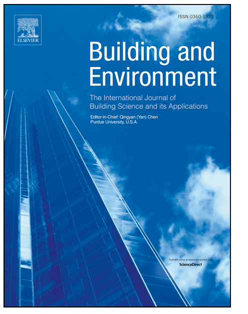

Case-based reasoning approach for decision-making in building retrofit: A review

Yangluxi Li, Hu Du, Satish Basavapatna Kumaraswamy

PII: S0360-1323(23)01057-0

DOI: https://doi.org/10.1016/j.buildenv.2023.111030

Reference: BAE 111030

To appear in: *Building and Environment*

Received Date: 11 July 2023

Revised Date: 20 October 2023

Accepted Date: 11 November 2023

Case-based Reasoning Approach for Decision-making in Building Retrofit: A Review

Yangluxi Li[a,*], Hu Du[b], Satish Basavapatna Kumaraswamy[c]

[a c.] Welsh School of Architecture, Cardiff University, Bute Building, King Edward VII Avenue, Cardiff CF10 3NB, Wales, United Kingdom

[b.] School of Civil Engineering and Built Environment, Liverpool John Moores University, Cherie Booth Building, Byrom St, Liverpool, L3 3AF, England, United Kingdom

Keywords:

Case-Based Reasoning; Decision-making; Building Retrofit; AI Models; Multi-criteria Attribute; Weight Factor

Abstract:

The rapid development of computer science has brought inspirations to building retrofit. Artificial intelligence (AI) provides more possibilities in decision-making for building retrofit, could be regarded as an alternative strategy compared to the abundant research time spent in the early decision-making stage of traditional retrofit approaches. This paper reviews the application of the statistic algorithm and AI approach, including CBR, in building retrofit decision-making, and the essential process of CBR, such as workflow, similarity degree calculation method, weight factors correction manner, and input or output content using building design to provide a synthetic overview of CBR utilisation in the building retrofit realm. Among those different models, Case-Based Reasoning (CBR) is valuable in providing references and avoiding possible failures, which is a promising approach for building retrofit. Yet, current research mainly focused on its utilisation to solve specific issues. There is still a lack of systematically summarized research on Case-Based Reasoning solution. Therefore, this study analyses the methods used for CBR approach in the field of building retrofit decision-making process, aiming to find the characteristics of internal commonness. It concludes that CBR has two significant impact factors: similarity attribute type and similarity calculation manner, which determines the judgement process. The results show that the CBR solution has great application potential in further building retrofit design.

Highlights:

- A review and comparison of AI models and algorithms used in multi-criteria decision-making for building retrofit
- A systematic review of Case-Based Reasoning approach in building retrofit
- A summary of the weight calculating during the Case-Based Reasoning process

1 Introduction

1.1 Background

With the acceleration of social development, about 40% of the world's annual CO2 emissions are generated by buildings (1). As the amount of building stocks tends to be saturated worldwide, building energy retrofit receive increasing attention, which is regarded as an efficient building energy efficiency method. The US government plans to invest a trillion dollars in energy-efficiency retrofitting of buildings (2). This action aids in diminishing about 616 million metric tons of CO2 emissions per year (3). In the construction sector, especially in Europe, a large number of investigations have been carried out on reducing energy use and carbon emissions. The Climate Change Act 2008 (4) set the 2050 Net-Zero target, requiring the UK government to reduce greenhouse emissions by 100% relative to 1990 levels by 2050. In order to further achieve this target, approximately 27 million (5) existing residential buildings in the UK will need to be retrofitted. The targets in retrofit are raised for at least a 32% share of renewable energy and at least a 32.5% improvement in energy efficiency (6).

Architects and building owners are often face challenges in selecting the appropriate retrofit approaches, especially when considering multiple objectives as many of them are complicated and conflicting (7), such as costs, construction time, energy collection or performance, etc. The decision-making process could broadly be classified into traditional design approaches and emerging design approaches. In Deb and Schlueter's research, they summarised these two ways as "Bottom-up approach" and "Top-down approach" (8).

The traditional design approach refers to the "Bottom-up approach" as it requires the measurement and analysis of fundamental details for individual target that lead into a specific retrofit strategy. It is a typical workflow that commonly used in building retrofit, which ensures the accuracy of the targeted case but requires sufficient work in the early design stage for not only survey and project setup but also energy auditing and performance assessment (9). On the other hand, the emerging design strategy, the "Top-down approach", benefits from the significant development from AI machine learning and data mining (8). It often employs algorithms to manipulate input parameters to achieve certain objectives. As the traditional Bottom-up approach is limited by experiences of experts who determine the trade-offs (7), so parameter design methods and decision-making tools, which can avoid this limitation, increasingly attract the attention of designers. However, some relative professionals criticise this kind of approach as it ignores the subjective feeling of the observer. Meanwhile, the traditional design method is also criticised as the reference case selection lacks scientific (10). Implementing the Net-zero energy goal by 2050 (4) is a global challenge, and building retrofit plays an essential role in it. Under the recent international affairs that happened in 2022, the escalation of energy consumptions, costs, and the scarcity of energy especially in European, urges the development of new approaches or tools to accelerate building retrofit and energy reduction. In this case, some solutions related to AI should be proposed to fill the gap to help others, including unprofessional and untrained people, to rapidly understand the potential retrofit solutions close to their demands. This paper analyses one of the AI solution, Case-Based Reasoning (CBR), utilised during building retrofitting, to coordinate with the traditional design scheme.

1.2 CBR as a Proposed Methodology for Early Stage Building Retrofit Strategy

It is generally accepted that the strategy adopted at the beginning of a building retrofit plays a decisive role in the entire process (10) (11). With the emphasis on energy efficiency retrofit, the cases of retrofit projects are also increasing. The finished projects can provide valuable experiences for supporting further building retrofitting decisions (12) (13). As the decision-making in building energy efficiency retrofit is a complex process, researchers believe that CBR is suitable for unstructured and complex problems (10) (14) (15).

Case-Based Reasoning (CBR) is an experience-based approach based on artificial intelligence (AI) and machine learning, firstly proposed in 1971 by Kling (16). CBR means using previous experiences or existing cases to solve new similar problems (17). Currently, it has been widely implemented in many fields to support decision-making, such as the graph recognition (18) (19) (20) (21), medical science (22) (23) (24) (25) (26) ,etc. But in terms of its application to buildings, especially in retrofit, not enough attention has been paid to it. Relative research has been done so far mainly focused on specific building issues such as construction cost, case search, etc. (11) (27). Nevertheless, CBR contains many details in the calculation section that directly influences the final output precision. Existing investigations adopt various approaches to correct the CBR process to improve accuracy (28) (29) (30) (31). In this case, there is a lack of a summary for the different solutions used during the CBR process that illustrates the work principle and workflow.

Most CBR models are mining the similar cases, through the widely recognised "4R" principle (17) of "Retrieve, Reuse, Revise and Retain", or the amended "R5" theory (32) (33) of identifying "Represent" at the beginning, to provide references for decision-making. This type of workflow is considered as the basic CBR model.

Based on the Plan of Work from RIBA, the most suitable stage to use this CBR model is stage 2, Concept Design. Shown in Fig. 1. The goal of this stage is to determine an architectural concept that could be admitted by the clients (34).

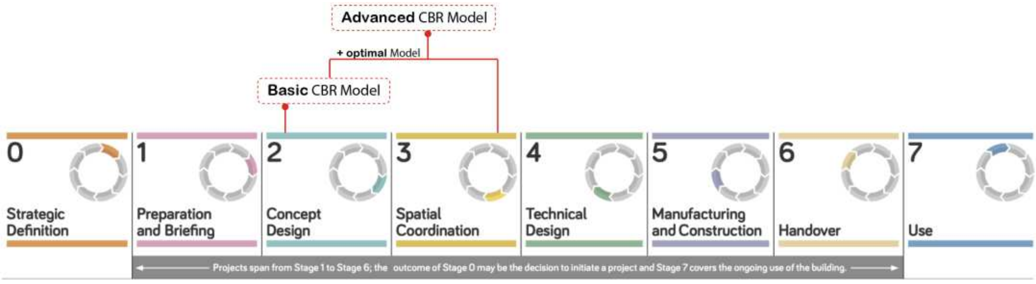

Fig. 1 CBR model applications during RIBA stage (image modified from RIBA Plan of Work)

Clients and designers are the main participants during this phase, who would need to review the concept design and consent to the design that is consistent with the budget, strategies, etc. for formulating the further detailed design programme (34). There is a lot of uncertainty at this stage, as amendments would be made to align with the feedback from the participants. In addition, RIBA also suggests that a "pragmatic review" (34) is essential to support determining the outline specification. Thus, the basic CBR models could fulfill the goals and provide a solution for these tasks.

For the basic CBR models, the whole process blelongs to the concept design stage. As the outcomes are sorted based on the user's input weight demands, which result in the combination of possible solutions that prioritise users' needs for building retrofit. This decision-making process involves both professionals and non-professionals, making the basic CBR a convenient decision-making support tool.

Yet for a consensus to be reached for leading the detailed design in stage 3, a further calculation of the optimal solution is mandatory. Stage 3 is about "testing and validating" (34) the outcome from stage 2. Professional design teams play a key role in this stage, clients are involved here for coordination. Hence, there were also 2 research tried to combine optimasition into the CBR cycle, Koo et al. (35) and Hong et al. (36) developed the "Advanced CBR(A-CBR) model", which was based on the 4R theory of basic CBR model and integrate with another optimisation model together for the extra evaluation process. Such proposed A-CBR model is considered to run through stages 2 and 3, as shown in Fig. 1. Not only indicating the possible solutions in the early concept design stage, but also undertaking the detailed analysis and test of the potential schemes. To make sure the outcome from stage 2 could be translated into stage 4 for manufacture details. This is a different trial, yet, the optimisation section is another important subject that may have better alternatives to be studied. At present, the basic CBR models would be more consistent with the common understanding of the CBR principle, which is the research target for this study as well.

2 Methodology

2.1 Research Strategy

Regarding the investigation purposes of reviewing the CBR method in building energy renewable retrofit, how to find the most match case is the core problem of review based on the decision makers' demands. The keywords of literature research are divided into three categories: "Building Retrofit", "CBR" and "Decision-making Model". The words and phrases related to these 3 categories are selected as search clues. The most ideal literature should contain all three parts, but individual studies can also be viewed. Besides the main goal of reviewing "Case-based Reasoning", other well-known machine learning algorithms used for decision making, for instance, K-nearest neighbors(KNN), can also be used as keywords to retrieve other research results that may relate to building retrofit for comparison.

To ensure the timeliness of the paper, the period after 2000 limits the time range of the literature. The reason for setting this time limit is due to the rapid renewal of computational applications and the limitation of mature research of building retrofit before 2000. As a result, most articles accord with the concept of this research present the latest findings in the range from 2010 to 2022.

It should be noted that all the above machine learning and decision-making methods are not always in the domain of architecture or building retrofit. But this type of solution can be used to analyse some architecture-related problems. Therefore, it is necessary to review these studies, which can also provide us with effective reference solutions and ideas. Although the literature covers a variety of methods in different fields of investigation, it is expected to select the most appropriate research in the field of building retrofit. The purpose of this study is to review relevant scholarly articles. By summarising the main reasons and specific solutions for each case study, it helps to find the most effective judgment method, study the significant gaps, and establish new contemporary methods with a systematic approach.

Fig. 2 presents the whole workflow for this investigation.

Fig. 2 Investigation Workflow for Literature Review

2.2 Method of Selecting Research Work

There were around 566 studies related to the topic gathered at the first stage. After quickly browsing the abstracts and reviewing the methods, the amount was narrowed down to 429 articles that related to building retrofit with a multi-criteria decision-making model. In this stage, some valuable in terms of investigated method and highly relevant research were filtered to review furtherly instead of all papers. To further analyse for the decision-making model, the methods commonly used were summaries into 4 categories, 237 records have remained to review for detailed information at this stage.

