samples.

Python framework: Hocine et al. [58] demonstrated that a Python-based framework can be employed to generate simulation data for building surrogate models.

Renovation: Asadi et al. [41] utilized a surrogate model to identify the most energy-efficient retrofitting measures for an old school building. Similarly, Arjomandnia et al. [59] employed a redesign approach aiming to minimize energy consumption through targeted renovations. When performing renovations, it is crucial to carefully calibrate the simulation tool to accurately represent the existing building.

Feature selection: Chen et al. [37] used relative weight analysis and FAST for prioritizing design parameters, while Li et al. [73] employed standardized regression coefficients. Chen et al. [61] combined methods such as FAST, Sobol, and correlation coefficients to identify critical parameters. Wang et al. [75] ranked the impact of design variables on housing thermal loads using sensitivity analysis. Similarly, Serbouti et al. [74] introduced SAMOT, integrating sensitivity analysis with optimization.

Razmi et al. [76] applied Principal Component Analysis (PCA) to reduce dimensionality of design variables while Zhan et al. [60] used Least Absolute Shrinkage and Selection operator (LASSO) to select relevant features.

Passive energy elements: Passive heating and cooling elements are essential for reducing energy consumption and are commonly included in optimization processes. Gou et al. [33] covered shading and ventilation parameters during optimization. Specifically, they used dimensions of shades and fins, window opening factor, opening control type, and air mass flow coefficient to control the passive gains. Similarly, Li et al. [78] emphasized natural ventilation, focusing on parameters such as window structure, opening factor, and effective opening area.

Lin et al. [77] focused on green roofs, optimizing parameters such as vegetation height, leaf area index, reflectivity, substrate dimensions, and thermal conductivity. Alsharif et al. [34], on the other hand, focused solely on optimizing shading elements, using parameters exclusively related to shading dimensions.

Active energy elements: Like passive elements, active electrical equipment influences energy consumption. Studies reviewed here demonstrate how Tier 3 components can be selected or optimized at the design stage. Magnier et al. [79] incorporated various active element features, such as HVAC set points, relative humidity settings, supply airflow rates for heating and cooling, and thermostat delays, in their optimization process. Amini et al. [7] optimized the battery capacity and the sizes of thermal/electrical energy storage systems used for heating and cooling at the design stage.

Li et al. [8] compared two AC operation strategies to determine the optimal approach for energy efficiency, while Zong et al. [9] included parameters such as heater type, control system type, and heat generator type as inputs to the surrogate model. On the other hand, Xu et al. [80] included PhotoVoltaic (PV) system parameters, optimizing the tilt and azimuth angles of the PV system alongside the building envelope design.

Case studies: The studies in this category primarily focus on applying standard methodologies to address case-specific challenges, offering valuable insights for those particular contexts. To enhance readability, the studies are grouped by the type of building or scenario they address:

Residential Buildings: Gossard [81] analyzed optimal building designs for a single-story residential building under two French climates, Nancy (continental) and Nice (Mediterranean), demonstrating the need for climate-specific designs to meet distinct thermal performance requirements. Elbeltagi et al. [91] optimized energy performance for a single-family home in New Cairo, Egypt, using a Visual Basic interface to simplify the input of key design parameters.

Office or High-Rise Buildings: Zhao et al. [86] developed a web-based tool using WebGL and Three.js to visualize optimized windows for high-rise office buildings with box-like geometries, employing an FFNN model with a GA.

Educational and Institutional Buildings: Xu et al. [80] applied FFNN for energy load optimization tailored to a school building in China's cold climate. Wu et al. [89] conducted a case study on a university teaching building in Wuhan, China, using an RF-NSGA-III algorithm to explore energy-efficient design options.

Multi-Building Scenarios: Wang et al. [87] optimized the layout for 12 structures on a Beijing site, improving indoor visual comfort and outdoor thermal performance. Si et al. [85] utilized surrogate models to streamline computational demands in a complex multi-building scenario, focusing on the tourist center at Niushou Mountain Park.

Prefabricated and Retrofit Buildings: Ji et al. [36] examined the optimization process for a prefabricated house, focusing on how optimization can be tailored to construction using pre-manufactured components. Prada et al. [83] applied surrogate modeling to optimize retrofit configurations for existing buildings, addressing the unique requirements of retrofitting scenarios.

General Optimization Frameworks: Li et al. [82] implemented an optimization framework with NSGA-II and MOPSO algorithms in MATLAB, demonstrating its application across various scenarios. Saryazdi et al. [90] combined GA coded in MATLAB with an ANN model for ANN-based optimization, showcasing its adaptability to diverse case studies.

In addition to serving as references for specific use cases, these studies offer insights applicable to similar climatic conditions and building designs. While these studies considered real-world scenarios and carried out extensive analysis, they did not follow through with the implementation and analysis of the studied cases.

5.4. Design optimization with tuned model

In this section, we review studies that integrate surrogate models with performance tuning techniques to enhance accuracy and then employ the model for design optimization. These studies can also be seen as two-stage optimization processes: first, the surrogate model itself is optimized through techniques such as hyperparameter tuning or adaptive sampling, and second, the optimized model is used to derive optimal building design parameters. The three main categories here are:

RL optimization: Pan et al. [92] trained an FFNN surrogate model, employing Bayesian optimization to tune its hyperparameters. They then used RL to optimize design variables, where the RL agent adjusted parameters based on building performance and received rewards for improvements predicted by the surrogate model.

Parameter optimization: Here, we review works that emphasize advanced tuning strategies to enhance model accuracy in the design optimization process. Sun et al. [93] used the Octopus plug-in, which applies an evolutionary algorithm to refine FFNN hyperparameters, while Garcia et al. [94] applied a GA to optimize the FFNN hyperparameters. Liu et al. [95] and Wu et al. [97] utilized GWO to tune key hyperparameters—specifically, the number of trees and the number of random features—in a RF model. Following a similar strategy, Shen et al. [96] and Khan et al. [98] employed Bayesian optimization to fine-tune the LightGBM hyperparameters. Six ML models-ridge regression, RF, XGBoost, SVR, K-NN regression, and MLP-were optimized in the study by Si et al. [99] using various tuning techniques including Bayesian optimization, random search, GA, and PSO. XGBoost, optimized using PSO achieved the highest model precision.

Adaptive sampling: Determining the optimal sample set for surrogate model training is challenging, so adaptive sampling iteratively expands the dataset, focusing on regions needing refinement to improve performance efficiently. An adaptive sampling based on Augmented LHS was used to train an FFNN-based surrogate model for multi-objective building performance optimization [100]. Starting with an initial LHS sample, authors progressively expanded samples using the augmentLHS technique. Similarly, augmentLHS method was used by Yue et al. [125] and You et al. [104].

