Multi-disciplinary and multi-objective optimization problem re-formulation in computational design exploration

A case of conceptual sports building design

Yang, Ding; Ren, Shibo; Turrin, Michela; Sariyildiz, Sevil; Sun, Yimin

DOI
10.1016/j.autcon.2018.03.023

Publication date
2018

Document Version
Final published version

Published in
Automation in Construction

Automation in Construction 92 (2018) 242–269

Contents lists available at ScienceDirect

Automation in Construction

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

Multi-disciplinary and multi-objective optimization problem re-formulation in computational design exploration: A case of conceptual sports building design

Ding Yang[a,b], Shibo Ren[c], Michela Turrin[b], Sevil Sariyildiz[b], Yimin Sun[a,*]

[a] School of Architecture, State Key Laboratory of Subtropical Building Science, South China University of Technology, Guangzhou, China
[b] Chair of Design Informatics, Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, The Netherlands
[c] Arup, Amsterdam, The Netherlands

ARTICLE INFO

Keywords:
Multi-disciplinary optimization
Multi-objective optimization
Computational design exploration
Knowledge extraction
Statistical analysis techniques
Optimization problem re-formulation
Sports buildings
Architectural performance
Climate performance
Structural performance

ABSTRACT

The benefits of applying multi-objective optimization (MOO) in building design have been increasingly recognized in recent decades. The existing or traditional computational design optimization (CDO) approaches mostly focus on optimization problem solving (OPS), as they often conduct optimizations directly by assuming the optimization problems in question are good enough. In contrast, the computational design exploration (CDE) approaches defined in this research mainly focus on optimization problem formulation (OPF), which are considered more essential and aim to achieve or ensure appropriate optimization problems before conducting optimizations. However, the application of the CDE is very limited especially in conceptual architectural design. The necessity of re-formulating original optimization problems and its potential impacts on optimization results are often overlooked or not emphasized enough.

This paper proposes a new CDE approach that highlights the knowledge-supported re-formulation of a changeable initial optimization problem. It improves upon the traditional CDO approach by introducing a changeable initial OPF and inserting a CDE module. The changeable initial OPF allows expanding the dimensionality of an objective space and design space being investigated, and the CDE module can re-formulate the changeable optimization problem using the information and knowledge extracted from statistical analyses. To facilitate designers in achieving the proposed approach, an improved computational platform is used which combines parametric modeling software (including simulation plug-ins) and design optimization software. Assisted by the platform, the proposed approach is applied to the conceptual design of an indoor sports building that considers multi-disciplinary performance criteria (including architecture-, climate- and structure-related criteria) and a wide range of geometric variations. Through the case study, this paper demonstrates the use of the proposed approach, verifies its benefits over the traditional method, and unveils the factors that may affect the behaviour of the proposed approach. Besides, it also shows the suitability of the computational platform used.

1. Introduction

Nowadays, multi-objective optimization (MOO), coupled with building performance simulation and parametric modeling, has been increasingly used to improve overall building performance [1–4]. However, the importance of optimization problem formulation (OPF) or re-formulation is often overlooked in conceptual architectural design. Most existing studies are only interested in optimization problem solving (OPS), i.e. running various algorithms to search for optimal solutions based on already formulated or initially formulated optimization problems, without sufficiently demonstrating how the problems are formulated and how they may affect the optimal results.

It is through the OPF that a design task can be partially converted to an optimization problem. Key components of the OPF include at least two aspects: (1) the formulation of objective space - selecting objective and constraint variables (i.e. output variables) and constraint values; (2) the formulation of design space - selecting design variables (i.e. input variables) and their domains. The former determines all performance goals and constraints to be achieved; while, the latter determines all possible design alternatives that can be searched from.

In fact, the OPF is more essential than the OPS. If an optimization problem is formulated in meaningless way, it makes no sense to solve it. An improperly formulated objective space may lead to entirely wrong results; and, an improperly formulated design space may provide a poor

* Corresponding author.
E-mail address: arymsun@scut.edu.cn (Y. Sun).

https://doi.org/10.1016/j.autcon.2018.03.023
Received 17 October 2017; Received in revised form 10 March 2018; Accepted 20 March 2018
Available online 25 April 2018

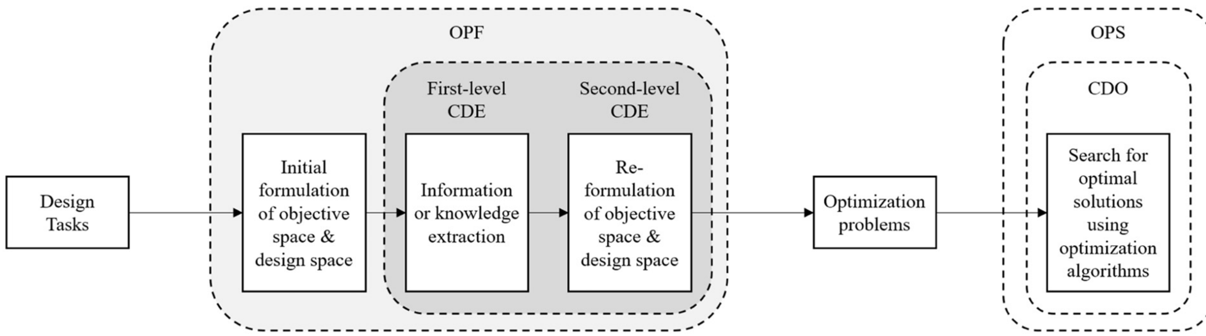


Fig. 1. The relationships between the CDO, CDE, OPS and OPF.

"design alternative pool" to search from. Apparently, it is not wise for designers to dive directly into the OPS, without properly considering the OPF. This is especially true for conceptual architectural design optimization. During the OPF, designers usually have large freedom in defining the objective space and design space, which may lead to improper definitions. The initial OPF is often unstable and poorly defined, due to the "ill-structured" nature of design tasks and the limited knowledge support (see Section 2). Thus, it indicates the need of re-formulating or revising the initial OPF with more sufficient information and knowledge support, which we consider as computational design exploration (CDE), a crucial step prior to computational design optimization (CDO).

Specifically, we define the CDE as the process of extracting useful information or knowledge (i.e. first-level CDE), and of applying it to re-formulate the original optimization problem (i.e. second-level CDE). The aim of the CDE is to achieve a good OPF before diving into the OPS. In contrast, the CDO is defined as the process that is only keen on the OPS. The aim of the CDO is to search for optimal solutions for a given or fixed optimization problem. The relationships between the CDO, CDE, OPS and OPF are summarized in a diagram (Fig. 1).

In response to the need of knowledge-supported re-formulation, this paper proposes a new holistic approach, emphasizing the CDE in which relevant information and knowledge are extracted to support the re-formulation of the initial optimization problem in a more informed manner. Statistical analysis techniques, such as correlation analysis, cluster analysis and sensitivity analysis are used for the knowledge extraction. An improved computational platform is also used for achieving the proposed approach, which integrates parametric modeling software (including simulation plug-ins) and design optimization software. With a focus on the conceptual design of indoor sports buildings, the proposed approach is applied to a complex real-world project which considers multi-disciplinary performance criteria (including architecture-, climate- and structure-related criteria) and a wide range of geometric variations. Through the case study, this paper demonstrates the use of the proposed approach, verifies its benefits over the traditional method, and unveils the factors that may affect the behaviour of the proposed approach. Besides, it also shows the suitability of the computational platform used.

2. Optimization problem (re)formulation and knowledge support

2.1. Initial formulation of an optimization problem

Due to the "ill-structured" nature of design tasks, the initial formulation of an optimization problem is usually unstable. As first defined by Simon [5], a building design task is ill-structured (i.e. lack of definition) in a number of respects; and, it seemed to reach a consensus, among researches in the late 1990s, that most of real-world tasks, in particular design tasks, are ill-structured [6–16]. This is especially true in the conceptual design stage. In this stage, there are no definitive goals and constraints, since the goals are usually vague and many performance criteria maybe unknown; and there are no definitive solutions either, because a wide range of different solutions can be valid responses to the goals and constraints [17]. Thus, the initially formulated objective space and design space are usually unstable; they are subject to change (i.e. re-formulation) once more information and knowledge becomes available.

Due to the limited knowledge support, the initial formulation of an optimization problem is often poorly defined. At the very beginning of a conceptual design, the designers are usually not able to perceive every aspect of the design task, since they have to rely on their limited knowledge (e.g. educated guesses and/or intuition). For converting the design task to an optimization problem, they have to answer: what are the most important design issues and performance criteria; and what kinds of solutions most probably manage to solve these issues? According to Logan and Smithers [9], the designers' answers to these questions are often subjective and highly context dependent; not surprisingly, the initial expression of the design task is often misleading. From the perspective of the OPF, the initial objective space and design space are probably poorly defined.

2.2. Re-formulation of an optimization problem

Given the limitations above, the re-formulation of the initial optimization problem is inevitable in conceptual architectural design. It requires a balance between reducing computational cost and increasing design creativity, i.e., between variable screening and variable adding. Here, design variable screening refers to the process of screening out unimportant design variables (that contribute the least to the variation of objective variables), and design variable adding refers to the process of introducing new design variables (that create new design variations). Objective variable screening refers to the process of identifying the most meaningful performance criteria to be considered as final objectives, and objective variable adding refers to the process of introducing new objective variables.

This balance is challenging due to its conflicting nature; designers may struggle between reducing and increasing the dimensionality of a design space and of an objective space. Specifically, for the re-formulation of a design space, the decision whether or not to include more design variables has to be made. From the perspective of increasing design creativity, the incorporation of new design variables is crucial for the creative design [18]; while from the perspective of reducing computational cost, the best model is usually the simplest one [19], and entities should not be multiplied beyond necessity according to the principle of "Occam's razor" [20]. Similarly, for the re-formulation of an objective space, the decision whether or not to include more objective variables has to be made. The incorporation of new objective variables may be beneficial for a more holistic assessment, while it also means the increase of computational cost. In this regard, the total number of final objective variables is often limited to less than or equal to three, given the challenges of handling many-objective optimization problems [21].

2.3. Information and knowledge support

To properly support the variable screening and variable adding, substantial information and knowledge are needed. The information desired for supporting the variable screening includes output-output and input-output relations. The former refers to the inter-correlations between pairs of objective variables; while the latter includes the impact of design variables on objective variables, and the relative importance ranking of all design variables with respect to objective variables. The knowledge is derived from designers' interpretation of the information in disciplinary contexts.

For large-scale projects (e.g. indoor sports buildings), the variable relations become complex when the number of the variables increases. First, many performance criteria from various disciplines need to be considered for large-scale projects, and some of them are often in conflict with each other. For instance, the maximization of daylight availability (in climate design) conflicts with the minimization of operational energy [22–24]; the minimization of maximum displacement (in structural design) conflicts with the minimization of structural weight [25,26] and hence embodied energy; the geometrical preference or aesthetics (in architectural design) may conflict with engineering performances [27] etc. When all these performance criteria are considered simultaneously, their inter-correlations become more complex. Second, a large number of design variables are often needed to define the complex geometries of large-scale projects. For some design concepts, the number of design variables is fixed; while for other concepts, it may be changeable. In the latter scenario that we are interested in, the changeable number of design variables facilitates the definition of geometries with different levels of complexity; or in other words, the number of design variables may readily increase, which complicates the relations between design variables and objective variables.

In this context, statistical data analysis and visualization techniques can be helpful for extracting the useful information and knowledge that are not known (or not clear) in the initial formulation but relevant to the re-formulation. The interrelations between the initial formulation, re-formulation, information and knowledge support are summarized in a diagram (Fig. 2).

3. Literature review

A series of studies that apply MOO techniques to building design are reviewed in this section. They are categorized into two groups: computational design optimization and computational design exploration. According to their definitions, the former mainly focuses on solving already formulated or initially formulated optimization problems by using various search algorithms; while, the latter focuses on formulating good optimization problems before solving them via the knowledge-supported re-formulation.

3.1. Computational design optimization

Studies related to CDO dominate the reviewed literature. Most of them assume that the already formulated optimization problems are good enough for performing optimizations, and do not demonstrate sufficiently how the problems are formulated, let along how the problem formulations may affect the optimization results. They may involve different building disciplines, such as climate design, structural design etc.

Typical MOO studies for climate design are reviewed, focusing on the geometrical optimization of building envelopes with respect to daylighting, thermal, energy and cost criteria etc. Lartigue, B., Lasternas, B. and Loftness, V. [22] optimized a simple building envelope according to the triple objective of heating load, cooling load and daylighting, by using brute-force search. Window to wall ratio and window type were selected as the only two design variables, which were considered as strongly impacting the objective functions. Manzan, M. and Clarich, A. [23] optimized the geometry of an external shading device according to energy and daylighting criteria, by using a fast algorithm that combines response surfaces and genetic algorithms. Three geometrical design variables were selected according to the given design concept. Futrell, B.J., Ozelkan, E.C. and Brentrup, D. [24] optimized the envelope of a single-zone classroom according to energy and daylighting criteria, by using a Hooke Jeeves and Particle Swarm Optimization algorithm. Five geometrical design variables were selected, together with six material-related variables; and some reasons of choosing them were briefly given. Kasinalis, C. et al. [28] utilized MOO to assess the performance potential of seasonally adaptable facades. For different seasonal scenarios, the envelope of a single-person south facing office was optimized according to energy and thermal criteria by using NSGA-II. Only one geometrical design variable (i.e. window to wall ratio) was chosen, together with five material-related variables. Brownlee, A.E.I. and Wright, J.A. [29] applied several variants of NSGA-II to optimize the envelope of a mid-floor of a small commercial office according to energy and construction cost criteria. Out of fifty design variables selected (for five thermal zones), half of them are geometry related but only for defining window to wall ratios and overhang shadings. Negendahl, K. and Nielsen, T.R. [30] focused on the optimization of the overall geometry of an office building by using SPEA2. Four objectives were considered, namely, building energy use, capital cost, daylight distribution and thermal indoor environment; and three design variables were selected to define the overall geometry based on a “folding envelope” concept.

Typical MOO studies for structural design are reviewed, focusing on the topological optimization of discrete structures with respect to deflection and weight (or cost, embodied energy) criteria etc. Kicinger, R. and Arciszewski, T. [25] minimized the total weight and maximum horizontal displacement of the steel structural system of a tall building, by using a weighted-sum approach. In the experimental design, a fixed-length genome consisting of 220 genes (i.e. design variables) was used

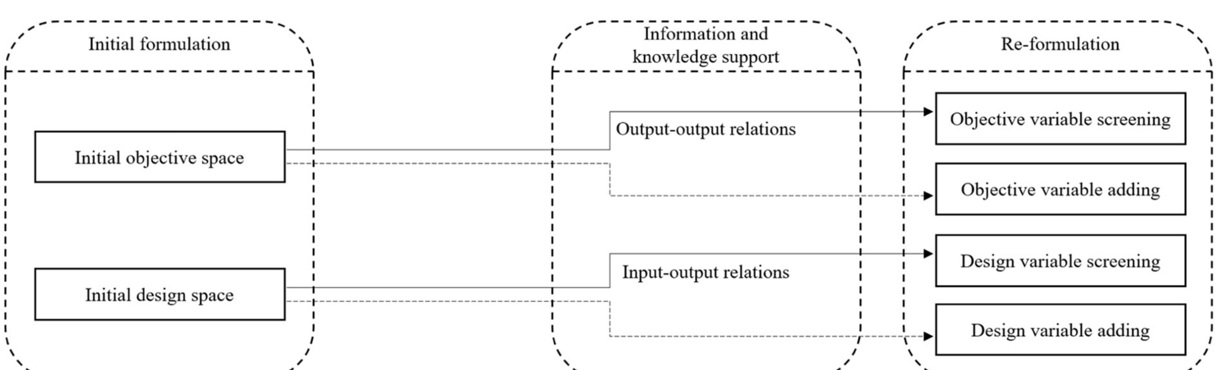


Fig. 2. The interrelations between the initial formulation, re-formulation, information and knowledge support.

to define the structural system (in terms of the types of bracing elements, beams and supports). Later, Kicinger, R., Obayashi, S. and Arciszewski, T. [26] extended the previous study by using a more advanced algorithm, SPEA2, for solving the same bi-objective optimization problem. Similarly, Richardson, J.N. et al. [31] minimized the deflection and cost of the X-bracing system of four steel facades by using MOGA. A fixed-length chromosome (determined by the number of possible positions for bracing cables) was used to define the bracing topology in each facade. Winslow, P., Pellegrino, S. and Sharma, S.B. [32] developed a novel method for synthesis of optimal grid structures (consisting of a repeating unit cell on free-form surfaces) based on MOGA. In this method, the selection of design variables depends on the choice of "design points" on a given surface. In an example case, twelve geometrical design variables (i.e. six design points) were selected, and the structure was optimized for minimizing the deflections under two different combinations of load cases.

There are very few MOO studies that simultaneously consider multi-disciplinary objectives (e.g. objectives from both climate design and structural design) in the literature of building design. They are known as multi-disciplinary optimization (MDO), and reviewed here. Flager, F. et al. [33] optimized the envelope and structure of a rectangular classroom by using process integration and design optimization (PIDO) software which is originally used in the aerospace industry. The objective is to minimize the capital cost of the steel structure and the operational energy cost; the constraints include criteria about structural safety, daylighting and architectural space; the design variables include building length, orientation, window to wall ratio, and structural element sections. Brown, N.C. and Mueller, C.T. [34] applied MOO to the conceptual design of three long span structures. For all the three cases, structural and energy performances were considered simultaneously, and a set of geometrical design variables was used to define the structure and envelope. Mueller, C.T. and Ochsendorf, J.A. [27] developed a MOO approach for considering both quantitative goals (e.g. structural efficiency, cost, embodied energy) and qualitative requirements (e.g. aesthetics, designer intent) in conceptual design. It extends existing interactive evolutionary algorithms for the increased inclusion of designer preferences.

In addition, some researchers used the term "exploration" to express the investigation of performance driven geometry using optimization techniques. They may concern knowledge extraction, but the knowledge extracted was not applied to modify the original optimization problems. Turrin, M., von Buelow, P. and Stouffs, R. [35] and their later study [36] optimized the geometry of a long span roof by using an interactive GA, for minimizing incident radiation and maximizing daylight factor in summer, and for maximizing both in winter. They emphasized the knowledge extraction from sub-optima solutions, rather than focusing only on optimal solutions. Janssen [37] optimized the configuration of an apartment for minimizing solar radiation and maximizing daylight. His main goal of using MOO techniques is not finding optimal solutions, but rather discovering unexpected design configurations and useful knowledge feedback.

3.2. Computational design exploration

Studies related to CDE are limited, compared with the other kind. Nevertheless, there is a series of early studies that emphasize the importance of design exploration theoretically. In these studies, the notion of design exploration was defined differently based on varying design models proposed. Gero [38] proposed a Function-Behavior-Structure (FBS) model of design (i.e. design prototype); based on which he [11] characterized exploration in design as a process of creating new or modifying existing design state spaces (consisting of three subspaces of function, behaviour and structure). Smithers et al. [39] developed an Exploration-Based model of design; based on the model, Smithers, Corne and Ross [12] considered that an exploration process involves the construction and incremental extension of problem statements and associated solutions. Maher, Poon and Boulanger [14,15] defined exploration as a phenomenon in design where problem spaces interact and evolve with solution spaces over time, based on a Problem-Design Exploration model.

Although the definitions of design exploration differ in detail, they essentially agreed on the re-formulation or co-evolution of design spaces and/or objective spaces, as well as on the dynamic or iterative nature. Gero and Kannengiesser [40,41] articulated three types of re-formulation of design state spaces, and indicated the dynamic nature of designing. Logan and Smithers [9] pointed out that the formulation of a design problem at any stage is not final; it needs to be re-defined in an iterative manner, until the knowledge obtained has become insignificant and the designer has reached the limits of his or her understanding on the problem. Maher, Poon and Boulanger [14,15] stated that design is an iterative interplay to "fix" a problem from the problem space and to "search" plausible solutions from the corresponding solution space. In addition, Jonas [10] considered that finding and solving design problems are a dynamic, cyclically self-sustaining process, instead of a static and linear process; Recently, the iterative re-formulation process was revisited by Arora [42] and demonstrated with some simple engineering examples.

Other related issues, such as knowledge extraction, design creativity etc. are among the early researchers' concerns. For the knowledge extraction, Smithers et al. [39] believed that knowledge about the nature of a design space should be obtained before goals can be well formulated; Gero [38] suggested to bring together all the necessary knowledge appropriate to a design situation in a conceptual schema (i.e. design prototype) to provide the basis for the start and continuation of the design. For the design creativity, Gero and Maher [43] stated that creative design occurs when new design variables are introduced in the design process; on the other hand, the introduction of new criteria may be also beneficial for achieving creative design, according to Navinchandra [7]. Moreover, Dorst and Cross [44] identified aspects of creativity in design related to the formulation of design problems. And, a set of studies in modeling creativity and knowledge-based creative design were provided in [45].

Except for the early theoretical studies as mentioned above, there are very few applications of CDE. The most typical ones are the applications of sensitivity analysis in building energy performance [46,47]. These studies aim to simplify (or re-formulate) the original design space by screening out unimportant design variables. But, most of them did not conduct the consequent optimization based on the simplified design space, and often focused on a late design stage where most of the design variables investigated are not geometry-related. Heiselberg et al. [48] conducted a sensitivity analysis of primary energy use for an office building. Various design variables, including very limited geometry-related variables, were ranked according to their relative importance. Shen and Tzempelikos [49] presented a global uncertainty and sensitivity analysis of five performance metrics to seven selected design variables for a private office. The information about the relative sensitivity of each design variable was extracted for further study.