The statistic hybrid algorithm is a research hotspot every year. Shown in Fig. 3**.** Since the statistical approach is a mature and applicable technology, it could be reformed easily forming new computational methods based on traditional statistical solutions. While questionnaire method indicates the smallest research as it is difficult to investigate the objective level and convenience.

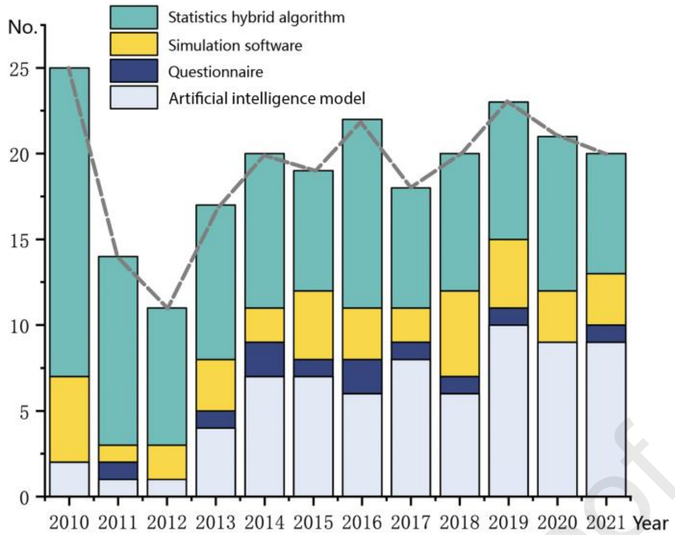

Fig. 3. Research relevant to 4 different common ways used in decision-making

In the aspect of artificial intelligence algorithms, especially in recent years, there is an obvious growth trend. This phenomenon shows that artificial intelligence algorithm is gradually applied to solve multi-criteria decision-making problems. This is due to significant developments in the field of artificial intelligence research, providing innovative solutions for machine learning. Therefore, according to the current research status, AI technology will be more and more applied in the field of decision research. It is necessary to review the research of artificial intelligence algorithms. Among the artificial intelligence algorithms category, the proportion of research combined with CBR has gradually increased over the past decade. Thus, based on the filtered literature review, around 30 relevant articles about CBR method implementation specifically in the architectural field are selected for intensive reading and analysis. Shown in Fig. 4**.**

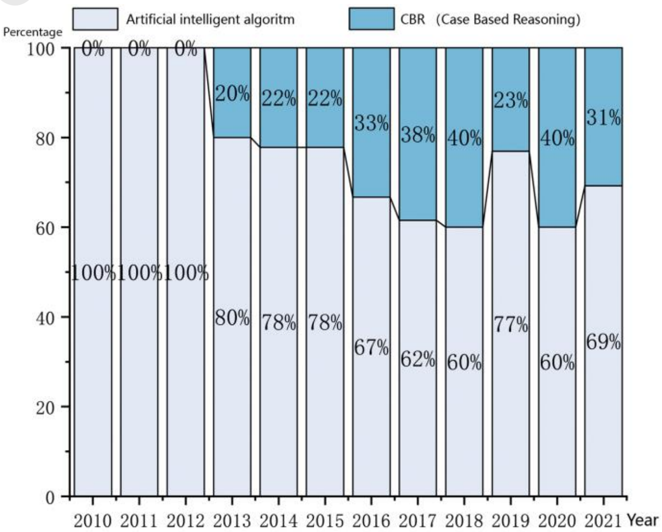

Fig. 4 CBR investigations among AI algorithms for Building Retrofit

The increasing utilisation of CBR in recent years is because the method has simple computational principles to manipulate the entire model structure. On this basis, the internal structure of the model is simplified to facilitate the integration with other weight determination methods and further improve the accuracy. As an effective solution for case investigation, this method has been widely used in other fields. (10) Yet, the CBR decision system has not been widely established in the architectural realm, especially in building retrofit.

3 Multi-criteria decision-making approaches for building retrofit

3.1 State-of-art

According to the reviewed literature, the commonly adopted methods of multi-criteria decision support for building retrofit are summarised into 4 categories: artificial intelligence (AI) models, questionnaires, simulation software and statistics hybrid algorithms. Therefore, with the popularity and development of AI in recent years, there is a new trend of combining artificial intelligence algorithms for the decision-making of building retrofit. AI models could be considered a more holistic approach. The utilisation of statistical algorithm and simulation software could be only a part of the AI modelling process. The questionnaire method has been sifted out from the scope of this article as its insufficient feature of convenience and precision.

There is a challenge to develop the methods that can not only speed up the retrofit procedure, but also assist the decision-makers who are either professionals or non-professionals to understand the potential solutions rapidly at the early design stage (7). Although simulation software and statistical hybrid algorithm have been developed and widely applied for a long time, they tend to be used for independent projects and requires certain professional skills (37) (38) (39) (40) (41). AI models, in comparison, have the potential to provide the straight-forward and comprehensive schemes to whom does not have sufficient knowledge of building retrofit.

On the other hand, different approaches are mainly targeted at different stages. For example, statistical algorithms are generally used at the early design stage, which can be used independently to generate the research data and the work for briefing. The application of simulation software is mainly used in the detailed design stage, such as the technical design. The simulation could test the feasibility of the proposal and predict the actual effect. AI models tend to cover a wider range of stages because they often include either statistical algorithms or tools during its process.

Differing from the linear processing of most statistical algorithms, AI models are considered as the comprehensive methods to comprise its own database. In recent years, there are few research projects have attempted to establish the databases of building retrofit approaches that can be further applied to data clustering and regression (7) (8). As this is an innovative direction, there are different attempts at AI models used for retrofit or building methods. For instance, Cecconi et al. (42) propose an AI model with ANN and GIS to only simulate the potential in energy efficiency retrofit but not consider other multi-objectives.

Thus, it would be tedious to distinguish or analyse the construction approaches according to various specific detail attributes among those cases. Amer et al. (43) propose a computer-aided decision-making solution with the Non-dominate Sorting Differential Evolution (NSDE) and Adaptive Sparrow Search Optimization Algorithm (ASSOA), which are both integrated with the Genetic Algorithm (GA) to determine the retrofit solution in specific objective. While Khansari and Hewitt (44) utilise the concept of an Agent-Based Model (ABM) to build a mathematical model in a traditional way to assist decision-making.

Indeed, those AI models or integrated methods can be used to analyse building reconstruction cases and datasets with multiple indexes in a quantitative path. However, those attempts were considering objective problems to find the optimal solution, the process of reanalysing cases and datasets is necessary if encountering different demands. Furthermore, even though those different studies of AI models are designed for decision-making, some of them work for the detailed design stage and professional involvement is required.

Selecting the right renovation strategy is crucial for the success of renovation projects. As a result, researchers have developed various decision tools to assist decision-makers in making informed choices. For example, Jafari and Valentin introduced a decision matrix that considers investor types and potential returns to guide the selection of renovation strategies (45). Similar research includes Mejjaouli and Alzahrani, who developed a decision support model that considers factors, for instance, lifecycle costs, budgets, thermal comfort, and lighting levels to help residential building owners choose the best energy-efficient renovation strategy (46). Juan, Gao, and their team focused on renovating office buildings and created a comprehensive decision support system that balances renovation costs, building quality, and environmental impact (47).

However, real retrofit projects are often complex and unique. Traditional mathematical models may not provide efficient solutions when the specific conditions are not the same. Therefore, for certain energy efficiency retrofit issues, sometimes it is more effective to draw on previous experiential cases, especially those similar to successful cases, rather than relying solely on decision-making models.

To facilitate this, establishing quick and accurate matching relationships with past renovation cases becomes crucial. In this context, Case-Based Reasoning (CBR) is considered a valuable tool for improving decision-making efficiency and drawing insights from past experiences (48).

Given this problem, the CBR enables decision-making fully to refer to other reference cases (49), and provides suggestions or guidance for a broader range of users. In the past, due to the lack of similar reference cases for research projects, this approach has not received sufficient attention. As there are many records of building retrofit cases that have been done in the past two decades, especially for problems with many referenced cases, the CBR method has a broader application prospect (27) (32). The CBR approach can be an alternative method to reduce the duration of the research process in the early design stage, which is a promising solution for decision-making support in building retrofitting.

Due to this solution has not attracted enough attention from designers, there is not as much literature reviewed relevant building research on CBR currently. Some review descriptions can only be found in a few research papers (27). Ahn et al. (27) summarized 10 relevant investigations and information on various steps such as distance calculation and weight determination of the CBR system. Chen et al. (50) reviewed the application of some case-based studies in the field of building construction safety. Cheng and Ma (49) concentrated on the specific "4R" steps of the theory and workflow for the CBR concept. Those research studies mainly focus on the general working steps or some specific principles of CBR.

Currently, the CBR research in the architectural realm are more inclined to the use of multi-criteria decision tools to support the selection of optimal building strategies through mathematical models (11) (51). The focus on retrofit construction is insufficient. An et al. (52) pointed out the current application fields of CBR, mainly focusing on the construction period and/or cost estimation system, bidding decision system, method selection system and management system. For instance, Gero et al. (53) developed a multi-criteria model to seek the balance between building thermal performance and other criteria. Carol Menassa (54) used economic analysis tools and other risk assessment tools to find the optimal retrofitted alternatives. Goodacre et al. (55) analysed the heating and hot water energy renewal efficiency of English building stock through a cost-benefit analysis system. Blondeau et al. (56) used a multi-criteria solution to judge the optimal ventilation strategy for university buildings from the perspective of the human behaviour.

Although these studies have analysed CBR from multiple perspectives, the internal indicators and comparison to other decision-making support approaches have not been fully studied for building retrofit (7). There is still a lack of systematic summaries of the internal details between the different methods used for decision-making, and the reason that CBR is more advantageous in early decision-making support for building retrofit compared to other approaches.

3.2 The Common Methods Used for Decision-making Support

In the field of artificial intelligence area, various algorithms and software are proposed to deal with the optimization of energy efficiency in buildings. It is worth emphasizing that the AI models, including CBR, are comprehensive decision-making models that normally contain the statistical algorithms and simulation software during the simulation or calculation process. According to the different development goals it can be composed of more than one algorithm or software during the modelling process. Statistical algorithms can be stand-alone, but AI models are hybrid.

In other words, there might not be a clear demarcation line between the AI models and the statistics hybrid algorithms in most cases. For instance, Delgarm et al. (57) proposed a mono-objective and Multi-Objective Particle Swarm Optimization (MOPSO) algorithm coupled with Energy Plus to assess the energy consumption performance. The results show that the proposed optimization method can find the optimal solution in the form of an objective function in a short time. Figueiredo et al. (58) employed AHP to achieve the sustainable material choice by integrating the BIM system. To extend the range of AHP algorithm employment, Haruna et al. (37) built a BIM model for developing sustainable building utilizing the enhanced AHP algorithm named ANP. Akaa et al. (59) developed a hybrid multi-criteria decision analysis tool based on the combination of Geometric Mean Method (GMM), AHP and TOPSIS to solve the optimisation between stakeholder's opinion and the design for fire-prove steel-frame building. To achieve different goals, AI models could adapt different algorithms in line with the specialises.

Similarly, combining with other algorithms is an essential procedure for CBR to implement the entire process. There are a variety of different methods that can be used for decision-making support, but the characteristics they excel at are different.