Work by Bamdad et al. [101] carried out initial sampling with LHS, followed by a committee of surrogate models to predict not simulated cases. High-variance predictions were selected for further simulation, with a focus on low-energy cases. Lahmar et al. [105] tested four adaptive sampling methods including error-based sampling, Lola-Voronoi, sample minimum strategy, and expected improvements: error-based and sample minimum methods showed faster convergence and better accuracy. An active learning-based adaptive sampling approach for building energy optimization, using a metamodel called MEVO, was employed by Batres et al. [103]. This approach iteratively selects new sampling points based on model optimization needs.

6. Challenges and opportunities

The previous sections reviewed the surrogate modeling process and methodologies used to enhance the efficiency and accuracy of surrogate-assisted design optimization. This section identifies and discusses the trends and opportunities related to surrogate model-assisted design optimization studies.

Limited reusability: One of the major challenges with surrogate models is their limited reusability. In most studies, surrogate models were developed for specific sites and often discarded after optimization. In contrast, Westermann et al. [40] enhanced their surrogate model by incorporating features extracted from weather data, making it applicable across different geographical locations. Similarly, Zheng et al. [57] and Himmetoglu [70] included weather-related features to generalize the model for various climate conditions. However, there remains an opportunity for further research to develop models that are reusable across diverse climate conditions and building types, requiring minimal retraining with only a few additional simulations.

Wind modeling: Wind significantly affects a building's energy consumption, and the wind patterns around a building are influenced by external factors such as surrounding trees and neighboring structures. These external factors, such as thees, also impact the direct sunlight and further affect energy use. Mateusz et al. [63] employed a CNN model to capture the impact of surrounding buildings on the energy consumption of the main building. Further research is needed for modeling the effects of wind blockers, such as nearby buildings and vegetation, and integrating these factors into the optimization process. Emerging techniques from the generative AI category, including Generative Adversarial Networks (GANs) and diffusion models, are gaining prominence across a diversity of domains and could be beneficial for this purpose.

Building shape: The shape of a building influences its aerodynamics and the wind flow around it, affecting the building's energy consumption. However, making substantial changes to the building shape between simulations can be time-consuming and remains underexplored. Storcz et al. [68] utilized building geometry descriptors in their model to approximate energy simulations, which captures the impact of shape. Compactness, a commonly used feature in many studies, also implicitly accounts for shape effects. Recent advancements in DL could be leveraged to modify building shapes more efficiently between simulations, allowing the shape to be directly included as a feature in the optimization process.

Model explainability: FFNNs remain the most dominant models for surrogate modeling. However, tree-based models are gaining popularity due to their explainability, which has become increasingly important in recent studies. Shen et al. [96] utilized SHAP to provide insights into how the models arrive at their predictions, addressing the black-box nature of ML models. Similarly, Khan et al. [98] employed Local Interpretable Model-Agnostic Explanations (LIME) to explain the model's predictions. As explainability becomes more important in ML applications, adopting better explainable practices in surrogate modeling is becoming a must.

Active learning and selective sampling: LHS is commonly used for generating samples in surrogate modeling. However, adaptive sampling methods have also proven effective in determining the optimal number of samples required for model training, improving both efficiency and accuracy. Further exploration of active learning techniques could improve the model's performance by selectively acquiring additional samples where the model is uncertain, thereby refining predictions and optimizing resource use.

Ethical impact: When employing AI and ML in surrogate-assisted optimization, it is important to consider challenges such as potential biases in algorithms, data privacy, and the impact on employment within the building design and construction sectors. Future studies should also consider these ethical implications to ensure responsible and equitable adoption of optimization techniques.

To provide researchers and practitioners with the guidance to select appropriate methodologies for specific use cases, we present a guidance framework in Fig. 10. The first step is to determine if the intention of modeling is for a single-use case or multi-scenario adaptability. Based on that, the process continues to the core aim of the case study such

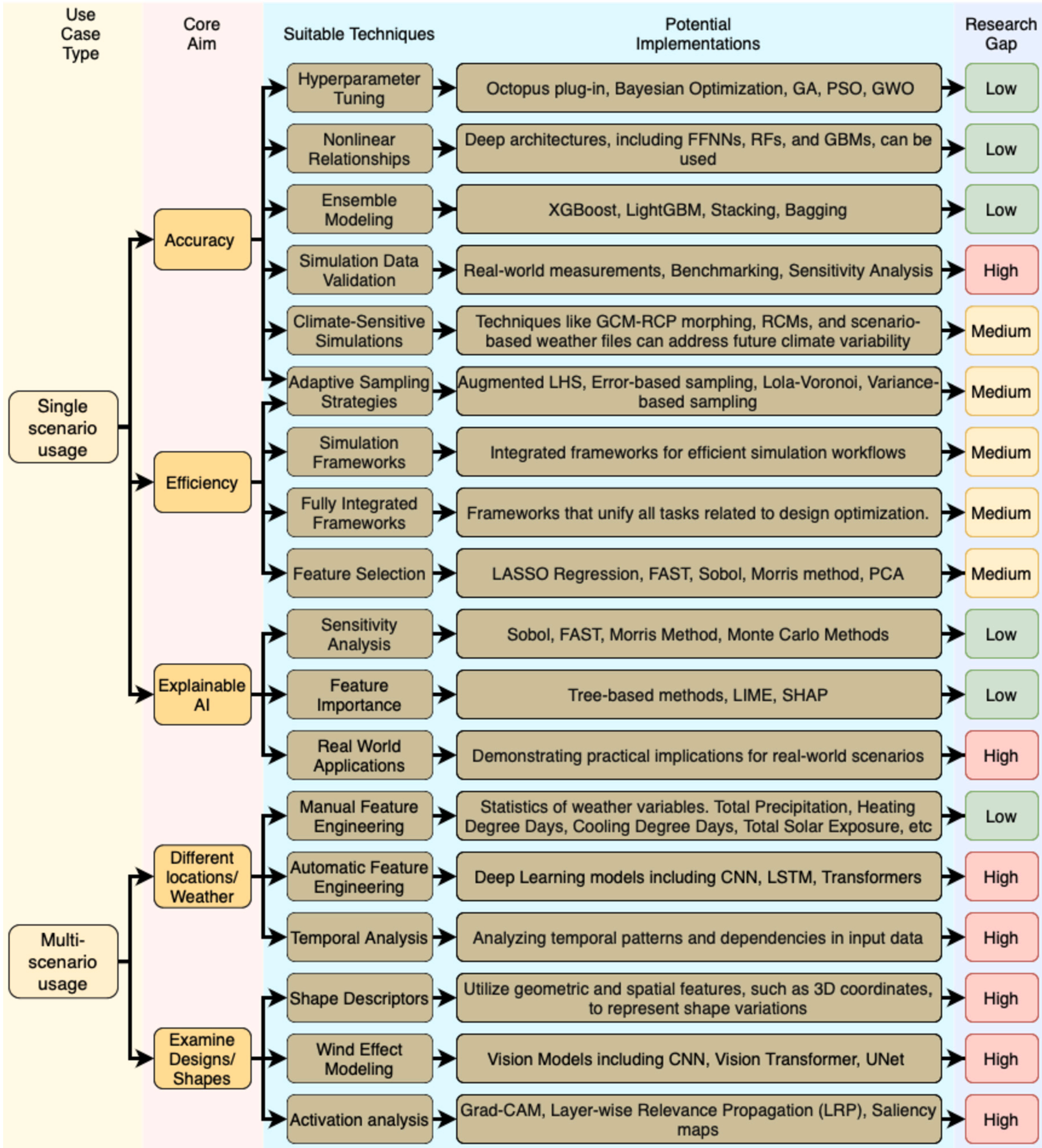

Fig. 10. Framework providing guidance for researchers and practitioners in selecting appropriate methodologies for specific use cases. Additionally, the research gaps are highlighted.