3.3. Current challenges

As indicated by the literature review, research challenges exist in the application of the CDE to building design optimization. On one hand, the early studies focusing on the theoretical development of design exploration models lack sufficient demonstration or application of the proposed models via real-world building design projects. On the other hand, the current studies applying sensitivity analysis to simplify original design spaces are limited in terms of dealing with geometric complexity and utilizing the simplified design spaces for the consequent optimizations. When focusing on the conceptual architectural design optimization, the application of the CDE is even less. The importance of

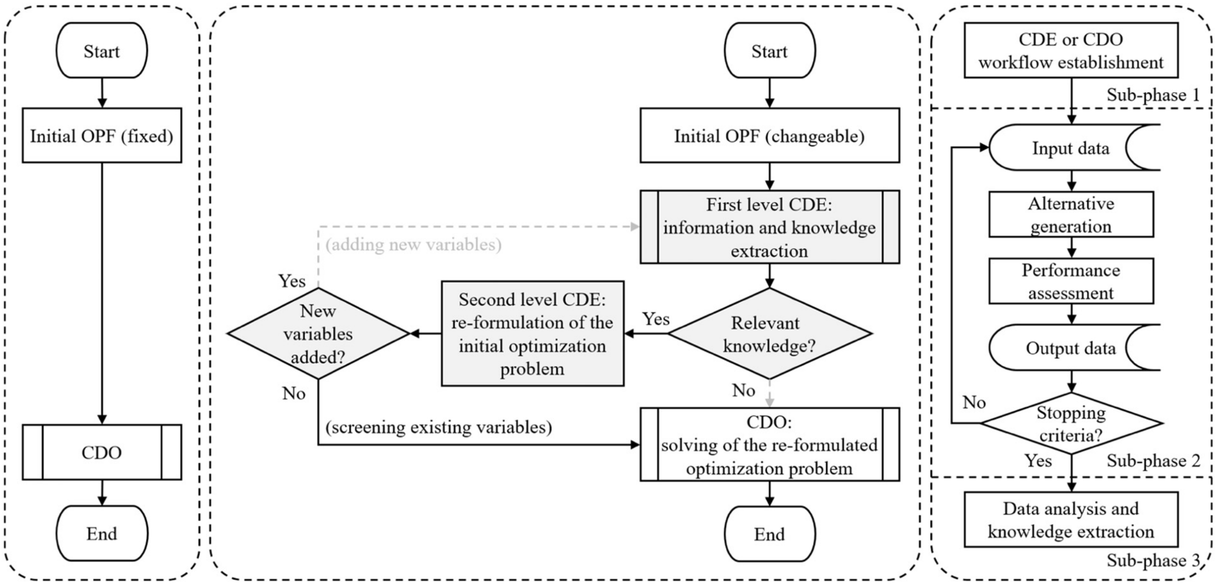


Fig. 3. The traditional CDO approach (left); the proposed CDE approach (middle); the sub-phases of the first-level CDE and of the CDO (right).

optimization problem re-formulation, and further, its possible impacts on the optimization results are often overlooked or not emphasized enough.

4. Computational approach and platform

4.1. Computational approach

In response to the challenges, a new CDE approach is proposed that highlights the knowledge-supported re-formulation of a changeable initial optimization problem. It is applicable to the conceptual design optimization of large-scale buildings (e.g. indoor sports buildings) involving multi-disciplinary criteria and complex geometries.

The traditional CDO approach (Fig. 3, left) often considers a fixed initial OPF which is assumed to be good and hence directly pushed into the problem-solving. In contrast, the proposed CDE approach (Fig. 3, middle) introduces a changeable initial OPF and inserts a CDE module (i.e. shaded components), aiming to ensure a good OPF prior to the problem-solving. The proposed overall approach can be implemented in different ways, depending on the decisions made during the CDE process. One possible way identified by solid arrows (Fig. 3, middle) is the focus of this paper. It emphases the screening of existing variables during the re-formulation process in order to narrow down design possibilities; thus, it is particularly suitable for the case with relatively large number of variables (like the case we are interested in this paper). With a focus on the way identified, the proposed overall workflow including the sub-phases of the first-level CDE and of the CDO (Fig. 3, right) is described below.

4.1.1. Initial formulation of a changeable optimization problem

The workflow starts with the initial formulation of a changeable design space and objective space (and the creation of parametric models and simulation models) based on the a priori knowledge of designers.

For defining the changeable design space, a hierarchical structure of input variables can be used. In this structure, there are three types of input variables, i.e. high-level (Type I), normal (Type II), and low-level (Type III) variables. The high-level variable determines the selection of low-level variables. In this way, the dimensionality of the design space being investigated is changeable. As an example, for the case in Section 5, the high-level variable "RoofSteps" determines the selection of low-level variables in the "Ridge Division" and "Front Row Division". The higher value of the "RoofSteps", the more dimensions of the design space, and vice versa.

For defining the initial objective space, initial objective variables which constitute the dimensions of the space need to be specified. They can be derived from many performance criteria, given the complexity of multi-objective and multi-disciplinary design tasks. That is, the most relevant criteria, the designers believe, are treated as initial objective variables; while, the remaining criteria are treated as initial constraints. And, the initial objective variables and constraints can be converted to each other, once more information becomes available. Thus, the dimensions of the initial objective space are also subject to change. In addition, it is worth noting that the initial objective variables are also called "candidate" objective variables, as final objective variables which constitute a final objective space can be selected from them; and that the number of the candidate objective variables can be larger than three, while the number of the final objective variables are often at most three (as mentioned in Section 2.2).

4.1.2. First-level CDE: information and knowledge extraction

To better understand the performances of the initial optimization problem, the designers can perform a parametric study and extract useful information and knowledge in the first-level CDE. The parametric study involves three sub-phases (Fig. 3, right): (1) specifying proper domains of the input variables and a design of experiments (DoE) sampling strategy; (2) automating the geometry generation, simulation run and data collection for each design sample; and (3) analysing the data collected and extracting useful information and knowledge. Note that, if the data obtained in the second sub-phase is not satisfying (e.g. too many unfeasible designs), the domains of input variables could be adjusted, thus the whole parametric study procedure is restarted. And, in the third sub-phase, correlation analysis (using Pearson Correlation [50]) is applied to identify the output-output relations, thus the most necessary objective variables; cluster analysis (using Hierarchical Clustering [51]) and sensitivity analysis (using Smoothing Spline ANOVA [52]) are applied to identify the input-output relations, thus the most promising clusters of alternatives and the most important design variables. Once the relevant information and knowledge are extracted, the designers can apply them to the re-formulation of initial optimization problem; or, they can also enter CDO directly, if the information and knowledge extracted indicate that the initial optimization problem has already well formulated. In this paper, we are interested in the former scenario.

4.1.3. Second-level CDE: re-formulation of the initial optimization problem

Assisted by the extracted information and knowledge, the designers can re-formulate the initial optimization problem in a more informed manner in the second-level CDE. During the re-formulation, new variables (that are not included in the initial OPF) can be conceived and added for increasing design creativity; and existing variables (that are included in the initial OPF) can be screened out for reducing computational cost. In this paper, we are interested in the later scenario as mentioned before. Thus, based on the results of the previous statistical analyses, unnecessary objective variables identified are treated as constraints; non-promising clusters of alternatives identified are left out from further consideration; and unimportant design variables identified are treated as constants. Note that, a prioritization or balance between quantitative and qualitative goals may be needed for the identification of promising clusters; and selecting the initial generation from the promising clusters can be helpful for obtaining desired solutions. Once the initial optimization problem is re-formulated and the initial generation is selected, the designers can conduct CDO without entering a next CDE iteration (as there are no new variables to be added and no new associated knowledge to be investigated in the case we are concerned with).

4.1.4. CDO: solving of the re-formulated optimization problem

In the CDO, the re-formulated optimization problem is solved using simulation-based optimization [3]. The optimization problem solving follows a similar workflow as the first-level CDE (Fig. 3, right). The major differences are that the CDO workflow implements optimization algorithms (rather than systematic evaluations); and that the information and knowledge extracted from the CDO are mainly about optimal solutions (rather than design samples including sub-optimal solutions). This paper compares the results of several re-formulated optimization problems, aiming to verify the benefits of the proposed approach over the traditional approach, and understand other factors that may affect the behaviour of the proposed approach.

4.2. Computational platform

4.2.1. Software selection

The computational platform (Fig. 4) used in this research was previously developed based on a collaboration between TUDelft and ESTECO. It is meant to facilitate designers in achieving the proposed approach. To form the platform, two software (and related plug-ins) are selected considering their appealing features. They are McNeel's Rhinoceros (Rhino) [53] and Grasshopper (GH) [54] for parametric modeling and ESTECO's modeFRONTIER (MF) [55] for design optimization and exploration.

Grasshopper (integrated with Rhinoceros) is one of the most popular parametric modeling environments among architectural design professionals, given its intuitive way of exploring complex geometries. It includes various plug-ins for integrating building performance simulation engines, such as, the environmental analysis plug-ins (e.g. Ladybug and Honeybee [56]) for integrating daylight and energy simulation engines, and the structural analysis plug-ins (e.g. Karamba [57]) for conducting finite element calculations. Daysim [58] is a daylight simulation engine developed based on the concept of daylight coefficients [59] and the Perez sky luminance model [60], and validated by Reinhart and Walkenhorst [61]. It can simulate indoor illuminance under arbitrary sky condition. EnergyPlus [62] is a widely used energy simulation engine developed by the U.S. Department of Energy, which can model the energy consumption for heating, cooling, ventilation, lighting and equipment loads etc.

modeFRONTIER is a process integration and automation platform for multi-objective and multi-disciplinary design optimization. It allows the integration with a variety of third party CAD/CAE tools; provides efficient DoE sampling strategies and optimization algorithms; supports parallel computing; and especially offers a number of easy-to-use data analysis tools and user-friendly interfaces.

4.2.2. GH-MF integration

The version of the platform used in this research is an improved one. As compared to previous versions, the new platform allows a direct communication between the two software, i.e. a direct GH-MF integration. In fact, the development of the GH-MF integration went through several stages where the communication between GH and MF was improved or simplified. In previous studies [63–65], the necessity and potentials of the GH-MF integration in supporting the CDO and CDE processes were shown, assisted by a prototype plugin (though it involves manual initiation and unstable behaviour); and later in [66], a new GH-MF integration plugin was developed which overcomes some limitations of the prototype and enables indirect communication between GH and MF via external files. In this paper, an improved version of the precedent is used to better support the proposed approach. It allows MF to directly detect or interact with the input and output variables in GH via an API, without the need to specify external file templates by GH (as required in the previous version). Thus, it much simplifies the preparation of the GH file.

The direct GH-MF integration is key to the computational platform, which facilitates designers in achieving the proposed approach. Fig. 4

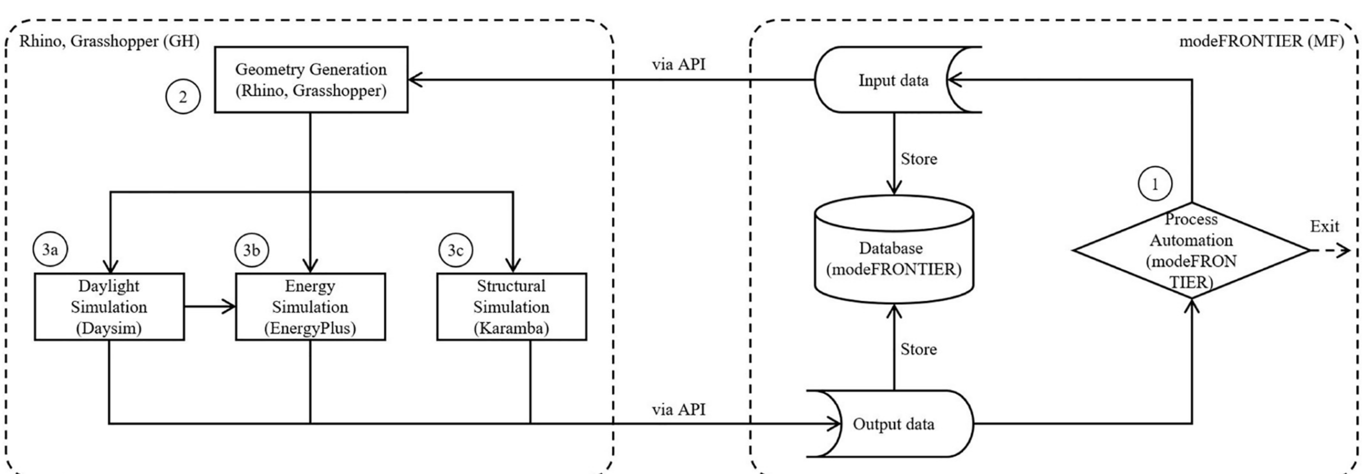


Fig. 4. The computational platform integrating Grasshopper and modeFRONTIER.

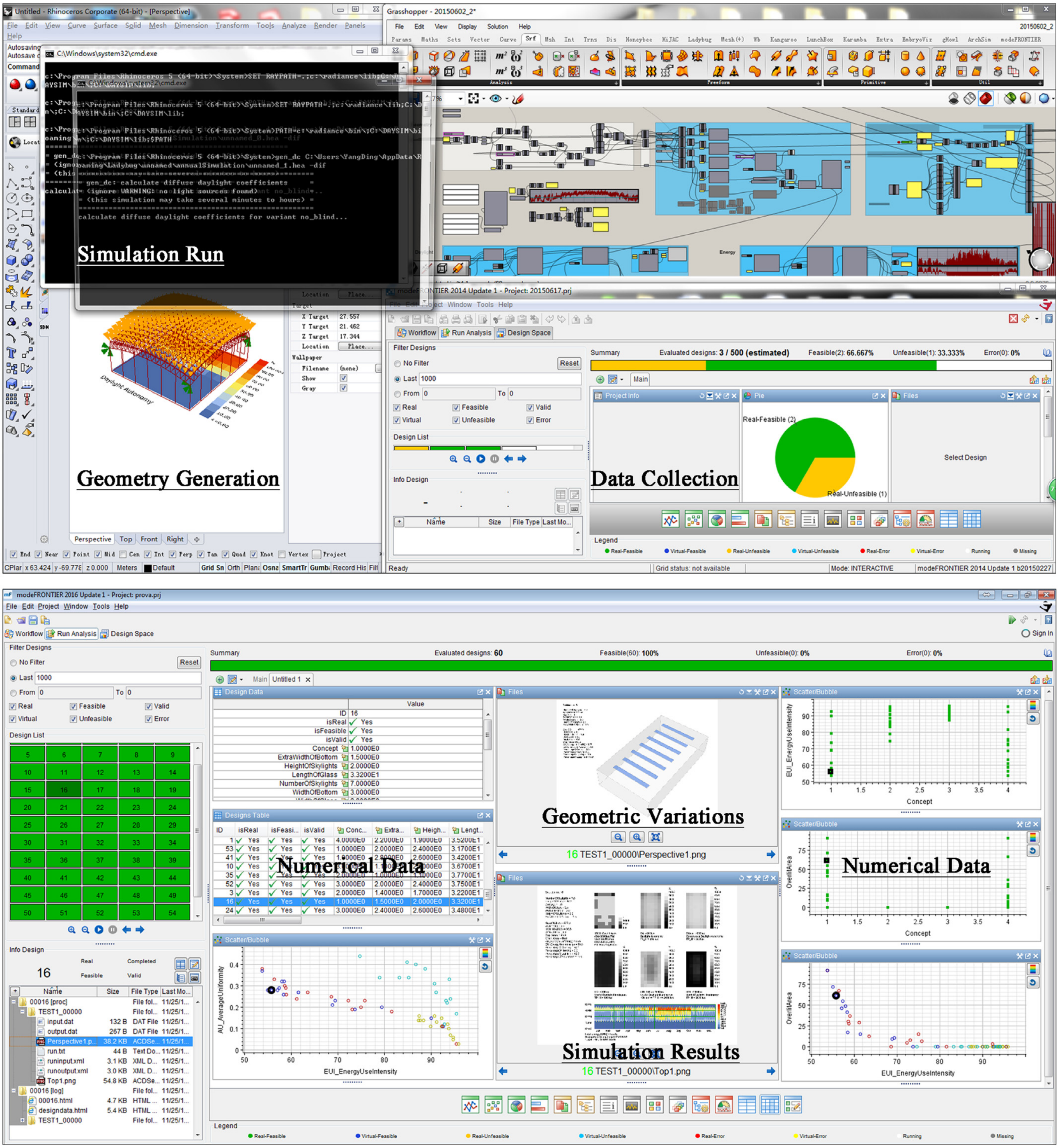


Fig. 5. Interface for process automation, i.e. automating the iterative processes of geometry generation, simulation run and data collection (top); interface for post-processing, i.e. monitoring geometric variations while exploring numerical data and other simulation results (bottom).

illustrates how the GH-MF integration assists the automated iterative processes of geometry generation, simulation run and data collection; Fig. 5 (top) shows the process automation interface. In this automation, modeFRONTIER is the driver which initiates the input data of design samples and determines when to stop; Grasshopper generates geometries and the related simulation engines calculate performances to obtain the output data. All the data (including numerical data, images of geometries and simulation results) are stored in the database for further analysis. Moreover, the GH-MF integration leverages other advantages of both GH and MF. For instance, the modeling and simulation tools help to create parametric simulation models featuring a changeable design space; the efficient DoE sampling strategies facilitate to generate proper sets of design samples; the easy-to-use post-processing tools (like correlation analysis, cluster analysis, sensitivity analysis and various data visualization charts) help to extract relevant information and knowledge; and the user-friendly interface, shown in Fig. 5 (bottom),

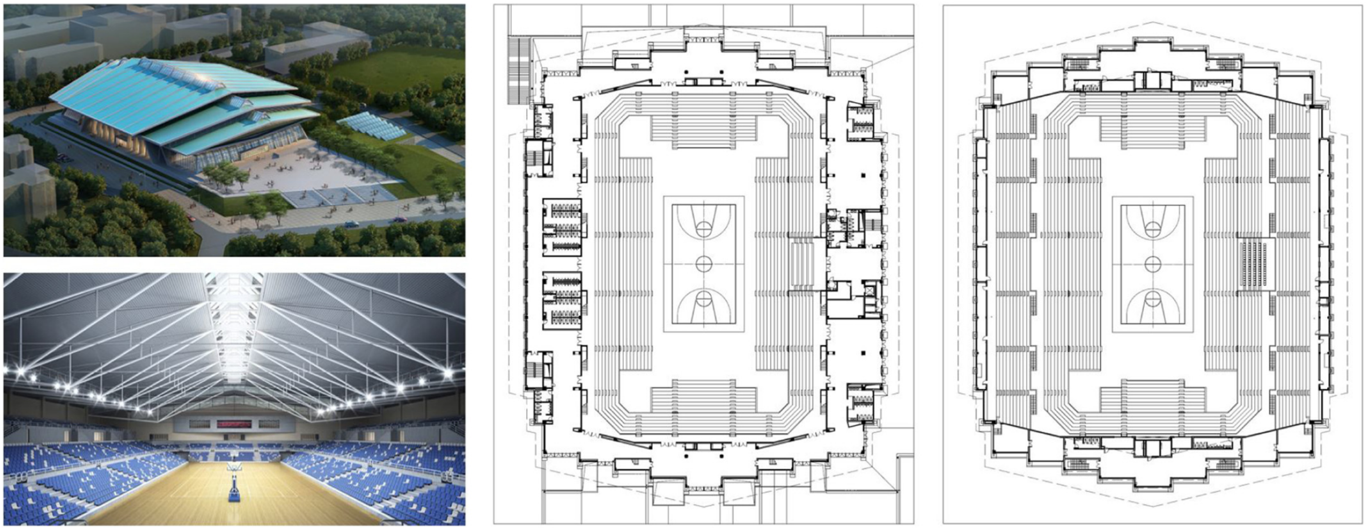


Fig. 6. Exterior and interior perspectives of the indoor sports building project in Wuhan University in China (left); first level plan (middle); second level plan (right).

allows to monitor geometric variations while exploring numerical data and other simulation results, which is beneficial for the balance between quantitative and qualitative goals etc.

5. Case study

In this section, the proposed approach is applied to a real-world project as a case study, assisted by the GH-MF platform. The background of the project is introduced in Section 5.1; and the formulation of the initial and changeable optimization problem is presented in Sections 5.2 and 5.3.

5.1. Project description

The project is an indoor sports building in Wuhan University in China (Fig. 6), designed by Sun Yimin Studio of the Architectural Design and Research Institute of South China University of Technology. The site is located in a historic district and in a subtropical climate zone. Given this context, the design team decided to respect the forms of Chinese traditional buildings and to take advantages of natural daylight; thus, a stair-like roof concept with clearstories was proposed at the very beginning of the conceptual design. The designers argued that proper architecture-, climate- and structure-related performances can be achieved by carefully designing the geometries associated with the concept. Several questions arose with this concept:

(1) What are the most meaningful performance criteria to be considered as final objectives?
(2) How many “steps” of the roof are more promising to ensure desired overall performance?
(3) Would the layout of the grandstand, the height of the volume, the division of the roof, the depth of overhang shadings, and the geometry of the roof structure affect the overall performance? If so, how and to what extent they may affect?
(4) How to obtain quantitatively well-performing solutions with desired qualitative performance?

To respond to these questions, the design team had to use rules of thumb, as this study began to be involved in the project when the concept had been finalized and was ready for the design development [67]. Nevertheless, it is worth investigating the potential improvement of the concept by applying the proposed approach, thus, the project is assumed being in the conceptual design phase. Moreover, this study focuses on the main competition hall where the main court and seating are accommodated (excluding auxiliary space like training hall, entrance hall, management offices etc.), and it avoids to use electric skylights which may be expensive and difficult to maintain on the rooftop.

5.2. Design variables and domains

There are a wide range of geometric variations, although focusing on one given concept (Fig. 7, top). The initial design variables and domains are listed in Table 1, based on which the corresponding parametric model is created (Fig. 7, bottom). All the design variables are geometry related and independent. They are grouped into nine families and parametrically define four integrated parts of the building, i.e. the geometry of the grandstand, building envelope, external shading and roof structure. Given the symmetry of the main competition hall along X and Y axes, a quarter of it is defined. Moreover, the benchmark configuration is also shown in Table 1, which is most like that of the real-world project.

For the geometric parameterization of the grandstand, a computational tool based on shape grammar [68] is used. Out of many possible design variables for the grandstand design, the number of maximum seat rows of the upper tier (i.e. the variable “SeatRows”) is selected. It determines how far the upper tier (and the eaves of the sub-roofs) extends horizontally and vertically. The total extension (in both horizontal and vertical directions) is equally divided by the number of sub-roofs, forming the jagged, saw-toothed outline of the upper tier (and of the sub-roofs).