From the reviewed research, some common methods are generated as follows:

Statistics hybrid algorithm/AI model:

- Case-Based Reasoning (CBR) is a "paradigm in artificial intelligence and cognitive science" (15). In areas where traditional rule-based or knowledge-based reasoning is relatively weak(68), CBR can provide solutions by analogy or referring to previous similar cases. (10) (18) (19) (20) (22) (30) (31) (49) (50)
- The original Mixed-Integer Linear Programming (MILP) is an improvement of a row relaxation problem, and the simplex method is continuously used to solve it. Branch solving by adding constraints until the integer optimal solution appears at a vertex of the new improved relaxation problem. (46)
- Agent-Based Model (ABM) simulates the action and interaction calculation model of autonomous agents, such as organizations/teams/etc. (44) The MILP model and the ABM are two pure mathematical models with high precision and complexity.
- Sensitivity Analysis, which finds out sensitive factors that have a vital impact on the economic benefit indicators of the investment project from multiple

uncertain factors and analyse and calculates the degree of influence and sensitivity on the economic benefit. (60) (61)

- Multiple Attribute Utility Technique (MAUT) and Sensitivity Analysis are theories in economics. Although the theory has a wide range of applications, its operation is complex with difficult that requires training in multi-attribute utility functions. (62) (63)
- Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS) is an objective evaluation method by detecting the distance between the evaluation object and the optimal or the worst solution carries out the ranking. If the evaluation object is the closest to the optimal solution and the furthest away from the worst solution, the object can be determined as the optimal one. It can be used widely in general, but not in some special cases. (64) (65) (66) (63) (59)
- Analytic Hierarchy Process (AHP) divides the various factors in complex issues into interconnected and orderly levels to make them organized. According to the subjective judgment structure of a certain objective reality (mainly a pairwise comparison), the expert opinions and the analyst's objective judgment results are directly combined to quantitatively describe the importance of elements at a level. (10) (64) (67) (58) (68) (59) (69) (70) (71) (72)
- ANP is a development method of AHP. To overcome the disadvantage of AHP, ANP can dispose of the relationships among criteria and sub-criteria. It has a great performance in decision-making when an extensive number of elements are involved. (37)
- Genetic Algorithm (GA) is an evolutionary algorithm that solves a population of individual solutions based on natural selection. (73)
- Enhanced Archimedes Optimization Algorithm (EAOA) is an enhanced algorithm for Archimedes' optimization algorithm. It overcomes traditional shortcomings like local optimization and premature convergence. EAOA outputs the optimum values of minimum, mean value and maximum value. In addition, it also has the minimum value of the standard deviation compared with other algorithms. (74)
- Decision-making Trial and Evaluation Laboratory (DEMATEL) and PROMETHEE II are variants of the AHP. But they significantly increase the difficulty and complexity. DEMATEL can calculate the degree of influence on other elements through the logical relationship between the elements in the system and the direct influence matrix. (75) The basic principle of PROMETHEE II is based on the pair-wise comparison of alternatives along each selected criterion. (68) (76)
- A Neural Network (ANN) is a new solution which can achieve many purposes. A neural network can be considered as either an AI model or an algorithm by itself, that can solve a series of problems by imitating the biological nervous system. However, as a complex internal structure, it was difficult for the architects to realize it. (77)

- Adaptive Sparrow Search Optimization Algorithm (ASSOA) is a new simulation-based optimization technique. It is a swarm intelligence optimization algorithm for sparrow foraging and evading predator behaviour proposed in 2020. Compared with the other optimization algorithms, ASSOA achieves the lowest amount of the functions that have the most certainty. (78)
- Non-dominated Sorting Genetic Algorithm II (NSGA-II) is a solid multi-objective algorithm by generates offspring using a specific type of crossover and mutation. Today it can be considered as an outdated approach. (11) (79) (80) (81) (82)
- K-Nearest Neighbour (KNN) is a non-parametric classifier. It is one of the first algorithms for data mining (83). It is commonly used for simple classification or regression problems as a "lazy learning approach". Yet, it also easily falls into the curse of dimensionality with the high-dimensional input of data. (84)

In terms of those analysed calculation approaches, KNN is rarely used recently as it has become increasingly inefficient due to its shortcomings in weight value. Besides of KNN, in fact, other solutions are all involve the weight calculation.

Simulation software:

- BECEREN is a tool developed by several companies focused on specialized issues rather than being widely applicable. (38)
- BIM-based Design Iteration Tool (BIM-DIT) can support the decision-making process by assisting the design team in the generation of design alternatives. (85) It helps decision-makers with precise knowledge of available options for achieving truly sustainable building projects. Yet, it is not suitable for non-professionals. (37) (41)
- Community VIZ GIS is a software focused on building intelligence, enabling a variety of functions. (86) The Construction Emission Evaluation tool is a tool used to evaluate the emissions level and impacts at different construction techniques and construction stages. (87) Both methods require experts to operate the software.

Besides those 3 simulation software, Open Studio, EnergyPlus, TRNSYS, DOE-2, ESP-R, eQuest, etc. are popular simulation packages that can be easily attached as well. These tools contain many features such as modelling and calculating energy consumption. However, the use of these tools requires professionals to limit their popularity. (39) (40)

All these methods can be used to support the decision-making. However, the operational difficulties vary. In addition, while a multi-criteria decision approach can be used to judge the performance of a retrofit strategy, users cannot maximize their selection of optimal cases that meet their specific needs. To this end, CBR mimics human reasoning that learns from the past experiences and adapts it to solve new problems (49), which could provide decision makers with an intuitive solution. Thus, compared with the advantages and disadvantages of other AI models and algorithms, the characteristics of CBR are more suitable in the early design stage.

Technically speaking, CBR can combine with most algorithms to fulfil the calculation and selection process, which completely depends on the purpose and ability of the designer. But in retrospect, one of the advantages of CBR is that it can provide an intuitive solution to people from different backgrounds, including non-professionals (10). Therefore, the concise algorithms or other simple data-processing methods would be definitely much more preferred. The advantages and disadvantages of those reviewed decision-making approaches are listed in Table 1.

Table 1 Pros and Cons of various decision-making approaches

Function	Name	Pros	Cons
Statistics Hybrid Algorithm/ Artificial Intelligent Model	CBR	Provide similar solutions referring to previous cases even if in areas of weak knowledge	Easily affected by the quality of the database of cases
	MILP	Able to pick up the limitation of the boundary for solutions	Only work for linear problems
	ABM	Suitable for complex systems and targets	Many parameters need to initialisa operations
	Sensitivity Analysis	Able to assess variables in precision	Require professional specialists to participate
	MAUT	Integrating multiple alternatives into a formula	Complex, too many calculation steps
	TOPSIS	Coupled objective factors into the decision process	Can't decide the optimal number of attributes
	AHP	Widely used, attributes defined by requirement	Subjective, can't generate s new case
	ANP	Great performance when an extensive number of elements are involved	Must be technically considered from the decider's perspective
	GA	Obtaining/guiding the optimal search without explicit rules, reduces the difficulty of code implementation	Involves optimization, relatively complicated
	EAOA	Avoid the local optimization and premature convergence issue	Require operation in many times improving precision level
	DEMATEL	Fuzzy evaluation model	Evaluation can't be made in quantitative
	PROMETHEE II	Less steps to calculate	Requires additional information provided by deciders
	ANN	Eliminating the noise disturb	Requires abundant training time and a large amount of basic data
	ASSOA	Achieves the lowest amount of the functions that have the most certainty	Limitations on data collection

	NSGA-II	Widely used in real-world applications	Need a solid benchmark to test against, considered out-of-date
	KNN	Simple and intuitive, easy to apply in data regression	No weight determination, crashes at high dimensions
Simulation software	BECEREN	The tool to calculate carbon emissions for varies steps	Only focuses on this specific environmental impact
	BIM-DIT	Provide knowledge of available options for achieving truly sustainable	Not suitable for non-professionals
	Communtiy VIZ GIS	Realize multiple functions based on requirements	Requires integrating into the software of GIS

Depending on the different building reference case datasets, some information hidden under statistics can be found. How to help customers quickly select the most suitable case for their needs as a reference case is very worthy of attention. This goal requires the customer to input corresponding demands, such as construction requirements, building information, etc.

Therefore, it is a necessary to develop a way to measure how similar a case is to the decision maker's demands. The best cases for the customer can then be identified and matched. To this end, Case-based reasoning (CBR) could attain this goal (88). In this method, similar cases are searched from the corresponding database to match potential project solutions. There were a few research fully applied the principle of the CBR approach to deal with the retrofit decision-making. For instance, in an Italian project "POI 2007-13" (77), the researchers built a database with 151 existing cases and used 2 ANN models to train the biological nervous system and compute the decision-making result. Zhao et al. (10) built a database of 71 retrofit cases in China to identify the attributes of the retrofitting buildings and implement the AHP algorithm for the CBR approach in a real case in Shanghai to realize the retrofit procedure. The results show that CBR helps identify valuable information and extract potential solutions from similar previous solutions, which not only simplifies the preliminary research process to a large extent, but also guide the decision makers to make decisions more easily. The whole principle and workflow are worthy to be promoted and referred for retrofit in the early stage.

4 Review of CBR approach in building retrofit

4.1 CBR Workflow

Case-Based Reasoning (CBR) differs from other AI approaches such as Knowledge-Based Systems (KBS) (89) in several ways. Rather than relying solely on general knowledge of the problem domain or correlating along general relationships between problem descriptors and conclusions, CBR uses specific knowledge of prior experience and specific problem situations. CBR also provides incremental, continuous learning, because each time a problem is solved, a new experience is retained and can be applied to future problems. The common understanding of the CBR concept is shown in Fig. 5.

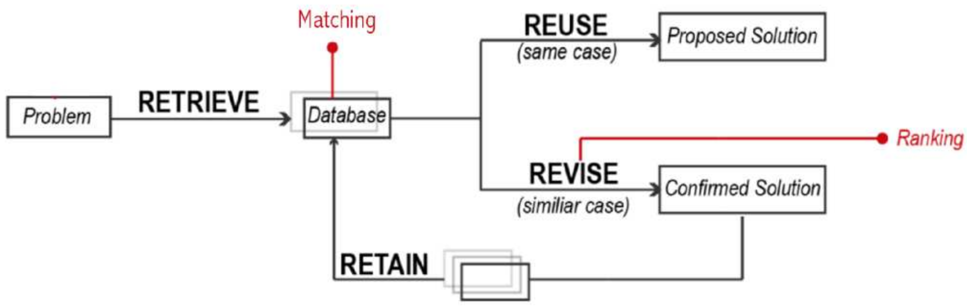

Fig. 5. Concept of CBR

For the benefit of architects, after comprehensively evaluating the performance of various cases, it is crucial to help decision makers select the most suitable case for their needs in terms of candidate building information. The core of the CBR method is to extract successful previous cases or solutions from the datasets by measuring the similarity level. Wang et al. (32) used CBR theory to create a Lesson Mining System (LSM) to avoid the possible recurrence similar problems caused by people during the process of urbanization. This LMS is based on their own developed curriculum database, allows policy makers who may not be fully trained in architecture to learn from existing experience effectively. Therefore, to provide an adequate reference scheme, a summary database must be established. Valuable cases from the past are placed in this dataset, waiting to be selected for matching the target cases. Four sections constitute the entire CBR system, as shown in Table 2.

Name	Purpose
Core Database	Store previous cases and solutions
Attributes database	Store case attributes
Measure method	Calculate similarity level
Modification method	Adjust the similarity computation method

Table 2 Four sub-sections of CBR system

The concept of CBR was first developed by an American cognitive and learning scientist Janet Kolodner in 1992 (17). Leake (90) first successfully applied the Case-Based Reasoning solution to coding a couple of years after. In Kolodner and Leake's point of view, CBR is considered as a learning loop of "remember, adapt and compare" (33). The common perception of CBR is origin from Kolodner and Leake's principle of "4R"— "Retrieve", "Reuse", "Revise" and "Retain" (17). This 4R theory is widely accepted and applied to decision-making support.

However, from the practical perspective, how to determine the problem and input the demands into the CBR system might also be ignorant. According to this problem, Finnie and Sun (33) raised an improved "R5" CBR model based on the original "4R", consisting of five steps: represent, retrieve, reuse, revise and retain. This redeveloped theory is also gaining acceptance from many researchers, since "Represent" is also a crucial part of this learning cycle that determine the problems and structure the information at the first stage (32).

Table 3 gives the names of individual steps and their corresponding effects. The most important stage among is the "Retrieve" stage, which is to match the case by evaluating similarity. The core is the attribute database that stores previous case information and the information for related retrofit buildings. In addition, the database retains case property information that is used to calculate similarity.