as high accuracy, efficiency, or explainability. Depending on this core objective, various techniques are available: the figure lists the suitable techniques with the main implementation options. Finally, the research gap column indicates the state of the research with respect to various techniques to provide researchers with insights into areas requiring further investigation or where existing knowledge is limited.

In summary, while surrogate modeling has made significant strides, challenges and opportunities for advancement remain. Key needs include developing reusable models adaptable to diverse climates and building types, improving wind and shape modeling for precise environmental impacts, and enhancing model explainability with interpretable models and tools. Additionally, active learning and adaptive sampling are promising methods for boosting model efficiency and accuracy. Addressing these areas can further refine surrogate-assisted design optimization for sustainable building design.

7. Conclusion

This review examined recent advancements in AI-driven surrogate models for building design optimization, focusing on their role in enhancing sustainability. By integrating strategies from Lechner's three-tier approach—covering fundamental design decisions, passive energy strategies, and mechanical and electrical system optimizations—AI technologies, particularly surrogate models, offer opportunities to improve sustainable building design processes. Our analysis highlighted methodologies, such as DL models, data generation frameworks, feature selection methods, and advanced sampling techniques, that have been employed to enhance the accuracy and applicability of surrogate-assisted optimization.

While surrogate models present promising advantages in reducing computational costs and expanding design exploration, challenges such as model reusability across different climates and contexts, and the need for improved interpretability, remain. Future research directions could focus on developing more generalizable models, leveraging advanced DL techniques to better account for factors such as building shape and external conditions, and integrating explainable AI practices to address the black-box nature of these models.

Overall, surrogate modeling represents a transformative approach in the field of sustainable building design, with continued innovation needed to maximize its potential. By focusing on these emerging opportunities and challenges, researchers and practitioners can further advance the application of AI to create more sustainable, efficient, and resilient buildings.

CRediT authorship contribution statement

Piragash Manmatharasan: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Girma Bitsuamlak:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Katarina Grolinger:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization.

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

During the preparation of this work, the authors used ChatGPT in order to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

This work was supported by the Climate Action and Awareness Fund under Grant EDF-CA-2021i018, Environnement Canada.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Katarina Grolinger reports financial support was provided by Environment and Climate Change Canada. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

Table A.1
Design variables related to wall and roof, along with their usage in the reviewed literature.

Publication	Wall thermal cond.	Wall specific heat	Wall heat inertia	Wall heat transfer coef.	Wall insulation thick.	Wall insulation type	Wall solar reflectance	Wall solar absorbance	Wall area	Wall width	Roof thermal cond.	Roof specific heat	Roof heat inertia	Roof heat transfer coef.	Roof insulation thick.	Roof insulation type	Roof solar reflectance	Roof solar absorbance	Roof area	Roof width
[81]	✗	✗									✗	✗								
[41]						✗										✗				
[54]			✗	✗									✗	✗						
[33]					✗	✗		✗							✗			✗		
[48]								✗										✗		
[40]					✗					✗					✗					
[93]						✗														✗
[100]						✗		✗								✗				
[101]					✗			✗										✗		
[51]					✗	✗		✗												
[102]						✗										✗				
[49]					✗		✗													
[77]						✗	✗									✗	✗			
[55]				✗				✗						✗				✗		
[94]						✗										✗				
[72]				✗				✗						✗				✗		
[7]				✗										✗						
[88]					✗										✗					
[35]				✗	✗			✗	✗					✗	✗			✗	✗	
[95]				✗				✗						✗				✗		
[44]						✗														
[50]									✗										✗	
[97]				✗				✗						✗				✗		
[78]						✗														
[96]					✗										✗					
[91]						✗										✗				
[53]				✗				✗						✗				✗		
[103]					✗										✗					
[62]	✗										✗									
[8]						✗														
[52]	✗																			
[36]					✗										✗					
[105]	✗										✗									
[99]					✗	✗									✗	✗				
[9]	✗										✗									
[60]	✗										✗									

Table A.1 (*continued*)

Publication	Wall thermal cond.	Wall specific heat	Wall heat inertia	Wall heat transfer coef.	Wall insulation thick.	Wall insulation type	Wall solar reflectance	Wall solar absorbtance	Wall area	Wall width	Roof thermal cond.	Roof specific heat	Roof heat inertia	Roof heat transfer coef.	Roof insulation thick.	Roof insulation type	Roof solar reflectance	Roof solar absorbtance	Roof area	Roof width
[92]	✗								✗		✗								✗	
[71]					✗										✗					
[43]					✗	✗									✗	✗				
[64]					✗										✗					
[104]						✗		✗								✗				
[90]					✗															
[61]	✗										✗									
[70]					✗	✗														
[66]						✗														
[83]					✗															
[37]		✗				✗														
[98]				✗																
[69]				✗										✗						
[57]				✗										✗						
[58]				✗				✗						✗				✗		
[59]				✗										✗						
[89]				✗				✗						✗				✗		
[67]		✗				✗						✗				✗				
[80]	✗				✗			✗												
[85]	✗				✗						✗				✗					
[65]									✗										✗	
[84]				✗										✗						
[82]					✗										✗					
[68]				✗										✗						

Table A.2
Design variables related to windows, compactness, shading, ventilation, air tightness, and active elements, along with their usage in the reviewed literature.