For the geometric parameterization of the stair-like sub-roofs, the length of the ridge of each sub-roof (denoted by L_{ri}) and the length of the upper tier front row under each sub-roof (denoted by L_{fi}) are determined by the variables in the “Ridge Division” and “Front Row Division” and the variable “RoofSteps”, according to Eqs. (1) and (2). And, the elevations of the ridges of the sub-roofs are calculated using the variables in the “Roof Height”, given the equal vertical distance between each ridge. All these variables (including “SeatRows”) define the geometries of the sub-roofs. Their bounds and intervals have been fine-tuned, to ensure rich variability of the roof, while avoiding a too small division in each end of the sub-roofs or a too small vertical distance between the sub-roofs (for clearstories).

$$L_{ri} = \frac{L_r * R_i}{\sum_{j=1}^{n} R_j} \tag{1}$$

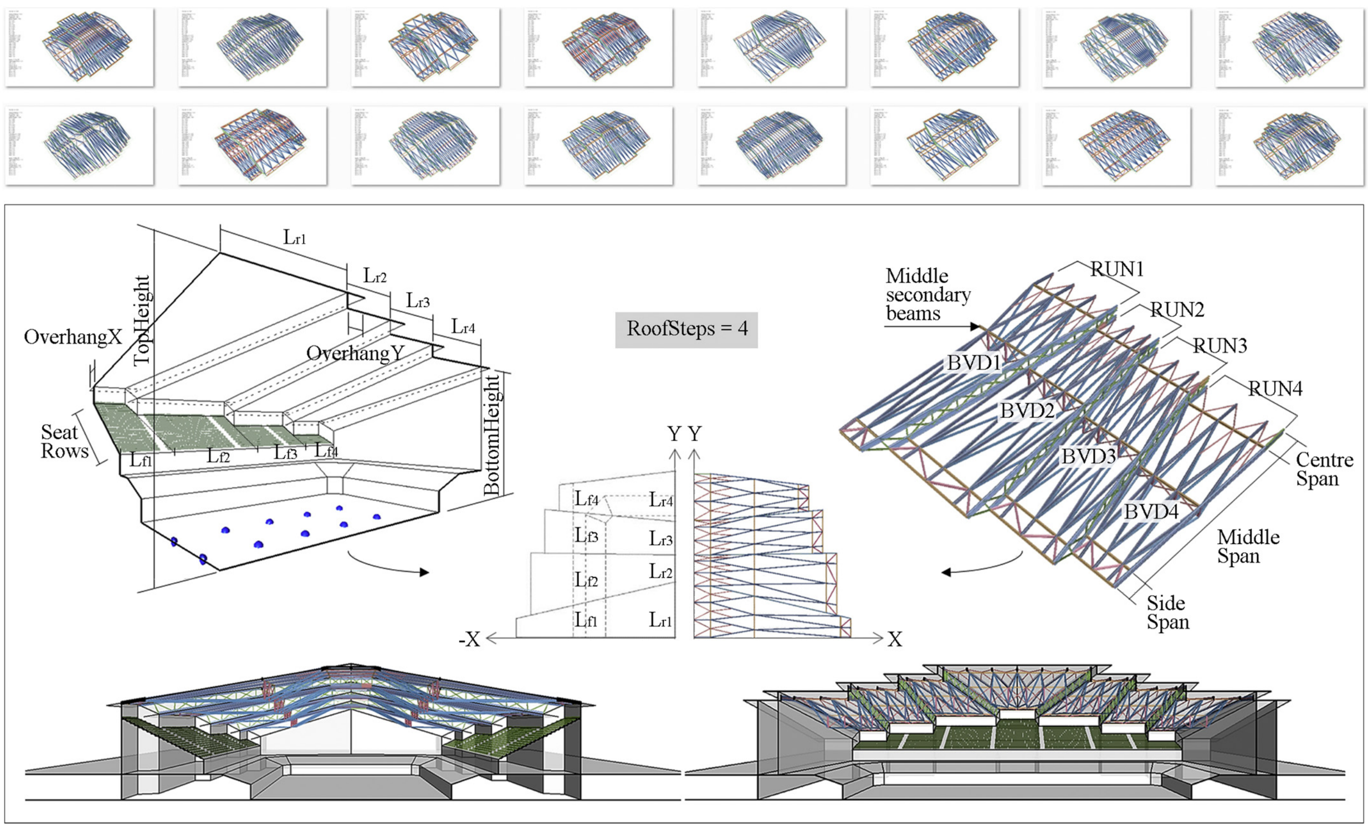


Fig. 7. Geometric variations based on the stair-like roof concept (top); the parametric model integrating the grandstand, building envelope, external shading and roof structure, when the "RoofSteps" equals to four (bottom).

$$L_{fi} = \frac{L_f * F_i}{\sum_{j=1}^{n} F_j} \tag{2}$$

where

L_r = half-length of the entire ridge (constant value: 48 m).

L_f = half-length of the entire upper tier front row (constant value: 44.5 m).

R_i or R_j = the portion of the ridge of sub-roof *i* or *j* (i.e. R1 to R5).

F_i or F_j = the portion of the upper tier front row under sub-roof *i* or *j* (i.e. F1 to F5).

n = number of roof steps (i.e. RoofSteps).

The hierarchical structure of design variables exists due to the inclusion of the Type I variable "RoofSteps" (Table 1, in dark gray). The value of this variable determines the selection of the Type III variables in the "Ridge Division" and "Front Row Division" (Table 1, in light gray). In this way, the geometric complexity of the roof can be readily changed, which can expand or shrink the design space and affect simulation time.

For the geometric parameterization of the roof structure, the half span is divided into three parts along X axis according to the structural principle. The horizontal length of each part can be changed by manipulating the variable "CentreSpan" or "SideSpan". The variable "MiddleSpan" defines the position of middle secondary beams (the lower bound represents the position closest to the centre and vice versa). Note that the "CentreSpan" and "SideSpan" are used identically for all sub-roofs; while the "MiddleSpan" is for the lowest sub-roof only (the middle secondary beams for the higher sub-roofs are forced to align with that for the lowest sub-roof on XY plane). The vertical distance between beams (i.e. BVD1 to BVD5) and the number of repeated units (i.e. RUN1 to RUN5) are defined separately for each sub-roof. The bounds and intervals of all the variables are set based on certain structural rules of thumb to avoid unfeasible design solutions. Note that the width of each repeated unit is constrained in between 3 m and 9 m, which prevents that a too large or too small repeated unit number is selected.

The Type I variable "RoofSteps" affects the geometric parameterization of the roof structure, as the main structural members are generated by taking sub-roof surfaces as reference. When the number of sub-roofs changes, the structural members attached to the sub-roofs can be added or removed. Accordingly, different sets of the Type III variables in the "Beam Vertical Distance" and "Repeated Unit Number" may be selected.

For the geometric parameterization of the external shading, the sub-roof surfaces are also taken as reference and extended outwards. Nevertheless, the variables in the "Shading Dimension" are not affected by the Type I variable "RoofSteps", as they are used identically for all the sub-roofs.

5.3. Objective functions and constraints

Multiple performance criteria from different building disciplines are considered simultaneously in this case. The initial objective functions and constraints are listed in Table 2, based on which the corresponding simulation models are created. Note that the criteria ensured by the parametric or simulation models are also listed. The disciplines involved include architecture, climate (daylight, thermal and energy) and structure. For each discipline, the corresponding simulation setup and the performance criteria considered are described below.

5.3.1. Architecture

For the conceptual architectural design of indoor sports buildings, both quantitative (or hard) and qualitative (or soft) criteria need to be considered. The quantitative criteria can be fulfilled via calculating the corresponding performances and comparing them with desired performance constraints; they can also be achieved by setting desired

Table 1
Initial design variables and domains (Type I, II and III variables are marked by dark, medium and light gray respectively).

	Variable family	Variable full name	Variable short name	Data type	Lower bound	Upper bound	Intervals	Benchmark
Grandstand	Seat row number	Number of maximum seat rows (of the upper tier)	SeatRows	Int.	15 (19)	20 (24)	1	11
Building envelope	Roof step number	Number of roof steps	RoofSteps	Int.	2	5	1	2
	Roof height	Height of the highest ridge (m)	TopHeight	Float	25.00 (27.00)	30.00 (32.00)	0.01	26.00
		Height of the lowest ridge (m)	BottomHeight	Float	15.00 (17.00)	20.00 (22.00)	0.01	24.00
	Ridge division	Portion of the ridge of sub-roof 1	R1	Float	0.20	0.90	0.01	0.9
		Portion of the ridge of sub-roof 2	R2	Float	0.20	0.90	0.01	0.2
		Portion of the ridge of sub-roof 3	R3	Float	0.20	0.90	0.01	–
		Portion of the ridge of sub-roof 4	R4	Float	0.20	0.90	0.01	–
		Portion of the ridge of sub-roof 5	R5	Float	0.20	0.90	0.01	–
	Front row division	Portion of the front row under sub-roof 1	F1	Float	0.20	0.90	0.01	0.9
		Portion of the front row under sub-roof 2	F2	Float	0.20	0.90	0.01	0.2
		Portion of the front row under sub-roof 3	F3	Float	0.20	0.90	0.01	–
		Portion of the front row under sub-roof 4	F4	Float	0.20	0.90	0.01	–
		Portion of the front row under sub-roof 5	F5	Float	0.20	0.90	0.01	–
External shading	Shading dimension	Overhang depth in X axis (m)	OverhangX	Float	0.10	3.00	0.01	3.80
		Overhang depth in Y axis (m)	OverhangY	Float	0.10	3.00	0.01	2.20
Roof structure	Span partition	Centre Span (m)	CentreSpan	Float	0.50	5.00	0.01	4.20
		Middle Span Partition (fraction)	MiddleSpan	Float	0.10	0.90	0.01	0.50
		Side Span (m)	SideSpan	Float	0.50	5.00	0.01	4.20
	Beam vertical distance	Beam vertical distance for sub-roof 1 (m)	BVD1	Float	2.00	7.00 (6.00)	0.01	4.60
		Beam vertical distance for sub-roof 2 (m)	BVD2	Float	2.00	7.00 (6.00)	0.01	2.00
		Beam vertical distance for sub-roof 3 (m)	BVD3	Float	2.00	7.00 (6.00)	0.01	–
		Beam vertical distance for sub-roof 4 (m)	BVD4	Float	2.00	7.00 (6.00)	0.01	–
		Beam vertical distance for sub-roof 5 (m)	BVD5	Float	2.00	7.00 (6.00)	0.01	–
	Repeated unit number	Repeated unit number for sub-roof 1	RUN1	Int.	1	5	1	5
		Repeated unit number for sub-roof 2	RUN2	Int.	1	5	1	1
		Repeated unit number for sub-roof 3	RUN3	Int.	1	5	1	–
		Repeated unit number for sub-roof 4	RUN4	Int.	1	5	1	–
		Repeated unit number for sub-roof 5	RUN5	Int.	1	5	1	–

performance constraints in parametric or simulation models. And, the fulfilment of the qualitative criteria (e.g. culture, beauty, emotions, etc.) often involves the integration of subjective human preferences.

To ensure the basic function of indoor sports buildings, some important quantitative criteria regarding the grandstand design and playing area need to be fulfilled in this case. For the grandstand design, C-value is a crucial criterion to ensure the view quality of spectators. Assisted by the parametric grandstand design tool [68], a C-value of 60 mm is set as an input to generate various variations of grandstands in this case. Thus, this criterion is always fulfilled during the generation process. And, the number of seats (i.e. the capacity) of the grandstand is calculated by the tool as one of the important performance feedback.

Table 2
Objective functions and constraints.

Disciplines		Performance criteria	Objective functions	Constraints (to be calculated)	Constraints (set in models)
Architecture		C-value	–	–	60 mm
		Number of seats in the upper tier	–	> 3600	–
		Minimum space check (SC)	–	> 15 m	–
Climate	Daylight	Modified Useful Daylight Illuminance (UDI_{mod})	Maximization	–	–
		Modified Uniformity Ratio (UR_{mod})	Maximization	–	–
	Thermal	Operative temperature	–	–	See Table 3
	Energy	Energy Use Intensity (EUI)	Minimization	–	–
Structure		Mass per square meter	Minimization	–	–
		Maximum utility check (UC)	–	< 0.9 (failed members < 2% of the total)	–
		Maximum displacement check (DC)	–	< 0.3 m	–

Given that the capacity of the lower tier is unchangeable, a constraint value is used for the upper tier to include > 3600 seats. Moreover, to ensure the clear space for holding various sports activities, the clear height above the court (i.e. from the bottom of the lowest structural elements to the court) is checked against the minimum requirement of 15 m.

To achieve the qualitative criteria, the subjective human preferences are considered in this case. Specifically, the preference on the number of roof steps from the aesthetics perspective will be integrated during the decision of the Type I variable "RoofSteps". This leaves sufficient flexibility for human designers to re-formulate optimization problems subjectively.

5.3.2. Climate (daylight, thermal and energy)

For climate design, reducing operational energy as much as possible while improving or maintaining daylight and thermal comfort are common concerns crucial for achieving the low-cost daily operation of indoor sports buildings. To obtain the climate-related performance feedback, annual hourly daylight and energy simulations are performed by Daysim [58] and EnergyPlus [62] respectively, via Ladybug and Honeybee [56] which connects GH geometries with the simulation engines.

To setup these simulations for this case, it requires some shared settings, including the same weather file of Wuhan derived from Chinese Standard Weather Data (CSWD) [69], the same geometry of the competition hall that has been converted to meshes properly to save simulation time, and the same occupancy schedule (Fig. 8) that considers both educational or recreational use (i.e. off-peak use from Monday to Saturday) and competition use (i.e. peak use on Sunday). Besides, for the daylight simulation, two lighting control zones are defined based on an analysis grid with a spacing of 6 m (Fig. 9, left), which cover the central and surrounding areas of the court respectively. In these zones where different activities may occur, different lighting control types and lighting power densities are assigned. And, for the energy simulation, EnergyPlus's Ideal Loads Air System is used to study the energy performance of the building without modeling a full HVAC system. It can be considered as an ideal unit that mixes air, and then adds or removes heat and moisture at 100% efficiency [70]. The boundary condition is set (Fig. 9, right), and the setpoint and setback temperatures are assigned. Moreover, other assumptions, such as optical and thermal material properties etc., are also made for the daylight and energy models based on related building codes or rules of thumb, as listed in Table 3. By running the simulations, the internal illuminances are calculated at each test point for 8760 h of a year; and the heating and cooling loads are calculated and scaled according to a generic heating system efficiency of 0.85 and a cooling system COP of 3.

Note that Daysim is utilized here in conjunction with EnergyPlus for different reasons. First, EnergyPlus has shown a significant limitation in calculating internal illuminances, as it tends to overestimate the amount of daylight in indoor environments [77]. Second, Daysim can model automated lighting control and provide hourly lighting schedules of different lighting zones for EnergyPlus to calculate the final energy use [24,78]. The lighting schedule can describe the control of a continuous dimming lighting system. It consists of a list of lighting power scalars (denoted by L) calculated depending on the availability of daylight, according to Eq. (3). In this case, a list of values ranging from 0.2 to 1 are obtained as the lighting schedule for each lighting zone.

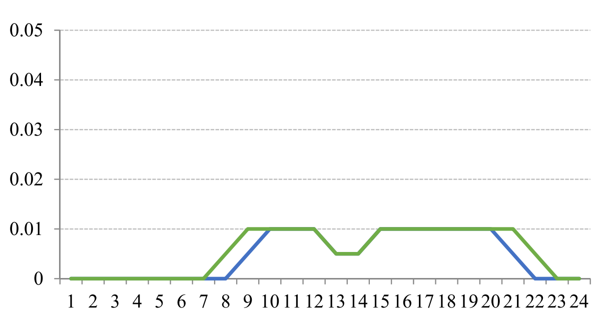
$$L = \begin{cases} BLF & \text{if } E_{min} \geq LS \\ BLF + (1 - BLF)*\left(1 - \frac{E_{min}}{LS}\right) & \text{if } E_{min} < LS \end{cases} \tag{3}$$

where

L = lighting power scalar.

BLF (*Ballast Loss Factor*) = percentage of peak energy used by a dimming system when fully dimmed down (constant value: 20%).

LS (*Lighting Setpoint*) = illuminance target (constant values: 300 lx



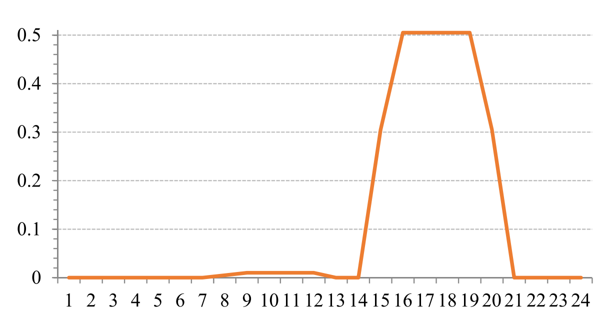


Fig. 8. Off-peak use without spectators from Mon. to Fri. (blue) and on Sat. (green); peak use with spectators on Sun. (Orange). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

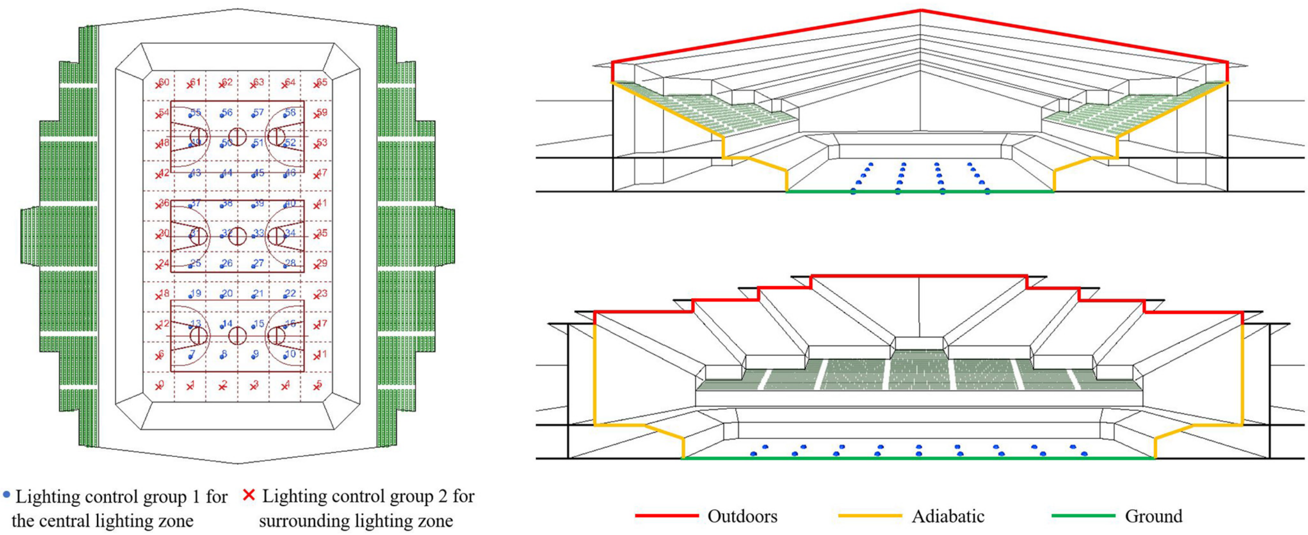


Fig. 9. Two lighting control zones of the court for daylight simulation (left); boundary conditions of the main competition hall for energy simulation (right). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Modeling assumptions for the daylight and energy models. Values from [71–76].

Daylight and energy model parameter	Value
Wall reflectance	0.55
Floor reflectance	0.30
Roof reflectance	0.75
Window Transmittance	0.40
Lighting control type (lighting control group 1)	Always on during occupied hours, automatic dimming, 300-lx target
Lighting control type (lighting control group 2)	Always on during occupied hours, automatic dimming, 200-lx target
Lighting power density (lighting control group 1)	15.00 W/m^2
Lighting power density (lighting control group 2)	9.00 W/m^2
Wall U-value	0.72 W/m^2 K
Ground floor U-value	3.70 W/m^2 K
Roof U-value	0.34 W/m^2 K
Window U-value	2.60 W/m^2 K
Window SHGC	0.37
Window VT	0.62
Cooling thermostat setpoint temperature	27 °C
Cooling thermostat setback temperature	30 °C
Heating thermostat setpoint temperature	17 °C
Heating thermostat setback temperature	14 °C
Occupancy density	0.92 person/m^2
Equipment power density	2 W/m^2
Ventilation rate	15 m^3/h person
Infiltration rate	4.5 m^3/h m^2

and 200 lx for lighting control group 1 and 2 respectively).

E_{min} = minimun daylight illuminance in the current lighting zone.

A modified Useful Daylight Illuminance (UDI$_{mod}$), or called spatial UDI, is used as a maximization objective function in this case. Daylight can be beneficial to sports halls if properly designed and well controlled. To measure the availability of daylight, the original Useful Daylight Illuminance (UDI) is often considered useful as it also attempts to incorporate factors related to overheating or glare risk. It is calculated based on the internal illuminances at all time steps (but at a specific analysis point); and, it represents the annual occurrence of the "useful" daylight illuminances that fall within the range of 100–2000 lx [79], that is, the percentage of time during occupied hours that an analysis point receives hourly illuminances between 100 and 2000 lx. Given that the original UDI is defined based on one specific analysis point, it is not sufficient to understand the overall daylighting condition of a large space (e.g. the 40 m ∗ 70 m court in question where 66 analysis points are needed). Thus, to consider multiple analysis points, UDI$_{mod}$ is defined as the percentage of floor area (represented by the percent of analysis points) that receives the "useful" illuminances (i.e. 100–2000 lx) for at least a specified percentage of occupied hours. Note that this percentage of occupied hours is initially set to 60% in this case (i.e. UDI$_{mod\text{-}60}$), which can be changed when needed. By using the UDI$_{mod}$, the overall daylighting condition (i.e. daylight availability) of the entire court is described by a single value, which facilitates the use of the criterion in optimization.

A modified Uniformity Ratio (UR$_{mod}$) of illumination, is also used as a maximization objective function in this case. The uniformity of illumination is crucial for sports as non-uniform conditions can make it more difficult to perceive fast-moving balls. To measure the uniformity, the original Uniformity Ratio (UR) is often used. It is calculated based on the internal illuminances at all analysis points (but at a specific time step) and expressed as the ratio of minimum to mean horizontal illuminance. To consider the daylight uniformity on an annual basis, the UR$_{mod}$ averages the UR values obtained at different time steps of a year (when daylight is available).

Energy Use Intensity (EUI), as a basic metric to benchmark a building's energy efficiency, is often used as a minimization objective function. It is defined as the annual energy consumption per unit of floor area (kWh/m^2), facilitating direct comparison with other buildings. Here, site energy for heating, cooling, lighting and equipment is considered. Moreover, operative temperature, as a simplified measure of human thermal comfort, is ensured by the settings of setpoint and setback temperatures in this case (See Table 3).

5.3.3. Structure

For large-span structural design, reducing the weight (or embodied energy) of the roof structure as much as possible while fulfilling other

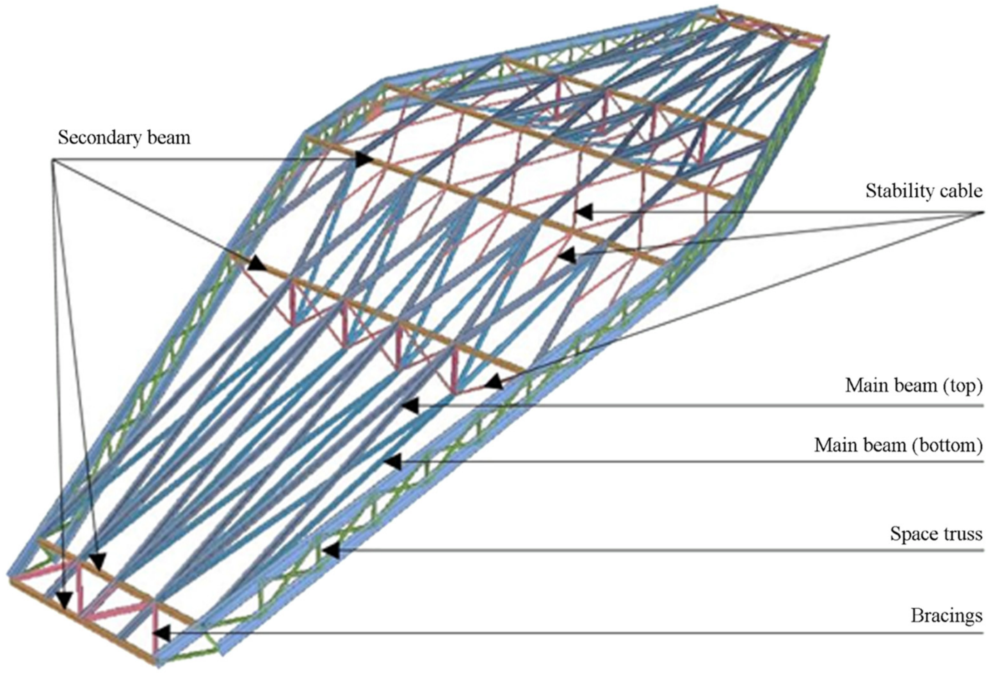


Fig. 10. Typical structure for one sub-roof.

structural performance constraints are traditional concerns crucial for reducing the initial investment of indoor sports buildings. To obtain the structure-related performance feedback, finite element analysis (FEA) is conducted by Karamba [57].

In this case, the large-span steel roof is the main load-bearing structure, which approximately spans over 91.6 m between the farthest supports (depending on the extension of the upper tie). Fig. 10 shows the typical structure for one sub-roof. The structural system features a one-way span steel frame in diamond patterns in two layers. Steel cables are applied in the lateral direction providing lateral stability at multiple locations. A space truss is used at the step area where two sub-roof surfaces at different elevations interface with each other. According to the structural function and practical engineering considerations, the structural elements are grouped into different types, as shown by the color coding.

For the FEA simulation, mechanical properties of structural elements and related analysis parameters are set up. The definition of sections is based on the element groups. Each group will be only assigned with one identical section to simplify the connection design and to reduce the number of different joints. For this, a list of standard steel section profiles (including HE beam and rod section profiles) is prepared for each group, from which the optimum section will be selected according to EN1993 by a local optimization module in Karamba; S355 steel is used as the steel grade for all the section profiles. In addition, buckling information is also defined based on the element groups. In order to increase the computational speed and simplify the model, some unimportant secondary members connecting the main beams are not modelled directly in the buckling calculation. To account for this, a reduction factor on the buckling length is introduced for the lateral torsional buckling and the minor axis bending buckling calculation of the main beams. Moreover, the single load case and load combinations are defined based on Eurocode, including the most typical loads such as the structure's self-weight, super imposed dead load, wind load and snow load.

The roof structure is designed to withstand the load combinations and satisfy the Ultimate Limit State (ULS) and Service Limit State (SLS) criteria. The corresponding performance constraints are applied to both criteria in order to ensure the structural safety. They are the maximum utility limit of 0.9 for ULS strength check and maximum vertical displacement limit of 0.3 m for SLS deformation control. Based on the constraints, the unity check and displacement check are performed respectively for each member. A code checking module in Karamba based on EN1993 is used for the unity check. It has been observed that the system will consider a design as an unfeasible solution even if there is only one member that fails the unity check. In fact, from a practical point of view, a small amount of the structural members that slightly go beyond the maximum utility limit are allowed, as this can be easily solved afterwards (e.g. by strengthening locally). Thus, to increase the total number of feasible solutions and not to miss many potential designs, a tolerance number of 2% (of the total members) has been introduced for the unity check.

6. CDE results

Having defined the initial and changeable optimization problem, two levels of CDE are conducted according to the proposed approach. The CDE results are presented in this section. Section 6.1 briefly describes the CDE workflow establishment and data collection (i.e. the first two sub-phases of the first-level CDE); Section 6.2 presents the results of information and knowledge extraction derived from data analysis (i.e. the last sub-phase of the first-level CDE); and Section 6.3 shows the re-formulated optimization problem based on the information and knowledge extracted (i.e. the second-level CDE).

6.1. CDE workflow establishment and data collection

To establish the CDE workflow (Fig. 11), Uniform Latin Hypercube (ULH) [80] sampling strategy is applied to generate 500 design samples, which guarantees a relatively broad and uniform distribution of samples over each input dimension. Then, the multidisciplinary simulations of each sample are automated sequentially by using the GH-MF integration, and the simulations are run on a 6-core and 12-thread CPU computer for about 51 h. All the numeric input and output data and images showing the geometries and simulation results are stored in a database.

According to the summary of the initial DoE data set (Table 4), unfeasible solutions account for a major portion (i.e. 87.6%) of the 500 design samples. They are mostly due to the violation of architectural

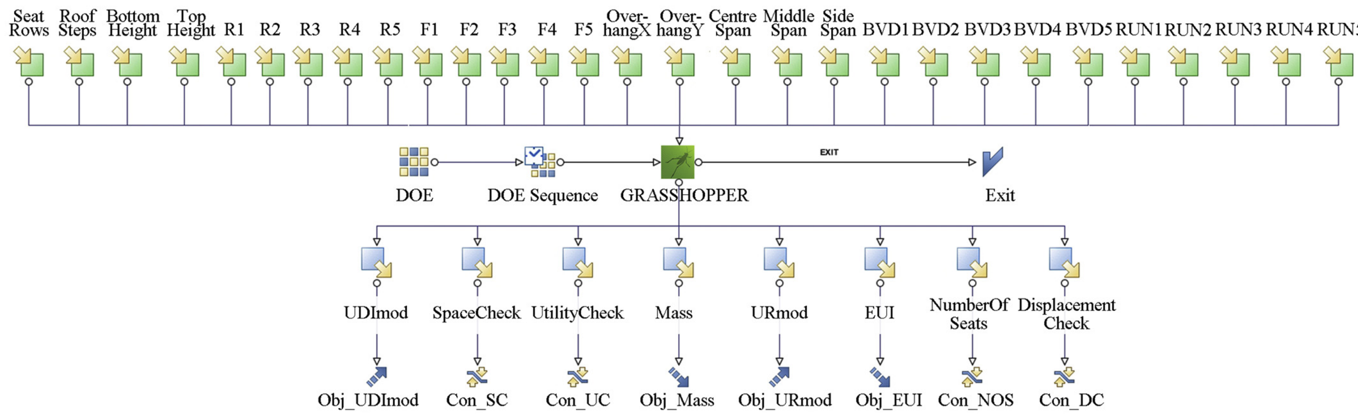


Fig. 11. CDE workflow for the case.

Table 4
Summary of DoE data sets.

	Feasible solutions	Unfeasible solutions	Violation of constraints			
			Con_NOS	Con_SC	Con_UC	Con_DC
Initial DoE data set	62 (12.4%)	438 (87.6%)	261	367	57	21
Second DoE data set	230 (46.0%)	270 (54.0%)	52	169	75	30

constraints, namely, the constraints for the number of seats in the upper tier (i.e. Con_NOS) and the clear height above the court (i.e. Con_SC). Given this fact, it is necessary to increase the portion of feasible solutions by fine-tuning the domains of some design variables. In this case, the domains of SeatRows, TopHeight, BottomHeight and BVD are adjusted (as indicated by the values in parentheses in Table 1), to obtain more feasible solutions that satisfy Con_NOS and Con_SC. Moreover, by observing the distribution of the initial output data, we find that the potential objective variable $\mathrm{UDI}_{\text{mod-60}}$ can be readily achieved, as most of its values are high (Fig. 12, left). This, in fact, indicates that a stricter criterion (or a higher goal) could be expected, thus the specified percentage of occupied hours (for which the "useful" illuminances are achieved) is changed to 65% (i.e. $\mathrm{UDI}_{\text{mod-65}}$).

A second DoE data set (Table 4) is obtained by repeating the previous processes. As expected, in this data set, unfeasible solutions are reduced significantly and the proportions of design samples violating different constraints become more even. Moreover, compared to the $\mathrm{UDI}_{\text{mod-60}}$ values, the $\mathrm{UDI}_{\text{mod-65}}$ values obtained are relatively low (Fig. 12, right), which leaves room for further improvement in later optimization. This data set will be used for further data analysis and knowledge extraction.

6.2. Data analysis and knowledge extraction

Data analysis, information and knowledge extraction are key to the first-level CDE. In this section, complex relationships between variables, including output-output, (Type I) input-output and (Type II & III) input-output relationships, are unveiled by correlation analysis, cluster analysis and sensitivity analysis respectively. Some of the information extracted is familiar to the designers and relatively easy to interpret in disciplinary contexts, while some is not. In the former scenario, the designers' educated guesses based on a prior knowledge can be confirmed quantitatively; while in the latter scenario, knowledge unfamiliar to the designers can be discovered.

6.2.1. Correlation analysis and output-output relationships

Knowing the output-output relationships is helpful for the objective variable screening (i.e. identifying the most meaningful performance criteria to be considered as final objectives). But, it is hard for designers to know the exact correlations between output variables quantitatively based on rules of thumb. What makes it even harder is that there may be many (i.e. more than three) candidate objective variables being considered at the same time, and that the correlations are subjected to

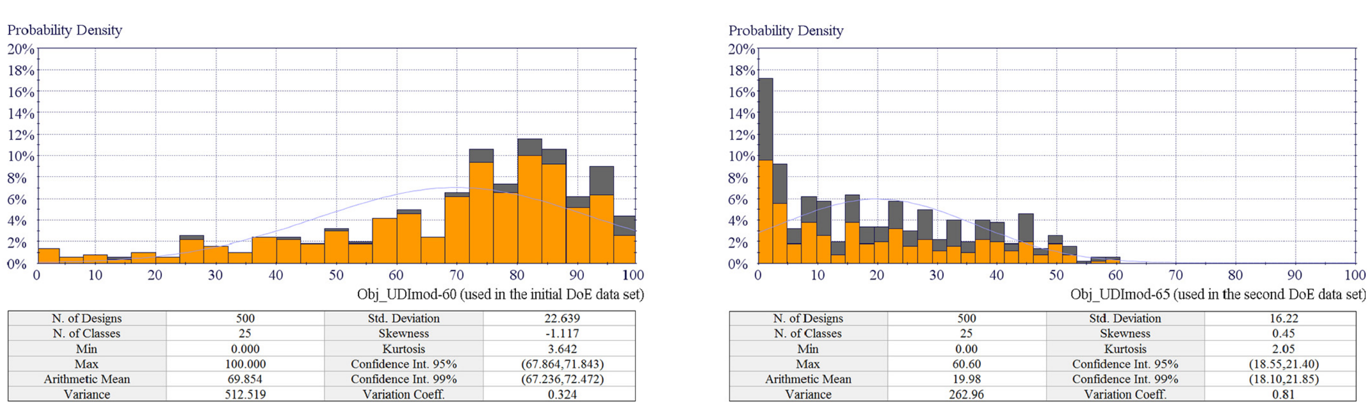


Fig. 12. Distribution of the $\mathrm{UDI}_{\text{mod-60}}$ values (left); distribution of the $\mathrm{UDI}_{\text{mod-65}}$ values (right).

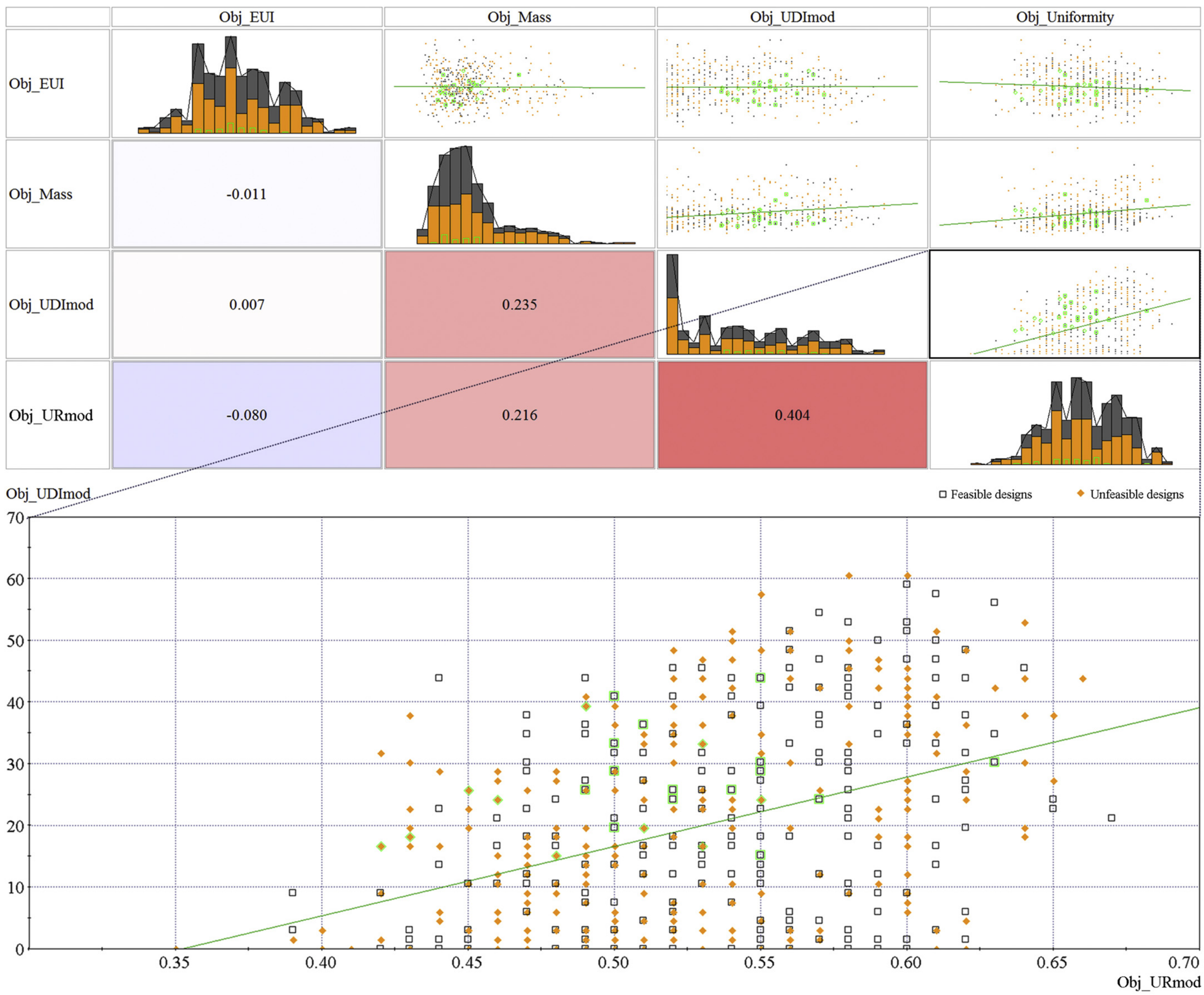


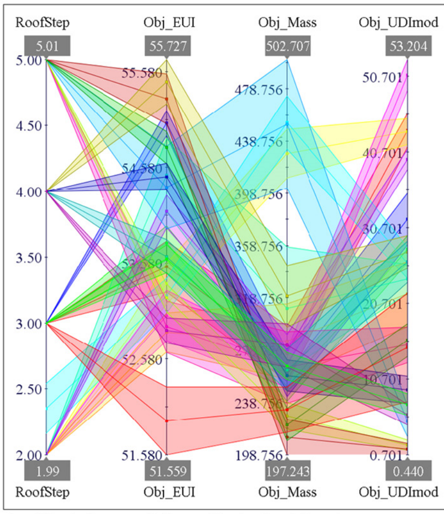
Fig. 13. Correlation matrix chart of the four candidate objective variables (top); 2D scatter plot of Obj_UDI_{mod-65} and Obj_UR_{mod} (bottom).

change due to the factors like the adjustment of input variable domains, the selection of sample points, and the definition of criteria etc.

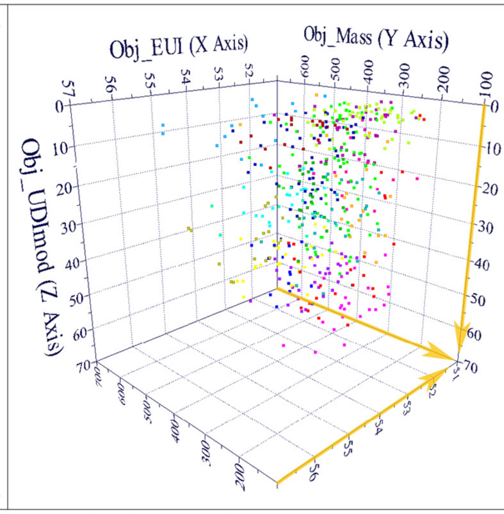
Correlation analysis (using Pearson Correlation [50]) is performed to investigate the output-output relationships in this case. For this analysis, four candidate objective variables (i.e. Obj_EUI, Obj_Mass, Obj_UDI_{mod-65}, Obj_UR_{mod}) are used. The inter-correlations between pairs of the candidate objective variables are visualized by using a correlation matrix chart (Fig. 13, top). In this chart, the correlation coefficients and corresponding 2D scatter plots are shown below and above the diagonal respectively; and the discrete probability density functions and related statistical summary are presented on the diagonal.

The following information is extracted from the correlation analysis results: (1) Obj_EUI does not correlate with the other three variables ($|$correlation coefficient$| < 0.1$); (2) Obj_Mass has a weak correlation with Obj_UDI_{mod-65} and Obj_UR_{mod} ($0.1 < |$correlation coefficient$| < 0.3$); (3) Obj_UDI_{mod-65} has a medium correlation with Obj_UR_{mod} ($0.3 < |$correlation coefficient$| < 0.5$); and (4) the dispersion of UDI_{mod-65} is relatively large in this case, as indicated by the probability density functions and variation coefficients (i.e. the standard deviation divided by the mean) of the four candidate objective variables. Moreover, by zooming into the 2D scatter plot of the most correlated variables (Fig. 13, bottom), it is showed that Obj_UDI_{mod-65} increases along with the increase of Obj_UR_{mod} to certain degree.

The information extracted from the correlation analysis is relatively easy to interpret in disciplinary contexts. First, objective variables from different disciplines can be not correlated or weakly correlated, as they may be the functions of many different design variables. In this case, the climate-related objective variables (i.e. Obj_EUI, Obj_UDI_{mod-65}, Obj_UR_{mod}) are the functions of variables defining the building envelope, while the structure-related objective variable (i.e. Obj_Mass) is calculated based on the variables defining the roof structure (though it also takes roof surfaces of the building envelope as reference). Thus, these two kinds of objective variables are probably not correlated or weakly correlated (as confirmed by the information obtained). Second, some objective variables from the same discipline can be correlated in some circumstances. In this case, the better the UDI_{mod-65} performance the more likely that the court can avoid too low illuminance levels (i.e. lower than 100 lx), which prevents to get uneven distribution of daylight according to the definition of UR_{mod}. In this sense, Obj_UDI_{mod-65} may be positively correlated with Obj_UR_{mod} to some extent (which is confirmed as well). Nevertheless, not all objective variables from the same discipline are necessarily correlated. For instance, good UDI_{mod-65} performance does not necessarily indicate good EUI performance. This may be because the energy saving from utilizing daylight can be offset



▲ Clustering parallel coordinate chart showing all clusters



▲ 3D scatter plot showing all clusters

CATEGORY	RoofStep	Obj_EUI	Obj_Mass	Obj_UDImod
CLUSTER_0	3.00	53.55	276.16	43.63
CLUSTER_1	3.00	53.80	250.48	27.93
CLUSTER_2	3.00	55.05	260.72	31.90
CLUSTER_3	2.00	53.37	429.42	43.48
CLUSTER_4	2.35	53.24	453.25	26.06
CLUSTER_5	2.00	53.04	264.74	50.92
CLUSTER_6	2.00	52.86	285.55	21.97
CLUSTER_7	2.00	54.50	310.70	23.95
CLUSTER_8	2.00	54.12	279.36	39.80
CLUSTER_9	5.00	52.92	265.62	15.26
CLUSTER_10	5.00	53.22	228.55	1.95
CLUSTER_11	5.00	54.37	452.07	5.09
CLUSTER_12	5.00	55.30	212.37	1.44
CLUSTER_13	5.00	54.79	221.89	15.88
CLUSTER_14	4.00	54.49	259.49	9.30
CLUSTER_15	4.00	55.48	320.24	27.54
CLUSTER_16	4.00	53.65	263.73	26.23
CLUSTER_17	4.00	52.88	282.57	5.98
CLUSTER_18	3.00	51.93	233.50	14.90
CLUSTER_19	3.00	53.63	266.65	7.37

▲ Mean values of RoofStep, Obj_EUI, Obj_Mass, Obj_UDImod of all clusters

Fig. 14.1. All twenty clusters created via the cluster analysis.

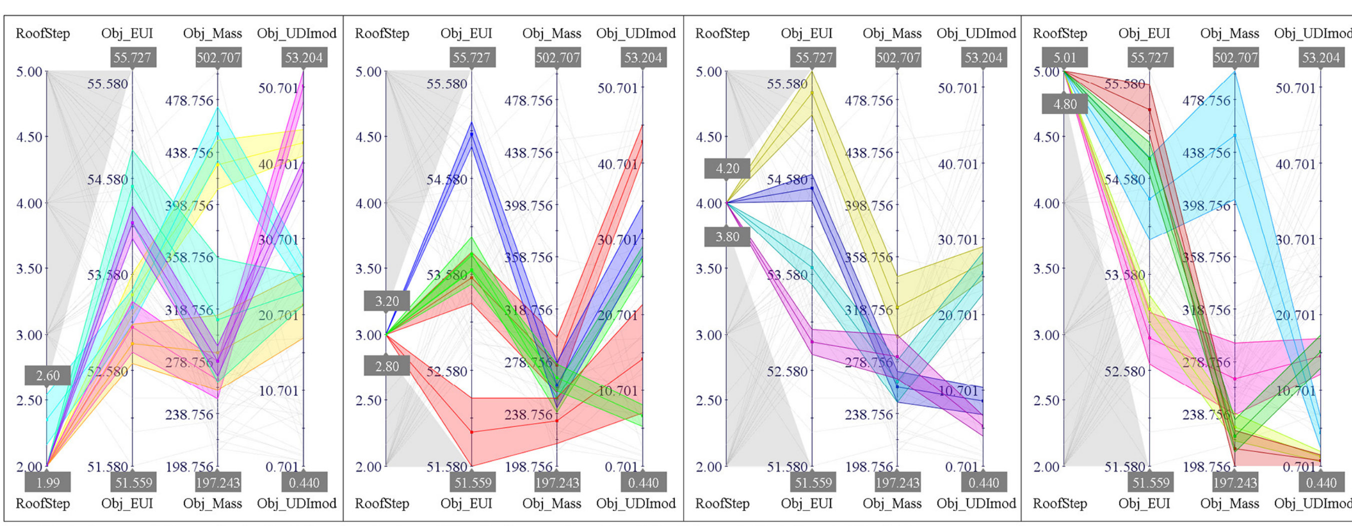
or overcompensated by the energy used for removing excessive heat gain, given the relatively wide range of the "useful" illuminance (i.e. 100–2000 lx).

6.2.2. Cluster analysis and (type I) input-output relationships

Given that the Type I design variable "RoofSteps" is crucial in this case (see Section 5.2), it is important to know its proper value in advance, i.e. how many "steps" of the roof are more promising to ensure desired overall performance. But, it is not easy to get an insight into the possible impacts of this variable by going through each alternative one by one (due to the large number of the alternatives). Instead, it would be helpful to group the alternatives into manageable and meaningful clusters (where the alternatives having similar configurations and performance trends gather together), namely, to simplify large amounts of multi-dimensional data.

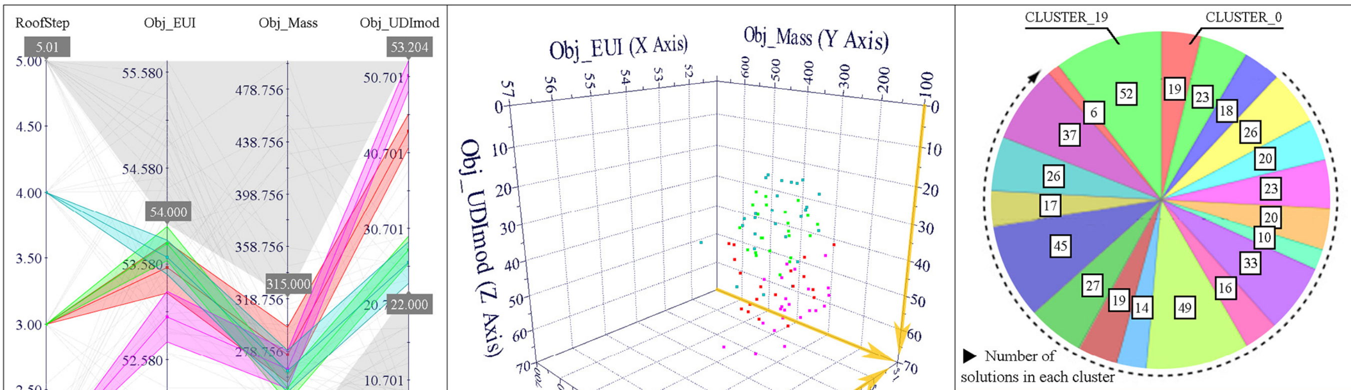
Cluster analysis (using Hierarchical clustering [51]) is performed to investigate the (Type I) input-output relationships in this case. Twenty clusters are created based on the "RoofSteps" and the Obj_EUI, Obj_Mass, Obj_UDI$_{mod-65}$ (which will be selected as final objective variables in Section 6.3). The distribution of the clusters is visualized by using a clustering parallel coordinate chart and a 3D scatter plot (Fig. 14.1). The chart represents each cluster with a centre line and a colored band which respectively indicates the mean (of the variables for the cluster in question) and the confidence interval (of the mean). The corresponding 3D scatter plot shows the clusters in the objective space using the same color coding as in the chart. The pie chart (in Fig. 14.3) shows the number of alternatives in each cluster.

The following information is extracted by manipulating the filters of the variables in question and observing the patterns of the concise clusters. First, by manipulating the filter of the "RoofSteps", the impacts of this variable on the overall performance can be captured quickly. As shown in Fig. 14.2, the clusters of alternatives having 2 and 3 roof steps perform relatively good in EUI and UDI$_{mod-65}$ performances; while the clusters of alternatives having 3 and 4 roof steps perform relatively good in Mass performance. Second, by manipulating the filters of the Obj_EUI, Obj_Mass and Obj_UDI$_{mod-65}$ towards desired directions, the clusters achieving good EUI, Mass and UDI$_{mod-65}$ performances can be filtered. Note that the closer a filter gets to the end of its desired direction, the more important the objective variable is. In this way, human preference on the relative importance of each quantitative goal



▲ Clustering parallel coordinate chart showing clusters with a similar number of roof steps; in each chart from left to right, the number of roof steps equals to (or approximately equals to) 2, 3, 4 or 5, respectively.

Fig. 14.2. The filtered clusters with a similar number of roof steps.



CATEGORY	RoofStep	Obj_EUI	Obj_Mass	Obj_UDImod
CLUSTER_0	3.00	53.55	276.16	43.63
CLUSTER_1	3.00	53.80	250.48	27.93
CLUSTER_5	2.00	53.04	264.74	50.92
CLUSTER_16	4.00	53.65	263.73	26.23

▲ Clustering parallel coordinate chart showing filtered clusters

▲ 3D scatter plot showing filtered clusters

▲ Mean values of RoofStep, Obj_EUI, Obj_Mass, Obj_UDImod of filtered clusters

Fig. 14.3. The filtered clusters with desired performance trends.

is integrated (without requiring transforming the problem into a single objective optimization problem, as in usual prior preference articulation approaches). As shown in Fig. 14.3, four clusters are filtered which achieve relatively low EUI and Mass values and high UDI_{mod-65} value; and most of the alternatives in the filtered clusters have 3 roof steps, while some alternatives have 2 or 4 roof steps. The above findings are consistent, indicating that the alternatives having 3 roof steps are generally more likely to achieve the desired quantitative goals. And, purely from the perspective of achieving the quantitative goals, some alternatives having 2 roof steps (i.e. CLUSTER_5) also perform very well, even better than some alternatives having 3 roof steps (e.g. CLUSTER_0) in all the three quantitative goals.

The information extracted from the cluster analysis seems unfamiliar to the designers in this case. Specifically, by relying on the educated guesses, the designers may only know very general information. For instance, the variation of the roof steps can change the locations of the clearstories and thus may affect the EUI and UDI_{mod-65} performances; it may also affect the Mass due to the significant change in structural topology. But the information provided by the cluster analysis gives a much clear view on how the value of the “RoofSteps” may affect the achievement of the three performance goals, namely, the relation between building shape and quantitative performances. In fact, to investigate this relation, some traditional parameters characterizing building shape can be used; and the related knowledge is known. For instance, “shape coefficient” (i.e. the ratio between the external surfaces and the inner volume of a building) was often used in energy studies; and it was believed that the energy use is inversely proportional to the compactness (i.e. weak shape coefficient) in a severely cold and scarcely sunny weather [81,82]. Nevertheless, in many real practices, various other design variables may be used as well to characterize building shape depending on specific design concepts, such as the “RoofSteps” in question. Architects tend to manipulate this kind of variable (that is directly associated with the concept) more often during the conceptual design; but lack of sufficient knowledge about the relation between the case-dependent design variable and quantitative performances. In this regard, the unfamiliar information extracted from the cluster analysis can possibly lead to new knowledge.

6.2.3. Sensitivity analysis and (type II and III) input-output relationships

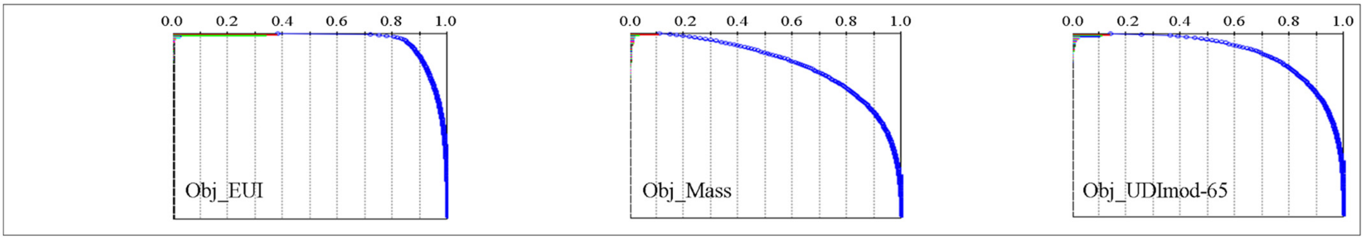
Knowing the (Type II and III) input-output relationships is helpful for the design variable screening (i.e. screening out unimportant design variables that contribute the least to the variation of objective variables). But, it is hard for designers to know the relative importance of each design variable on objective variables quantitatively based on rules of thumb, especially when the number of design variables is large and/or the interaction effects are considered.

Sensitivity analysis (using Smoothing Spline ANOVA [52]) is performed to investigate the (Type II and Type III) input-output relationships in this case. For this analysis, the design variables excluding BVD4, BVD5, F4, F5, R4, R5, RUN4, RUN5, RoofSteps are used as “factors” (given that RoofSteps = 3); and the Obj_EUI, Obj_Mass, Obj_UDI_{mod-65} are used as “responses”. Both main effects and interaction effects are considered. A factor with a main effect or a pair of factors with an interaction effect is called a “term”. The relative importance (i.e. the percentage of effect or contribution) of each term to the global variance of a response is visualized by a column in an effect column chart; and the curve in the chart represents the cumulative effect of the terms in question (Fig. 15).

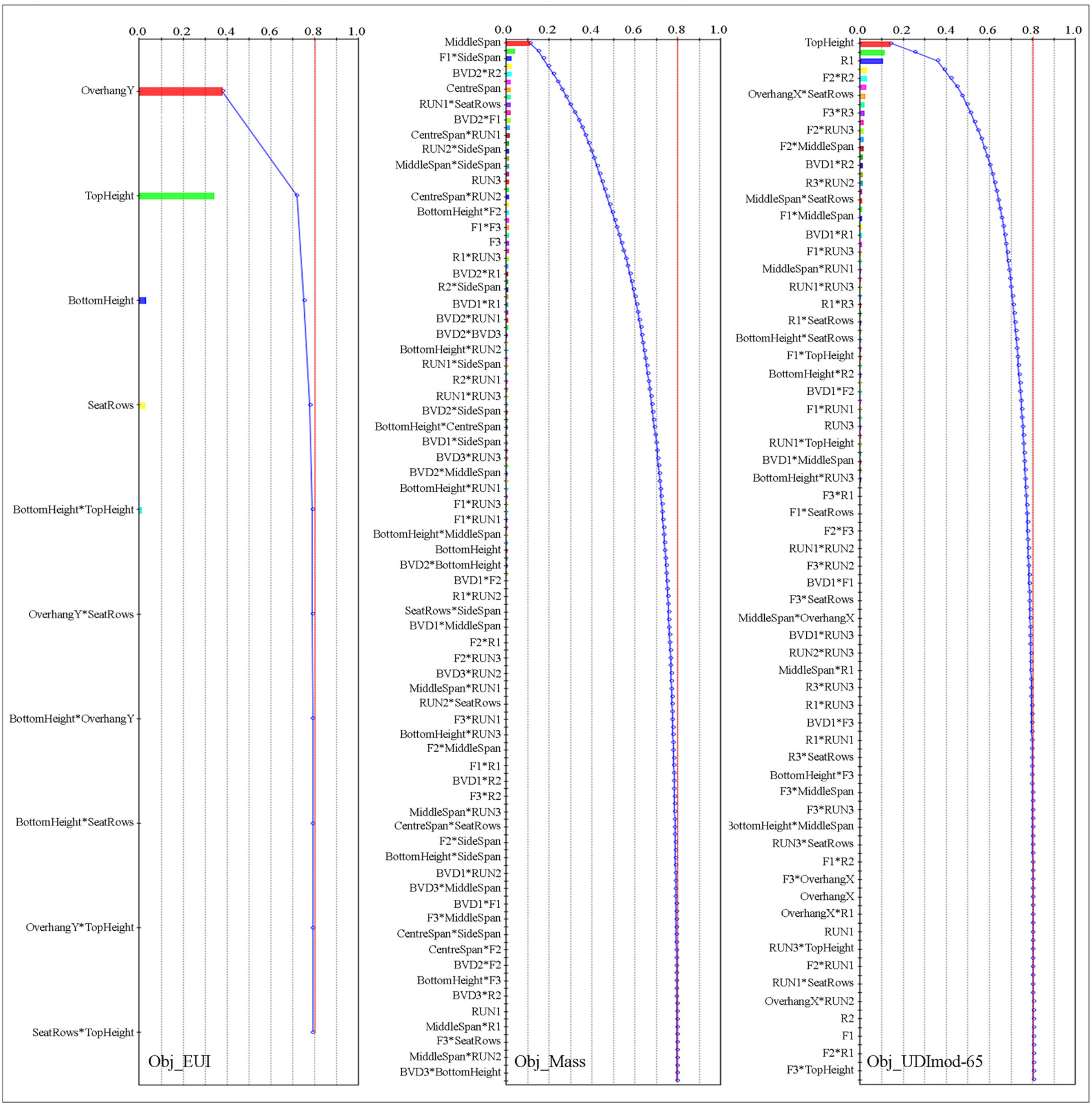
The following information is extracted from the sensitivity analysis results: (1) the contributions of the terms on Obj_EUI are the most diverse, the most important two design variables (i.e. OverhangY and TopHeight) are responsible for the major portion (i.e. 37.9% and 33.9% respectively) of the variation of Obj_EUI; (2) the contributions of the terms on Obj_Mass are the least diverse, the most important design variable (i.e. MiddleSpan) is only responsible for a small portion (i.e.10.8%) of the variation of Obj_Mass, while the remaining terms account for the major portion as a whole; (3) the pattern of the contributions on Obj_UDI_{mod-65} is more similar to that on Obj_Mass, the most important three design variables (i.e. TopHeight, BottomHeight and R1) account for 14.2%, 11.2% and 10.5% of the overall contribution respectively, while the remaining terms as a whole account for a larger portion (although each of them contributes a very small portion). These findings are indicated by the different curvatures of the cumulative effect curves (Fig. 15, top). Moreover, the interaction effects are non-negligible, especially for Obj_Mass and Obj_UDI_{mod-65}, as indicated by the many interaction terms (Fig. 15, bottom).

By filtering out less important terms to a degree that maintains a cumulative effect of about 80% (Fig. 15, bottom), some design variables important for each response emerge (as indicated in gray in Table 5). Note that if an interaction effect is important, all the involved design variables are considered as important even if their main effects are not significant. In addition, the importance of the Type I variable “RoofSteps” is also confirmed by the sensitivity analysis (that includes the variable as one of the factors).

The information extracted from the sensitivity analysis is logical in disciplinary contexts. From an energy point of view, the OverhangY



▲ Effect column charts showing all main and interaction terms that maintain a cumulative effect of 100%
(note: all the names of the terms are hidden, as the focus of the charts is to show the complete cumulative effects)



▲ Effect column charts showing the main and interaction terms that maintain a cumulative effect of about 80%
(note: some names of the terms are hidden, due to the limited space)

Fig. 15. All terms that maintain a cumulative effect of 100% (top); the terms that maintain a cumulative effect of 80% (bottom).

Table 5
Important design variables for each response and their main effects.

	Seat rows	Bottom height	Top height	R1	R2	R3	F1	F2	F3	Over-hang X	Over-hang Y	Centre Span	Middle Span	Side Span	BVD1	BVD2	BVD3	RUN1	RUN2	RUN3
Obj_EUI	0.026	0.032	0.339	0.026	0.002	0.000	0.012	0.000	0.000	0.004	0.379	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.001	0.000
Obj_Mass	0.013	0.002	0.002	0.019	0.000	0.000	0.003	0.000	0.000	0.003	0.002	0.020	0.108	0.021	0.012	0.012	0.000	0.000	0.039	0.012
Obj_UDI$_{mod\text{-}65}$	0.009	0.112	0.142	0.105	0.000	0.005	0.000	0.000	0.017	0.000	0.000	0.000	0.000	0.001	0.004	0.006	0.001	0.000	0.000	0.003

Table 6
Re-formulated optimization problems and the execution.

No.	Item	Optimization problem (re-)formulation				Execution		
		Objective var. number	(Type I) design var. number	(Type II & III) design var. number	Initial generation	Actual evaluated designs	Total time	Time per design
1	Traditional	4	1 (RoofSteps = 2,3,4,5)	16–28	ULH	462	37 h:43 m	4.90 m
2	Proposed	3	0 (RoofSteps = 3)	20	Promising cluster[a]	463	42 h:02 m	5.45 m
3	Factor1	2	0 (RoofSteps = 3)	20	Promising cluster[a]	453	40 h:51 m	5.41 m
4	Factor2	3	0 (RoofSteps = 2)	16	Promising cluster[b]	455	32 h:01 m	4.22 m
5	Factor2	3	0 (RoofSteps = 4)	24	Promising cluster[c]	449	51 h:41 m	6.91 m
6	Factor3	3	0 (RoofSteps = 3)	18	Promising cluster[d]	463	42 h:52 m	5.56 m
7	Factor4	3	0 (RoofSteps = 3)	20	ULH	465	44 h:09 m	5.70 m

[a] The initial generation is selected from the promising clusters CLUSTER_0 and CLUSTER_1.
[b] The initial generation is selected from the promising cluster CLUSTER_5.
[c] The initial generation is selected from the promising cluster CLUSTER_16.
[d] The initial generation is selected from the promising clusters CLUSTER_0 and CLUSTER_1 in which the R3 and RUN3 are treated as constants.

defines the overhang depth for south- and north-facing clearstories, it can determine the sun exposure in the indoor space thus affect the energy use. And, the TopHeight can affect the energy use via changing the volume of the indoor space. From a structural point of view, the MiddleSpan defines the location of the maximum vertical distance between upper and lower main beams, thus it can change the load-bearing capacity and hence the mass. From a daylight point of view, the TopHeight and BottomHeight define the vertical locations and sizes of clearstories, they may affect the illuminance levels of the court. And, the R1 helps to define the horizontal locations of clearstories that are closest to the centre of the court, thus it may also affect the illuminance levels.

6.3. Optimization problem re-formulation

With the relevant information and knowledge extracted, the designers can re-formulate the original optimization problem in a more informed manner. In this section, the initial objective variables and design variables are re-formulated in different manners for different purposes (as shown in Tables 6 and 7). They are re-formulated according to the proposed method, thus forming the optimization problem No. 2. Meanwhile, they are kept unchanged when following the traditional method, thus forming the optimization problem No. 1. Additionally, to investigate the factors that may affect the behaviour of the proposed method (i.e. Factor1 to Factor4), the optimization problems No. 3–No. 7 are also formulated by changing each factor at a time. The underlines in Table 6 identify the factors being changed, in order to facilitate quick understanding of the formulation of the optimization problems No. 3-No. 7. The (re-)formulation processes are described in more detail below.

During the objective variable screening, the Obj_UR$_{mod}$ is screened out for the optimization problem No. 2. This is based on the information obtained from the correlation analysis. Specifically, given that the Obj_UR$_{mod}$ and Obj_UDI$_{mod\text{-}65}$ vary in a similar direction, one of them can be removed from the four candidate objective variables and treated as a constraint during optimization. In this case, the UDI$_{mod\text{-}65}$ remains as an objective as its values are more widely spread out; and the Obj_UR$_{mod}$ is treated as a constraint > 0.58 (i.e. its third quartile in the DoE data set). Moreover, to understand how the overscreening of the objective variables (i.e. Factor1) may affect the optimization results, the optimization problem No. 3 is formulated. In this formulation, the Obj_UDI$_{mod\text{-}65}$ is also screened out and treated as a constraint > 32.55% (i.e. its third quartile in the DoE data set).

Regarding the decision on the (Type I) design variable value, the "RoofSteps" equals to three, thus the BVD4, BVD5, F4, F5, R4, R5, RUN4, RUN5 are eliminated accordingly for the optimization problem No. 2. This decision is based on the information obtained from the cluster analysis. Given that the alternatives having three roof steps (i.e. CLUSTER_0 and CLUSTER_1) account for the major portion of the filtered alternatives, they are considered more promising from the perspective of achieving quantitative goals. Meanwhile, these alternatives are also considered subjectively preferred from the perspective of achieving qualitative goals (like aesthetics). Thus, the expected "RoofSteps" values from both perspectives are the same (i.e. three) in the optimization problem No. 2. However, the expected "RoofSteps" values can be in conflict with each other. For instance, the designers may subjectively prefer the alternatives having two or four roof steps (i.e. CLUSTER_5 or CLUSTER_16) which are quantitatively less promising. In this case, they may need to balance between the quantitative and qualitative goals before determining the "RoofSteps" value (thus the promising clusters). If the qualitative goals are considered dominant over the quantitative goals, the "RoofSteps" can be two or four. In this sense, it is meaningful to understand how the human preference on the qualitative goals (i.e. Factor2) may affect the optimization results. For this, the optimization problems No. 4 and 5 are formulated in which the "RoofSteps" equals to two and four, respectively, due to the dominant qualitative goals.

During the (Type II & III) design variable screening, none of the remaining design variables is further screened out for the optimization problem No. 2. This is based on the information obtained from the

Table 7
Lists of re-formulated objective variables and design variables.

No.	1	2	3	4	5	6	7
Item	Traditional	Proposed	Factor1	Factor2	Factor2	Factor3	Factor4
Objective variables	Obj_EUI	Obj_EUI	Obj_EUI	Obj_EUI	Obj_EUI	Obj_EUI	Obj_EUI
	Obj_Mass	Obj_Mass	Obj_Mass	Obj_Mass	Obj_Mass	Obj_Mass	Obj_Mass
	$Obj\_UDI_{mod\text{-}65}$	$Obj\_UDI_{mod\text{-}65}$	–	$Obj\_UDI_{mod\text{-}65}$	$Obj\_UDI_{mod\text{-}65}$	$Obj\_UDI_{mod\text{-}65}$	$Obj\_UDI_{mod\text{-}65}$
	Obj_UR_{mod}	–	–	–	–	–	–
Design variables (Type I)	RoofSteps	–	–	–	–	–	–
Design variables (Type II & III)	SeatRows	SeatRows	SeatRows	SeatRows	SeatRows	SeatRows	SeatRows
	BottomHeight	BottomHeight	BottomHeight	BottomHeight	BottomHeight	BottomHeight	BottomHeight
	TopHeight	TopHeight	TopHeight	TopHeight	TopHeight	TopHeight	TopHeight
	R1	R1	R1	R1	R1	R1	R1
	R2	R2	R2	R2	R2	R2	R2
	R3	R3	R3	–	R3	–	R3
	R4	–	–	–	R4	–	–
	R5	–	–	–	–	–	–
	F1	F1	F1	F1	F1	F1	F1
	F2	F2	F2	F2	F2	F2	F2
	F3	F3	F3	–	F3	F3	F3
	F4	–	–	–	F4	–	–
	F5	–	–	–	–	–	–
	OverhangX	OverhangX	OverhangX	OverhangX	OverhangX	OverhangX	OverhangX
	OverhangY	OverhangY	OverhangY	OverhangY	OverhangY	OverhangY	OverhangY
	CentreSpan	CentreSpan	CentreSpan	CentreSpan	CentreSpan	CentreSpan	CentreSpan
	MiddleSpan	MiddleSpan	MiddleSpan	MiddleSpan	MiddleSpan	MiddleSpan	MiddleSpan
	SideSpan	SideSpan	SideSpan	SideSpan	SideSpan	SideSpan	SideSpan
	BVD1	BVD1	BVD1	BVD1	BVD1	BVD1	BVD1
	BVD2	BVD2	BVD2	BVD2	BVD2	BVD2	BVD2
	BVD3	BVD3	BVD3	–	BVD3	BVD3	BVD3
	BVD4	–	–	–	BVD4	–	–
	BVD5	–	–	–	–	–	–
	RUN1	RUN1	RUN1	RUN1	RUN1	RUN1	RUN1
	RUN2	RUN2	RUN2	RUN2	RUN2	RUN2	RUN2
	RUN3	RUN3	RUN3	–	RUN3	–	RUN3
	RUN4	–	–	–	RUN4	–	–
	RUN5	–	–	–	–	–	–

sensitivity analysis. Specifically, design variables can be eventually screened out when they are considered unimportant for all the performance goals in question. But, in this case, all the remaining design variables are important for at least one of the three quantitative performance goals, thus they are all kept. Moreover, to understand how the overscreening of the design variables (i.e. Factor3) may affect the optimization results, the optimization problem No. 6 is formulated. In this formulation, the R3 and RUN3 are further screened out (at the cost that the cumulative effects of the remaining terms on Obj_Mass and $Obj\_UDI_{mod\text{-}65}$ are 60%); and they are treated as constants (i.e. their central values).

In addition, selecting an initial generation from promising clusters is believed a good start for the consequent optimization. Therefore, for the optimization problem No. 2, the initial generation is selected from the promising clusters which consist of alternatives having three roof steps (i.e. CLUSTER_0 and CLUSTER_1). Similarly, for the optimization problems No. 3–No. 6, the initial generations are also selected from promising clusters, but the clusters being considered as promising are different and depend on the "RoofSteps" values, which is indicated by the footnotes of Table 6. Moreover, to understand how the initial generation (i.e. Factor4) may affect the optimization results, the optimization problem No. 7 is formulated. In this formulation, the initial generation is selected using the ULH sampling strategy.

7. CDO results and comparison

Having re-formulated the original optimization problem, the designers can enter the CDO directly without further conducting CDE iterations (given that the focus of this case is to narrow down design possibilities as mentioned in Section 4.1). In Section 7.1, the results of the optimization problems No. 1 and No. 2 are visualized and compared, to verify the benefits of the proposed approach over the traditional method; In Section 7.2, the results of the optimization problem No. 2 are compared with the results of the optimization problems No. 3–No. 7, respectively, to further understand how the four factors may affect the behaviour of the proposed approach.

To ensure the comparability, all the optimizations listed in Table 6 were executed using the same optimization algorithm (i.e. NSGA-II [83]) and the same optimization settings (i.e. population size of 25 and 20 generations); and they were run on the same 6-Core (12-Thread) machine. For each optimization, the number of actual evaluated designs is similar and < 500 because of occasional skips; the execution time is limited to around two days; and the average time for evaluating each design is calculated and recorded in Table 6. The design summary bars at the top of Figs. 16, 17, 19, 20, 21, 22 and 23 keep track of each design evaluation. They show feasible designs in green, unfeasible designs in yellow and occasional skips in gray. The optimization results shown in these figures include both quantitative and qualitative performances of Pareto solutions. All these results to be compared are summarized in Table 8, including the number of Pareto solutions, of unfeasible designs, and of broken designs (that violate certain constraints). The underlines in Table 8 identify the quantitative performances superior to that derived from the optimization problem No. 2, in order to facilitate quick understanding of the relative goodness of the results derived from other optimization problems.

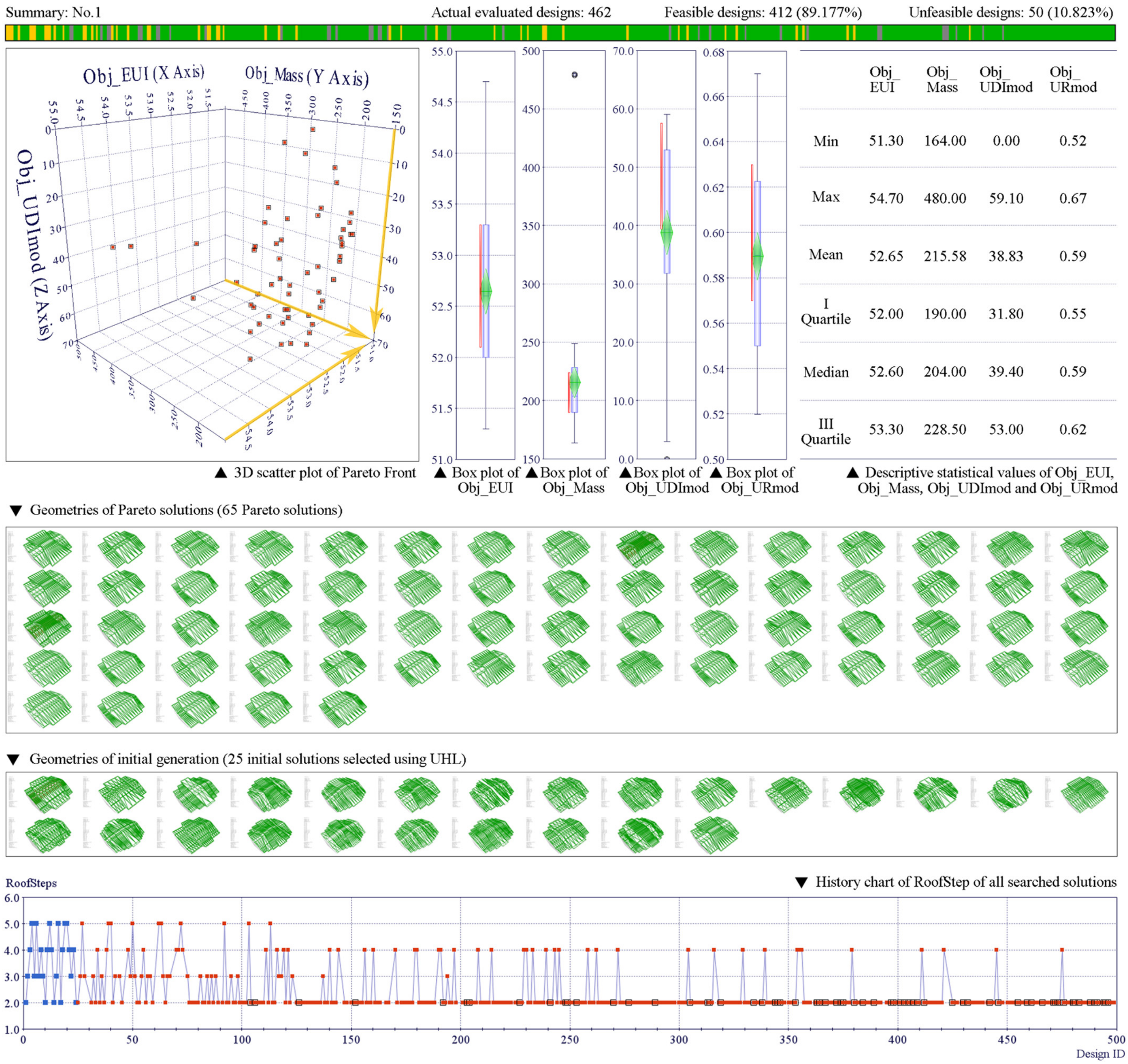


Fig. 16. Results of the optimization problem No. 1 (top); history chart of the “RoofSteps” in the optimization problem No. 1 (bottom).

7.1. Comparison between proposed and traditional methods

The results of the optimization problems No. 1 and No. 2 are visualized in Fig. 16 (top) and Fig. 17 (top). The Pareto Front derived from each optimization is plotted in the same 3D space composed of Obj_EUI, Obj_Mass and Obj_$UDI_{mod\text{-}65}$ dimensions (although the UR_{mod} is treated as an objective in the optimization problem No. 1); the EUI, Mass, $UDI_{mod\text{-}65}$ and UR_{mod} performance data of Pareto solutions are described by box-whisker plots and tables; and the geometries of initial generations and of Pareto solutions are also presented. By comparing all these results, the relative advantages of the proposed approach are indicated, as described below.

From the perspective of quantitative performances, the Pareto solutions derived from the optimization problem No. 2 generally perform better than those derived from the optimization problem No. 1, although the total number of unfeasible designs during the optimization process is slightly larger. They are considered better, given the following facts. In the optimization problem No. 2, the EUI and Mass median performance values are relatively lower; the $UDI_{mod\text{-}65}$ and UR_{mod} median performance values are relatively higher; the interquartile performance ranges are relatively more concentrated; and the total number of the Pareto solutions is proper. These desired features facilitate to obtain quantitatively high-performing solutions that better meet the designers' preference on the quantitative goals. Note that although the performance ranges are relatively concentrated, the Pareto solutions are still distributed in a diverse manner, which allows the designers to balance between the conflicting quantitative goals during the final decision making.

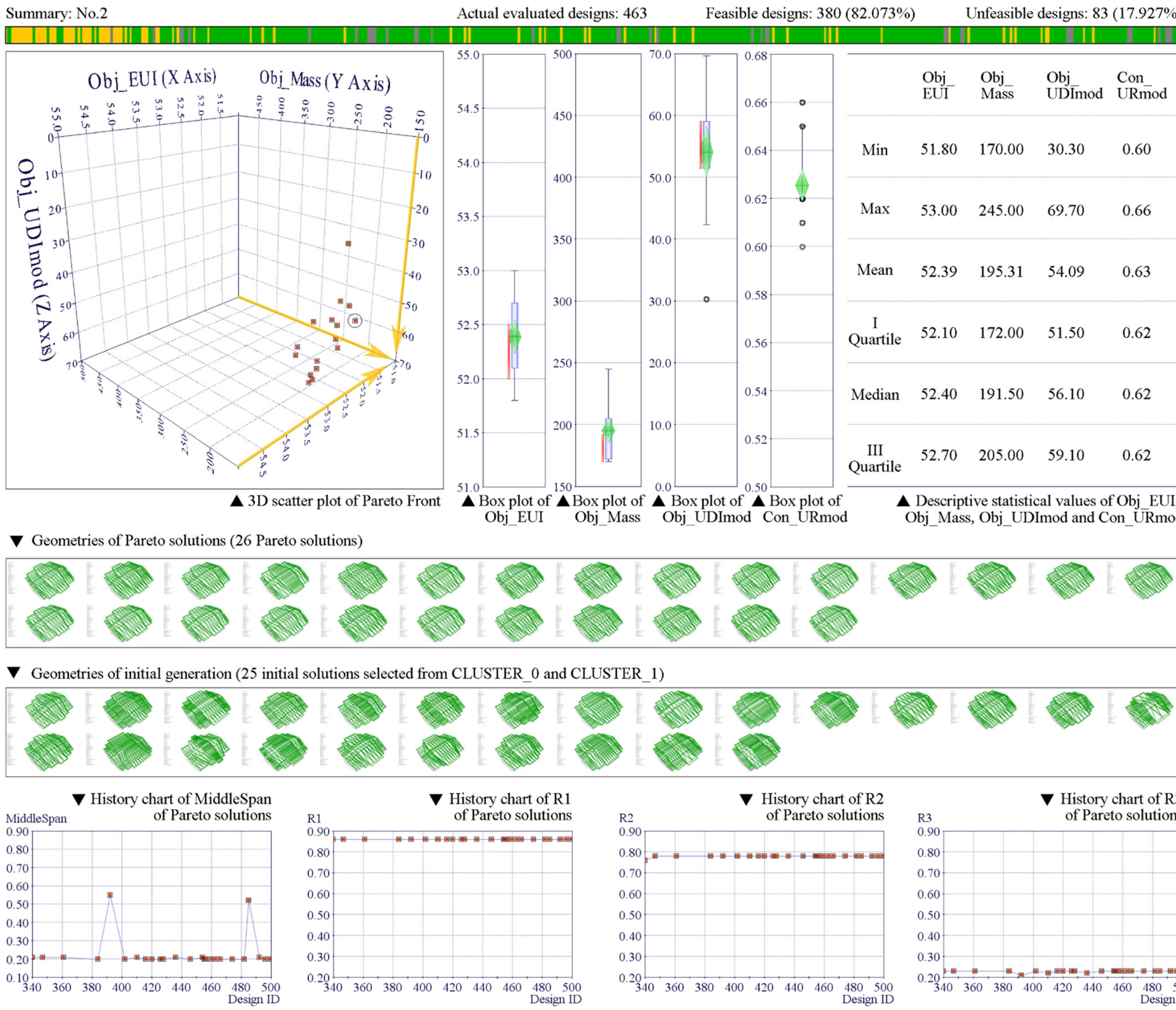


Fig. 17. Results of the optimization problem No. 2 (top); history charts of the "MiddleSpan", "R1", "R2", "R3" in the optimization problem No. 2, showing the Pareto solutions only (bottom).

It is also interesting to compare the quantitative results derived by using the proposed method with that of the real-world project. For this, a Pareto solution randomly selected from the optimization problem No. 2 (marked by a circle in Fig. 17), and the benchmark solution most like the real-world project (as mentioned in Table 1) are shown in Fig. 18. It is proved that all the EUI, Mass, $UDI_{mod\text{-}65}$ and UR_{mod} performances of the former are better than that of the latter. Especially, the $UDI_{mod\text{-}65}$ performance can improve from 0 to 51.5. This potential improvement confirms the necessity of adding additional skylights on the roof top in the real-world project, to increase daylight availability (as shown in Fig. 6).

From the perspective of qualitative performances, the geometries of the Pareto solutions derived from the optimization problem No. 2 are aesthetically preferred, rather than those from the optimization problem No. 1. This is associated with whether the subjective human preference on qualitative goals is considered. Specifically, in the optimization problem No. 2, the aesthetic preference on three roof steps is integrated during the decision of the Type I design variable "RoofSteps". The "RoofSteps" is considered as a constant (i.e. 3) during the optimization, thus, all the Pareto solutions derived have three roof steps as desire. On the contrary, in the optimization problem No. 1, there is no preference on the number of roof steps. The "RoofSteps" is considered as an integer variable (i.e. ranging from 2 to 5) during the optimization; thus, the Pareto solutions derived (i.e. boxed solutions in Fig. 16, bottom) do not necessarily include alternatives having three roof steps. Furthermore, the history chart of the "RoofSteps" (Fig. 16, bottom) also shows that, as the optimization No. 1 proceeds, the algorithm searches more among the alternatives having two roof steps, although they are not aesthetically preferred. It also explains why the average time for evaluating per design in the optimization problem No. 1 is relatively shorter (given that the alternatives having less roof steps often consume less simulation time).

Another observation regarding the geometries of the Pareto solutions is that the geometries derived from the optimization problem No. 2 are more like each other compared to those derived from the optimization problem No. 1, which may indicate knowledge unfamiliar to

Table 8
Comparison of the optimization results.

No.	Item	Pareto solution number	Quantitative performances of Pareto solutions				Qualitative performances of Pareto solutions		Unfeasible design number	Broken design number
			EUI	Mass	$UDI_{mod\text{-}65}$	UR_{mod}	Geometric preference	Geometric similarity		
1	Traditional	65	52.60 (1.30)	204.00 (38.50)	39.40 (21.20)	0.59 (0.07)	RoofSteps = 2,3,4,5	*	50 (10.8%)	21 (Con_UR_{mod}) 21 (Con_SC)
2	Proposed	26	52.40 (0.60)	191.50 (33.00)	56.10 (7.60)	0.62 (0.00)	RoofSteps = 3	***	83 (17.9%)	72 (Con_UR_{mod}) 10 (Con_SC)
3	Factor1	4	52.65 (1.05)	179.00 (10.50)	46.25 (7.55)	0.65 (0.02)	RoofSteps = 3	***	90 (19.9%)	67 (Con_UR_{mod}) 33 (Con_$UDI_{mod\text{-}65}$) 11 (Con_SC)
4	Factor2	40	52.10 (0.85)	178.50 (28.50)	47.75 (34.05)	0.60 (0.02)	RoofSteps = 2	**	144 (31.6%)	97 (Con_UR_{mod}) 72 (Con_UC)
5	Factor2	20	53.10 (1.15)	214.00 (34.50)	38.65 (9.85)	0.61 (0.04)	RoofSteps = 4	**	186 (41.4%)	168 (Con_UR_{mod}) 32 (Con_SC) 10 (Con_DC)
6	Factor3	19	53.20 (0.85)	204.00 (62.25)	34.80 (35.55)	0.59 (0.01)	RoofSteps = 3	**	302 (65.2%)	294 (Con_UR_{mod}) 37 (Con_SC)
7	Factor4	29	53.00 (0.93)	215.00 (79.25)	43.90 (17.07)	0.62 (0.03)	RoofSteps = 3	**	165 (35.3%)	154 (Con_UR_{mod}) 26 (Con_SC)

In the 4th, 5th, 6th, 7th column, median performance values and interquartile ranges are shown without and with parentheses respectively.
In the 9th column, the number of stars represents the degree of geometric similarity; the more stars the more similar.
In the last column, major broken constraints (i.e. those violated by > 10 designs) are listed; some designs may violate multiple constraints.

the designers. The geometric similarity in the optimization problem No. 2 can be confirmed by the history charts of the related design variables (Fig. 17, bottom). The charts show that most of the “MiddleSpan” values of Pareto solutions are around 0.2, and the dominant values of the “R1”, “R2”, “R3” are 0.86, 0.78, 0.23. These values indicate the knowledge that, for the well-performing solutions, the position of middle secondary beams tends to be close to the centre of the building, and the half-ridge tends to be divided in the proportion of 0.86:0.78:0.23. Note that despite the geometric similarity, the Pareto solutions still maintain a sufficient degree of geometric diversity, which allows the designers to balance between quantitative and qualitative goals during the final decision making.

7.2. Behaviour of the proposed method

The benefits of the proposed approach (represented by the optimization problem No. 2) may derive from different factors during CDE, such as the objective variable screening, integration of human preference on qualitative goals, design variable screening, and selection of an initial generation. In other words, the behaviour of the proposed method may be affected by the use or abuse of these factors. Thus, it would be helpful to further understand the magnitudes of the possible impacts of these factors. For this, the results of the optimization problems No. 3–No. 7 (visualized in Figs. 19–23 and Table 8) are compared with the results of the optimization problem No. 2 respectively. Useful messages about the behaviour of the proposed approach are indicated, as described below.

In the optimization problem No. 3, the quantitative and qualitative performances, the unfeasible design and broken design numbers are relatively close to those in the optimization problem No. 2; while the Pareto solution number is much smaller than that in the optimization problem No. 2. This information indicates that (1) the overscreening of objective variables may lead to similar performances of the Pareto solutions, in comparison with the proper screening, as long as the objective variables being screened out are treated as constraints with proper limits, such as, the third quartiles of the variables in the DoE

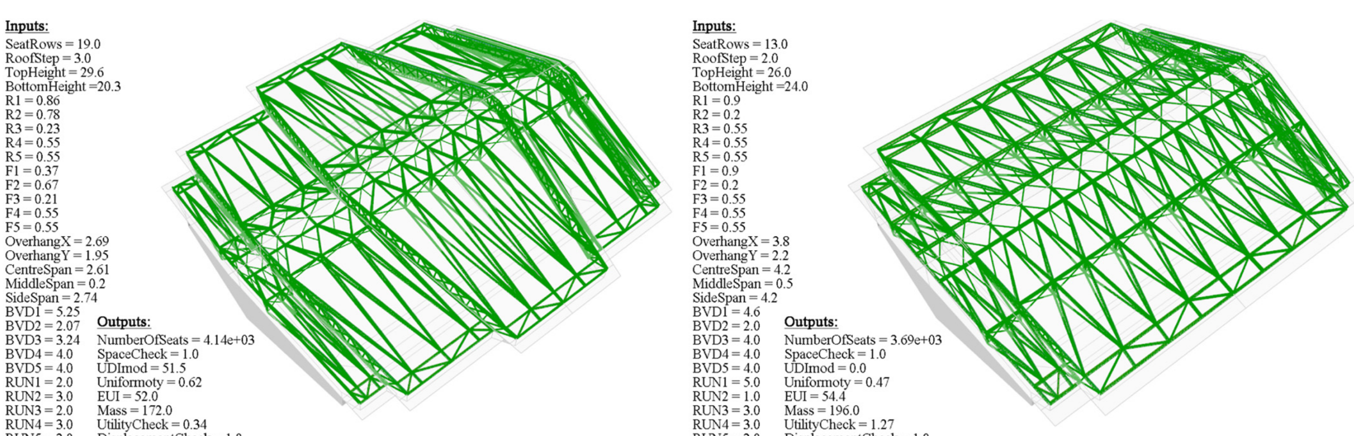


Fig. 18. a Pareto solution randomly selected from the optimization problem No. 2 (left); the benchmark solution most like the real-world project (right).

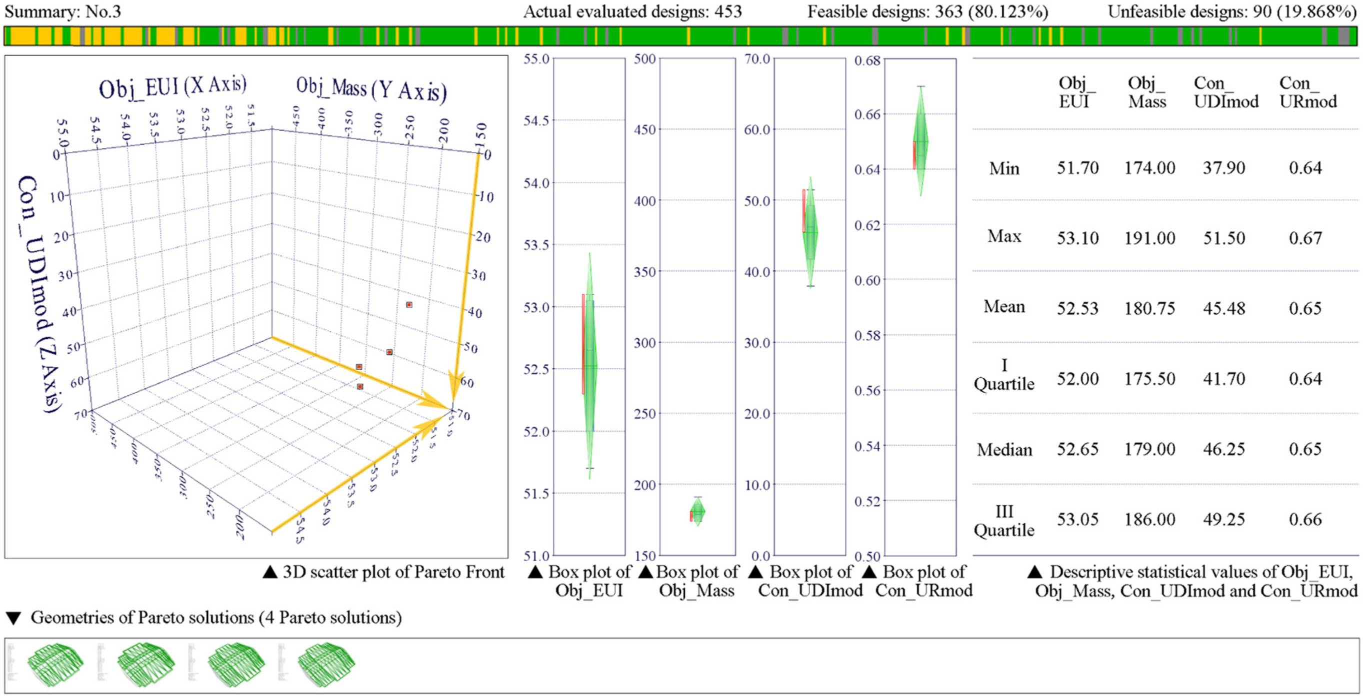


Fig. 19. Results of the optimization problem No. 3.

data set; and (2) it may also lead to the sharp decrease of the Pareto solution number, which may not be preferable for the final decision making (if too few Pareto solutions are left).

In the optimization problems No. 4 and No. 5, only some of the Pareto solutions (indicated by the ellipse in Fig. 20) have similar quantitative performances as in the optimization problem No. 2; while most of them deviate from the desired direction (i.e. the bottom right corner of the objective space). This may be associated with the different "promising" clusters (i.e. CLUSTER_5 and CLUSTER_16) from which the initial generations are selected. Regarding the qualitative

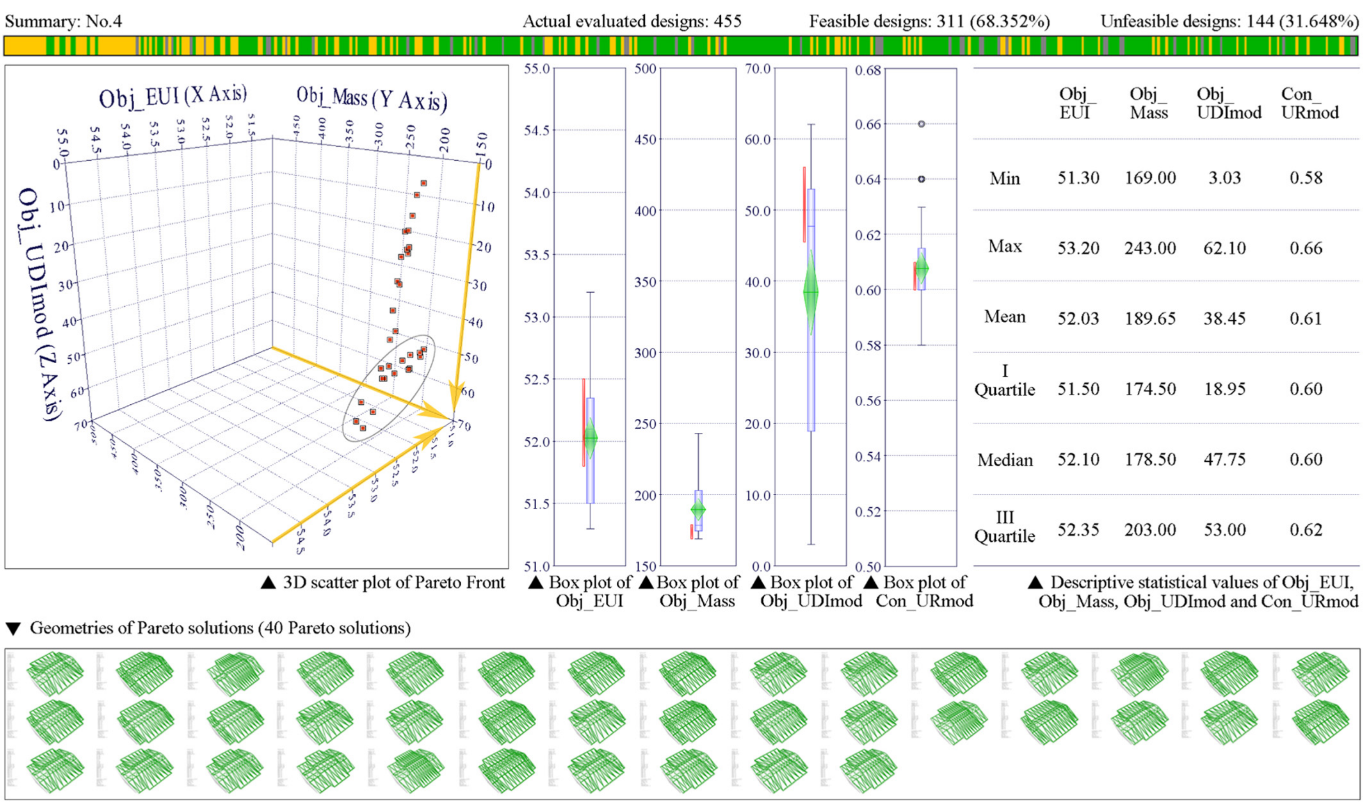


Fig. 20. Results of the optimization problem No. 4.

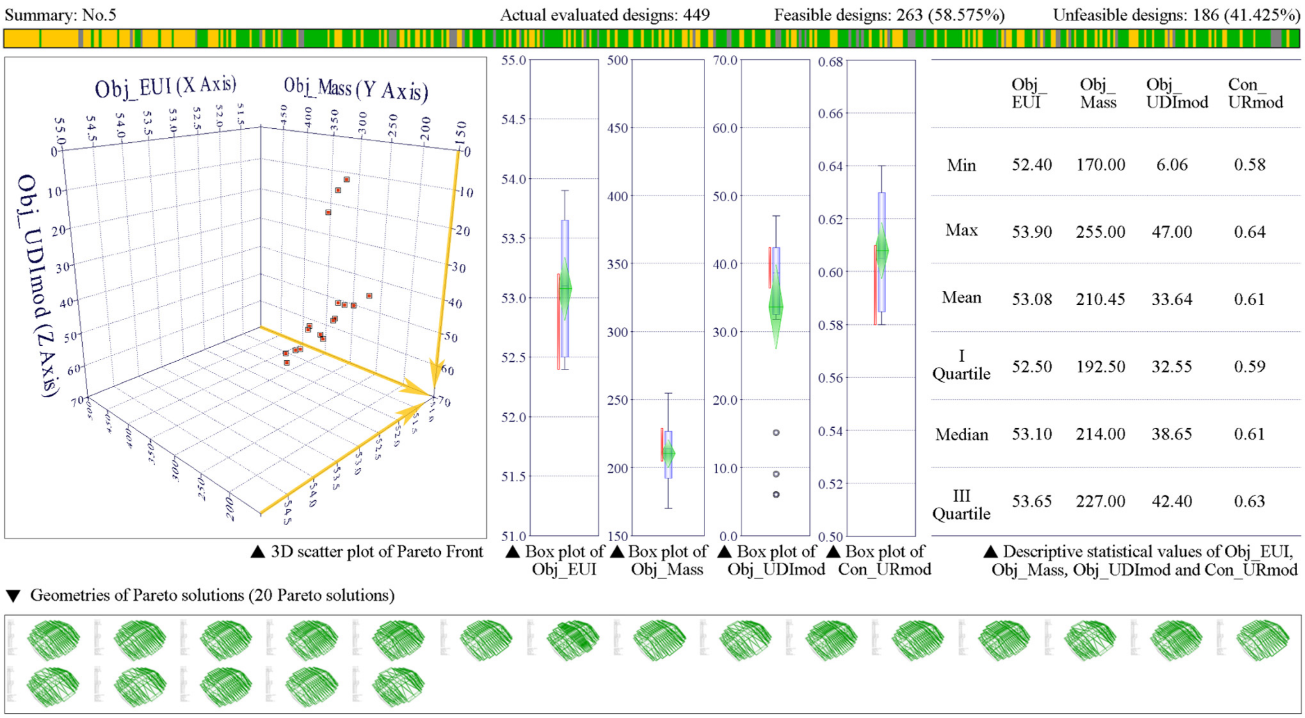


Fig. 21. Results of the optimization problem No. 5.

performances, both optimizations can achieve aesthetically preferred geometries with desired numbers of roof steps; while, the geometries derived are less similar with each other compared to those derived from the optimization problem No. 2. This can be confirmed by the values of some design variables like the "RoofSteps", "MiddleSpan", "R1", "R2" etc. Moreover, the unfeasible design and broken design numbers increase considerably in the optimization problem No. 4 and No. 5. The above information indicates that the integration of dominant qualitative goals (like aesthetics) can lead to aesthetically preferred geometries but probably with compromise in the quantitative performances.

In the optimization problems No. 6 and No. 7, the quantitative performances deviate from the desired direction obviously; the geometric similarity is relatively low; and the unfeasible design and broken design numbers increase dramatically, especially in the optimization

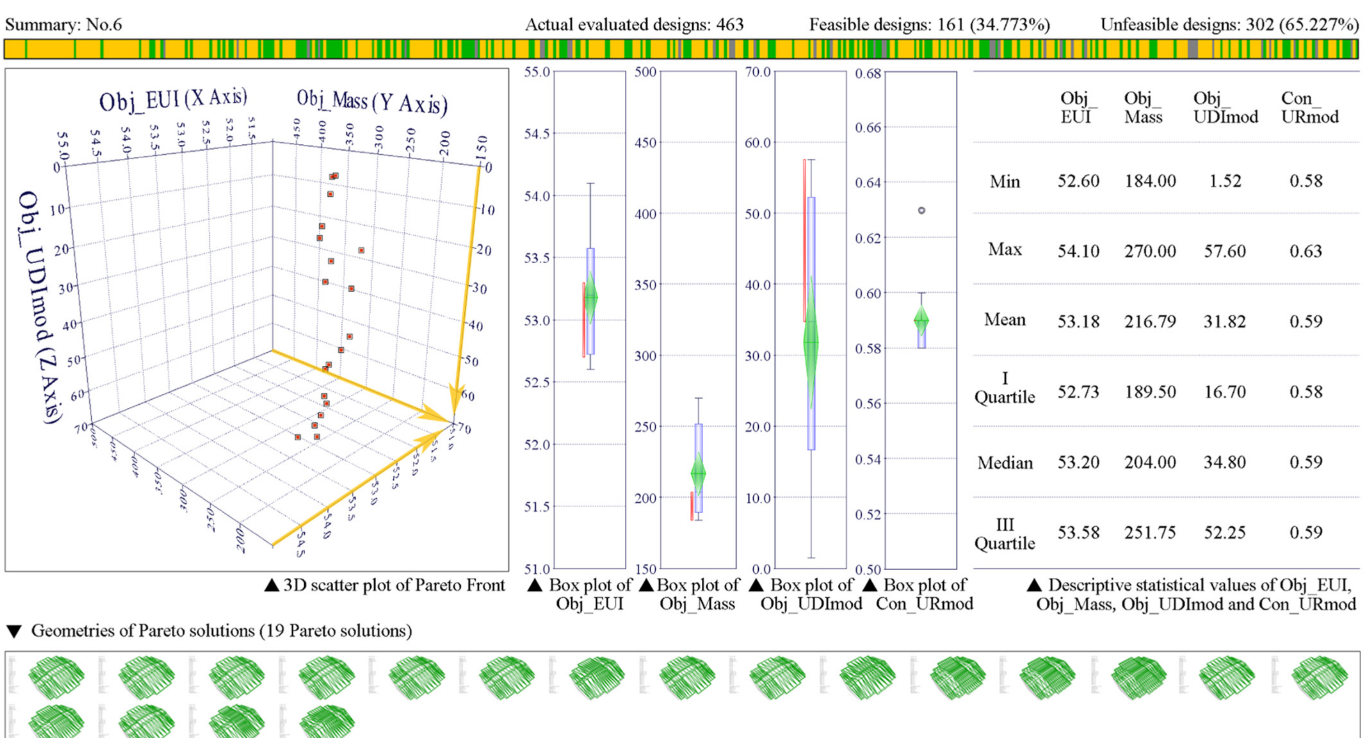


Fig. 22. Results of the optimization problem No. 6.

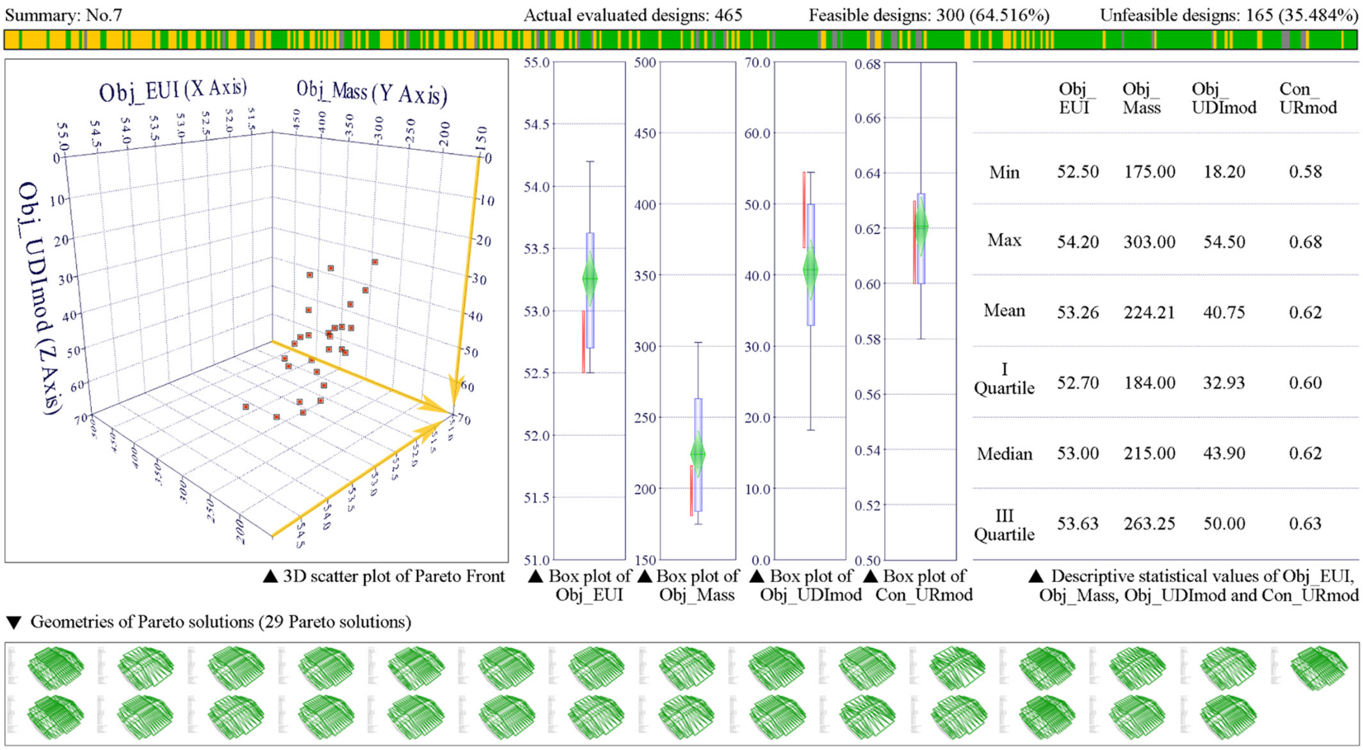


Fig. 23. Results of the optimization problem No. 7.

problem No. 6. This information indicates that (1) the overscreening of (Type II & III) design variables may deteriorate the quantitative performances of Pareto solutions significantly, in comparison with the proper screening where the cumulative effects of the remaining terms are high enough, such as 80%; and (2) the selection of an initial generation does affect the optimization results significantly, specifically, selecting the initial generation using the ULH sampling strategy is not as beneficial as selecting it from promising clusters.

8. Conclusion

8.1. Discussion and summary of contributions

This paper has proposed and demonstrated a new CDE approach which is developed for the conceptual design optimization of large-scale buildings involving multi-disciplinary criteria and complex geometries (e.g. indoor sports buildings). The proposed approach improves upon the traditional method by introducing a changeable initial OPF and inserting a CDE module. The changeable initial OPF allows or facilitates the expansion of the dimensionality of an objective space and design space being investigated, which is valuable for encouraging the designers' creative intentions. The CDE module can re-formulate the changeable optimization problem by screening out unnecessary objective variables, unimportant design variables, and by focusing on promising clusters of alternatives, which is beneficial for using the computational resources more wisely. Moreover, the proposed approach highlights the role designers play throughout the CDE process, which is crucial for integrating subjective human preferences. In particular, it allows the designers to prioritize quantitative goals during the cluster filtering (as described in Section 6.2.2), thus, the human preferences on the relative importance of each quantitative goal can be integrated. It also allows the designers to prioritize between quantitative and qualitative goals during the determination of promising clusters (as described in Section 6.3), thus, the human preferences on the qualitative goal can be considered as well. In this way, the human involvement on the front end (i.e. in the CDE process) can make the approach more flexible and not constraining to convergent thinking only as in the traditional methods. Thus, this approach fits conceptual design tasks requiring significant creativity.

This paper has verified the benefits of the proposed approach over the traditional one, and has unveiled the factors that may affect the behaviour of the proposed approach, by comparing the results of a series of optimization problems formulated for the same design task.

The comparison results concerning the proposed and traditional approaches show that (1) the Pareto solutions derived from the proposed approach are quantitatively more promising and qualitatively more preferred in general, although their total number is relatively smaller; (2) they are more concentrated in the objective space (i.e. quantitatively similar) and their geometries look more convergent (i.e. qualitatively similar), which are valuable for focusing on the quantitatively high-performing solutions that match well with the designers' preference, and for indicating new knowledge about the relations between geometries and quantitative performances; and (3) despite the similarity, they still maintain a sufficient degree of diversity, which allows the designers or decision makers to select the most satisfying solution from among them.

The comparison results concerning the behaviour of the proposed approach show that (1) the overscreening of objective variables can decrease the total number of Pareto solutions dramatically, while may not affect too much the overall performance; (2) integrating dominant qualitative goals can lead to aesthetically preferred geometries, but probably with compromise in the quantitative performances; (3) the overscreening of (Type II & III) design variables may deteriorate the quantitative performances of Pareto solutions significantly; and (4) selecting the initial generation from promising clusters is more beneficial for achieving quantitatively high-performing Pareto solutions. These results provide a good basis for the proper use of the proposed approach.

In addition, this paper also showed the suitability of the computational platform used. By combining the parametric modeling software (including simulation plug-ins) and the design optimization software, the platform can leverage their advantages, so that it can be used to create the parametric simulation model featuring a changeable design space, establish the CDE and CDO workflows, automate the geometry generation, simulation run, data collection processes, and especially, facilitate the knowledge extraction and optimization problem re-formulation (via the use of post-processing tools). Among these tools, the clustering analysis and the corresponding charts (Fig. 14.1–3) and the interface for browsing geometries and numerical data (Fig. 5, bottom) can effectively support the integration of subjective human preferences.

8.2. Future work and concluding remarks

This research could be extended in several aspects. The work presented in this paper focuses on the "variable screening" (i.e. screening out unnecessary or unimportant variables from the initial OPF, as mentioned in Section 4.1.3). But, the "variable adding" (i.e. inspiring and adding new variables that are not originally included in the initial OPF) could be another important advantage of the proposed approach, and valuable for increasing design creativity. Thus, future work should aim to expand the approach to include the variable adding, forming an iterative CDE process that can bring about new design possibilities. Furthermore, the current work compares different OPFs based on one given design concept only. To further encourage divergent thinking for the conceptual design, multiple design concepts could be compared in terms of the OPF. For this, the hierarchical structure of design variables can be helpful. For instance, a Type I design variable called "Concept" can be used to label different sets of design variables for different concepts to be compared. In addition, although the proposed approach is already applied within a rather broad context of multi-disciplinary and multi-objective design optimization, it could be still interesting to include additional objective variables (e.g. cost) and/or additional design variables (e.g. envelope attributes), achieving a more integrated design in the early design stages.

In conclusion, although various MOO methods have been developed for the conceptual architectural design, most of them can be categorized as the traditional approach where the achievement of high-performing solutions only relies on advanced optimization algorithms and their improvements. Meanwhile, the importance of the OPF is often overlooked, which can actually affect the optimization results significantly. The proposed approach provides a crucial perspective for MOO by emphasizing the CDE. During the CDE process, relevant information and knowledge are extracted to support the re-formulation of the initial optimization problem, and hence to ensure a proper OPF before the OPS. Assisted by the platform used, this approach can be applied to the conceptual architectural design optimization involving multi-disciplinary goals and complex geometries, and can achieve quantitatively more promising and qualitatively more preferred Pareto solutions. This contribution has the potential to broaden the use of MOO in the sustainable conceptual design of complex projects.

Acknowledgement

This research was supported by the Urban System and Environment (USE) Joint Research Centre between South China University of Technology (SCUT) and Delft University of Technology (TU Delft), by the Key Project of National Natural Science Foundation of China (Grant No. 51138004), the 2015 Open Project of State Key Laboratory of Subtropical Building Science, SCUT (Grant No. 2015ZA01), and the TU Delft Sports Engineering Institute Development Fund. The work on the case study was supported by Arup Amsterdam (Grant No. IiA12456). We also thank ESTECO for the collaboration based on which the GH-MF integration platform was previously developed, and for the useful discussions during the case study.

References

[1] S. Attia, M. Hamdy, W. O'Brien, S. Carlucci, Assessing gaps and needs for integrating building performance optimization tools in net zero energy buildings design, Energ. Buildings 60 (5) (2013) 110–124, http://dx.doi.org/10.1016/j.enbuild.2013.01.016.
[2] R. Evins, A review of computational optimisation methods applied to sustainable building design, Renew. Sust. Energ. Rev. 22 (6) (2013) 230–245, http://dx.doi.org/10.1016/j.rser.2013.02.004.
[3] A.T. Nguyen, S. Reiter, P. Rigo, A review on simulation-based optimization methods applied to building performance analysis, Appl. Energy 113 (1) (2014) 1043–1058, http://dx.doi.org/10.1016/j.apenergy.2013.08.061.
[4] S. Stevanović, Optimization of passive solar design strategies: a review, Renew. Sust. Energ. Rev. 25 (9) (2013) 177–196, http://dx.doi.org/10.1016/j.rser.2013.04.028.
[5] H.A. Simon, The structure of ill structured problems, Artif. Intell. 4 (3–4) (1973) 181–201, http://dx.doi.org/10.1016/0004-3702(73)90011-8.
[6] T. Smithers, W. Troxell, Design is intelligent behaviour, but what's the formalism? AI EDAM 4 (2) (1990) 89–98, http://dx.doi.org/10.1017/S0890060400002286.
[7] D. Navinchandra, Exploration and Innovation in Design: Towards a Computational Model, Springer-Verlag New York, Inc., 1991 (ISBN-13: 978-1-4612-7802-3).
[8] T. Smithers, Design as exploration: puzzle-making and puzzle-solving, https://www.academia.edu/22991882/Design_as_Exploration_Puzzle-Making_and_Puzzle-Solving , Accessed date: 24 December 2017.
[9] B. Logan, T. Smithers, Creativity and design as exploration, in: J.S. Gero, M.L. Maher (Eds.), Modeling Creativity and Knowledge-based Creative Design, Lawrence Erlbaum Associates, Inc., 1993, pp. 139–173 (ISBN: 0-8058-1153-2).
[10] W. Jonas, Design as problem-solving? Or: here is the solution—what was the problem? Des. Stud. 14 (2) (1993) 157–170, http://dx.doi.org/10.1016/0142-694X(93)80045-E.
[11] J.S. Gero, Towards a model of exploration in computer-aided design, in: J.S. Gero, E. Tyugu (Eds.), Formal Design Methods for CAD, Elsevier Science, Inc., 1994, pp. 315–336 (ISBN: 0-444-81970-3).
[12] T. Smithers, D. Corne, P. Ross, On computing exploration and solving design problems, in: J.S. Gero, E. Tyugu (Eds.), Formal Design Methods for CAD, Elsevier Science, Inc., 1994, pp. 293–313 (ISBN: 0-444-81970-3).
[13] V. Goel (Ed.), Sketches of thought, Massachusetts Institute of Technology, 1995 (ISBN 0-262-07163-0).
[14] M.L. Maher, J. Poon, S. Boulanger, Formalising design exploration as co-evolution: a combined gene approach, in: J.S. Gero, F. Sudweeks (Eds.), Advances in Formal Design Methods for CAD, Springer Science + Business Media Dordrecht, 1996, pp. 3–30 (ISBN: 978-1-4757-4428-6).
[15] M.L. Maher, J. Poon, Modeling design exploration as co-evolution, Comput. Aided Civ. Inf. Eng. 11 (3) (1996) 195–209, http://dx.doi.org/10.1111/j.1467-8667.1996.tb00323.x.
[16] G. Goldschmidt, Capturing indeterminism: representation in the design problem space, Des. Stud. 18 (4) (1997) 441–455, http://dx.doi.org/10.1016/S0142-694X(97)00011-2.
[17] H. Zhu (Ed.), Software Design Methodology: From Principles to Architectural Styles, Elsevier Ltd., 2005 (ISBN: 978-0-7506-6075-4).
[18] J.S. Gero, Creativity, emergence and evolution in design, Knowl.-Based Syst. 9 (7) (1996) 435–448, http://dx.doi.org/10.1016/S0950-7051(96)01054-4.
[19] G.S. Parnell, P.J. Driscoll, D.L. Henderson (Eds.), Decision Making in Systems Engineering and Management, John Wiley & Sons, Inc., 2011 (ISBN: 978-0-470-90042-0).
[20] Wikipedia, Occam's razor, https://en.wikipedia.org/wiki/Occam%27s_razor , Accessed date: 24 December 2017.
[21] S. Chand, M. Wagner, Evolutionary many-objective optimization: a quick-start guide, Surv. Oper. Res. Manag. Sci. 20 (2) (2015) 35–42, http://dx.doi.org/10.1016/j.sorms.2015.08.001.
[22] B. Lartigue, B. Lasternas, V. Loftness, Multi-objective optimization of building envelope for energy consumption and daylight, Indoor Built Environ. 23 (1) (2014) 70–80, http://dx.doi.org/10.1177/1420326X13480224.
[23] M. Manzan, A. Clarich, FAST energy and daylight optimization of an office with fixed and movable shading devices, Build. Environ. 113 (2) (2017) 175–184, http://dx.doi.org/10.1016/j.buildenv.2016.09.035.
[24] B.J. Futrell, E.C. Ozelkan, D. Brentrup, Bi-objective optimization of building enclosure design for thermal and lighting performance, Build. Environ. 92 (10) (2015) 591–602, http://dx.doi.org/10.1016/j.buildenv.2015.03.039.
[25] R. Kicinger, T. Arciszewski, Multiobjective evolutionary design of steel structures in tall buildings, Proceedings of the AIAA 1st Intelligent Systems Technical Conference, International Conference, Chicago, USA, 2004, http://dx.doi.org/10.2514/6.2004-6438.
[26] R. Kicinger, S. Obayashi, T. Arciszewski, Evolutionary multiobjective optimization of steel structural systems in tall buildings, in: S. Obayashi, K. Deb, C. Poloni, T. Hiroyasu, T. Murata (Eds.), Evolutionary Multi-criterion Optimization, Springer Berlin Heidelberg, 2007, pp. 604–618 (ISBN: 978-3-540-70927-5).
[27] C.T. Mueller, J.A. Ochsendorf, Combining structural performance and designer preferences in evolutionary design space exploration, Autom. Constr. 52 (4) (2015) 70–82, http://dx.doi.org/10.1016/j.autcon.2015.02.011.
[28] C. Kasinalis, R. Loonen, D. Cóstola, J.L.M. Hensen, Framework for assessing the performance potential of seasonally adaptable facades using multi-objective optimization, Energ. Buildings 79 (8) (2014) 106–113, http://dx.doi.org/10.1016/j.enbuild.2014.04.045.
[29] A.E.I. Brownlee, J.A. Wright, Constrained, mixed-integer and multi-objective

optimisation of building designs by NSGA-II with fitness approximation, Appl. Soft Comput. 33 (8) (2015) 114–126, http://dx.doi.org/10.1016/j.asoc.2015.04.010.
[30] K. Negendahl, T.R. Nielsen, Building energy optimization in the early design stages: a simplified method, Energ. Buildings 105 (10) (2015) 88–99, http://dx.doi.org/10.1016/j.enbuild.2015.06.087.
[31] J.N. Richardson, G. Nordenson, R. Laberenne, R.F. Coelho, S. Adriaenssens, Flexible optimum design of a bracing system for façade design using multiobjective genetic algorithms, Autom. Constr. 32 (7) (2013) 80–87, http://dx.doi.org/10.1016/j.autcon.2012.12.018.
[32] P. Winslow, S. Pellegrino, S.B. Sharma, Multi-objective optimization of free-form grid structures, Struct. Multidiscip. Optim. 40 (1) (2010) 257–269, http://dx.doi.org/10.1007/s00158-009-0358-4.
[33] F. Flager, B. Welle, P. Bansal, G. Soremekun, J. Haymaker, Multidisciplinary process integration and design optimization of a classroom building, J. Inf. Technol. Constr. 14 (8) (2009) 595–612 http://www.itcon.org/paper/2009/38 , Accessed date: 24 December 2017.
[34] N.C. Brown, C.T. Mueller, Design for structural and energy performance of long span buildings using geometric multi-objective optimization, Energ. Buildings 127 (9) (2016) 748–761, http://dx.doi.org/10.1016/j.enbuild.2016.05.090.
[35] M. Turrin, P. von Buelow, R. Stouffs, Design explorations of performance driven geometry in architectural design using parametric modeling and genetic algorithms, Adv. Eng. Inform. 25 (4) (2011) 656–675, http://dx.doi.org/10.1016/j.aei.2011.07.009.
[36] M. Turrin, P. Von Buelow, A. Kilian, R. Stouffs, Performative skins for passive climatic comfort: a parametric design process, Autom. Constr. 22 (3) (2012) 36–50, http://dx.doi.org/10.1016/j.autcon.2011.08.001.
[37] P. Janssen, Dexen: a scalable and extensible platform for experimenting with population-based design exploration algorithms, AI EDAM 29 (4) (2015) 443–455, http://dx.doi.org/10.1017/S0890060415000438.
[38] J.S. Gero, Design prototypes: a knowledge representation schema for design, AI Mag. 11 (4) (1990) 26–36, http://dx.doi.org/10.1609/aimag.v11i4.854.
[39] T. Smithers, A. Conkie, J. Doheny, B. Logan, K. Millington, M.X. Tang, Design as intelligent behaviour: an AI in design research programme, Artif. Intell. Eng. 5 (2) (1990) 78–109, http://dx.doi.org/10.1016/0954-1810(90)90004-N.
[40] J.S. Gero, Constructive memory in design thinking, http://papers.cumincad.org/cgi-bin/works/Show?be33 , Accessed date: 24 December 2017.
[41] J.S. Gero, U. Kannengiesser, The situated function–behaviour–structure framework, Des. Stud. 25 (4) (2004) 373–391, http://dx.doi.org/10.1016/j.destud.2003.10.010.
[42] J.S. Arora (Ed.), Introduction to Optimum Design, Fourth Edition, Elsevier Inc., 2016 (ISBN: 978-0-12-800806-5).
[43] J.S. Gero, M.L. Maher, Mutation and analogy to support creativity in computer-aided design, in: G.N. Schmitt (Ed.), CAADFutures '91 - Computer Aided Architectural Design Futures - Education, Research, Applications, Friedrich Vieweg & Sohn Verlagsgesellschaft mbH, 1991, pp. 241–249 (ISBN: 3-528-08821-4).
[44] K. Dorst, N. Cross, Creativity in the design process: co-evolution of problem–solution, Des. Stud. 22 (5) (2001) 425–437, http://dx.doi.org/10.1016/S0142-694X(01)00009-6.
[45] J.S. Gero, M.L. Maher (Eds.), Modeling Creativity and Knowledge-based Creative Design, Lawrence Erlbaum Associates Inc., 1993 (ISBN: 0-8058-1153-2).
[46] W. Tian, A review of sensitivity analysis methods in building energy analysis, Renew. Sust. Energ. Rev. 20 (4) (2013) 411–419, http://dx.doi.org/10.1016/j.rser.2012.12.014.
[47] T. Østergård, R.L. Jensen, S.E. Maagaard, Building simulations supporting decision making in early design–a review, Renew. Sust. Energ. Rev. 61 (8) (2016) 187–201, http://dx.doi.org/10.1016/j.rser.2016.03.045.
[48] P. Heiselberg, H. Brohus, A. Hesselholt, H. Rasmussen, E. Seinre, S. Thomas, Application of sensitivity analysis in design of sustainable buildings, Renew. Energy 34 (9) (2009) 2030–2036, http://dx.doi.org/10.1016/j.renene.2009.02.016.
[49] H. Shen, A. Tzempelikos, Sensitivity analysis on daylighting and energy performance of perimeter offices with automated shading, Build. Environ. 59 (1) (2013) 303–314, http://dx.doi.org/10.1016/j.buildenv.2012.08.028.
[50] Wikipedia, Pearson correlation coefficient, https://en.wikipedia.org/wiki/Pearson_correlation_coefficient , Accessed date: 24 December 2017.
[51] A.K. Jain, M.N. Murty, P.J. Flynn, Data clustering: a review, ACM Comput. Surv. 31 (3) (1999) 264–323, http://dx.doi.org/10.1145/331499.331504.
[52] L. Ricco, E. Rigoni, A. Turco, Smoothing spline ANOVA for variable screening, Dolomites Research Notes on Approximation, 6 (2013), pp. 130–139, http://dx.doi.org/10.14658/pupj-drna-2013-Special_Issue-13.
[53] Rhinoceros, https://www.rhino3d.com , Accessed date: 24 December 2017.
[54] Grasshopper, http://www.grasshopper3d.com , Accessed date: 24 December 2017.
[55] modeFRONTIER, http://www.esteco.com/modefrontier , Accessed date: 24 December 2017.
[56] Ladybug + Honeybee, http://www.grasshopper3d.com/group/ladybug , Accessed date: 24 December 2017.
[57] Karamba, http://www.karamba3d.com , Accessed date: 24 December 2017.
[58] Daysim, http://daysim.ning.com , Accessed date: 24 December 2017.
[59] P.R. Tregenza, I.M. Waters, Daylight coefficients, Light. Res. Technol. 15 (2) (1983) 65–71, http://dx.doi.org/10.1177/096032718301500201.
[60] R. Perez, R. Seals, J. Michalsky, All-weather model for sky luminance distribution—preliminary configuration and validation, Sol. Energy 50 (3) (1993) 235–245, http://dx.doi.org/10.1016/0038-092X(93)90017-I.
[61] C.F. Reinhart, O. Walkenhorst, Validation of dynamic RADIANCE-based daylight simulations for a test office with external blinds, Energ. Buildings 33 (7) (2001) 683–697, http://dx.doi.org/10.1016/S0378-7788(01)00058-5.
[62] EnergyPlus, https://energyplus.net/ , Accessed date: 24 December 2017.
[63] D. Yang, Y. Sun, M. Turrin, P. von Buelow, J.C. Paul, Multi-objective and multidisciplinary design optimization of large sports building envelopes: a case study, Proceedings of the International Association for Shell and Spatial Structures Symposium 2015, International Conference, Amsterdam, The Netherlands, 2015, https://repository.tudelft.nl/islandora/object/uuid:86cfb715-8f7a-46a0-8e29-1a216a0441b1/?collection=research , Accessed date: 24 December 2017.
[64] R. Šileryte, A. D'Aquilio, D. Di Stefano, D. Yang, M. Turrin, Supporting exploration of design alternatives using multivariate analysis algorithms, in: R. Attar, A. Chronis, S. Hanna, M. Turrin (Eds.), 2016 Proceedings of the Symposium on Simulation for Architecture and Urban Design (SimAUD 2016), International Conference, London, UK, 2016 (ISBN: 978-1-365-05872-1).
[65] D. Yang, Y. Sun, D. Di Stefano, M. Turrin, S. Sariyildiz, Impacts of problem scale and sampling strategy on surrogate model accuracy: an application of surrogate-based optimization in building design, Proceedings of 2016 IEEE Congress on Evolutionary Computation (CEC 2016), International Conference, Vancouver, Canada, 2016, http://dx.doi.org/10.1109/CEC.2016.7744323 (ISBN: 978-1-5090-0623-6).
[66] D. Yang, Y. Sun, D. DI Stefano, M. Turrin, A computational design exploration platform supporting the formulation of design concepts, in: M. Turrin, B. Peters, W. O'Brien, R. Stouffs, T. Dogan (Eds.), 2017 Proceedings of the Symposium on Simulation for Architecture and Urban Design (SimAUD 2017), International Conference, Toronto, Canada, 2017 (ISBN: 978-1-365-88878-6).
[67] D. Yang, M. Turrin, S. Sariyildiz, Y. Sun, Sports building envelope optimization using multi-objective multidisciplinary design optimization (M-MDO) techniques: case of indoor sports building project in China, Proceedings of 2015 IEEE Congress on Evolutionary Computation (CEC 2015), International Conference, Sendai, Japan, 2015, http://dx.doi.org/10.1109/CEC.2015.7257165 (ISBN: 978-1-4799-7492-4).
[68] Y. Sun, L. Xiong, P. Su, Grandstand grammar and its computer implementation, in: R. Stouffs, S. Sariyildiz (Eds.), eCAADe 2013 Computation and Performance - Proceedings of the 31st International Conference on Education and Research in Computer Aided Architectural Design in Europe (eCAADe 2013), International Conference, Delft, The Netherlands, 2013 (ISBN: 978-94-91207-04-4).
[69] EnergyPlus, Weather data sources, https://energyplus.net/weather/sources , Accessed date: 24 December 2017.
[70] Bigladdersoftware, Group – zone forced air units, http://bigladdersoftware.com/epx/docs/8-0/input-output-reference/page-032.html , Accessed date: 24 December 2017.
[71] Standard for daylighting design of buildings (GB 50033-2013), http://www.codeofchina.com/standard/GB50033-2013.html , Accessed date: 24 December 2017.
[72] Standard for lighting design of buildings (GB 50034-2013), http://www.codeofchina.com/standard/GB50034-2013.html , Accessed date: 24 December 2017.
[73] Design code for sports building (JGJ 31-2003), http://codeofchina.com/standard/JGJ31-2003.html , Accessed date: 24 December 2017.
[74] Design standard for energy efficiency of public buildings (GB50189-2015), http://www.codeofchina.com/standard/GB50189-2015.html , Accessed date: 24 December 2017.
[75] Code for thermal design of civil building (GB 50176-2016), http://www.codeofchina.com/standard/GB50176-2016.html , Accessed date: 24 December 2017.
[76] Graduations and test methods of air permeability, watertightness, wind load resistance performance for building external windows and doors (GB/T 7106-2008), http://www.codeofchina.com/standard/GBT7106-2008.html , Accessed date: 24 December 2017.
[77] G. Ramos, E. Ghisi, Analysis of daylight calculated using the EnergyPlus programme, Renew. Sust. Energ. Rev. 14 (7) (2010) 1948–1958, http://dx.doi.org/10.1016/j.rser.2010.03.040.
[78] E.L. Didoné, F.O.R. Pereira, Integrated computer simulation for considering daylight when assessing energy efficiency in buildings, Proceedings of Building Simulation 2011 - 12th Conference of International Building Performance Simulation Association, International Conference, Sydney, Australia, 2011, http://www.ibpsa.org/?page_id=105 , Accessed date: 24 December 2017.
[79] A. Nabil, J. Mardaljevic, Useful daylight illuminances: a replacement for daylight factors, Energ. Buildings 38 (7) (2006) 905–913, http://dx.doi.org/10.1016/j.enbuild.2006.03.013.
[80] ESTECO SpA, modeFRONTIER user guide, http://www.esteco.com/modefrontier , Accessed date: 3 September 2018.
[81] P. Depecker, C. Menezo, J. Virgone, S. Lepers, Design of buildings shape and energetic consumption, Build. Environ. 36 (5) (2001) 627–635, http://dx.doi.org/10.1016/S0360-1323(00)00044-5.
[82] F. Qi, Y. Wang, A new calculation method for shape coefficient of residential building using Google Earth, Energ. Buildings 76 (6) (2014) 72–80, http://dx.doi.org/10.1016/j.enbuild.2014.02.058.
[83] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Comput. 6 (2) (2002) 182–197, http://dx.doi.org/10.1109/4235.996017.