Therefore, considering that each attribute has different important characteristics, it is necessary to introduce a weight coefficient to improve the accuracy of similarity measurement. The weight value is combined with the similarity calculation to generate the final project that best meets the decision maker's needs.

Step	Function
Represent	Identify the problems and the demands for outputs
Retrieve	Pick out a similar case from the database
Reuse	Use the chosen case as a target reference
Revise	Adjust solution to adapt to new condition
Retain	Store new solution and corresponding cases in the database

Table 3 Five significant steps constituting CBR system

4.2 Characteristics of Each Step in CBR Cycle

The database in the CBR cycle contains attributes and related information for the projects that are worth learning from. In the following part of the weight grading scheme, according to the retrofit goals and demands, appropriate statistical methods are used to sort various situations. Therefore, to compensate for the shortcomings of the ranking method, the CBR system focuses on searching for suitable cases based on the general information of the target building, such as year/type/size/climate/cost, etc (10).

These attributes determine the result of similarity calculation. The characteristics of each step are summarized below:

1. Represents:

The goal of CBR is to find cases matching the target cases at a high level. So, the first step is to set a clear goal. It's entirely up to the decision maker. It is important to note that the various attributes of the target must be the same as the case in the database, otherwise the attributes matching the target cannot be calculated.

This step is considered as the structure of the database. The structure of the database is very relevant and very specific to the needs of the user. In fact, the first step of the CBR is to determine how the cases are organized in the database. Generally speaking, the main content of the database is a series of events, events should contain a description of their results, and at the same time, events need to be indexed to ensure that people can find the corresponding events (17). To build a database is to organize the past cases in a structured way. Past situations can be reused in the future, and accordingly, a new case is a description of a new problem to be solved. This database roughly covers a range of problems that arise in one domain. Both of success and failure cases should all be included.

2. Retrieval:

Attributes are used to represent cases in the database. They need to be defined to summarize the case. On the other hand, the indexes in the database are attributes, and the differences in attributes represent the differences in the case. Different researchers will set attributes based on their own understanding of the problem. For example, in the issue of green promotion, six attributes including green grade, project type, owner type, total area, total property area and location can be used (14), or more attributes can be used to represent a case.

Attributes are the source of input, and when looking for a particular case, it is not necessary to use all attributes, but to input some more specific attributes. Thus, we need to use the precise vocabulary to select the appropriate index for the new case. The accessibility of all indexes is essential when we add to the database.

The retrieval phase is the most important part of a CBR solution. Similarity measurements are needed to assess closeness. The concept of similarity includes three types: surface similarity, derivative similarity, and structural similarity (49) (91). Those three types are all proposed from the perspective of attribute form, without considering measurement methods. Surface similarity refers to the basic information of the targets. For example, the features of cases such as size, application, location, etc., are the basic data for calculating surface similarity. The derivative similarity is calculated between the deductive attribute value and the target. Deductive statistics are generated from basic information such as the area obtained by the product of side lengths. However, this kind of data is usually produced by simple manipulation of surface data and only changes in surface information. Conversely, another analogous concept called structural similarity derives from complex calculations, such as graph measures and first-order terms (91). In this case, the structural properties of the case need to be determined first, and then the corresponding similarity level calculated. Other functions and algorithms such as neural networks are usually integrated into the process. Table 4 shows the comparison of the above three similarity qualities.

Name	Concept	Relative Parameters	Difficulty
Surface similarity	Surface information similarity	Case basic attribute	Low
Derivative similarity	Derived info generated from surface information similarity	Simple operation of case basic attributes	Low
Structural similarity	Internal case structural similarity	High order operation of case internal data	High

Table 4 Comparison of surface, derived and structural attributes

During this phase of the CBR model, a corresponding database should be first established to support the similarity measurement. Then, depending on the implementing demands, the appropriate algorithm will be combined to determine the weight precision for realizing the functionality needs. For instance, according to the aforementioned algorithms in Section 3, Kim et al. (92) utilized a CBR structure with weight decision method of genetic algorithm (GA) to predict budget level under inputting some basic attributes of bridge such as width, location etc. It achieved the cost estimation of bridge construction based on previous data collection. Another example is a CBR solution proposed by Zhao et al. (10) in 2019 was regarded as the specific method used in future research. In this article, the authors adopted the CBR method to extract the best matched building retrofit case from the collection database including previous sustainable building retrofit plans. In addition, the weight value was determined by an AHP solution which could be validated via a consistency checking process, in which the precision of weight calculation was guaranteed.

3. Reuse, Revise and Retain:

The final part of the CBR process can be understood as a combination of those three steps. Application of computed result by pre-similarity calculation is realized in this part. In the reuses section, the selection case is chosen to solve the issue, but in some cases, this stage could also go back to aid in enhancing model performance (91). Revise section adapts the issue proposed process situation after reusing process which is commonly integrate into the reuse step. The last section of retaining is to store the research outcome to the database under special format. However, database establishment should consider its simplicity and efficiency features ensuring the value of this dataset serving for decision makers. The space for storage also limits the dataset to some extent, simultaneously. Consequently, some solutions have been proposed to filter and remove useless cases from the dataset (93).

Following Table 5 presents relatively major information on weight determination solutions used in CBR research related to building design in recent years.

4.3 Weight Determination Solutions in CBR Model

CBR cycle essentially is similarity calculation, which computes the weight coefficients for diverse cases to find the most similar case. Consequently, how to calculate this indispensable value of weight is the core of the CBR studied solution.

Similarity calculation of CBR is generally classified into two types of weight factor and non-weight factor computation. In terms of the non-weight factor computational approach, it is an originally investigated manner that simply measures the mathematic distance number without any corrections, such as KNN (83) (84). Although this is a simple solution to manipulate, the diverse features of the input attributes are neglected. Therefore, final precision would be impacted significantly (94).

Due to the characteristic of KNN is non-weight calculation that normally cannot be used independently in the cycle of CBR if the datasets are complex in dimensionality.

The condition of using KNN for CBR is in combination with other algorithms and involves optimisation, which could be considered as another direction for further work. In Cheng and Ma's research (49), the CBR cycle is built based on an ANN model, which completes the calculation process to filter the most similar cases. The KNN concept here was used for the "reuse" step based on a "trial-and-error" process, which needs certain work of repeat computing, to test out the optimal case. Faia et al.'s (95) research follows a similar practice aiming at optimisation. Similar results were obtained by repeated calculations using KNN, and the Particle Swarm Optimization (PSO) was combined to optimise the selection of the variables. Therefore, once related to weight determination, KNN's weaknesses are obvious.

To cope with this issue, weight factors are integrated into the system to improve the accuracy and calculation procedure. Table 5 analysed the weight determination solutions used for the CBR model in architectural related research.

Abbreviations for Table 5

AER	Absolute Error Ratio
AHP	Analytic Hierarchy Process
ANN	Artificial Neural Network
GA	Genetic Algorithm
GDM	Gradient Descent Method
GMM	Geometric Mean Method
KNN	K-Nearest Neighbours
MAD	Mean Absolute Deviation
MAE	Mean Absolute Error
MAER	Mean Absolute Error Rate
MAPE	Mean Absolute Percentage Error
MER	Modulation Error Ratio
MRA	Multiple Regression Analysis
MSD	Mean Standard Deviation
PSO	Particle Swarm Optimization
RL	Reinforcement Learning
SER	Standard Error Rate
SHAP	SHapley Additive exPlanation

Table 5 Relative information about CBR investigations

Weight determination solution	Application	Integration with other methods	Validation	Time	Author
AHP	Method improvement	GDM	MAER	2007	(52)
	Prediction	No	No	2008	(96)
	Prediction	No	MAE	2009	(97)
	Prediction	No	No	2010	(50)
	Method improvement	No	No	2014	(94)
	Prediction	No	No	2017	(14)
	Selection	No	No	2017	(98)
	Prediction	No	MAPE	2017	(99)
	Selection	No	No	2019	(32)

	Selection	No	Black-box/Experts	2019	(10)
	Prediction	No	MAPE/MSD/MAD	2020	(27)
	Selection	GMM	No	2020	(59)
	Selection	No	No	2021	(100)
GA	Prediction	ANN	SER	2010	(101)
	Prediction	No	MAER	2010	(92)
	Prediction	MRA/ANN	SE	2011	(35)
	Selection	No	No	2012	(102)
	Selection	No	MAPE	2015	(103)
	Prediction	MRA/ANN	MAPE	2015	(36)
	Selection	No	No	2017	(48)
	Prediction	No	MAPE	2020	(104)
	Prediction	No	MER	2020	(105)
	Prediction	No	No	2020	(106)
	Prediction	No	MAPE	2021	(107)
MRA	Prediction	No	No	2012	(108)
KNN	Prediction	PSO	No	2017	(95)
RL	Selection	No	No	2022	(109)
SHAP	Selection	4 Approaches in parallel	No	2023	(7)
ANN	Prediction	No	AER	2011	(110)
	Selection	KNN	Boolean	2015	(49)
	Prediction	No	No	2017	(77)

As mentioned earlier, there are very little research implement CBR approach in architectural realm, especially building retrofit. It can be seen from the Table 5, that around 2/3 research was done after 2015.

In the field of architectural research, the applications of the CBR model mainly focus on prediction, and selection in the second place. Shown in Fig. 6. Some CBR models may contain the combination of two or more algorithms that would be defined by the primary algorithm shown in the first column in Table 5.

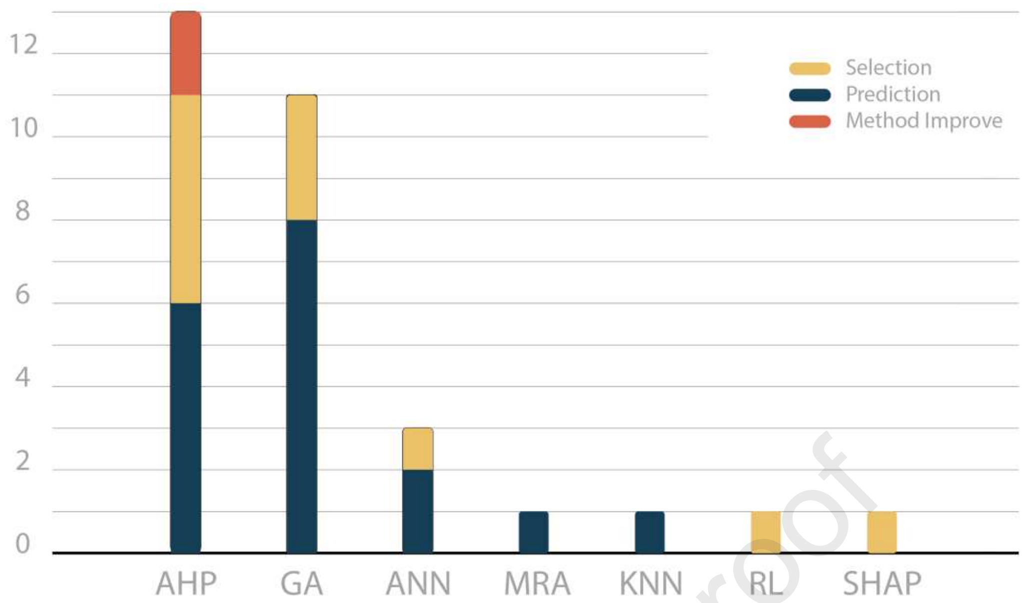

Fig. 6 Percentage of Application in Algorithms

The application of prediction pays attention to cost-estimation or risk evaluation rather than retrofit. It is important to emphasize that even though the contents of retrieval function among some studies may not be as much as predictive research, each study includes the process of retrieving the matched cases from a database, which is the core part of CBR. For example, Ahn et al. (27) use CBR to extract past empirical cases and improve the accuracy of construction budget estimation, the prediction was based on five normalized methods including interval, Gaussian distribution-based, Z-score, ratio, and logical function-based, which pre-process multiple attributes. Wang et al. (96) utilised a CBR model to replace the traditionally intuitive estimation method, the result showed this new CBR solution could not only reduce the time for reviewing the budget but also predict the cost effectively. Chen et al. (50) collected 133 guilty verdicts from the court of architectural fatal construction occupational accidents (COA), which used AHP to classify and layer the problem and solution attributes, and then weighted those attributes for determining responsibility and sentencing. This CBR model breaks the knowledge barrier for professionals by offering the judgement rules during construction, simultaneously, serving as a reference to the law attorneys for possible similar judgements in the future. Koo et al. (101) regarded the sensitivity coefficients of ANN as the weight factors to compute mathematic distance and integrated with GA to predict the budget and construction duration of multi-family housing in line with specific features. Offering a clear indication while there still are limitations and uncertainties. Likewise, due to the uncertainty, Chang et al. (106) built a multi-objective decision model, using GA, to evaluate the feasibility of the retrofit. This provides a guideline to the decision maker and benefits the framework for sustainable retrofit.