Publication	Orientation	Window Dimensions	Glazing type	Window to wall ratio	U value	SHGC	Window transmitivity	Number of floors	Floor dimensions	Shading transmittance	shading device dimension	Window area ventilation	Window opening factor	Ventilation flow rate	Infiltration	People activity	Equipment load	HVAC set points	Equipment selection
[79]		✗																✗	
[41]			✗																✗
[54]	✗			✗	✗			✗	✗										
[33]	✗			✗	✗	✗					✗				✗				
[73]																			
[48]	✗			✗		✗									✗	✗	✗		
[40]	✗			✗	✗	✗		✗							✗		✗	✗	
[63]																			
[93]		✗	✗						✗										
[100]	✗	✗	✗	✗							✗	✗			✗				
[101]	✗	✗									✗							✗	
[51]	✗	✗			✗	✗													
[102]				✗	✗	✗						✗	✗		✗				✗
[86]	✗			✗				✗	✗										
[49]	✗	✗					✗		✗										
[77]			✗						✗						✗				
[55]			✗						✗						✗				
[94]	✗		✗								✗				✗		✗	✗	✗
[72]			✗						✗						✗				
[7]					✗	✗			✗									✗	
[76]	✗			✗					✗										
[88]	✗		✗	✗															
[35]			✗		✗				✗						✗				
[95]				✗	✗													✗	
[44]	✗		✗	✗				✗	✗										
[50]	✗	✗							✗			✗							
[97]				✗	✗										✗			✗	
[78]	✗		✗	✗								✗	✗						

Table A.2 (*continued*)

Publication	Orientation	Window Dimensions	Glazing type	Window to wall ratio	U value	SHGC	Window transmitivity	Number of floors	Floor dimensions	Shading transmittance	shading device dimension	Window area ventilation	Window opening factor	Ventilation flow rate	Infiltration	People activity	Equipment load	HVAC set points	Equipment selection
[96]			✗	✗		✗									✗	✗	✗	✗	
[91]	✗			✗	✗	✗	✗		✗									✗	
[53]	✗			✗	✗	✗	✗							✗			✗	✗	
[103]					✗	✗													
[62]					✗	✗											✗	✗	
[34]											✗								
[8]	✗		✗	✗							✗								
[52]				✗	✗	✗			✗										
[36]	✗		✗												✗				
[105]	✗		✗	✗															
[99]	✗	✗			✗	✗				✗								✗	
[9]									✗										✗
[60]			✗	✗															
[92]				✗	✗	✗								✗	✗			✗	
[71]				✗	✗	✗													
[43]				✗	✗	✗						✗			✗				✗
[64]	✗			✗		✗					✗				✗		✗		
[104]	✗	✗	✗						✗	✗		✗							
[90]							✗							✗			✗	✗	
[61]	✗			✗	✗	✗								✗	✗		✗	✗	
[66]					✗	✗									✗			✗	
[83]			✗																✗
[37]	✗			✗	✗	✗	✗				✗				✗				
[98]				✗	✗	✗			✗				✗		✗			✗	
[69]	✗		✗	✗							✗		✗	✗	✗			✗	
[57]	✗			✗	✗	✗		✗									✗		
[58]			✗	✗															
[89]				✗	✗	✗									✗			✗	
[67]	✗						✗								✗				
[80]				✗	✗	✗	✗				✗								
[85]			✗								✗							✗	
[65]	✗	✗							✗								✗		
[84]				✗	✗	✗			✗								✗		
[82]			✗	✗														✗	
[68]	✗			✗	✗				✗										

Data availability

No data was used for the research described in the article.

References

[1] U. N. E. Programme, Building materials and the climate: constructing a new future, https://wedocs.unep.org/20.500.11822/43293, 2023.

[2] N. Lechner, Heating, Cooling, Lighting: Sustainable Design Methods for Architects, John Wiley & Sons, 2014.

[3] K. Negendahl, Building performance simulation in the early design stage: an introduction to integrated dynamic models, Autom. Constr. 54 (2015) 39–53, https://doi.org/10.1016/j.autcon.2015.03.002.

[4] J. Westerman, Surrogate models for building sustainability: applications and impact, Energy Build. 198 (2019) 170–182, https://doi.org/10.1016/j.enbuild.2019.05.057.

[5] I. Elwy, A. Hagishima, The artificial intelligence reformation of sustainable building design approach: a systematic review on building design optimization methods using surrogate models, Energy Build. 323 (2024) 114769, https://doi.org/10.1016/j.enbuild.2024.114769.

[6] A. Cruz, L. Caldas, V. Mendes, J. Mendes, L. Bastos, Multi-objective optimization based on surrogate models for sustainable building design: a systematic literature review, Build. Environ. 266 (2024) 112147, https://doi.org/10.1016/j.buildenv.2024.112147.

[7] H.A. Toosi, M. Lavagna, F. Leonforte, C.D. Pero, N. Aste, A novel LCSA-machine learning based optimization model for sustainable building design-a case study of energy storage systems, Build. Environ. 209 (2022) 108656, https://doi.org/10.1016/j.buildenv.2021.108656.

[8] C. Li, Y. Chen, Modeling and optimization method for building energy performance in the design stage, J. Build. Eng. 87 (2024) 109019, https://doi.org/10.1016/j.jobe.2024.109019.

[9] C. Zong, X. Chen, F. Deghim, J. Staudt, P. Geyer, W. Lang, A holistic two-stage decision-making methodology for passive and active building design strategies under uncertainty, Build. Environ. 251 (2024) 111211, https://doi.org/10.1016/j.buildenv.2024.111211.

[10] J. Chen, X. Gao, Y. Hu, Z. Zeng, Y. Liu, A meta-model-based optimization approach for fast and reliable calibration of building energy models, Energy 188 (2019) 116046, https://doi.org/10.1016/j.energy.2019.116046.

[11] A. Kusiak, M. Li, F. Tang, Modeling and optimization of HVAC energy consumption, Appl. Energy 87 (2010) 3092–3102, https://doi.org/10.1016/j.apenergy.2010.04.008.

[12] A. Kusiak, F. Tang, G. Xu, Multi-objective optimization of HVAC system with an evolutionary computation algorithm, Energy 36 (2011) 2440–2449, https://doi.org/10.1016/j.energy.2011.01.030.

[13] P.M. Ferreira, A.E. Ruano, S. Silva, E.Z. Conceição, Neural networks based predictive control for thermal comfort and energy savings in public buildings, Energy Build. 55 (2012) 238–251, https://doi.org/10.1016/j.enbuild.2012.08.002.

[14] A. Afram, F. Janabi-Sharifi, Theory and applications of HVAC control systems - a review of Model Predictive Control (MPC), Build. Environ. 72 (2014) 343–355, https://doi.org/10.1016/j.buildenv.2013.11.016.

[15] A. Afram, F. Janabi-Sharifi, A.S. Fung, K. Raahemifar, Artificial Neural Network (ANN) based Model Predictive Control (MPC) and optimization of HVAC systems: a state of the art review and case study of a residential HVAC system, Energy Build. 141 (2017) 96–113, https://doi.org/10.1016/j.enbuild.2017.02.012.

[16] H. Huang, L. Chen, E. Hu, A new model predictive control scheme for energy and cost savings in commercial buildings: an airport terminal building case study, Build. Environ. 89 (2015) 203–216, https://doi.org/10.1016/j.buildenv.2015.01.037.

[17] W. Kim, Y. Jeon, Y. Kim, Simulation-based optimization of an integrated daylighting and hvac system using the design of experiments method, Appl. Energy 162 (2016) 666–674, https://doi.org/10.1016/j.apenergy.2015.10.153.

[18] A. Ghahramani, K. Zhang, K. Dutta, Z. Yang, B. Becerik-Gerber, Energy savings from temperature setpoints and deadband: quantifying the influence of building

and system properties on savings, Appl. Energy 165 (2016) 930–942, https://doi.org/10.1016/j.apenergy.2015.12.115.
[19] W. Zhang, F. Liu, R. Fan, Improved thermal comfort modeling for smart buildings: a data analytics study, Int. J. Electr. Power Energy Syst. 103 (2018) 634–643, https://doi.org/10.1016/j.ijepes.2018.06.026.
[20] X. Dai, J. Liu, X. Zhang, A review of studies applying machine learning models to predict occupancy and window-opening behaviours in smart buildings, https://doi.org/10.1016/j.enbuild.2020.110159, 2020.
[21] L. Bereketeab, A. Zekeria, M. Aloqaily, M. Guizani, M. Debbah, Energy optimization in sustainable smart environments with machine learning and advanced communications, IEEE Sens. J. 24 (2024) 5704–5712, https://doi.org/10.1109/JSEN.2024.3355229.
[22] Y. Yan, J. Cai, Y. Tang, L. Chen, Fault diagnosis of HVAC AHUs based on a BP-MTN classifier, Build. Environ. 227 (2023) 109779, https://doi.org/10.1016/j.buildenv.2022.109779.
[23] J. Bi, H. Wang, E. Yan, C. Wang, K. Yan, L. Jiang, B. Yang, AI in HVAC fault detection and diagnosis: a systematic review, Energy Rev. 3 (2024) 100071, https://doi.org/10.1016/j.enrev.2024.100071.
[24] F. Hodavand, I.J. Ramaji, N. Sadeghi, Digital twin for fault detection and diagnosis of building operations: a systematic review, Buildings 13 (2023), https://doi.org/10.3390/buildings13061426.
[25] Y. Zhao, T. Li, X. Zhang, C. Zhang, Artificial intelligence-based fault detection and diagnosis methods for building energy systems: advantages, challenges and the future, Renew. Sustain. Energy Rev. 109 (2019) 85–101, https://doi.org/10.1016/j.rser.2019.04.021.
[26] D. Dey, B. Dong, A probabilistic approach to diagnose faults of air handling units in buildings, Energy Build. 130 (2016) 177–187, https://doi.org/10.1016/j.enbuild.2016.08.017.
[27] M. Dey, S.P. Rana, S. Dudley, A case study based approach for remote fault detection using multi-level machine learning in a smart building, Smart Cities 3 (2020) 401–419, https://doi.org/10.3390/smartcities3020021.
[28] S.R. West, Y. Guo, S.R. West, X.R. Wang, J. Wall, Automated fault detection and diagnosis of hvac subsystems using statistical machine learning, in: 12th International Conference of the International Building Performance Simulation Association, Sydney, Australia, 2011.
[29] J. Liang, R. Du, Model-based fault detection and diagnosis of HVAC systems using support vector machine method, Int. J. Refrig. 30 (2007) 1104–1114, https://doi.org/10.1016/j.ijrefrig.2006.12.012.
[30] Z. Wang, J. Guo, S. Zhou, P. Xia, Performance evaluation of chiller fault detection and diagnosis using only field-installed sensors, Processes 11 (2023) 3299, https://doi.org/10.3390/pr11123299.
[31] L. Sehovac, K. Grolinger, Deep learning for load forecasting: sequence to sequence recurrent neural networks with attention, IEEE Access 8 (2020) 36411–36426, https://doi.org/10.1109/ACCESS.2020.2975738.
[32] A. L'Heureux, K. Grolinger, M.A.M. Capretz, Transformer-based model for electrical load forecasting, Energies 15 (2022) 4993, https://doi.org/10.3390/en15144993.
[33] S. Gou, V.M. Nik, J.-L. Scartezzini, Q. Zhao, Z. Li, Passive design optimization of newly-built residential buildings in Shanghai for improving indoor thermal comfort while reducing building energy demand, Energy Build. 169 (2018) 484–506, https://doi.org/10.1016/j.enbuild.2017.09.095.
[34] R. Alsharif, M. Arashpour, E. Golafshani, A. Rashidi, H. Li, Multi-objective optimization of shading devices using ensemble machine learning and orthogonal design of experiments, Energy Build. 283 (2023) 112840, https://doi.org/10.1016/j.enbuild.2023.112840.
[35] Z. Li, Y. Zhao, H. Xia, S. Xie, A multi-objective optimization framework for building performance under climate change, J. Build. Eng. 80 (2023) 107978, https://doi.org/10.1016/j.jobe.2023.107978.
[36] Y. Ji, J. Lv, H.X. Li, Y. Liu, F. Yao, X. Liu, S. Wang, Improving the performance of prefabricated houses through multi-objective optimization design, J. Build. Eng. 84 (2024) 108579, https://doi.org/10.1016/j.jobe.2024.108579.
[37] Y. Chen, L.K. Norford, H.W. Samuelson, A. Malkawi, Optimal control of HVAC and window systems for natural ventilation through reinforcement learning, Energy Build. 169 (2018) 195–205, https://doi.org/10.1016/j.enbuild.2018.03.051.
[38] R.Z. Homod, K.S.M. Sahari, H.A. Almurib, Energy saving by integrated control of natural ventilation and HVAC systems using model guide for comparison, Renew. Energy 71 (2014) 639–650, https://doi.org/10.1016/j.renene.2014.06.015.
[39] N. Ramessur, M. Gooroochurn, Automated passive measures: the next step in reducing the carbon footprint of our buildings, in: 2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), Brisbane, Australia, 2021.
[40] P. Westermann, M. Welzel, R. Evins, Using a deep temporal convolutional network as a building energy surrogate model that spans multiple climate zones, Appl. Energy 278 (2020) 115563, https://doi.org/10.1016/j.apenergy.2020.115563.
[41] E. Asadi, M.G. da 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, https://doi.org/10.1016/j.enbuild.2014.06.009.
[42] Q. He, Z. Li, W. Gao, H. Chen, X. Wu, X. Cheng, B. Lin, Predictive models for daylight performance of general floorplans based on CNN and GAN: a proof-of-concept study, Build. Environ. 206 (2021) 108346, https://doi.org/10.1016/j.buildenv.2021.108346.