In the view of selection, the purpose is mainly about building retrofit or knowledge learning. CBR has the great advantage of selecting the similar past cases to reduce the work of research. In the research of Okudan et al. (100), the Risk Management (RM) process is usually integrated with multiple indicators, they developed a tool named CBRisk to support the RM processes as it is a knowledge-intensive process that requires effective related experience and knowledge, which bridged the gap between professional knowledge with the public. Another risk management research by Akaa et al. (59) combined GMM and AHP to study the portal-framed building cases, and support formulating the RM guideline based on AS/NZS ISO 31000:2009, to avoid the possible design of steel-framed buildings might expose to fires. Wang et al. (32) also adopted this method in developing a Lessons Mining System (LMS) to search for the most appropriate urban planning case for the decision maker as reference, which can help them to break the knowledge barrier, foresee and avoid the recurrence of potential problems. Xiao et al. (98) implemented the CBR manner to build a model named Green Building Experience-Mining (GBEM), without weight factor correction, to perform green building retrofit design scheme based on the past renovation solutions. Jafari and Valentin (48) designed a decision-making framework by CBR, which learns the Life Cycle Cost (LCC) of past cases to consider a comprehensive economic goal for energy retrofits. Hong et al. (102) investigated 362 cases in Seoul and used CBR to select the multi-family housing complex that has the effect energy saving potential.

In addition, the method improvement of how to assign values with high precision, is one of the research directions. In Kolodner's (17) principle, the weight values for CBR attributes should be determined by experts. While An et al. (52) considered the knowledge of experts were highly relied on personal experiences, thus, they integrated AHP with the Gradient Descent Method (GDM) for the CBR model to determine the specific weight in terms of perfume cost estimation through computational process. With the same goal, Ahn et al. (94) developed an attribute weight-assessing method based on CBR model to critically measure the values, which improves the accuracy and efficiency of cost estimation in the computational procedure.

Among the research for those 3 applications of CBR, the algorithm is used independently in the majority of situation as a straight-forward way to get. Thereinto, AHP and GA are the most widely used. As AHP has the advantage of layering attributes (67) (68) (69), GA optimizes the ideal case considering multiple complex attributes based on similarity (92) (101).

Apart from AHP and GA, Jin et al. (108) also introduced MRA into the CBR cycle to improve the accuracy of final cost prediction. However, due to the large number of independent variables, the calculation is rather troublesome, so statistical software is generally used in practice. Guerrero et al. (109) implied RL to train a "trial and error mechanism". However, its shortcoming of requiring certain human engineering makes it hard to popularise. Generally speaking, these two complex solutions are only suitable for multi-attribute determination problems. However, such a complex approach is costly and claims professionalism, which is not necessary for some simple building optimization projects.

Furthermore, to achieve multiple functions or goals, other algorithms can be combined within CBR cycle due to their simple internal logic and easy programming. ANN has the advantage of being integrated within CBR process. Based on the information from the big dataset, ANN can predict the future results in a large range. Such as the afore-mentioned model of ANN and KNN combination by Cheng and MA (49), they implemented the advanced non-linear solution instead of the traditional linear solution to generate a new building LEED certification level based on the previous LEED case database. Koo et al. (35) integrates the prediction process with MRA and ANN, uses GA to optimize the optimization process of the CBR model, and realises the cost prediction function of early-stage construction projects based on 101 previous projects.

In terms of validation, most evaluation processes are combined with prediction as an indicator, to achieve cost estimation. Shown in Fig. 7. Please note that this evaluation process is not mandatory for the CBR model, in fact, most CBR models used for retrofitting design do not include this evaluation component.

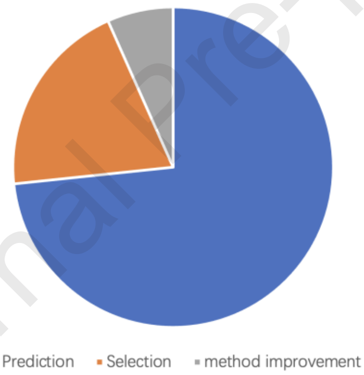

Fig. 7 Purpose for Validation

Several validation performance indicators are used to evaluate the errors during the procedure. Table5 shows that MAPE is a commonly used evaluation indicator, the same as the MAER principle (92). Ahn et al. (99) disposed that the weighted Mahalanobis distance solution is used to process the covariance effect of similarity measure into the engineering cost estimation based on the CBR-based MAER evaluation loss function. Hong et al. (36) combined MAPE to evaluate the outcomes and compare the results with the basic CBR model, which shows the advanced CBR model has more accuracy. Other methods, such as MSD, MAD, etc., only target on some specific problems (27).

Thus, the key point, to develop a CBR model for selecting potential retrofit solutions, is to determine the weighting factor. In the process of artificial algorithm development, a lot of research on solving weight factors has been carried out. In line with the results summarised above, the following section analyses and compares the primary algorithms used to determine weight factors for building retrofits.

4.3.1 Analytic Hierarchy Process (AHP)

An American operational research scientist Thomas L. Saaty (111) invented the analytic Hierarchy process in 1970. The purpose of this method is to compare the significance degree for various cases based on multiple attributes. Contraposing to some qualitative standards, AHP could establish a hierarchy model to transfer the qualitative indicators into number patterns so that calculate weight for different properties. Pairwise comparison is the core solution for achieving the importance measurement. Through the method of pairwise comparison, the factors and properties of cases were compared to explore the relationship between them (111).

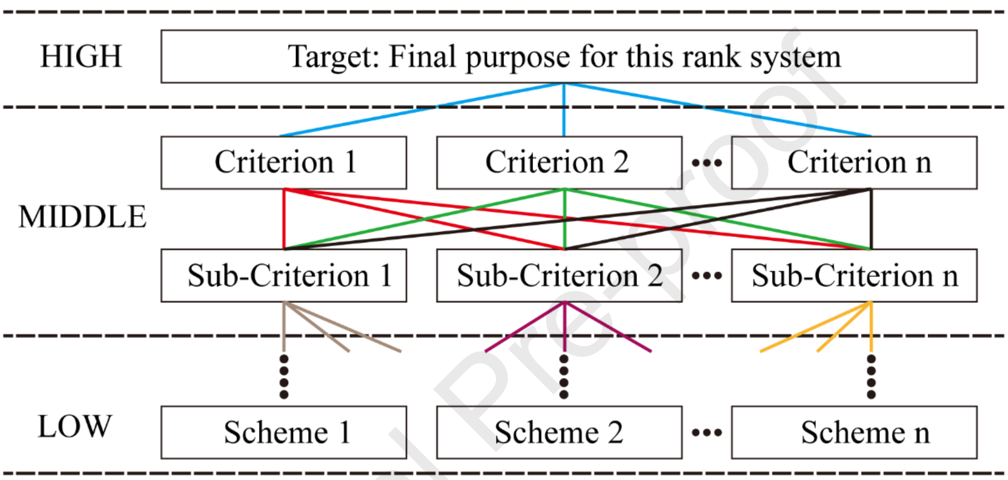

Fig. 8 Construction of hierarchy for AHP

The first step of AHP is to establish a hierarchical model of the relationship between various factors. In general, this model consists of three layers: high, middle, and low. Shown in Fig. 8. The higher level determines the lower-level elements. That is, the final result requires the product of the weights from each layer. After the model is established, the core step of weight calculation is to build the judgement matrix. Under this circumstance, all non-number elements can be converted into a number pattern. This matrix means to perform pairwise comparisons of criterions. It should be noted that, the degree of relative importance for each element is assigned entirely according to human subjectivity. In addition, apart from the numerical transformation method, the level of the whole model is significant as well, because the weight of the computed results refers to the weight of the lower criterion against the upper one. In other words, the weight achieved each time is only the weight for this layer, and the result of the scheme is the product of the results for each layer. As mentioned, in Wang et al.'s research (96), they adopted the AHP method to generate the weight value of similarity calculation and estimate the retrofit budget of historical buildings. Chou et al. (97) prove that AHP has the best performance in the aspect of new construction cost estimation and achieves final architectural budget estimation. Zhao et al. (10) present a comprehensive study of the AHP with the interior model structure. They innovatively integrated AHP method with an entropy solution to search for appropriate green building retrofit cases. Under this circumstance, the disadvantage issue of AHP in subjective could be revised via the entropy manner.

At present, this algorithm has been frequently used in the reviewed studies. Its main advantages are as follows: first, the algorithm is intuitive, and the programming calculation is relatively simple. Second, users can assess or decide the weight order subjectively, which is in line with the differentiated hypothesis of user demands. Different from GA, which requires a professional evaluation to eliminate impossible factors in advance to achieve the optimised solution. Although the result of AHP may not be the best option, it can ensure the results match the user's demands. Throughout the research process, it is important to provide users with an approximate result that meets their desired needs, even if the result is not optimal. In most cases, matching is not equal to optimisation. As mentioned earlier, the study of optimal solutions is an optimisation problem and can be regarded as another big theme.

4.3.2 Genetic Algorithm (GA)

As the most used optimization algorithm in statistics, the Genetic Algorithm (GA) is a computational model of the biological evolution process that simulates natural selection and the genetic mechanism of Darwin's biological evolution (73). In essence, it is an approach to searching for the optimal solution by simulating the natural evolution process. Compared with other optimization methods, GA adopts the probabilistic optimization method, and the optimal search space can be obtained and guided automatically without definite rules, which decreases the code-achieved difficulty.

The significant point of GA is to determine the constraint rule first and then eliminate the weight factors not meeting the relative rules. That is to say, the best result of the weight coefficient is generated after excluding other bad outcomes.

As mentioned, Hong et al. (104) integrated MAPE as a validation indicator during the calculation process. GA is used as the basic algorithm for the CBR model, which obtains the weight factors of individual attributes and forecasts the dynamic operational rating of residential buildings. The purpose of combining GA with MAPE is to enhance the optimisation and improve the accuracy. Koo et al. (101) claimed that the implementation of GA with CBR can improve the accuracy of optimal results and easy to manipulate for changing attributes during the process. In another research by Koo et al. (103), the CBR model was optimised by GA based on two criteria, RAW attribute weight range and MCAS, and the final prediction results were obtained.

In brief, the key point of GA is to determine constraint rules and exclude impossible weight factors in advance, which requires the participation of experts with professional backgrounds or rich experiences. As this algorithm is usually used to deal with optimization problems, which is relatively complicated.

4.3.3 Artificial Neutral Network (ANN)

As the most widely used data-driven algorithm, ANN is, as Koo et al. declared, the “most superior among the methodologies for calculating the weight factors” (101). ANN aims to seek the potential relationships between data hidden in the database by imitating the structure of neurons in the human brain (110). This kind of network depends on the complexity of the system and achieves the purpose of processing information by adjusting the interconnection among a large number of nodes (77).

In other words, ANN could adjust its own parameters to enable the best results without re-constructing the entire model. According to the different logic frameworks of the model, the neural network could be classified into multiple algorithms such as ANN, BPNN, CNN etc. (77) (110) ANN is a complex network structure formed by the interconnection of a large number of processing units (neurons), which is an abstraction, simplification and simulation of the human brains’ organizational structure and operating mechanism.

It is an information processing system based on the structure and function of brain neural network and simulates the activity of neurons through a mathematical model. Shown in Fig. 9.

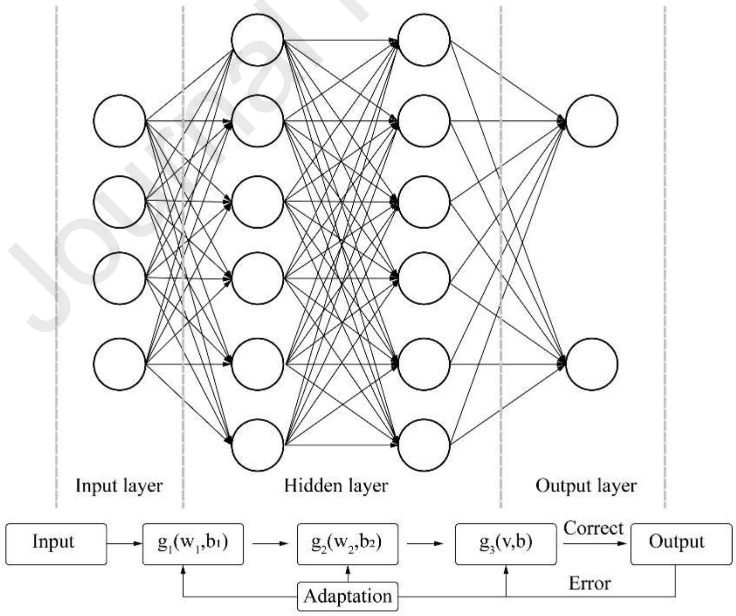

Fig. 9 Typical structure of neural network and information transmission direction

In terms of determining the weight coefficients in CBR, ANN usually trains the similarity distance immediately instead of searching for the optimal weight value, which is different from GA and AHP. However, among all weight factor determination methods, ANN is rarely used due to its complex internal structure, which is extremely unfriendly toward non-professionals.

4.4 Input and Output of CBR Model

The input is entirely dependent on the demands of users. As summarised in section 4.2, input mainly refers to surface similarity (49) (91). For the CBR system, the surface similarity determines the characteristics of the building and represents the specific features of the reference building. In this case, the input data is the basis of code recognition. In general, the input data relates to the studied objectives and often expresses its multiple attributes. In line with the summarised results, two types of input information, basic construction data and objective data, cover the whole features needed for a building. Koo et al. (35) implement this kind of data to perform cost estimation investigation in a CBR manner. Other objective data are more relevant to the ultimate purpose of the investigation. These objective data usually directly reflect the attributes related to research goals, such as building energy consumption, building retrofit costs, LEED evaluation, etc. Faia et al. (95) apply the equipment parameters as the input data, to estimate the relative building energy consumption. The combination of these two types of data forms the input that is used to locate a similar reference case in the CBR system. Cheng and Ma (49) proposed 6 types of basic building information that recognized by the U.S.Green Building Council(USGBC) as their input attributes for easier obtained values.

The output indicates the result of CBR utilisation. Through the review of the literature results, it can be concluded that the final output results include various forms, which include and not limited to specific case examples, cost, credits, criteria, laws, etc. All these patterns could be classified into one form of weight value. This is attributed that despite some research exporting target cases or other outcomes, all the results were constructed in line with the calculated scores under the CBR method. Consequently, the current output of CBR is essentially calculating the scores of different cases to pick out scenarios that meet the requirements.

5 Discussion

5.1 Beneficiaries and Objectives

According to the literature review, the beneficiaries of the CBR approach for architectural relevant issues mainly focus on two types of users: architects and stakeholders. For architects, the CBR method could assist them by providing multiple reasonable cases that reduce the efforts spent on research. For stakeholders, it could contribute to afford an intuitive understanding and foresee the possible building operational performance such as energy consumption, cost, façade exterior, etc.

In terms of objectives, cost estimation is the most significant target of relevant investigations at present (27) (52) (92) (94) (97) (99) (35) (105) (110). This is mainly because in general, the historical data related to the construction budget is sufficient to facilitate the establishment of the basic database.

Apart from this, sustainable building retrofit is another focus of attention. However, compared to cost prediction, the sustainable building retrofit investigation requires more details on buildings in line with disparate aspects to construct the reference datasets. Such complex information demands limit the development of CBR applications in building retrofit. Because of this, for other objectives, insufficient reliable reference data could lead to the impreciseness of the CBR approach. Therefore, database-based performance determines how well a CBR solution runs.

5.2 Limitations

The scientists acknowledged the advantages and disadvantages of CBR. On the positive side, remembering past experiences can help learners avoid repeating previous mistakes, and decision makers can identify which features of a problem are important to focus (49) (88). Another benefit is that the system learns by fetching new cases, which makes maintenance easier (52) (94). CBR also enables the decision makers to quickly propose solutions to problems without being fully trained in the profession and explain open and ill-defined concepts (49) (14).

On the negative side, some critics (88) claim that the main premise of the CBR cycle is based on the anecdotal evidence, which adapts elements of one case to another. This process can be complex and lead to inaccuracies. However, recent work has enhanced the CBR model with a statistical framework. This makes it possible for case-based predictions to have a higher degree of confidence and accuracy.

Besides that, the CBR input indicators reviewed for making retrofit are tending to choose the basic building information for surface similarity (49) (91), which users can easily provide. However, the inputs that involve performance indicators such as energy consumption, carbon emission or equipment performance, etc., would be unfriendly to the unprofessional users. Therefore, it is necessary to further study how to realise a system that can dynamically express the energy status of buildings with the change of input parameters. This could translate the professional understanding of performance indicators along with the input of basic surface similarity.

5.3 Future Work

In summary, there are main directions that could be further studied: (1) The sufficient and high-quality database is the guarantee of the CBR's implementation. Some architectural datasets have been established to provide reference cases for architects in all respects. With the increasing utilisation of the CBR model, each research team could consider the open access of the research database to promote the accuracy with massive datasets established. (2) The optimisation process is currently not considered in most CBR models, the concept from the mentioned A-CBR model (35) (36) could be further investigated to better support the determination of the design scheme.

6 Conclusion

This study carried out a systematic review of the CBR model in decision-making support for building retrofit. The current decision-making methods in the field of architecture have been classified and compared. The advantages of the CBR principle applied in the early decision-making for building retrofit are analysed. On this basis, this paper provides an overview of CBR approach utilisation in the building retrofit field.

Firstly, the interior-specific structure of the CBR model is reviewed and explains each step's content. In general, the CBR cycle contains five processes: represent, retrieve, reuse, revise and retain. Each phase is responsible for a unique function.

Secondly, as a data analysis method, the CBR model has not been utilized widely in the architectural realm. It can be obtained that in the building research realm, most investigations using the CBR model mainly focus on prediction and selection. What needs to be emphasized is that despite the retrieve function study being less than prediction investigations, each of this kind of research must contain the process of retrieving optimal cases from the database which is the core section of the CBR model. For the retrieving stage, how to calculate the distance between the case and target and the weight determination method are the most significant issues, which is also the difference among various approaches.

Thirdly, the weight calculation in the CBR cycle is generally classified into two types: weight factor and non-weight factor computation. The weight factor method refers to utilising some small numbers to revise the similarity computation process. Concerning weight coefficient determination in CBR, GA, AHP and ANN are the three most used weight determination solutions. Thus, the AHP method is the easiest to implement and combine with other methods. For CBR chosen system, in line with the review literature, two significant impact factors of similarity attribute type and similarity calculation control the judgement process. As the similarity calculation only relates to building basic information, the surface and derived similarity attribute could satisfy the research needs.

Fourth, given statistical data, the quality of the inputs from users determines the accuracy of the reference case. The subjective user demand preferences and the objective information for architecture cover the whole characteristics needed for inputs. The change of order will also greatly affect the outcomes.

The result of this review indicates that the CBR solution has great potential in utilising in the field of building design as reviewed in the above content. Especially in the era of big data, the amount of reference cases dataset could efficiently aid architects in conducting design in this way.

Acknowledgement

The work was conducted in the first author's Ph.D. programme as a full-time Post Graduate Researcher at the Welsh School of Architecture, Cardiff University.

Competing interests

The authors declare no competing interests.

Author information

Affiliations

Yangluxi Li[a,*], Hu Du[b], Satish Basavapatna Kumaraswamy[c]

[a c.] Welsh School of Architecture, Cardiff University, Bute Building, King Edward VII Avenue, Cardiff CF10 3NB, Wales, United Kingdom

[b.] School of Civil Engineering and Built Environment, Liverpool John Moores University, Cherie Booth Building, Byrom St, Liverpool, L3 3AF, England, United Kingdom

Email

Yangluxi Li (Corresponding author)	liy209@cardiff.ac.uk
Hu Du	h.du@ljmu.ac.uk
Satish Basavapatna Kumaraswamy	satish.bk@cardiff.ac.uk

CRediT authorship contribution statement

Yangluxi Li: Conceptualisation, Writing – Original draft, Methodology, Investigation, Data visualisation.

Hu Du: Conceptualisation, Supervision.

Satish Basavapatna Kumaraswamy: Supervision.

Bibliography

1. Why the building sector? *Architecture2030.* [Online] https://architecture2030.org/why-the-building-sector/.

2. Fulton , Mark and Grady , Heather. *United States Building Energy Efficiency Retrofits Maket Sizing and Financing Models.* Frankfurt : Deutsche Bank AG, 2012.

3. U.S. Census Bureau. US Census Bureau Construction Spending Survey, U.S. Department of Commerce. [Online] 2018. [Cited: 24 9 2018.] https://www.census.gov/construction/c30/h istorical_data.html..

4. Climate Change Act 2008 part 1. The target for 2050. *Legislation.gov.uk.* [Online] https://www.legislation.gov.uk/ukpga/2008/27/part/1.

5. Ministry of Housing, Communities & Local Goverment. English Housing Survey 2016 to 2017: headline report. *Gov.UK.* [Online] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/705821/2016-17_EHS_Headline_Report.pdf.

6. Du, Hu, Huang, Puxi and Jones, Phillip. Modular facade retrofit with renewable energy technologies: The definition and current status in Europe. *Energy and Buildings.* 2019, Vol. 205, p. 109543.

7. Ma, Dingyuan, et al. An intelligent retrofit decision-making model for building program planning considering tacit knowledge and multiple objectives. *Energy.* 2023, Vol. 263, Part B, p. 125704.

8. Deb, Chirag and Schlueter, Arno. Review of data-driven energy modelling techniques for building retrofit. *Renewable and Sustainable Energy Reviews.* 2021, Vol. 144, p. 110990.

9. Ma, Zhenjun, et al. Existing building retrofits: Methodology and state-of-the-art. *Energy and Buildings.* 2012, Vol. 55, pp. 889-902.

10. Zhao, Xue, et al. Case-based reasoning approach for supporting building green retrofit decisions. *Building and Environment.* 2019, Vol. 160, p. 106210.

11. Asadi, Ehsan, et al. Multi-objective optimization for building retrofit strategies: A model and an application. *Energy and Buildings.* 2012, Vol. 44, pp. 81-87.

12. Michael, Maria, Zhang, Lijun and Xia, XIaohua. An optimal model for a building retrofit with LEED standard as reference protocol. *Energy and Buildings.* 2017, Vol. 139, pp. 22-30.

13. Perino, Marco, Quenard, Daniel and Wakili, Karim Ghazi. Building retrofit: Energy conservation, comfort and sustainability. *Energy and Buildings.* 2017, Vol. 152, p. 40.

14. Shen, Liyin, et al. An integrated system of text mining technique and case-based reasoning (TM-CBR) for supporting green building design. *Building and Environment.* 11 2017, Vol. 124, pp. 388-401.