[43] N. Yue, M. Caini, L. Li, Y. Zhao, Y. Li, A comparison of six metamodeling techniques applied to multi building performance vectors prediction on gymnasiums under multiple climate conditions, Appl. Energy 332 (2023) 120481, https://doi.org/10.1016/j.apenergy.2022.120481.
[44] V. Venkatraj, M. Dixit, W. Yan, S. Caffey, P. Sideris, A. Aryal, Toward the application of a machine learning framework for building life cycle energy assessment, Energy Build. 297 (2023) 113444, https://doi.org/10.1016/j.enbuild.2023.113444.
[45] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, Preferred reporting items for systematic reviews and meta-analyses: the prisma statement, PLoS Med. 6 (2009) e1000097, https://doi.org/10.1371/journal.pmed.1000097.
[46] D. Meng, S. Yang, A.M.D. Jesus, T. Fazeres-Ferradosa, S.-P. Zhu, A novel hybrid adaptive kriging and water cycle algorithm for reliability-based design and optimization strategy: application in offshore wind turbine monopile, Comput. Methods Appl. Mech. Eng. 412 (2023) 116083, https://doi.org/10.1016/j.cma.2023.116083.
[47] D. Meng, H. Yang, S. Yang, Y. Zhang, A.M.D. Jesus, J. Correia, T. Fazeres-Ferradosa, W. Macek, R. Branco, S.-P. Zhu, Kriging-assisted hybrid reliability design and optimization of offshore wind turbine support structure based on a portfolio allocation strategy, Ocean Eng. 295 (2024) 116842, https://doi.org/10.1016/j.oceaneng.2024.116842.
[48] I.G. Kerdan, D.M. Gálvez, Artificial neural network structure optimisation for accurately prediction of exergy, comfort and life cycle cost performance of a low energy building, Appl. Energy 280 (2020) 115862, https://doi.org/10.1016/j.apenergy.2020.115862.
[49] Y. Han, L. Shen, C. Sun, Developing a parametric morphable annual daylight prediction model with improved generalization capability for the early stages of office building design, Build. Environ. 200 (2021) 107932, https://doi.org/10.1016/j.buildenv.2021.107932.
[50] W. Cai, X. Wen, C. Li, J. Shao, J. Xu, Predicting the energy consumption in buildings using the optimized support vector regression model, Energy 273 (2023) 127188, https://doi.org/10.1016/j.energy.2023.127188.
[51] Y. Zou, S. Lou, D. Xia, I.Y. Lun, J. Yin, Multi-objective building design optimization considering the effects of long-term climate change, J. Build. Eng. 44 (2021) 102904, https://doi.org/10.1016/j.jobe.2021.102904.
[52] Y. Shen, Y. Hu, K. Cheng, H. Yan, K. Cai, J. Hua, X. Fei, Q. Wang, Utilizing interpretable stacking ensemble learning and NSGA-III for the prediction and optimisation of building photo-thermal environment and energy consumption, Build. Simul. 17 (2024) 819–838, https://doi.org/10.1007/s12273-024-1108-7.
[53] R. Chen, Y.-S. Tsay, T. Zhang, A multi-objective optimization strategy for building carbon emission from the whole life cycle perspective, Energy 262 (2023) 125373, https://doi.org/10.1016/j.energy.2022.125373.
[54] W. Yu, B. Li, H. Jia, M. Zhang, D. Wang, Application of multi-objective genetic algorithm to optimize energy efficiency and thermal comfort in building design, Energy Build. 88 (2015) 135–143, https://doi.org/10.1016/j.enbuild.2014.11.063.
[55] B. Chegari, M. Tabaa, E. Simeu, F. Moutaouakkil, H. Medromi, Multi-objective optimization of building energy performance and indoor thermal comfort by combining artificial neural networks and metaheuristic algorithms, Energy Build. 239 (2021) 110839, https://doi.org/10.1016/j.enbuild.2021.110839.
[56] R. Chaganti, F. Rustam, T. Daghriri, I. de la Torre Díez, J.L.V. Mazón, C.L. Rodríguez, I. Ashraf, Building heating and cooling load prediction using ensemble machine learning model, Sensors 22 (2022) 7692, https://doi.org/10.3390/s22197692.
[57] Z. Zheng, J. Zhou, Y. Yang, F. Xu, H. Liu, Economic optimization of exterior wall insulation in Chinese office buildings by coupling artificial neural network and genetic algorithm, Therm. Sci. Eng. Prog. 50 (2024) 102582, https://doi.org/10.1016/j.tsep.2024.102582.
[58] B. Hocine, B. Abdelghani, B. Rafik, R. Zaid, Residential building design and optimization in arid climates using multiple objectives and ANN-GA, J. Adv. Res. Fluid Mech. Therm. Sci. 110 (2023) 14–31, https://doi.org/10.37934/arfmts.110.2.1431.
[59] R. Arjomandnia, M. Ilbeigi, M. Kazemidemneh, A.N. Hashemi, Renovating buildings by modelling energy–CO_2 Iran, Indoor Built Environ. 32 (2023) 1621–1637, https://doi.org/10.1177/1420326X231151244.
[60] J. Zhan, W. He, J. Huang, Comfort, carbon emissions, and cost of building envelope and photovoltaic arrangement optimization through a two-stage model, Appl. Energy 356 (2024) 122423, https://doi.org/10.1016/j.apenergy.2023.122423.
[61] R. Chen, Y.-S. Tsay, S. Ni, An integrated framework for multi-objective optimization of building performance: carbon emissions, thermal comfort, and global cost, J. Clean. Prod. 359 (2022) 131978, https://doi.org/10.1016/j.jclepro.2022.131978.
[62] B. Kubwimana, H. Najafi, A novel approach for optimizing building energy models using machine learning algorithms, Energies 16 (2023), https://doi.org/10.3390/en16031033.
[63] M. Płoszaj-Mazurek, E. Ryńska, M. Grochulska-Salak, Methods to optimize carbon footprint of buildings in regenerative architectural design with the use of machine learning, convolutional neural network, and parametric design, Energies 13 (2020) 5289, https://doi.org/10.3390/en13205289.
[64] S.K. Didwania, T.A. Reddy, M.S. Addison, Synergizing design of building energy performance using parametric analysis, dynamic visualization, and neural network modeling, J. Archit. Eng. 29 (2023), https://doi.org/10.1061/JAEIED.AEENG-1521.