15. Leake, D. B. Problem Solving and Reasoning: Case-based. [book auth.] James Wright. *International Encyclopedia of the Social & Behavioral Sciences.* Pergamon : Elsevier Ltd., 2001, pp. 12117-12120.

16. R. A., Kling. A Paradigm for Reasoning by Analogy. *Artificial Intelligence.* 1971, Vol. 2, 2, pp. 147-178.

17. Kolodner, Janet L. An Introduction to Case-based Reasoning. *Artificial intelligence review.* 1992, Vol. 6, 1, pp. 3-34.

18. Hamza, Hatem, Belaïd, Yolande and Belaïd, Abdel. A Case-Based Reasoning Approach for Invoice Structure Extraction. *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007).* 2007, Vol. 1, pp. 327-331.

19. *Using CBR learning for the low-level and high-level unit of an image interpretation system.* Perner, Petra. London : Springer, 1999. International Conference on Advances in Pattern Recognition. pp. 45-54.

20. *Case-based reasoning for invoice analysis and recognition.* Hamza, Hatem, Belaïd, Yolande and Belaïd, Abdel. Berlin, Heidelberg : Springer, 2007. pp. 404-418.

21. Zahed, Nurdiyana, et al. Classification of honey odor-profile using case-based reasoning technique (Cbr). *ARPN Journal of Engineering and Applied Sciences.* 2016, Vol. 11, 10, pp. 6675-6679.

22. *CBR in the Health Sciences.* Montani, Stefania, et al. Lyon : s.n., 2012. Proceedings of the CBR in the Health Sciences Workshop at the Twentieth International Conference on Case-Based Reasoning. pp. 3-6.

23. Holt, Alec, et al. Medical applications in case-based reasoning. *The Knowledge Engineering Review.* 2005, Vol. 20, 3, pp. 289-292.

24. *Advancements and Trends in Medical Case-Based Reasoning: An Overview of Systems and System Development.* Nilsson, Markus and Sollenborn, Mikael. 2004. FLAIRS Conference. pp. 178-183.

25. *How to Improve the Adaptation Phase of the CBR in the Medical Domain.* Pusztová, Ľudmila, et al. Cham : Spinger, 2019. International Cross-Domain Conference for Machine Learning and Knowledge Extraction. pp. 168-177.

26. El-Sappagh, Shaker and Elmogy, Mohammed Mahfouz. Medical case based reasoning frameworks: Current developments and future directions. *Virtual and Mobile Healthcare: Breakthroughs in Research and Practice.* 2020, pp. 516-552.

27. Ahn, Joseph, et al. Performance evaluation of normalization-based CBR models for improving construction cost estimation. *Automation in Construction.* 2020, Vol. 119, p. 103329.

28. *Assessing classification accuracy in the revision stage of a cbr spam filtering system.* Méndez, José Ramón, et al. Berlin, Heidelberg : Springer, 2007. International Conference on Case-Based Reasoning. pp. 374-388.

29. Ji, ChangYoon, Hong, TaeHoon and Hyun, ChangTaek. CBR revision

model for improving cost prediction accuracy in multifamily housing projects. *Journal of Management in Engineering.* 2010, Vol. 26, 4, pp. 229-236.

30. Fu, Yonggang and Shen, Ruimin. GA based CBR approach in Q&A system. *Expert Systems with Applications.* 2004, Vol. 26, 2, pp. 167-170.

31. Guo, Yuan, Hu, Jie and Peng, Yinghong. Research on CBR system based on data mining. *Applied Soft Computing.* 2011, Vol. 11, 8, pp. 5006-5014.

32. Wang, Jinhuan, et al. A lessons mining system for searching references to support decision making towards sustainable urbanization. *Journal of Cleaner Production.* 2019, Vol. 209, 1, pp. 451-460.

33. Finnie, Favin and Sun, Zhaohao. R5 model for case-based reasoning. *Knowledge-Based Systems.* 2003, Vol. 16, 1, pp. 59-65.

34. RIBA. Plan of Work 2020 Overview. *RIBA Architecture.* [Online] 2020. www.ribaplanofwork.com.

35. Koo, ChoongWan, Hong, TaeHoon and Hyun, ChangTaek. The development of a construction cost prediction model with improved prediction capacity using the advanced CBR approach. *Expert Systems with Applications.* 2011, Vol. 38, 7, pp. 8597-8606.

36. Hong, Taehoon, et al. An estimation methodology for the dynamic operational rating of a new residential building using the advanced case-based reasoning and stochastic approaches. *Applied Energy.* 2015, Vol. 150, pp. 308-322.

37. Haruna, Abdulrahman, Shafiq, Nasir and Monstasir, O. A. Building information modelling application for developing sustainable building (Multi criteria decision making approach). *Ain Shams Engineering Journal.* 2021, Vol. 12, 1, pp. 293-302.

38. Olsson, Stefan, Malmqvist, Tove and Glaumann, Mauritz. An approach towards sustainable renovationd A tool for decision support in early project stages. *Building and Environment.* 2016, Vol. 106, pp. 20-32.

39. Nik-Bakht, Mazdak, et al. Economy-energy trade off automation – A decision support system for building design development. *Journal of Building Engineering.* 2020, Vol. 30, p. 101222.

40. Boxer, Eric, Henze, Gregor P. and Hirsch, Adam I. A model-based decision support tool for building portfolios under uncertainty. *Automation in Construction.* 2017, Vol. 78, pp. 34-50.

41. Ahmada, Tayyab, Aibinua, Ajibade and Thaheemb, Muhammad Jamaluddin. BIM-based iterative tool for sustainable building design: A conceptual framework. *Procedia Engineering.* 2017, Vol. 180, pp. 782 – 792.

42. Cecconi, Re, Moretti, F. N. and Tagliabue, L. C. Application of artificial neutral network and geographic information system to evaluate retrofit potential in public school buildings. *Renewable and Sustainable Energy Reviews.* 2019, pp. 266-277.

43. Amer, Mohamed, et al. Methodology for design decision support of cost-optimal zero-energy lightweight construction. *Energy & Buildings.* 2020, Vol. 223,

p. 110170.

44. Khansari, Nasrin and Hewitt, Elizabeth. Incorporating an agent-based decision tool to better understand occupant pathways to GHG reductions in NYC buildings. *Cities.* 2020, Vol. 97, p. 102503.

45. Jafari, Amirhosein and Valentin, Venessa. Selection of optimization objectives for decision-making in building energy retrofits. *Building and Environment.* 2018, Vol. 130, pp. 94-103.

46. Mejjaouli, Sobhi and Alzahrani, Maha. Decision-making model for optimum energy retrofitting strategies in residential buildings. *Sustainable Production and Consumption.* 2020, Vol. 24, pp. 211-218.

47. Juan, Yi-Kai, Gao, Peng and Wang, Jie. A hybrid decision support system for sustainable office building renovation and energy performance improvement. *Energy and Buildings.* 2010, Vol. 42, 3, pp. 290-297.

48. Jafari, Amirhosein and Valentin, Vanessa. An optimization framework for building energy retrofits decision-making. *Building and Environment.* 2017, Vol. 115, pp. 118-129.

49. Cheng, Jack C.P. and Ma, Lucky J. A non-linear case-based reasoning approach for retrieval of similar cases and selection of target credits in LEED projects. *Building and Environment.* 11 2015, Vol. 93, pp. 349-361.

50. Chen, Wei Tong, et al. Developing a CBR-based adjudication system for fatal construction industry occupational accidents. Part I: Building the system framework. *Expert Systems with Applications.* 2010, Vol. 37, 7, pp. 4867-4880.

51. Pohekar, Sanjay D. and Ramachandra, M. Application of multi-criteria decision making to sustainable energy planning—A review. *Renewable and Sustainable Energy Reviews.* 2004, Vol. 8, 4, pp. 365-381.

52. An, Sung-Hoon, Kim, Gwang-Hee and Kang, Kyung-In. A case-based reasoning cost estimating model using experience by analytic hierarchy process. *Building and Environment.* 2007, Vol. 42, 7, pp. 2573-2579.

53. Gero, John S., D'Cruz, Neville A. and Radford, Antony D. Energy in context: A multicriteria model for building design. *Building and Environment.* 1983, Vol. 18, 3, pp. 99-107.

54. Menassa, Carol C. Evaluating sustainable retrofits in existing buildings under uncertainty. *Energy and Buildings.* 2011, Vol. 43, 12, pp. 3576-3583.

55. Goodacre, Christopher, Sharples, Stephen and Smith, Peter. Integrating energy efficiency with the social agenda in sustainability. *Energy and Buildings.* 2002, Vol. 34, 1, pp. 53-61.

56. Flourentzou, Flourentzos and Roule, Claude Alain. Elaboration of retrofit scenarios. *Energy and Buildings.* 2002, Vol. 34, 2, pp. 185-192.

57. Delgarm, Navid, et al. Multi-objective optimization of the building energy performance: A. *Applied Energy.* 2016, Vol. 170, 15, pp. 293-303.

58. Figueiredo, Karoline, et al. Sustainable material choice for construction projects: A Life Cycle Sustainability Assessment framework based on BIM and

Fuzzy-AHP. *Building and Environment.* 2021, Vol. 196, p. 107805.

59. Akaa, Obinna Ukeni, Abu, Anthony and Spearpoint, Michael. A case study of design decision-making for an exemplar steel portal-framed building under fire exposure using a hybrid analysis technique. *Journal of Building Engineering.* 2020, Vol. 28, p. 101093.

60. Gercek, Mumine and Arsan, Zeynep Durmuş. Energy and environmental performance based decision support process for early design stages of residential buildings under climate change. *Sustainable Cities and Society.* 2019, Vol. 48, p. 101580.

61. William, Micheal A., et al. Building envelopes toward energy-efficient buildings: A balanced multi-approach decision making. *International Journal of Energy Research.* 2021, Vol. 45, 15, pp. 21096-21113.

62. Mosalama, Khalid M., et al. Performance-based engineering and multi-criteria decision analysis for sustainable and resilient building design. *Structural Safety.* 2018, Vol. 74, pp. 1-13.

63. *A decision making support tool for selecting green building certification credits based on project delivery attributes.* Seyis, Senem and Ergen, Esin. 2017, Building and Environment, Vol. 126, pp. 107-118.

64. Mukhamet, Tileuzhan, et al. Ranking PCMs for building façade applications using multi-criteria decision-making tools combined with energy simulations. *Energy.* 2021, Vol. 215, p. 119102.

65. Sánchez-Garrido, Antonio J., Navarro, Ignacio J. and Yepes, Víctor. Multi-criteria decision-making applied to the sustainability of building structures based on Modern Methods of Construction. *Journal of Cleaner Production.* 2022, Vol. 330, p. 129724.

66. Ren, Fukang, Wei, Ziqing and Zhai, Xiaoqiang. Multi-objective optimization and evaluation of hybrid CCHP systems for different building types. *Energy.* 2021, Vol. 215, p. 119096.

67. Sangiorgio, Valentino, et al. AR-AHP to Support the Building Retrofitting: Selection of the Best Precast Concrete Panel Cladding. *New Approaches for Multi-Criteria Analysis in Building Constructions.* USA : Springer, Cham, 2021, pp. 83–101.

68. Laguna Salvadó, L., Villeneuve, E. and Masson, D. Decision Making in Near Zero Energy Building Refurbishment: A Technology Alternatives Ranking Tool. *IFAC PapersOnLine.* 2019, Vol. 52, 13, pp. 313-318.

69. Gade, Anne N., et al. REDIS: A value-based decision support tool for renovation of building portfolios. *Building and Environment.* 2018, Vol. 142, pp. 107-118.

70. *Multi-criteria decision making for adaptive reuse of heritage buildings: Aziza Fahmy Palace, Alexandria, Egypt.* Haroun, Hebatu-Allah Abdul Fattah, Bakr, Ali Fouad and Hasan, Asmaa El-Sayed. 2019, Alexandria Engineering Journal, Vol. 58, pp. 467-478.