[65] S. Seyedzadeh, F.P. Rahimian, P. Rastogi, I. Glesk, Tuning machine learning models for prediction of building energy loads, Sustain. Cities Soc. 47 (2019) 101484, https://doi.org/10.1016/j.scs.2019.101484.

[66] X. Li, R. Yao, A machine-learning-based approach to predict residential annual space heating and cooling loads considering occupant behaviour, Energy 212 (2020) 118676, https://doi.org/10.1016/j.energy.2020.118676.

[67] A. Barbaresi, M. Ceccarelli, G. Menichetti, D. Torreggiani, P. Tassinari, M. Bovo, Application of machine learning models for fast and accurate predictions of building energy need, Energies 15 (2022) 1266, https://doi.org/10.3390/en15041266.

[68] T. Storcz, G. Várady, I. Kistelegdi, Z. Ercsey, Regression models and shape descriptors for building energy demand and comfort estimation, Energies 16 (2023) 5896, https://doi.org/10.3390/en16165896.

[69] S. Alghamdi, W. Tang, S. Kanjanabootra, D. Alterman, Optimising building energy and comfort predictions with intelligent computational model, Sustainability 16 (2024) 3432, https://doi.org/10.3390/su16083432.

[70] S. Himmetoğlu, Y. Delice, E.K. Aydoğan, PSACONN mining algorithm for multi-factor thermal energy-efficient public building design, J. Build. Eng. 34 (2021) 102020, https://doi.org/10.1016/j.jobe.2020.102020.

[71] H. Yang, Z. Xu, Y. Shi, W. Tang, C. Liu, A. Yunusa-Kaltungo, H. Cui, Multi-objective optimization designs of phase change material-enhanced building using the integration of the stacking model and NSGA-III algorithm, J. Energy Storage 68 (2023) 107807, https://doi.org/10.1016/j.est.2023.107807.

[72] B. Chegari, M. Tabaa, E. Simeu, F. Moutaouakkil, H. Medromi, An optimal surrogate-model-based approach to support comfortable and nearly zero energy buildings design, Energy 248 (2022) 123584, https://doi.org/10.1016/j.energy.2022.123584.

[73] H. Li, S. Wang, R. Tang, Robust optimal design of zero/low energy buildings considering uncertainties and the impacts of objective functions, Appl. Energy 254 (2019) 113683, https://doi.org/10.1016/j.apenergy.2019.113683.

[74] A. Serbouti, M. Rattal, E.M. Oualim, A. Mouhsen, An innovative multicriteria decision-making tool for building performance optimization, Eng. Technol. Appl. Sci. Res. 11 (2021) 6603–6608, https://doi.org/10.48084/etasr.3953.

[75] S. Wang, N. Liu, J. Zhang, Thermal performance optimization for housing unit design in a cold region of China, J. Build. Perform. Simul. 14 (2021) 461–479, https://doi.org/10.1080/19401493.2021.1970811.

[76] A. Razmi, M. Rahbar, M. Bemanian, PCA-ANN integrated NSGA-III framework for dormitory building design optimization: energy efficiency, daylight, and thermal comfort, Appl. Energy 305 (2022) 117828, https://doi.org/10.1016/j.apenergy.2021.117828.

[77] Y. Lin, L. Zhao, X. Liu, W. Yang, X. Hao, L. Tian, Design optimization of a passive building with green roof through machine learning and group intelligent algorithm, Buildings 11 (2021) 192, https://doi.org/10.3390/buildings11050192.

[78] C. Li, Y. Chen, A multi-factor optimization method based on thermal comfort for building energy performance with natural ventilation, Energy Build. 285 (2023) 112893, https://doi.org/10.1016/j.enbuild.2023.112893.

[79] L. Magnier, F. Haghighat, Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and artificial neural network, Build. Environ. 45 (2010) 739–746, https://doi.org/10.1016/j.buildenv.2009.08.016.

[80] Y. Xu, G. Zhang, C. Yan, G. Wang, Y. Jiang, K. Zhao, A two-stage multi-objective optimization method for envelope and energy generation systems of primary and secondary school teaching buildings in China, Build. Environ. 204 (2021) 108142, https://doi.org/10.1016/j.buildenv.2021.108142.

[81] D. Gossard, B. Lartigue, F. Thellier, Multi-objective optimization of a building envelope for thermal performance using genetic algorithms and artificial neural network, Energy Build. 67 (2013) 253–260, https://doi.org/10.1016/j.enbuild.2013.08.026.

[82] K. Li, L. Pan, W. Xue, H. Jiang, H. Mao, Multi-objective optimization for energy performance improvement of residential buildings: a comparative study, Energies 10 (2017) 245, https://doi.org/10.3390/en10020245.

[83] A. Prada, A. Gasparella, P. Baggio, On the performance of meta-models in building design optimization, Appl. Energy 225 (2018) 814–826, https://doi.org/10.1016/j.apenergy.2018.04.129.

[84] X. Xu, G. Feng, D. Chi, M. Liu, B. Dou, Optimization of performance parameter design and energy use prediction for nearly zero energy buildings, Energies 11 (2018) 3252, https://doi.org/10.3390/en11123252.

[85] B. Si, J. Wang, X. Yao, X. Shi, X. Jin, X. Zhou, Multi-objective optimization design of a complex building based on an artificial neural network and performance evaluation of algorithms, Adv. Eng. Inform. 40 (2019) 93–109, https://doi.org/10.1016/j.aei.2019.03.006.

[86] S. Zhao, Using artificial neural network and WebGL to algorithmically optimize window wall ratios of high-rise office buildings, J. Comput. Des. Eng. 8 (2021) 638–653, https://doi.org/10.1093/jcde/qwab005.

[87] S. Wang, Y.K. Yi, N. Liu, Multi-Objective Optimization (MOO) for high-rise residential buildings' layout centered on daylight, visual, and outdoor thermal metrics in China, Build. Environ. 205 (2021) 108263, https://doi.org/10.1016/j.buildenv.2021.108263.

[88] Q. Xue, Z. Wang, Q. Chen, Multi-objective optimization of building design for life cycle cost and co2 emissions: a case study of a low-energy residential building in a severe cold climate, Build. Simul. 15 (2022) 83–98, https://doi.org/10.1007/s12273-021-0796-5.

[89] X. Wu, Z. Feng, H. Chen, Y. Qin, S. Zheng, L. Wang, Y. Liu, M.J. Skibniewski, Intelligent optimization framework of near zero energy consumption building performance based on a hybrid machine learning algorithm, Renew. Sustain. Energy Rev. 167 (2022) 112703, https://doi.org/10.1016/j.rser.2022.112703.

[90] S. Mohammad Ebrahimi Saryazdi, A. Etemad, A. Shafaat, A.M. Bahman, Data-driven performance analysis of a residential building applying Artificial Neural Network (ANN) and multi-objective Genetic Algorithm (GA), Build. Environ. 225 (2022) 109633, https://doi.org/10.1016/j.buildenv.2022.109633.

[91] E. Elbeltagi, H. Wefki, R. Khallaf, Sustainable building optimization model for early-stage design, Buildings 13 (2022) 74, https://doi.org/10.3390/buildings13010074.

[92] Y. Pan, Y. Shen, J. Qin, L. Zhang, Deep reinforcement learning for multi-objective optimization in BIM-based green building design, Autom. Constr. 166 (2024) 105598, https://doi.org/10.1016/j.autcon.2024.105598.

[93] C. Sun, Q. Liu, Y. Han, Many-objective optimization design of a public building for energy, daylighting and cost performance improvement, Appl. Sci. 10 (2020) 2435, https://doi.org/10.3390/app10072435.

[94] I.G. Kerdan, D.M. Gálvez, ANNEXE: an open-source building energy design optimisation framework using artificial neural networks and genetic algorithms, J. Clean. Prod. 371 (2022) 133500, https://doi.org/10.1016/j.jclepro.2022.133500.

[95] Y. Liu, T. Li, W. Xu, Q. Wang, H. Huang, B.-J. He, Building information modelling-enabled multi-objective optimization for energy consumption parametric analysis in green buildings design using hybrid machine learning algorithms, Energy Build. 300 (2023) 113665, https://doi.org/10.1016/j.enbuild.2023.113665.

[96] Y. Shen, Y. Pan, BIM-supported automatic energy performance analysis for green building design using explainable machine learning and multi-objective optimization, Appl. Energy 333 (2023) 120575, https://doi.org/10.1016/j.apenergy.2022.120575.

[97] X. Wu, X. Li, Y. Qin, W. Xu, Y. Liu, Intelligent multiobjective optimization design for NZEBs in China: four climatic regions, Appl. Energy 339 (2023) 120934, https://doi.org/10.1016/j.apenergy.2023.120934.

[98] A.M. Khan, M.A. Tariq, S.K.U. Rehman, T. Saeed, F.K. Alqahtani, M. Sherif, BIM integration with XAI using LIME and MOO for automated green building energy performance analysis, Energies 17 (2024) 3295, https://doi.org/10.3390/en17133295.

[99] B. Si, Z. Ni, J. Xu, Y. Li, F. Liu, Interactive effects of hyperparameter optimization techniques and data characteristics on the performance of machine learning algorithms for building energy metamodeling, Case Stud. Therm. Eng. 55 (2024) 104124, https://doi.org/10.1016/j.csite.2024.104124.

[100] F. Bre, N. Roman, V.D. Fachinotti, An efficient metamodel-based method to carry out multi-objective building performance optimizations, Energy Build. 206 (2020) 109576, https://doi.org/10.1016/j.enbuild.2019.109576.

[101] K. Bamdad, M.E. Cholette, J. Bell, Building energy optimization using surrogate model and active sampling, J. Build. Perform. Simul. 13 (2020) 760–776, https://doi.org/10.1080/19401493.2020.1821094.

[102] N. Yue, L. Li, A. Morandi, Y. Zhao, A metamodel-based multi-objective optimization method to balance thermal comfort and energy efficiency in a campus gymnasium, Energy Build. 253 (2021) 111513, https://doi.org/10.1016/j.enbuild.2021.111513.

[103] R. Batres, Y. Dadras, F. Mostafazadeh, M. Kavgic, MEVO: a metamodel-based evolutionary optimizer for building energy optimization, Energies 16 (2023) 7026, https://doi.org/10.3390/en16207026.

[104] X. You, G. Yan, M. Thwin, Applying modified coot optimization algorithm with artificial neural network meta-model for building energy performance optimization: a case study, Heliyon 9 (2023) e16593, https://doi.org/10.1016/j.heliyon.2023.e16593.

[105] S. Lahmar, M. Maalmi, R. Idchabani, Investigating adaptive sampling strategies for optimal building energy performance using artificial neural networks and kriging surrogate models, J. Build. Eng. 82 (2024) 108341, https://doi.org/10.1016/j.jobe.2023.108341.

[106] M.D. McKay, R.J. Beckman, W.J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics 21 (1979) 239–245, http://www.jstor.org/stable/1268522.

[107] F. Kharvari, An empirical validation of daylighting tools: assessing radiance parameters and simulation settings in ladybug and honeybee against field measurements, Sol. Energy 207 (2020) 1021–1036, https://doi.org/10.1016/j.solener.2020.07.054.

[108] C.F. Reinhart, O. Walkenhorst, Validation of dynamic RADIANCE-based daylight simulations for a test office with external blinds, Energy Build. 33 (2001) 683–697, https://doi.org/10.1016/S0378-7788(01)00058-5.

[109] M. Magni, F. Ochs, S. de Vries, A. Maccarini, F. Sigg, Detailed cross comparison of building energy simulation tools results using a reference office building as a case study, Energy Build. 250 (2021) 111260, https://doi.org/10.1016/j.enbuild.2021.111260.

[110] M. Sadeghipour Roudsari, M. Pak, Ladybug: a parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design, in: Proceedings of BS 2013: 13th Conference of the International Building Performance Simulation Association, 2013, pp. 3128–3135.

[111] KBOB, Domnées des écobilans dans la construction, IPB, KBOB, 2014.

[112] S. Lek, M. Delacoste, P. Baran, I. Dimopoulos, J. Lauga, S. Aulagnier, Application of neural networks to modelling nonlinear relationships in ecology, Ecol. Model. 90 (1996) 39–52, https://doi.org/10.1016/0304-3800(95)00142-5.

[113] J. Schmidhuber, Deep learning in neural networks: an overview, Neural Netw. 61 (2015) 85–117, https://doi.org/10.1016/j.neunet.2014.09.003.
[114] C. Gambella, B. Ghaddar, J. Naoum-Sawaya, Optimization problems for machine learning: a survey, Eur. J. Oper. Res. 290 (2021) 807–828, https://doi.org/10.1016/j.ejor.2020.08.045.
[115] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[116] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Stat. 29 (2001), https://doi.org/10.1214/aos/1013203451.
[117] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444, https://doi.org/10.1038/nature14539.
[118] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, D.J. Inman, 1D convolutional neural networks and applications: a survey, Mech. Syst. Signal Process. 151 (2021) 107398, https://doi.org/10.1016/j.ymssp.2020.107398.
[119] S. Katoch, S.S. Chauhan, V. Kumar, A review on genetic algorithm: past, present, and future, Multimed. Tools Appl. 80 (2021) 8091–8126, https://doi.org/10.1007/s11042-020-10139-6.
[120] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Comput. 6 (2002) 182–197, https://doi.org/10.1109/4235.996017.
[121] K. Deb, H. Jain, An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: Solving problems with box constraints, IEEE Trans. Evol. Comput. 18 (2014) 577–601, https://doi.org/10.1109/TEVC.2013.2281535.
[122] J. Kennedy, R. Eberhart, Particle swarm optimization, in: Proceedings of ICNN'95 - International Conference on Neural Networks, IEEE, 1995, pp. 1942–1948.
[123] M. Mersha, K. Lam, J. Wood, A.K. AlShami, J. Kalita, Explainable artificial intelligence: a survey of needs, techniques, applications, and future direction, Neurocomputing 599 (2024) 128111, https://doi.org/10.1016/j.neucom.2024.128111.
[124] A. Abusitta, M.Q. Li, B.C. Fung, Survey on explainable ai: techniques, challenges and open issues, Expert Syst. Appl. 255 (2024) 124710, https://doi.org/10.1016/j.eswa.2024.124710.
[125] F. Yu, J. Leng, Quantitative effects of glass roof system parameters on energy and daylighting performances: a bi-objective optimal design using response surface methodology, Indoor Built Environ. 30 (2021) 1268–1285, https://doi.org/10.1177/1420326X20941220.