71. Moghtadernejad, Saviz, Chouinard, Luc E. and Mirza, M. Saeed. Design strategies using multi-criteria decision-making tools to enhance the performance of building façades. *Journal of Building Engineering.* 2020, Vol. 30, p. 101274.

72. Ruggeri, Aurora Greta, et al. Planning energy retrofit on historic building stocks: A score-driven decision support system. *Energy & Buildings.* 2020, Vol. 224, p. 110066.

73. Mirjalili, Seyedali. Genetic algorithm. *Evolutionary algorithms and neural networks.* 2019, pp. 43-55.

74. She, Chen, et al. Life cycle cost and life cycle energy in zero-energy building by multi-objective optimization. *Energy Reports.* 2021, Vol. 7, pp. 5612-5626.

75. Yadegaridehkordi, Elaheh, et al. Assessment of sustainability indicators for green building manufacturing using fuzzy multi-criteria decision making approach. *Journal of Cleaner Production.* 2019, Vol. 277, p. 122905.

76. Abedi, Maysam, et al. PROMETHEE II: A knowledge-driven method for copper exploration. *Computers and Geosciences.* 2012, Vol. 46, pp. 255-263.

77. Beccali, Marco, et al. Artificial neural network decision support tool for assessment of the energy performance and the refurbishment actions for the non-residential building stock in Southern Italy. *Energy.* 2017, Vol. 137, pp. 1201-1218.

78. Liu, Bo and Rodriguez, Dragan. Renewable energy systems optimization by a new multi-objective optimization technique: A residential building. *Journal of Building Engineering.* 2021, Vol. 35, p. 102094.

79. He, Miaomiao, et al. Multi-objective optimization for a large scale retrofit program for the housing stock in the North East of England. *Energy Procedia.* 2015, Vol. 78, pp. 854-859.

80. Naderi, Ehsan, et al. Multi-objective simulation-based optimization of controlled blind specifications to reduce energy consumption, and thermal and visual discomfort: Case studies in Iran. *Building and Environment.* 2020, Vol. 169, p. 106570.

81. Martínez, Sandra, et al. Model calibration and exergoeconomic optimization with NSGA-II applied to a residential cogeneration. *Applied Thermal Engineering.* 2020, Vol. 169, p. 114916.

82. Lan, Lan, Wood, Kristin L. and Yuen, Chau. A holistic design approach for residential net-zero energy buildings: A case study in Singapore. *Sustainable Cities and Society.* 2019, Vol. 50, p. 101672.

83. K-Nearest Neighbors Algorithm. *IBM.* [Online] https://www.ibm.com/hk-en/topics/knn.

84. Zhang, Min-Ling and Zhou, Zhi-Hua. ML-KNN: A lazy learning approach to multi-label learning. *Pattern recognition.* 2007, Vol. 40, 7, pp. 2038-2048.

85. *BIM-based iterative tool for sustainable building design: A conceptual*

framework. Ahmad, Tayyab, Aibinu, Ajibade and Thaheem, Muhammad Jamaluddin. 2017. Procedia engineering 180. pp. 782-792.

86. Moghadam, Sara Torabi and Lombardi, Patrizia. An interactive multi-criteria spatial decision support system for energy retrofitting of building stocks using CommuntiyVIZ to support urban energy planning. *Building and Environment.* 2019, Vol. 163, p. 106233.

87. Sandanayake, Malindu, Zhang, Guomin and Setunge, Sujeeva. Estimation of environmental emissions and impacts of building construction – A decision making tool for contractors. *Journal of Building Engineering.* 2019, Vol. 21, pp. 137-185.

88. Kolodner, J.L. *Case based reasoning.* San Mateo : Morgan Kaufmann, 1993.

89. Roth, Frederick Hayes and Jacobstein, Neil. The state of knowledge-based systems. *Publication History.* 1994, Vol. 37, 3, pp. 26-39.

90. Leake, David B. Case-based reasoning. *The Knowledge Engineering Review.* 1994, Vol. 9, 1, pp. 61-64.

91. DE MANTARAS , RAMON LOPEZ, et al. Retrieval, reuse, revision and retention in case-based reasoning. *The Knowledge Engineering Review.* 2005, Vol. 20, 3, pp. 215-240.

92. Kim, Kyong Ju and Kim, Kyoungmin. Preliminary Cost Estimation Model Using Case-Based Reasoning and Genetic Algorithms. *Journal of Computing in Civil Engineering.* 2010, Vol. 24, 6.

93. Ontañón, Santiago and Plaza, Enric. Collaborative Case Retention Strategies for CBR Agents. *Case-Based Reasoning Research and Development.* 2003, pp. 392-406.

94. Ahn, Joseph, et al. The attribute impact concept: Applications in case-based reasoning and parametric cost estimation. *Automation in Construction.* 2014, Vol. 43, pp. 195-203.

95. Faia, Ricardo, et al. Case based reasoning with expert system and swarm intelligence to determine energy reduction in buildings energy management. *Energy and Buildings.* 2017, Vol. 155, pp. 269-281.

96. Wang, Huey-Jiun, Chiou, Chien-Wei and Juan, Yi-Kai. Decision support model based on case-based reasoning approach for estimating the restoration budget of historical buildings. *Expert Systems with Applications.* 2008, Vol. 35, 4, pp. 1601-1610.

97. Chou, Jui-Sheng. Web-based CBR system applied to early cost budgeting for pavement maintenance project. *Expert Systems with Applications.* 2009, Vol. 36, 2, pp. 2947-2960.

98. Xiao, Xue, Skitmore, Martin and Hu, Xin. Case-based Reasoning and Text Mining for Green Building Decision Making. *Energy Procedia.* 3 2017, Vol. 111, pp. 417-425.

99. Ahn, Joseph, et al. Covariance effect analysis of similarity measurement

methods for early construction cost estimation using case-based reasoning. *Automation in Construction.* 2017, Vol. 81, pp. 254-266.

100. Okudan, Ozan, Budayan, Cenk and Dikmen, Irem. A knowledge-based risk management tool for construction projects using case-based reasoning. *Expert Systems with Applications.* 2021, Vol. 173, p. 114776.

101. Koo , Choong-Wan, et al. A study on the development of a cost model based on the owner's decision making at the early stages of a construction project. *International Journal of Strategic Property Management.* 2010, Vol. 14, pp. 121-137.

102. Hong, Taehoon, Koo, Choongwan and Park, Sungki. A decision support model for improving a multi-family housing complex based on CO2 emission from gas energy consumption. *Building and Environment.* 2012, Vol. 52, pp. 142-151.

103. Koo, Choongwan and Hong, Taehoon. A dynamic energy performance curve for evaluating the historical trends in the energy performance of existing buildings using a simplified case-based reasoning approach. *Energy and Buildings.* 4 2015, Vol. 92, pp. 338-350.

104. Kwon, Nahyun, et al. Maintenance cost prediction for aging residential buildings based on case-based reasoning and genetic algorithm. *Journal of Building Engineering.* 2020, Vol. 28, p. 101006.

105. Jung, Sangsun, et al. Construction Cost Estimation Using a Case-Based Reasoning Hybrid Genetic Algorithm Based on Local Search Method. *Sustainability.* 2020, Vol. 12, 19, p. 7920.

106. Chang, Soowon, Castro-Lacouture, Daniel and Yamagata, Yoshiki. Decision support for retrofitting building envelopes using multi-objective optimization under uncertainties. *Journal of Building Engineering.* 2020, Vol. 32, p. 101413.

107. Liang, Yunping, Ashur, Baabak and Li, Mingshu. Forecasting the Construction Expenditure CashFlow for Transportation Design-Build Projectswith a Case-Based Reasoning Model. *Journal of Construction Engineering and Management.* 2021, Vol. 147.6, p. 04021043.

108. Jin, RunZhi, et al. MRA-based revised CBR model for cost prediction in the early stage of construction projects. *Expert Systems with Applications.* 4 2012, Vol. 39, 5, pp. 5214-5222.

109. Guerrero, J.I., Miró-Amarante, G. and Martín, A. Decision support system in health care building design based on case-based reasoning and reinforcement learning. *Expert Systems with Applications.* 2022, Vol. 187, p. 116037.

110. Ji, Sae-Hyun, Park, Moonseo and Lee, Hyun-Soo. Cost estimation model for building projects using case-based reasoning. *Canadian Journal of Civil Engineering.* 2011, pp. 570-581.

111. Saaty, Thomas L. How to make a decision: the analytic hierarchy process. *Interfaces.* 1994, Vol. 24, 6, pp. 19-43.

112. *Rapid multi-objective optimization with multi-year future weather*

condition and decision-making support for building retrofit. Shen, Pengyuan, et al. 2019, Energy, Vol. 172, pp. 892-912.

113. *Sustainable management of the building stock: A Delphi study as a decision support tool for improved inspections.* Jiménez-Pulido, Cristina, Jiménez-Rivero, Ana and García-Navarro, Justo. 2020, Sustainable Cities and Society, Vol. 61, p. 102184.

114. *Tools and procedures to support decision making for cost-effective energy and carbon emissions optimization in building renovation.* Mora, T. Dalla , et al. 15, 2018, Energy and Buildings, Vol. 167, pp. 200-215.

115. Asad, Ehsan, et al. Multi-objective optimization for building retrofit strategies: A model and an application. *Energy and Buildings.* 1 2012, Vol. 44, pp. 81-87.

116. The Danish Government. *Strategy for energy renovation of buildings.* 2014.

117. Mavromatidis, Lazaros Elias, Bykalyuk, Anna and Lequay, Hervé. Development of polynomial regression models for composite dynamic envelopes' thermal performance forecasting. *Applied Energy.* 2013, Vol. 104, pp. 379-391.

118. Roberti, Francesca, et al. Energy retrofit and conservation of a historic building using multi-objective optimization and an analytic hierarchy process. *Energy and Buildings.* 2017, Vol. 138, 1, pp. 1-10.

119. Gahrooei, Mostafa Reisi, et al. Timing residential photovoltaic investments in the presence of demand uncertainties. *Sustainable Cities and Society.* 2016, Vol. 20, pp. 109-123.

120. Guo, Rui, et al. Optimization of cool roof and night ventilation in office buildings:. *Renewable Energy.* 2020, Vol. 147, pp. 2279-2294.

121. Torres-Rivas, Alba, et al. Multi-objective optimisation of bio-based thermal insulation materials in. *Applied Energy.* 2018, Vol. 224, pp. 602-614.

122. Lim, Hyunwoo and Zhai, Zhiqiang (John). Influences of energy data on Bayesian calibration of building energy model. *Applied Energy.* 2018, Vol. 231, pp. 686-698.

123. Stevanovic, Sanja. Parametric study of a cost-optimal, energy efficient office building in Serbia. *Energy.* 2016, Vol. 117, pp. 492-505.

124. Yang, Sheng-Yuan. Developing an energy-saving and case-based reasoning information agent with Web service and ontology techniques. *Expert Systems with Applications.* 2013, 40, pp. 3351–3369.

125. Administration, U.S. Energy Information. *Emissions of Greenhouse Gases in the United States 2009.* s.l. : U.S. Energy Information Administration, 2011.

126. Ministry of Construction, PRC. *Design Standard for Energy Efficiency of Public Buildings GB50189-2019.* Beijing : Ministry of Construction, PRC;, 2019.

127. *Developing a Significant Nearest Neighbor Search Method for Effective Case Retrieval in a CBR System.* Tsai, CHieh-Yuan and Chiu, Chuang-Cheng.

Singapore : Spring Conference, 2009. 2009 International Association of Computer Science and Information Technology. pp. 262-266.

128. Costa-Carrapiçoa, Inês, Raslanb, Rokia and González, Javier Neila. A systematic review of genetic algorithm-based multi-objective optimisation for building retrofitting strategies towards energy efficiency. *Energy & Buildings.* 2020, Vol. 210, p. 109690.

Declaration of interests

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

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: