# Design exploration of quantitative performance and geometry typology for indoor arena based on self-organizing map and multi-layered perceptron neural network

Pan, Wang; Sun, Yimin; Turrin, Michela; Louter, Christian; Sariyildiz, Sevil

**DOI**
10.1016/j.autcon.2020.103163

**Publication date**
2020

**Document Version**
Final published version

**Published in**
Automation in Construction

Automation in Construction 114 (2020) 103163

Contents lists available at ScienceDirect

Automation in Construction

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

# Design exploration of quantitative performance and geometry typology for indoor arena based on self-organizing map and multi-layered perceptron neural network

Wang Pan[a,b,c], Yimin Sun[a,b,*], Michela Turrin[c], Christian Louter[d], Sevil Sariyildiz[c]

[a] School of Architecture, South China University of Technology, Guangzhou, China
[b] State Key Laboratory of Subtropical Building Science, South China University of Technology, Guangzhou, China
[c] Chair of Design Informatics, Department of Architectural Engineering and Technology, Delft University of Technology, Delft, the Netherlands
[d] Institute of Building Construction, Faculty of Civil Engineering, Technische Universität Dresden, Dresden, Germany

ARTICLE INFO

*Keywords:*
Comprehensive design exploration
Complex indoor arena
SOM-LLM (local linear mapping based on self-organizing map)
SOM-MLPNN (multi-layer perceptron neural network based on self-organizing map)

ABSTRACT

During the early design process, simulations allow numeric assessment and 3D models allow visual inspection for qualitative criteria. However, exploring different design alternatives based on both is challenging. To support the design exploration of quantitative performance and geometry typology of various design alternatives during the early design stages of indoor arenas, this paper proposed a novel design method of SOM-MLPNN by combing self-organizing map (SOM) and multi-layer perceptron neural network (MLPNN), based on the inspiration of local linear mapping based on self-organizing map (SOM-LLM). In SOM-LLM or SOM-MLPNN, the SOM can support designers to explore the whole design space according to geometry typologies and provides reference/labelled inputs for LLM/MLPNN to approximate multiple quantitative performance data for various design alternatives. Both SOM-LLM and SOM-MLPNN are applied and compared in a design of indoor arena. Besides the development of the method, original contributions include 1) proposing two operations (using a large size of SOM network and using a small amount of input data to train the SOM network) to save the computational time and increase the accuracy in data approximation and 2) proposing a series of data visualizations to interpret the results and support design explorations in different ways.

## 1. Introduction

Conceptual design, which is the early stage of the whole architectural design process, aims to generate promising concepts which satisfy a series of design requirements and can be developed in the following design processes [1]. These design requirements include both 1) quantitative design requirements which can be measured and assessed based on numeric data and are usually related to architectural functionality and engineering and 2) qualitative design requirements which are difficult to be measured and assessed based on data and are usually related to humanity and social science (e.g. aesthetics, culture, politics, etc.). So far, for qualitative design requirements, designers still tend to evaluate the overall form/geometry of the designs according to their knowledge and experience, based on visual inspection. To generate promising designs, conceptual design is usually performed in two steps: divergent step in which various concepts are generated (B in Fig. 1) and convergent step in which one or several concepts are selected (D in Fig. 1) [2]. To progress across the two steps, the information (related to both quantitative and qualitative design requirements) of the numerous design alternatives should be rapidly obtained and organized in an effective way, based on which designers can perform a design exploration to investigate and assess design alternatives (C in Fig. 1).

This process is especially important for the conceptual design of indoor sports arenas. For such building, during the conceptual design process, it is crucial to integrate the multi-functional space and long-span roof structure and to formulate proper building geometry, since these two elements are highly interrelated and determine the overall form of the building [3]. This process involves complex and challenging decision-making, which demands adequate and effectively-organized information of various design alternatives for designers to perform design exploration. The information includes multiple kinds of performance data related to quantitative design requirements (e.g. the requirements on spectators' views, acoustics, structural performance, etc.) as well as the overall form/geometries for visual inspection based on which designers can assess the designs alternatives according to

* Corresponding author at: School of Architecture, South China University of Technology, Guangzhou, China.
*E-mail address:* arymsun@scut.edu.cn (Y. Sun).

https://doi.org/10.1016/j.autcon.2020.103163
Received 22 July 2019; Received in revised form 28 February 2020; Accepted 5 March 2020

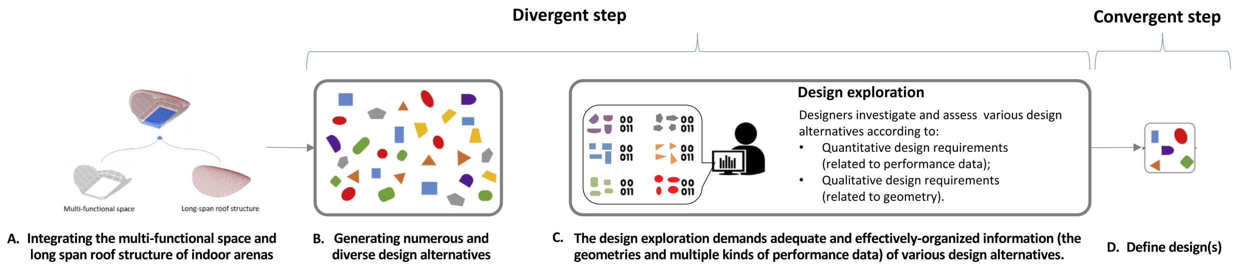



**Fig. 1.** A process of the conceptual design of indoor arenas with an emphasis on the integration of the multi-functional space and long-span roof structure.

qualitative design requirements (e.g. aesthetics). Moreover, the numeric and the visual information should be organized and visualized effectively, since it is difficult for human designers to deal with mass of information. Fig. 1 illustrates a process of architectural conceptual design for indoor arenas with an emphasis on the integration of the multi-functionality and long-span roof structure.

In this light, an efficient design method is needed to satisfy the aforementioned demands, therefore, to support the design exploration of indoor arenas, which is also the motivation of this paper. Nowadays, several computational design methods have been used to support design exploration of architectural conceptual design, including multi-objective optimization (MOO) [3–12], surrogate model based on supervised learning [13–16], and unsupervised clustering based on self-organizing map (SOM) [17–19]. Fig. 2 demonstrates the overall workflows of these methods. However, there are still limitations for these methods in satisfying the aforementioned demands and supporting the conceptual design of indoor arena with emphasis on the integration of multi-functional space and long-span roof structure.

In these methods, a parametric model should be firstly formulated based on the basic spatial composition of an indoor arena, in which various elements of the building are associated and controlled by parameters. By changing the values of the parameters (design inputs), various designs can be generated to compose a design space. Specific performance data (related to quantitative design requirements) of the designs can be obtained by specific building performance simulations. However, since the simulations are usually time-consuming, it is unpractical to use them to obtain the performance data of numerous designs. Based on parametric model and simulations, these methods support design explorations in different ways:

- MOOs iteratively search for 'well-performing' designs within the design space according to specific criteria by using a certain heuristic algorithm (e.g. genetic algorithm). However, a standard MOO only provides the 'well-performing' designs to designers. Besides, in general, MOOs can efficiently deal with the problems in which the design objectives are not more than three, but when the objectives are more than three, it is difficult to find optimal solutions [20].
- Surrogate models based on supervised learning can learn the relationships between the design inputs and performance data. They are used to rapidly approximate the performance data for numerous designs. Therefore, based on surrogate model and parametric model, it is possible to obtain both the geometries and performance data of

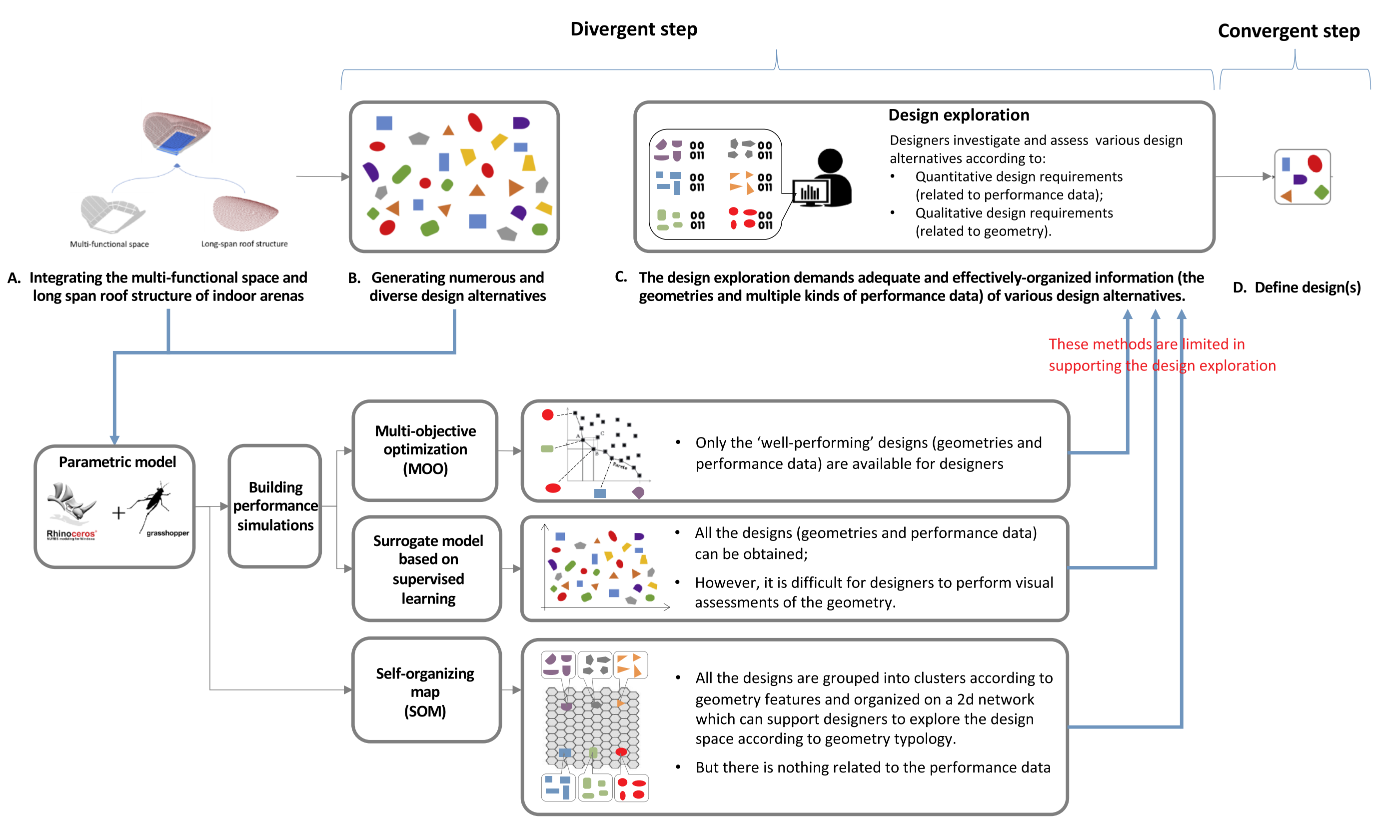



**Fig. 2.** The general workflows of three computational methods and their limitations in supporting the design exploration of indoor arenas with emphasis on the integration of the multi-functional space and long-span roof structure.

all the design alternatives within a discrete design space. However, it is unpractical for designers to investigate so many designs. It is necessary to efficiently organize the information about the geometries and performance data (related to quantitative aspects) of the numerous designs and to demonstrate the relationships between them.

- SOMs are used in design exploration to group numerous design alternatives into clusters according to their geometry features (indicated by the parameters/design inputs) and generated a node design for each cluster to represent all the designs within the cluster. Moreover, all the node designs are organized by a two-dimensional network and similar ones are close while different ones are far away, which can reflect the design space (but the effect of the reflection can become weak, as the dimensions of the design space increase, since the curse of dimensionality). Therefore, by using SOMs, designers can investigate various types of designs alternatives according to geometry features. However, this process does not deal with performance data (related to quantitative aspects) of designs.

Although there are limitations for these methods, combining the surrogate model based on supervised learning and SOM can be a way to overcome the limitations. Accordingly, this paper proposes a novel design method based on their combination. Among various supervised learning methods supporting surrogate model, this paper focuses on local linear mapping based on self-organizing map (SOM-LLM) and multi-layer perceptron neural network (MLPNN). For SOM-LLM, SOM has been combined with the supervised learning method of LLM [21,22], which can be directly used to support the aforementioned design exploration (Fig. 3) [23]. Besides, since MLPNN has been widely used in various fields for its capability of universal approximation [24], in this paper, MLPNN is combined with SOM to formulate SOM-MLPNN to support the aforementioned design exploration (Fig. 3). Specifically, during the development of the method, a series of challenges need to be overcome, including: 1) coupling the SOM and MLPNN to fulfill clustering and data approximation and verifying the effects, 2) saving the computational time of SOM and ensuring the accuracy of data approximation, and 3) visualizing the results and facilitate designers to explore the design space according to both quantitative and qualitative design requirements, 4) compare SOM-LLM and SOM-MLPNNN to determine the final method.

The workflows of both SOM-LLM and SOM-MLPNN are independent. The related aspects are reviewed in Section 2, and the workflows are elaborated in Section 3 and are applied in case studies in Section 4. The results of the case studies related to SOM-LLM and SOM-MLPNN are discussed and compared to define a final method in Section 5.

In this paper, the proposed method based on SOM-LLM or SOM-MLPNN is specifically developed with focus on the design exploration of the integrated design of indoor arena, in which the design inputs are directly related to the overall geometry of the building. In its current state, this method is limited for the designs of other types of buildings and is also limited for the studies of various parameters/design inputs which are not directly related to the overall form but are important for the building performance. Nevertheless, potentially the method can be generalized, despites that there are a series of challenges. The limitations and the challenges in the generalization of this method are discussed in Section 5.

## 2. Background information

### 2.1. Obtaining performance data for numerous designs by surrogate model

A surrogate model can approximate a high-fidelity but time-consuming function in reasonable accuracy, based on sampled/labelled data obtained by design of experiments (DoE) of the high-fidelity function [15,25]. This method has been applied in building designs [13–16].

The formulation of a surrogate model can be considered as a process

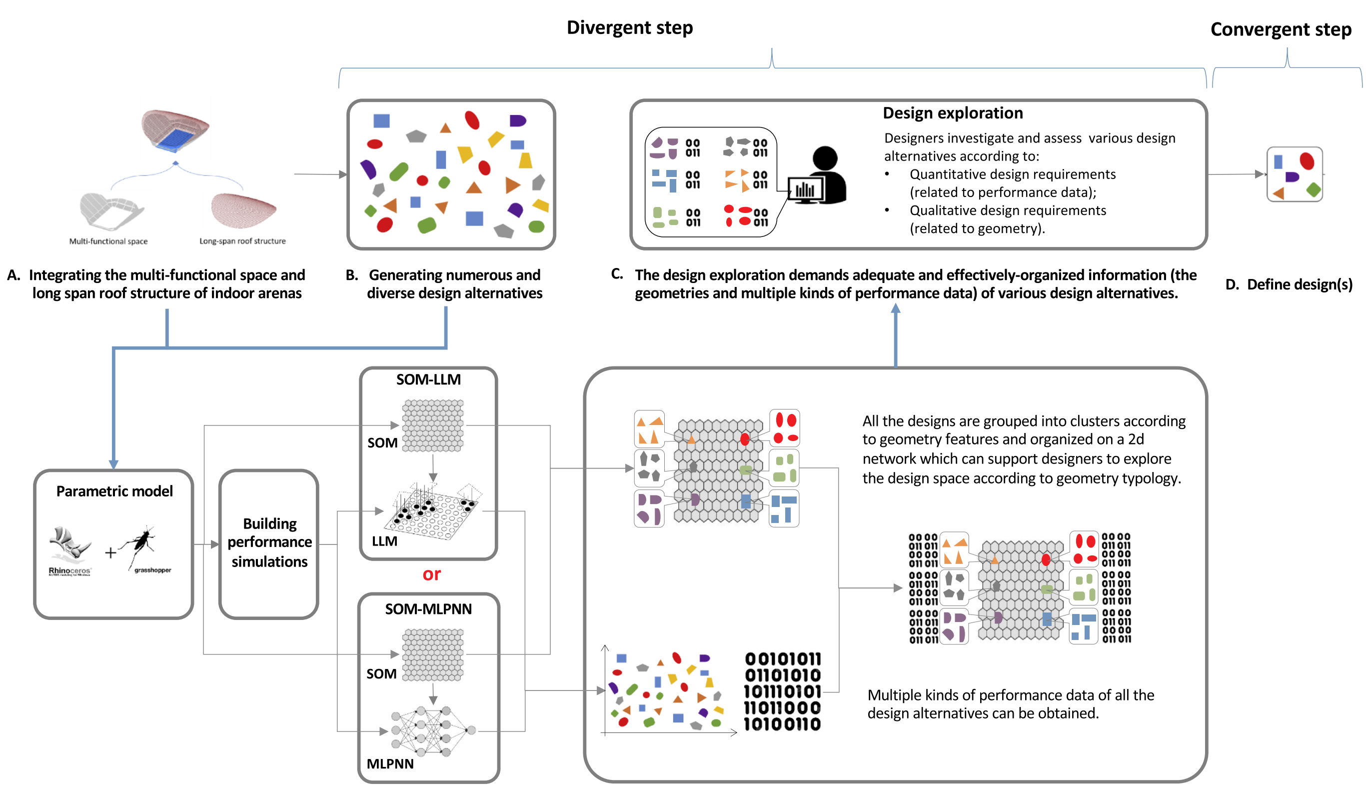



**Fig. 3.** The scheme of using SOM-LLM or SOM-MLPNN to support design exploration of indoor arena with emphasis on the integration of multi-functional space and long-span roof structure.

of supervised learning. Supervised or predictive learning is a process that learns a mapping between the inputs and outputs of a system, based on a labelled set of input-output pairs [26]. In general, the process is fulfilled in five steps: 1) formulating an initial model; 2) sampling the input space to obtain a number of labelled inputs; 3) obtaining the labelled outputs (corresponding to the labelled inputs) by design of experiments (DoE); 4) training, validating, and testing the model by the labelled data; 5) using the tested model to approximate new data.

There are various supervised learning methods to achieve a surrogate model. Some of the methods have been used in engineering or architectural designs, including poly-nominal regression and response surface method (RSM) [16,27–29], multi-layer perceptron neural network (MLPNN) [15,29] [30,31], random forest (RF) [15,32,33], radial basis function network (RBFN) [14,15], kriging [15,34,35]. In [15], these methods are used to support surrogate models for a long-span building design focusing on structural self-weight and energy. The results in [15] indicated that the MLPNN has the fastest speed and smallest errors in the data approximation of structural weight and energy consumption for the design example.

#### 2.1.1. Multi-layer perceptron neural network (MLPNN)

A MLPNN is composed of neural networks of an input layer and an output layer as well as one or multiple hidden layers between them [36]. The input layer is related to the input data and the number of the neurons equals the dimensions of the input data, while the output layer is related to the output data. Between them, one or multiple hidden layers connecting the input and output layers are used to learn a mapping between the inputs and outputs according to labelled data. The connection between the neurons of two adjacent layers is based on the calculation related to activation function, bias, and the weighted sum of the values of the anterior neural layer [26,36].

A MLPNN model learns the mapping between the inputs and outputs by adjusting the weights and bias for each neuron to minimize the error function, which is obviously an optimization process. The error function is usually the mean squared error (MSE). Back propagation is used to feed back the error to the neural networks, therefore, can accelerate and improve the optimization process [37,38]. The details of MLPNN can be found in [26,36].

For MLPNNs, the capability of universal approximation is verified in [24] and they also used as universal function approximators in recent years [36]. Specially, for the building designs, MLPNNs are widely used for the predictions of energy consumptions [39–43], structural analysis and design [13,44,45], and integrated design [11]. It worth noting that for the applications of MLPNN, the structures of neural networks (the number of the hidden layers, the amount of the neurons on each hidden layer) can be different. In fact, to define a proper network to obtain promising performance of data approximation is one of the main challenges of the applications of MLPNNs, and using growing neural networks as well as pruning technique are two of the ways to find proper networks for specific problems [36,46]. Besides, the uncertainty of MLPNNs in data approximation is another main challenge of the applications of MLPNNs, which includes input uncertainty, parameter uncertainty, and structure uncertainty [47]. A series of methods are proposed to quantify the uncertainties, therefore, to help users to evaluate the networks [47].

#### 2.1.2. Local linear map based on self-organizing map (SOM-LLM)

Besides the methods mentioned above, interpolation is also used for data approximation to achieve surrogate model, for its simplicity [23], ability in limiting interference [23,48], and transparency. Among various interpolation methods, local linear mapping based on self-organizing map (SOM-LLM) is considered to have good accuracy and take less computation resource [23,49]. Comparing with MLPNNs, a two-dimensional problem had been used as an example to verify that the SOM-LLM has equivalent performance in functional approximation [21]. More importantly, in SOM-LLM, self-organizing map (SOM) proposed by Kohonen [50,51] can group data objects according to their similarity in specific features. Based on this capability, SOM has been used to support design explorations of geometry typology [17–19] (the details are elaborated in Section 2.2).

In general, a linear interpolation inserts new data points between some known reference data points (sampled/labelled data), to calculate the outputs of the new interpolated points according to the distance between the inputs of the interpolated and the reference data points. However, the related errors can be quite large if the problem is complex. To decrease the errors, there are other methods applied in different fields to meet specific requirements, including polynomial interpolation, splines and B-splines techniques, Kriging, and natural neighbor method using Voronoi tessellation. Nevertheless, most of these methods are computationally expensive and are limited in dealing with high-dimensional problems [17].

To increase the accuracy of data approximation and avoid expensive computation, local linear mapping based on self-organizing map (SOM-LLM) was proposed in [21,23,49] based on the original local linear mapping (LLM). LLM is formulated based on linear interpolation and additional weights [48]. In LLM, an interpolated data point is considered to be mainly related to the nearby reference data points. Hence, in order to save the computation time, the output of an interpolated data point is calculated only based on its nearby reference data points. To find these nearby reference data points, self-organizing map is used [23].

In a SOM, a network (usually two-dimensional) with nodes/neurons is predefined in the input space of the interpolated data points. Each of the interpolated data points is captured iteratively by one of the nodes/neurons on the network and is grouped into a cluster represented by the node/neuron, according to the SOM algorithm (details can be found in [50,51]). For each cluster, the related node can be considered as the reference data point for LLM. Therefore, for a certain interpolated data point, the node of its cluster and the nodes of the neighbouring clusters can be considered as the nearby reference data points. The inputs and outputs (obtained by design of experiments) of these reference data points then are used to calculate the output of this interpolated data point, according to the algorithm of LLM (details can be found in [23]).

### 2.2. Exploring designs according to geometry typologies by self-organizing map (SOM)

As mentioned above, a SOM can group data points into different clusters according to their similarity measured by a distance function (e.g. Euclidean distance). Each of the clusters is represented by a node/neuron organized by a two-dimensional network. On the network, similar nodes are close while the different ones are far away, which can be used to reflect the intrinsic topology of the data set [52,53].

SOMs have been applied in architectural design to support designers to explore the design alternatives within design space according to geometry types [17–19]. In these applications, the design parameters (related to the building geometry of each design alternative within the design space) are used as the inputs to train a predefined SOM network. Therefore, the design alternatives with similar geometries are grouped in the same clusters.

For each cluster, the vector of the node is provided by the SOM, based on which the geometry of the node design can be generated by the parametric model. Each node design represents all the design alternatives within the related cluster. Moreover, all the node designs are presented on the SOM network. On this network, similar node designs are close and different ones are far away, which can reflect the design space. Therefore, designers can have a quick glimpse of the whole design space and explore all the design alternatives according to geometry typology [17].

It worth noting that the process of SOM clustering can be also considered as a dimensionality reduction in which a high-dimensional design space is projected on a two-dimensional network [52,53]. As the

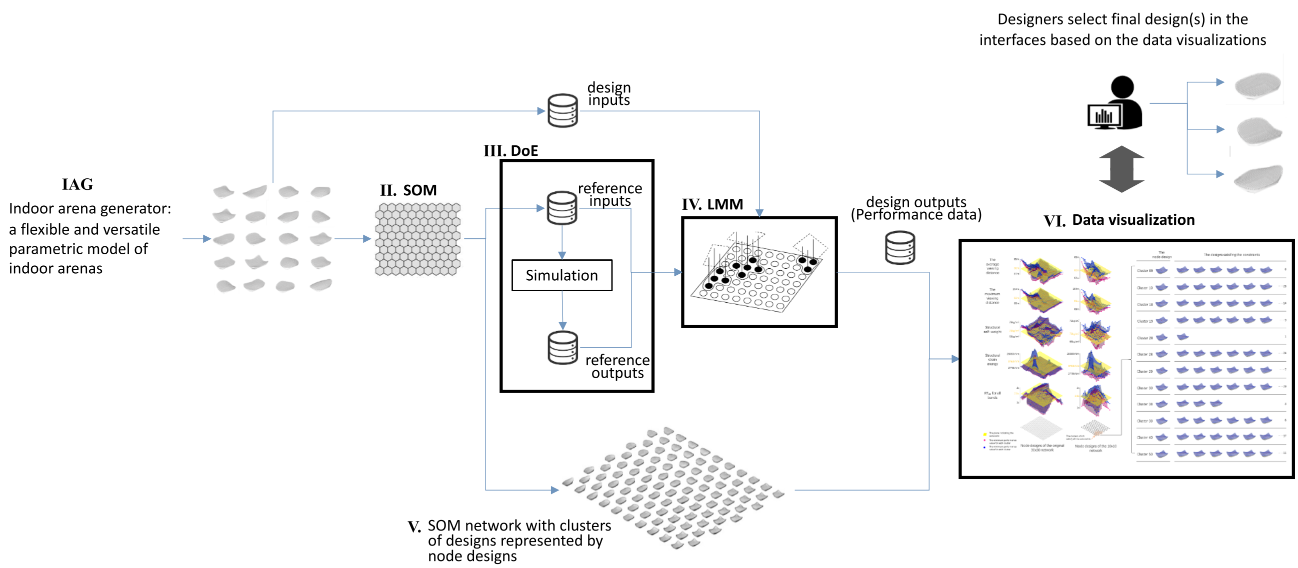



**Fig. 4.** The workflow of SOM-LLM.

dimensions of the design space increase, it becomes difficult for a SOM network to reflect the original design space, since the curse of dimensionality. Hence, when using SOM to support design exploration, the dimensions of the design space (equalling the dimensions of the design inputs) should be limited. In [17], an experiment shows that a nine-dimensional design space can be effectively reflected by a two-dimensional network.

## 3. Methodology

This section elaborates the workflows of LLM-SOM (Fig. 4) and MLPNN-SOM (Fig. 5). These workflows are similar, except the step IV in which LLM and MLPNN are respectively used for data approximations. The IAG used in this method is proposed in [3] and based on the software of Rhinoceros 3D [54] and its plugin grasshopper [55], the simulation of structure is based on Karamba3D [56], a plugin of Rhinoceros 3D. The SOM is based on the toolbox of self-organizing maps in MATLAB [57], the LLM is achieved by the codes written by the authors in MATLAB [58], and the MLPNN is based on the toolbox of feedforward neural network in MATLAB [59]. The details of the method are elaborated in the following subsections.

### *3.1. Defining design space based on indoor arena generator (IAG)*

In step I of the two workflows, indoor arena generator (IAG, a versatile and flexible parametric model of indoor arena) is used. The IAG, which is proposed by the authors in [3], can integrate the multi-functional space and long-span roof structure of indoor arena and generate various design alternatives. In [3], thirty main design parameters of IAG are listed, designers can select some of them as design variables. The values of the variables can be changed to generate various designs while other design parameters are fixed in specific values. Based on the definition of design variables (defining the interval and range for each variable), a specific design space is formulated, which includes a limited number of design alternatives. In the following steps of both LLM-SOM and MLPNN-SOM, these design variables are used as the design inputs for the generation of the geometries of design alternatives, the data approximations of multiple performance data, and the clustering of designs according to geometry typology.

The amount of the design alternatives within the design space is determined by the amount, intervals, and ranges of the design variables. The amount of the variables equals the dimensions of the design space. In practice, the values of design variables are usually discrete,

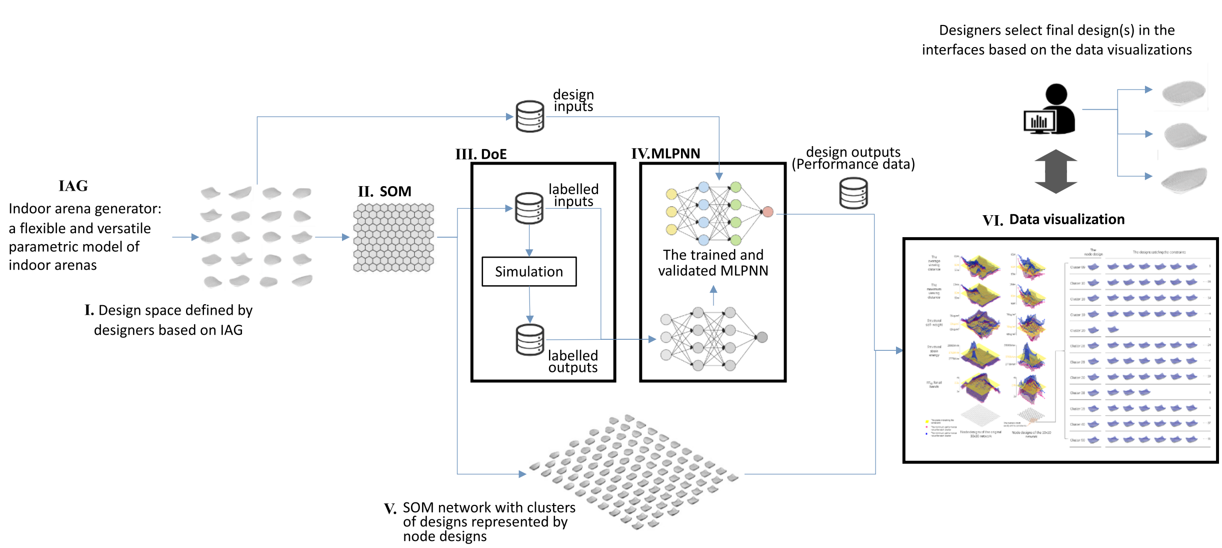



**Fig. 5.** The workflow of SOM-MLPNN.

and designers can define the interval for each variable within a specific range. The ranges of variables are usually defined according to the regulations of urban planning, design codes, scale of the building, etc.

This paper focuses on the overall building forms/geometries of indoor arenas (which are generated by integrating the multi-functional spaces and long-span roof structures) and the relationships between the overall forms and the building performance (related to quantitative requirements). Therefore, only the design parameters directly related to the overall form of the building are selected as design variables in the proposed method (the parameters in step 2 and 3 in the table 2 in [3]). Since the design variables are used as the design inputs for the SOM to cluster design alternatives according to their geometry features, it can influence the effect of the clustering if the design variables which are not directly related to the overall form of the building (e.g. the cross sections of the structural elements). However, the parameters which are not directly related to geometry can be important for building performance. Hence, it is one of the limitations of the proposed method to exclude these design parameters during design exploration. Besides, as mentioned in sub-section 2.2, since the effect of SOM clustering in reflecting the design space can reduce as the dimensions of the design space increase (since the curse of dimensionality), in this method, the amount of the design variables (which equals the dimensions of the design space) should be limited within nine (an experiment in [17] shows a nine-dimensional design space can be effectively reflected by a SOM network). A definition of design variables based on IAG is demonstrated in the case studies in Section 4.

### 3.2. SOM (self-organizing map) clustering

In step II of the two workflows, the defined design variables of the design alternatives are used as design inputs for SOM clustering which aims to 1) sample reference/labelled data for DoE and LLM/MLPNN by using the nodes of SOM network and 2) cluster all design alternatives into groups, according to their geometry features indicated by the design variables. To use SOM clustering, it is necessary to define the size of the SOM network (how many rows and columns of nodes on the network) and the amount of the design inputs which are used to train the SOM network.

For the definition of the size of the SOM network, there is a trade-off. A large size of network contains more nodes, which increases the computation time of SOM as well as the processing time of the design of experiments (DoE) in step III and the data approximation of LLM/MLPNN in step IV. Moreover, it also influences the exploration of designs about their geometries in step V, since it is difficult for designers to investigate too many node designs on the SOM network. Nevertheless, a large network generating more node designs, which provides more reference/labelled data for LLM/MLPNN, therefore, can potentially increase the accuracy of data approximation. This paper applies both a large and a small SOM network for SOM-LLM and SOM-MLPNN in the case studies in Section 4 to make a comparison, then users can select a proper one according to the results. An operation is used to solve the problem that too many node designs on a large size of SOM network impeding designers to explore different types of designs. In this operation, a large network can be shrunk by simply combing several adjacent clusters into a new bigger cluster represented by a new node design, which makes the network flexible and decreases the number of the node designs (details can be found in Section 4.2).

For the amount of the design inputs which are used to train the SOM network, there is also a trade-off. Using a large number of inputs increase the computation time of SOM, while using a small number of inputs can save the time, but it may influence the effect of SOM in sampling the design space, therefore influences the accuracy of data approximation. This paper uses a large and a small sets of design inputs (which are generated by setting different intervals for each variables) to train the SOM network in the case studies in Section 4 to make a comparison (see Table 1 in Section 4).

These two operations (using SOM-networks in different sizes and using different amount of design inputs to train the SOM-networks) are applied in cases studies to make comparison. For clustering, normalization of the design inputs is necessary. This paper uses a min-max normalization (or 0–1 normalization), to project all the design inputs into the range between 0 and 1.

### 3.3. Design of experiments (DoE)

In step III of the two workflows, design of experiments (DoE) are used to obtain the outputs for the reference/labelled inputs, therefore, to compose the reference/labelled data which are used for data approximation in step IV. In DoE, first, the reference/labelled inputs (which are the vectors of the node designs generated by SOM clustering in step II) are used to generate related geometries of the node designs based on IAG (the parametric model). Then multiple kinds of performance data of these geometries (which are the labelled/reference outputs) can be obtained by specific building performance simulations. In [3], a framework of performance assessments criteria with related simulation tools for indoor arenas are proposed. The framework, which focuses on the capacity of multiple activities, spectators' viewing, acoustics, and structural performance of indoor arena, contains various indicators and the related simulation tools. Based on the framework, designers can select multiple indicators to assess the design alternatives. Correspondingly, the related simulations tools are used in DoE to obtain the reference/labelled outputs.

### 3.4. Data approximation based on LLM and MLPNN

In step IV, based on the labelled data obtained by design of experiments (DoE), data approximation model is trained to predict the performance data of all the design alternatives. LLM and MLPNN are used to perform data approximation in step IV of SOM-LLM and SOM-MLPNN, respectively. The results are compared in the case studies (Section 4.1).

For MLPNN, the composition of the networks and the activation function applied for each neuron are usually predefined empirically, which can impact the performance of the approximation results. This paper sets three hidden layers of networks (6-6-10) between the input and output layers and applies sigmoid and Levenberg-Marquardt (a backpropagation optimization algorithm) as the activation function and training function, respectively. Mean squared errors (MSE) is used as the cost/error function which should be minimized during the training process. This MLPNN model is also used in [15] for a long-span building to predict structural self-weight and energy consumption. To overcome the problems related to overfitting and generalization, a validation and test process is used. The labelled data are divided into three sets: training set, validation set, and test set. Within an iteration of the training process, the training set is used to train the MLPNN model and the validation set are used to validate the trained model by measuring the difference of the two errors obtained by these two data sets. The iterations will stop until both the errors (related to the training and validation sets, respectively) of the trained model and the difference between these two errors are small enough. After iterations, the final trained model will be tested by the test set. If the error is also accepted, then the trained model will be used for data approximation. It is worth noting that one MLPNN model can be only trained for one kind of performance data. To predict multiple kinds of performance, the same number of MLPNN models are needed.

Differing from MLPNN, LLM approximates data according to the reference data and the distribution of the input data space (reflected by the trained SOM network). Therefore, there is no training and validation process for LLM, the calculation is mainly related to the input data. The details of the calculations are elaborated in [23]. It is worth noting that, most of the calculations of SOM-LLM (the calculations of equations 4 to 5 in [23]) deal with the relationship between the reference inputs

**Table 1**
Design parameters and performance indicators of an indoor arena (with 14,000 to 15,000 fixed seats).

| | | | |
|---|---|---|---|
| Design parameters (variable) | L-X | Length in X-axis: 80 m to 132 m; | Interval: 2 m (for the large set of inputs), 4 m (for the small set of inputs); |
| | L-Y | Length in Y-axis: 94 m to 166 m; | Interval: 2 m (for the large set of inputs), 4 m (for the small set of inputs); |
| | Cp | Corner position: 0 to 10; | Interval: 1 (for the large set of inputs), 2 (for the small set of inputs); |
| | CenH-roof | Height of the headroom of the centre of the pitch: 18 m to 40 m; | Interval: 1 m (for the large set of inputs), 2 m (for the small set of inputs); |
| Design parameters (fixed) | Cuv-BO | Curve type of the building outline: 3 (curve); | |
| | Cuv-X-roof | Curve type of the roof in X axis: 3 (curve); | |
| | Cuv-Y-roof | Curve type of the roof in Y axis: 3 (curve); | |
| | H-CPi-bdr | The height of the ith control point of the structural boundary: 0 | |
| | StruType | Structural type: SF (space frame); | |
| | GridSize-roof | Size of the grid: 4 m; | |
| | StruDpth-ctr | Structural depth in the centre: 2 m; | |
| | StruDpth-bdr | Structural depth on the boundary: 2 m; | |
| | Cross-section | Cross-section of structural elements: - The shape of cross-section: circle hollow; - Upper and bottom chords: diameter of 200 mm and thickness of 12 mm; - Web: diameter of 80 mm and thickness of 8 mm; - - Material: S355 (steel). | |
| Performance indicators | $VD_{avr\text{-}p}$ | Average viewing distance (m) | Obtained by the measurement in Rhino [54] and Grasshopper [55] |
| | $VD_{max\text{-}p}$ | Maximum viewing distance (m) | Obtained by the measurement in Rhino [54] and Grasshopper [55] |
| | SW | Structural self-weight (kg/m$^2$) | Obtained by Karamba3D [56] |
| | SE | Strain energy (kN·m) | Obtained by Karamba3D [56] |
| | $RT_{60}$ | The reverberation time of all octave band frequencies (s) | Obtained by Sabine equation [60] |
| | $RT_{60\text{-}1K}$ | The reverberation time of octave band frequency in 1 k Hz (s) | Obtained by Sabine equation [60] |

and the interpolated inputs (the design inputs in this method), which means for the predictions of different kinds of performance values for a certain design, the calculations are the same, except the last step of the process (the calculations of equation 1 in [23]).

### 3.5. Organizing different types of designs by SOM network

Based on the SOM clustering in step II, all the design alternatives within the design space can be grouped into clusters according to their similarity in geometry features (indicated by the design variables) and are represented by the related node designs. In step V, the geometries of the node designs are generated by using the parametric model (IAG) and are organized by the SOM network on which similar ones are close while different ones are far away. Therefore, designers can explore different types of designs and have an overview of the whole design space. Moreover, for each cluster, besides the node design, the geometries of the design alternatives within the cluster can be also generated by parametric model (IAG), which can be used as an index system to search for various designs.

### 3.6. Design exploration based on data approximation

Based on the performance data of all the design alternatives obtained in step IV and the node designs on the SOM network obtained in step V, a series of data visualizations are proposed (based on Rhino [54] and Grasshopper [55]) to support the design exploration in four ways (the details are demonstrated in the case studies in Section 4.2):

- Exploring different types of designs with the related performance data:

In practice, studying the performance data of different types of geometries is crucial for designers to investigate the relationships between performance and form/geometry. The data visualization supporting this exploration aims to present the performance data of the node design for each cluster (of design alternatives) and to demonstrate how the performance data change for different node designs (which represent different types of geometries).

- Exploring different types of designs according to design objectives related to extreme performance data:

In practice, some design objectives usually require the designs obtaining the minimum or maximum values of specific performance indicators. Optimization can search for these designs within the whole design space, but it cannot support designers to study other design alternatives. The data visualization here aims to display the geometry of the design which obtains the extreme performance values in each cluster and demonstrate how the extreme performance values change for different types of geometries.

- Exploring different types of designs according to design constraints related to multiple performance indicators:

Besides design objectives, design constraints are also important,

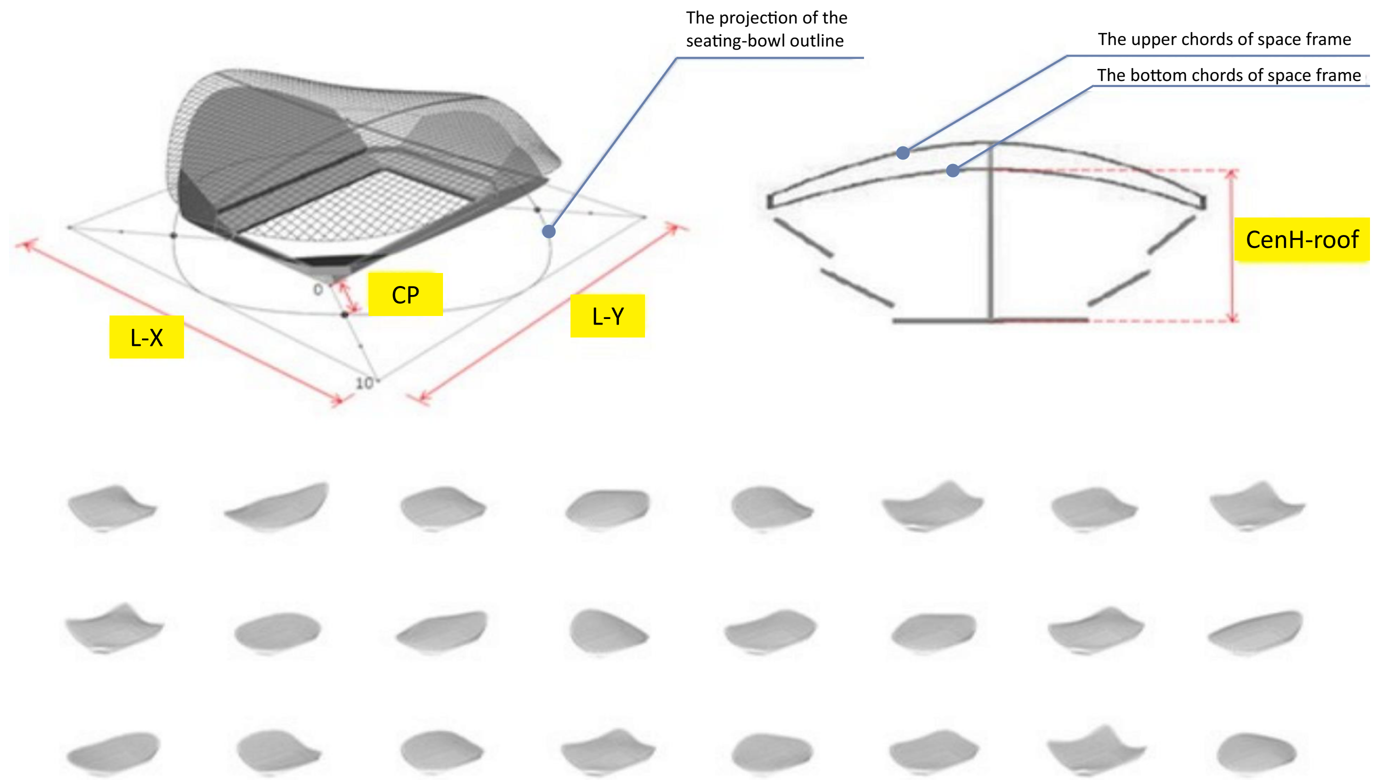



**Fig. 6.** The diagram of the proposed parametric model (IAG) and some design alternatives in the design space. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

**Table 2**
Settings of SOM, LLM, and MLPNN.

| SOM | | LLM | | MLPNN | |
|---|---|---|---|---|---|
| Neural network(s) | Triangle grid:<br>10 × 10<br>20 × 20<br>30 × 30 | The width of the fully activated region around the neighbouring clusters | $\mu = 0.3$ | Neural networks (the hidden layers) | 6-6-10 |
| Iteration times | Ordering: 10,000<br>Tuning: 100,000 | the width of the area of the neighbouring influence kernel | $\gamma = 0.8$ | Ratios of training, validation, and test sets | 70%, 15%, and 15% |
| Learning rate | Ordering: 0.8<br>Tuning: 0.8 | | | Activation function | Sigmoid |
| Initial neighborhood | 5 | | | Training algorithm | Levenberg-Marquardt |

which determine whether a design is feasible. The data visualization here aims to 1) display all the feasible designs within each cluster according to specific design constraints provided by designers, 2) demonstrate how each cluster of designs satisfy the design constraints, and 3) present the performance/output space corresponding with the design/input space.

- Exploring the geometries and the related performance values of the preferred designs:

In practice, among diverse types of designs, designers may focus on several preferred ones selected based on their experience and knowledge about qualitative aspects (e.g. aesthetics). Hence, it is necessary to explore the multiple performance values of the preferred types of designs. Moreover, it is also crucial to compare these preferred designs to other designs in the design space, according to multiple performance indicators. Therefore, the data visualization here aims to 1) aid designers to select the preferred designs according to geometry typology, 2) highlight the preferred types of designs among all the design alternatives in the design space, and 3) demonstrate and compare the performance values of the preferred designs and other designs.

## 4. Case studies

The case studies are divided into two parts. The first part (Section 4.1) aims to compare the performance of SOM-LLM and SOM-MLPNN in data approximations. In this part, for the workflows of both SOM-LLM and SOM-MLPNN, an experiment is applied, in which the operations proposed in Section 3.2 (using SOM networks in different sizes for LLM and MLPNN as well as using different amount of design inputs to train the SOM network) are used. The experiment studies whether these operations can save computation time and ensure acceptable accuracy of data approximation. Moreover, in the experiment, the proposed SOM-MLPNN is compared to SOM-LLM according to the accuracy of the data approximations, to select a proper workflow and model for the design explorations in practice. The second parts (Section 4.2), based on the selected workflow and the related model, aims to use the proposed data visualizations to support the design exploration based on numeric

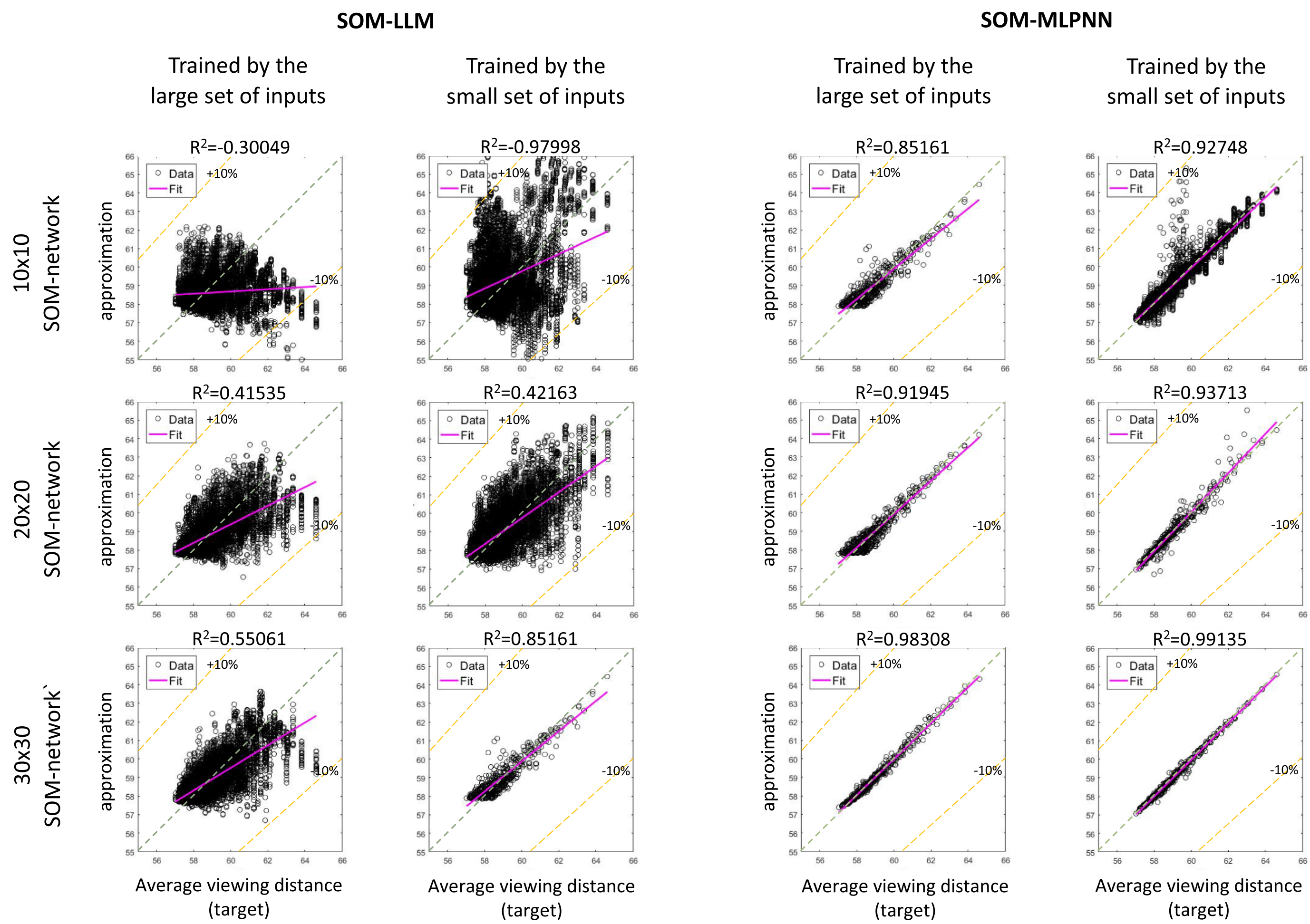



**Fig. 7.** The data approximation of the average viewing distance of spectators.

data (related to multiple quantitative requirements) and visual inspections on different types of geometries (related to qualitative requirements) of indoor arenas.

A hypothetical design of a typical multi-functional indoor arena is used as an example for the case studies. The arena has 14,000 to 15,000 fixed seats and mainly caters for professional basketball games and concerts. Therefore, for the quantitative design requirements, three aspects are highlighted in the design (including the viewing of spectators for basketball court, acoustics for both basketball game and concerts, and structural performance of the long-span roof). The design parameters and performance indicators related to quantitative requirements are listed in Table 1. The first four design parameters, which are directly related to the overall geometry, are selected as design variables which are also the design inputs data of SOM and data approximation.

Based on the different intervals for the design variables in Table 1, two sets of design inputs can be obtained. There are 252,747 design alternatives in the large set and 20,748 ones in the small set. By using the IAG proposed in [3], the geometries of the design alternatives are generated based on the four design variables (labelled in yellow in the upper chart of Fig. 6). Some of the geometries are randomly selected and demonstrated in the bottom chart in Fig. 6. However, according to the requirements on the number of fixed seats (14,000 to 15,000) in the arena, some design alternatives which have too many or too less fixed seats are automatically weeded out by IAG. Finally, there are 10,511 design alternatives in the large set and 1381 ones in the small set. Both the large and small sets of design inputs are used for the experiment about SOM-LLM and SOM-MLPNN.

This example is used for the early design stage of indoor arenas, in which designers mainly focus on the overall form of the building (which is defined by several key design variables). Hence the dimensionality of design variables is low (four dimensions), and correspondingly, the amount of the related design alternatives is small (10,511). Nevertheless, the example is still acceptable for testing and verifying the proposed method. For some practical designs of other types of buildings, in which the overall forms can be more complex and are defined by high dimensional design variables (correspondingly, there are more design alternatives), further tests are necessary to exam the effect of this method in future work (which are included in Section 5.2).

### *4.1. Comparison and experiment of SOM-LLM and SOM-MLPNN in data approximations*

In the experiment, the SOM-LLM and SOM-MLPNN are used to perform the data approximations of three building performance indicators selected in Table 1 (the average viewing distance, structural self-weight, and the reverberation time of all octave band frequencies).

For both SOM-LLM and SOM-MLPNN, according to the operation proposed in Section 3.2, three different sizes of SOM networks (10 × 10, 20 × 20, and 30 × 30) trained by two different sets of input data (10,511 and 1381 design inputs mentioned above) are applied. Hence, there are six models for SOM-LLM and SOM-MLPNN, respectively. These twelve models are trained to approximate the aforementioned performance indicators for all the 10,511 design alternatives within the design space. To assess the approximations, the actual values of the 10,511 designs are obtained by simulations, these high-fidelity

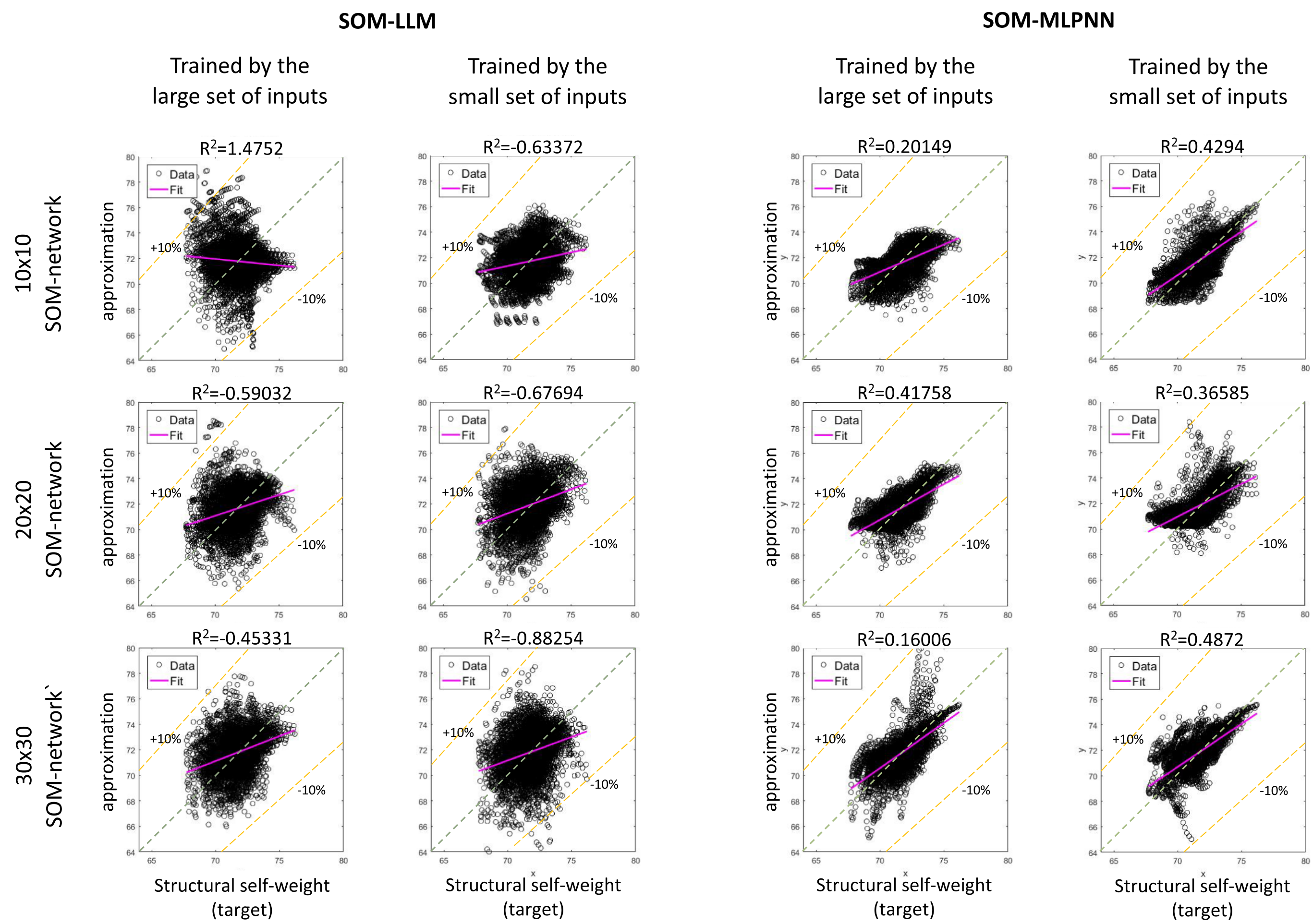



**Fig. 8.** The data approximation of structural self-weight. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

data will be used to assess the predicted data generated by LLM and MLPNN. The settings of SOM, LLM, and MLPNN are lists in Table 2. The results are illustrated in Figs. 7 to 9.

According to the results illustrated in Figs. 7, 8, 9, obviously, for the performance of the data approximations of the three selected indicators, SOM-LLM is not as well as SOM-MLPNN. The reason can be that LLM is based on piecewise interpolation, in which the computation is much simpler than that of MLPNN. This simpler computation process makes SOM-LLM can prediction different kinds of performance data by a fixed calculation process, which save computational time, but it also influences the performance in data approximations. Nevertheless, the errors of the three predictions supported by SOM-LLM are still acceptable for practice (most of the errors are < 10%).

Moreover, for the data approximations supported by both SOM-MLPNN and SOM-LLM, as the size of the SOM network increases from 10 × 10 to 20 × 20 and 30 × 30, the accuracy improves. The reason can be that a larger size of SOM network has more nodes/neurons, which generates more reference/labelled data for LLM/MLPNN.

For the 30 × 30 SOM network, the SOM-MLPNN based on the SOM network trained by the small set of inputs (the 4th charts of the last rows of Figs. 7, 8, 9) obtains higher accuracy, comparing to the counterpart related to the large set of inputs (the 4th charts of the last rows of Figs. 7, 8, 9). The reason can be that it may be easier for a SOM network to capture a small sets of data points and reflect the intrinsic topology of the data space.

Hence, in this paper, SOM-MLPNN is selected as the workflow of the proposed method to support the aforementioned design exploration, and the MLPNN based on the 30 × 30 SOM-network trained by the small set of design inputs is selected as the model. This model is used to approximate all the six performance indicators listed in Table 1 for all the 10,511 design alternatives in the design space. Fig. 10 demonstrates the training, validations, tests, and generalizations of the model. For the data approximations of viewing distances and reverberation time (the 1st, 2nd, 5th, and 6th rows in Fig. 10), the fitting of the training, validation, and tests sets are ideal, since the related correlations of determination ($R^2$) are quite close to or equal 1, and the data points are almost on the diagonal lines of the correlate charts which means the values of the approximated data almost equal the actual values obtained by simulations. The related generalizations are also ideal (the 1st, 2nd, 5th, and 6th rows of the right column in Fig. 10). The correlations of determination ($R^2$) are fixed or slightly decrease, comparing to those of the training, validation, and test sets. The majority of the data points also lay along the diagonal lines of the coordinate charts. In comparison, the accuracies of the data approximations of structural self-weight and strain energy (the 3rd and 4th rows in Fig. 10) are low. For the training, validation, and test sets, although the correlations of determination ($R^2$) are still close to 1 (between 0.91 and 0.99), the data points do not lay along the diagonal lines. The accuracies even decrease in the generalizations (the 3rd and 4th chart in the right column in Fig. 10). The correlations of determination ($R^2$) decrease from 0.93 and 0.99 to 0.49 and 0.79, respectively, and the data points are scattered.

These phenomena can be caused by the uncertainty of the trained models and the complex relationships between the design inputs and these two indicators related to structural performance. Although dealing with the uncertainty is not the focus of this paper, a series of methods can be used to quantify the uncertainty [47] and using deeper

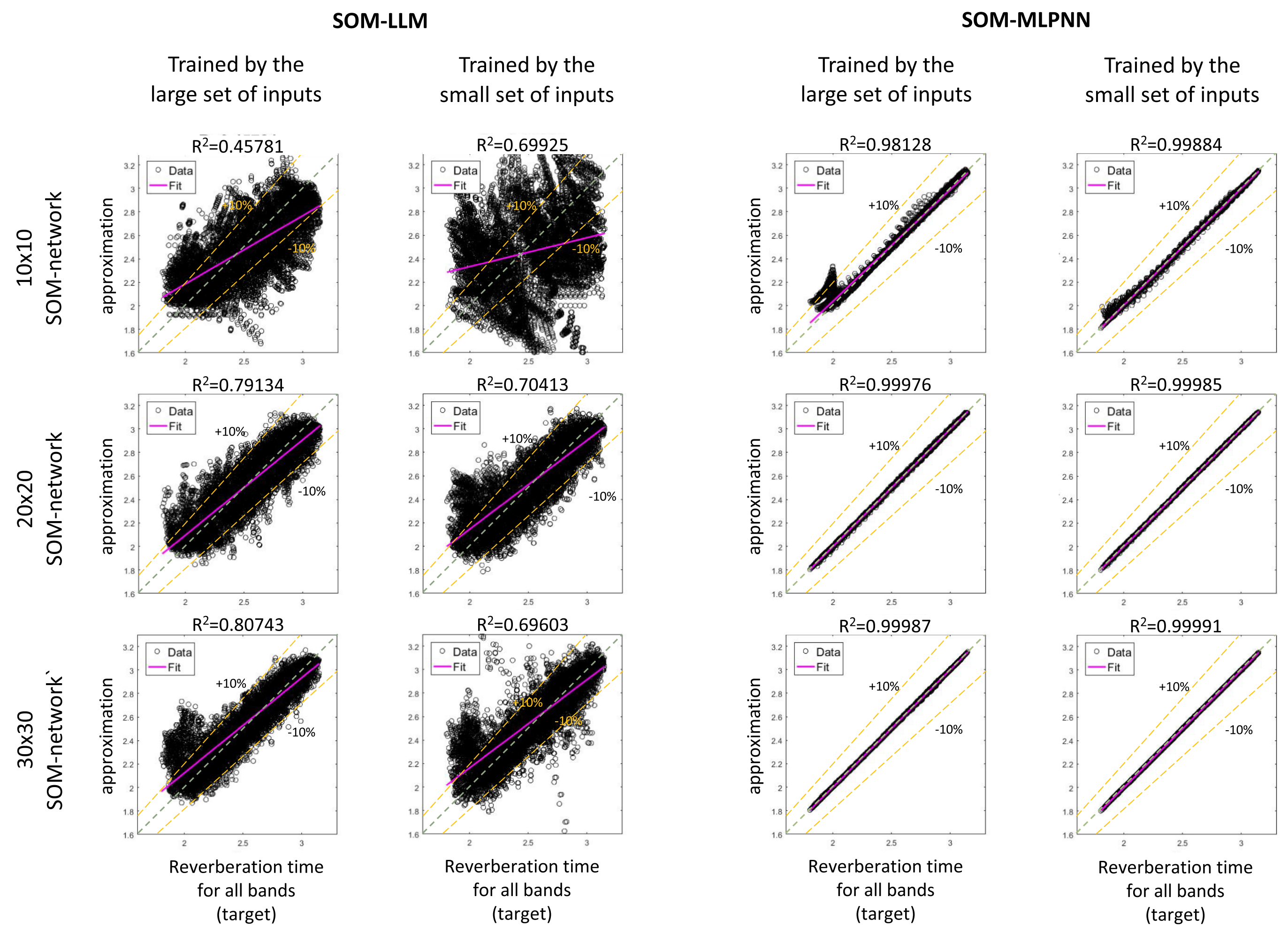



**Fig. 9.** The data approximation of reverberation time.

MLPNN or using cross-validation to fully exploit the labelled data can be possible ways to improve the performance. Nevertheless, for conceptual design process, except those related to structural strain energy, the errors of all the data approximations are acceptable, which are < 10% (see the data points and the yellow dash lines in the charts of Fig. 10). The decrease of the accuracy from the training, validation, and test to the generalization demonstrates the limitation of MLPNN in data approximations. In this situation, designers may trust the trained model according to the results in training, validation, and test, but they cannot know the accuracy of its generalization in practice. It is still a problem in the applications of MLPNNs.

### 4.2. Comprehensive design exploration supported by SOM-MLPNN

For the selected model of SOM-MLPNN, since the 30 × 30 SOM network is trained by the 1380 design inputs but not all the design alternatives, only these 1380 designs are grouped into the 900 clusters. Other design alternatives of the 15,011 ones are grouped into the nearest clusters according to the distances (between the design inputs of these alternatives and the vectors of the neurons related to the 900 clusters).

Based on the results of the SOM, for each one of the 900 clusters, the node/neural design, which represents all the designs within the cluster, can be generated by parametric model. All the node designs are distributed on the network, and similar ones are close while different ones are far away, which reflects the original design space (Fig. 11 left). Therefore, designers can view the design space and explore different types of designs based on geometry typologies.

However, a large size SOM network can make the exploration being difficult (as mentioned in Section 2.1), since there are too many node designs and the adjacent ones are quite similar which impedes designers to efficiently explore different types of designs. To overcome this, a smaller size network with fewer node designs can be generated by simply combining a group of adjacent clusters into a new one. Within each group of the adjacent clusters on the original network, the node design of the central cluster can be considered as the node design of the new combined cluster on the new network. In this paper, the 30 × 30 SOM network is transformed into a 10 × 10 network (Fig. 11 right) by combining every nine adjacent clusters into one, which can make the design exploration more efficiently. Other sizes of networks that are smaller than 30 × 30 can be also generated based on this approach. It worth noting that this process is not a clustering process but an operation on the results of the trained SOM network.

Based on the SOM network and the results of data approximations for the six kinds of performance data, a series of data visualizations are proposed to support design exploration in four ways mentioned in Section 3.6. It worth noting that the visualizations can be only presented in static figures in this paper, but they are proposed as interfaces in practice, based on which designers can obtain information by dynamically interacting with the objects in the interfaces.

#### 4.2.1. Exploring different types of designs with the related performance values

To support this design exploration, the related data visualization is used to present the performance value of the node design and of all the designs alternatives for each cluster, therefore, to demonstrate how the

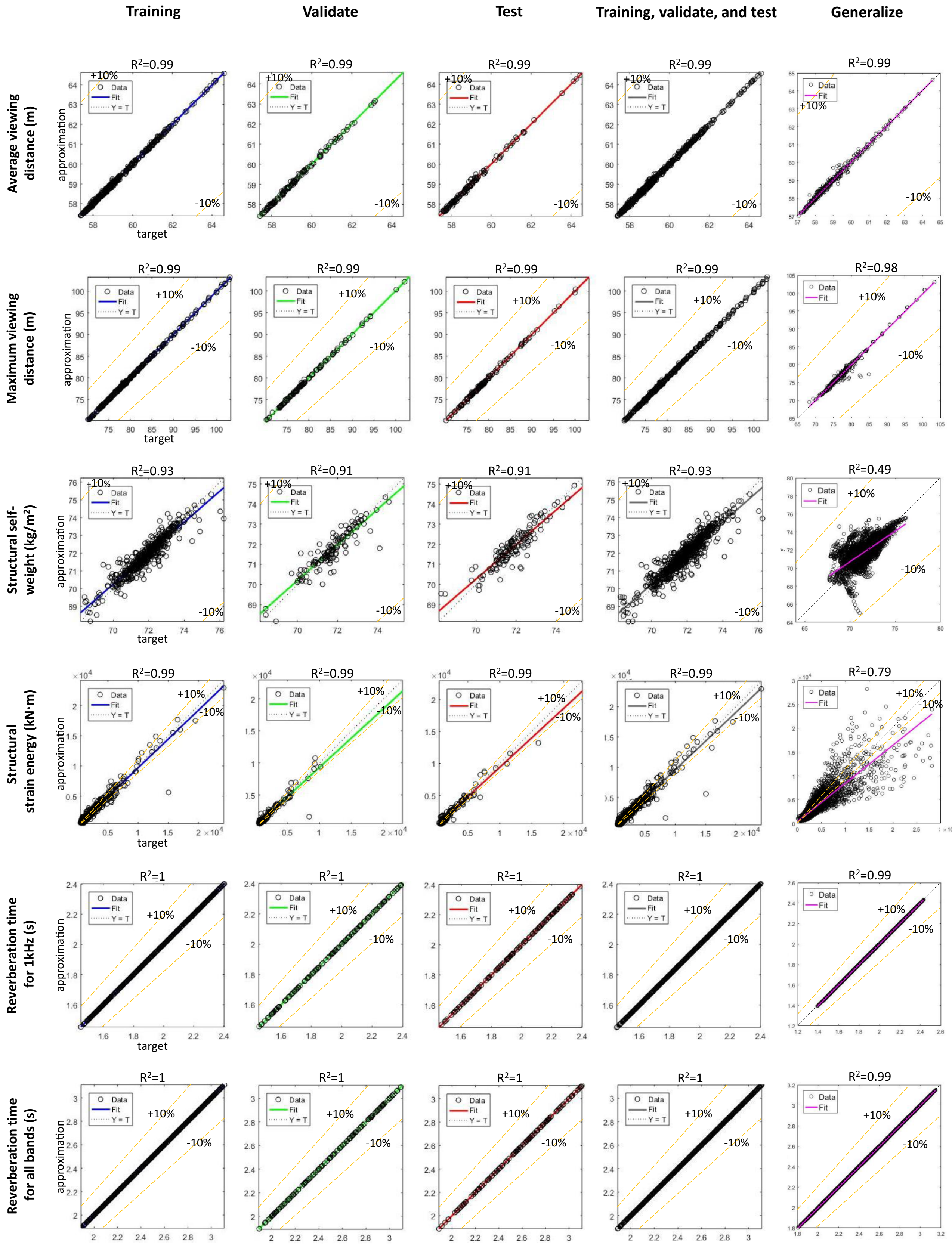



**Fig. 10.** The training, validation, test, and generalization of the selected model (a SOM-MLPNN model based on the 30 × 30 SOM network trained by the small set of design inputs) for the approximation of multiple performance data.

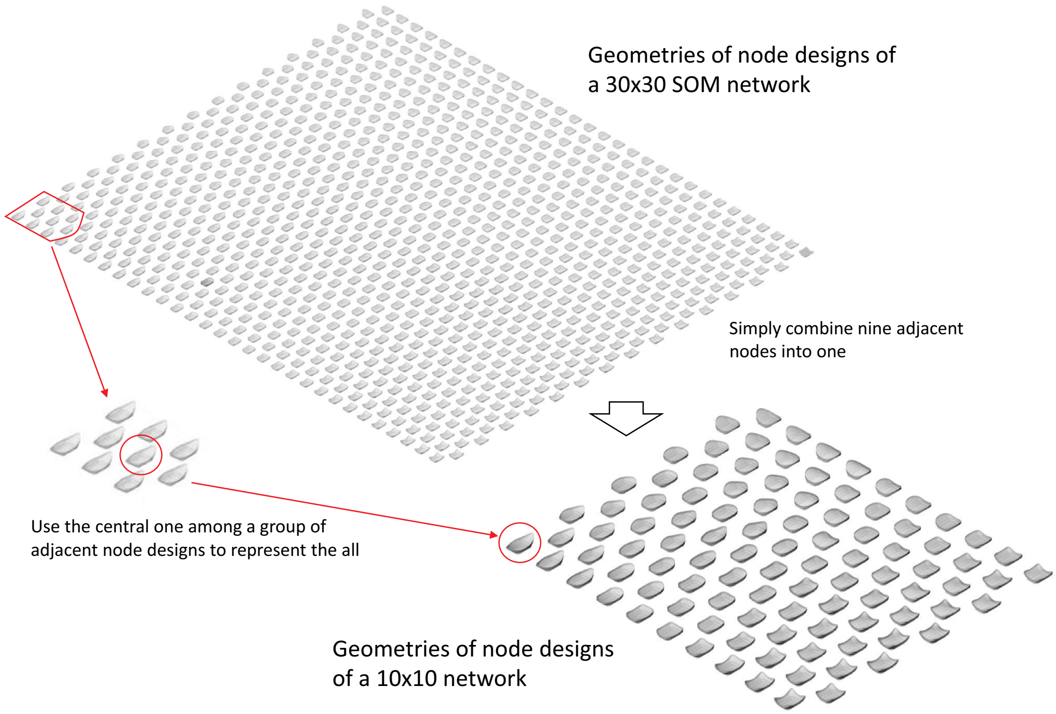



**Fig. 11.** Simply shrink a large size of SOM network into a small one.

performance values change for different node designs (which represent different types of geometries) and to show the output space corresponding to the input space reflecting by the SOM network.

Fig. 12 demonstrates the visualization. For each node design on the SOM network, a blue dot is assigned and located above it (Fig. 12). The altitude of the dot indicates the performance value of the node design, which can be read according to the vertical axes in the left (Fig. 12). These dots illustrate how the performance value changes among different types of designs.

Besides, for each kind of performance indicators, two meshes (highlighted in red and green in Fig. 8) are generated based on the designs obtaining the maximum and minimum performance values in each cluster (see the red and green dots in the middle columns in Fig. 12). The space between the meshes is the output space corresponding to the input space reflecting by the SOM network, which is crucial for designers to study the relationships between design inputs (geometries) and outputs (performance data).

Furthermore, designers can investigate any designs in each cluster. A series of grey dots are assigned to all the designs in each cluster, and the related performance values are also indicated by the vertical axes. In the right part of Fig. 12, the output space of the 100th cluster are magnified as an example, for each kind of performance indicators, a series of grey dots are assigned to all the designs in this cluster (the middle column in Fig. 12), the designs obtaining the maximum and minimum values and the node designs are highlighted (in red, green, and blue dots, respectively). The related geometries as well as the ID numbers in design space can be also obtained (the right column in Fig. 12). Moreover, designers can investigate any designs according to performance values and geometries. As an example, for each kind of performance indicators, two designs in the 100th cluster are randomly selected (highlighted in black dots in the middle column in Fig. 12).

#### 4.2.2. Exploring different types of designs according to design objectives

To support this design exploration, the data visualization aims to display the geometries of the design which obtains the extreme performance values in each cluster and to demonstrate how the extreme performance values changes for different types of geometries. This exploration can aid designers to find the optimal designs (which obtain the extreme performance values) not only in the whole design space (like what MOOs do) but also in different clusters. Moreover, it also aids designers to quickly understand how each type of designs satisfy the related performance requirements, therefore can further study the relationships between geometry and performance.

Fig. 13 demonstrates the visualization. Three kinds of performance indicators (the average viewing distance, structural self-weight, and reverberation time) are selected as examples. For each kind of the performance indicators, the SOM network shows the design obtaining the minimum performance values in each cluster (see the upper part of Fig. 13), above which a dot is assigned to each design and its altitude indicate the performance value. Based on the dots, a mesh is generated to show how the extreme performance value changes among different types of designs. Two other dots are also assigned to each design, to indicate the other two performance values. The designs obtaining the minimum values of the three kinds of performance indicators in five clusters (the 1st, 10th, 55th, 91st, 100th clusters) are magnified (the bottom part of Fig. 13) as an example to show how designers can investigate the geometries and performance data of these designs.

#### 4.2.3. Exploring different types of geometries according to design constraints

To support this design exploration, the data visualization aims to display all the feasible designs within each cluster according to specific design constrains provided by designers, therefore, to demonstrate how each cluster of designs satisfy the design constraints.

Fig. 14 illustrates an example. Five performance indicators are selected, and the related constraints are supposed to be set by designers ($VD_{avr\text{-}p} \leq 60$ m, $VD_{max\text{-}p} \leq 83$ m, $SW \leq 70$ kg/m$^2$, $SE \leq 9763$kNm, $RT_{60} \leq 2.8$ s). Two meshes (like those in Fig. 12) are presented to visualize the output/performance spaces. The related constraints are visualized by the yellow planes cutting the meshes (the left part of Fig. 14). Correspondingly, the clusters containing feasible designs (which satisfy all the constraints) are highlighted on the SOM network below. Then, the geometries of all the feasible designs are presented with the related node designs (the right part of Fig. 14).

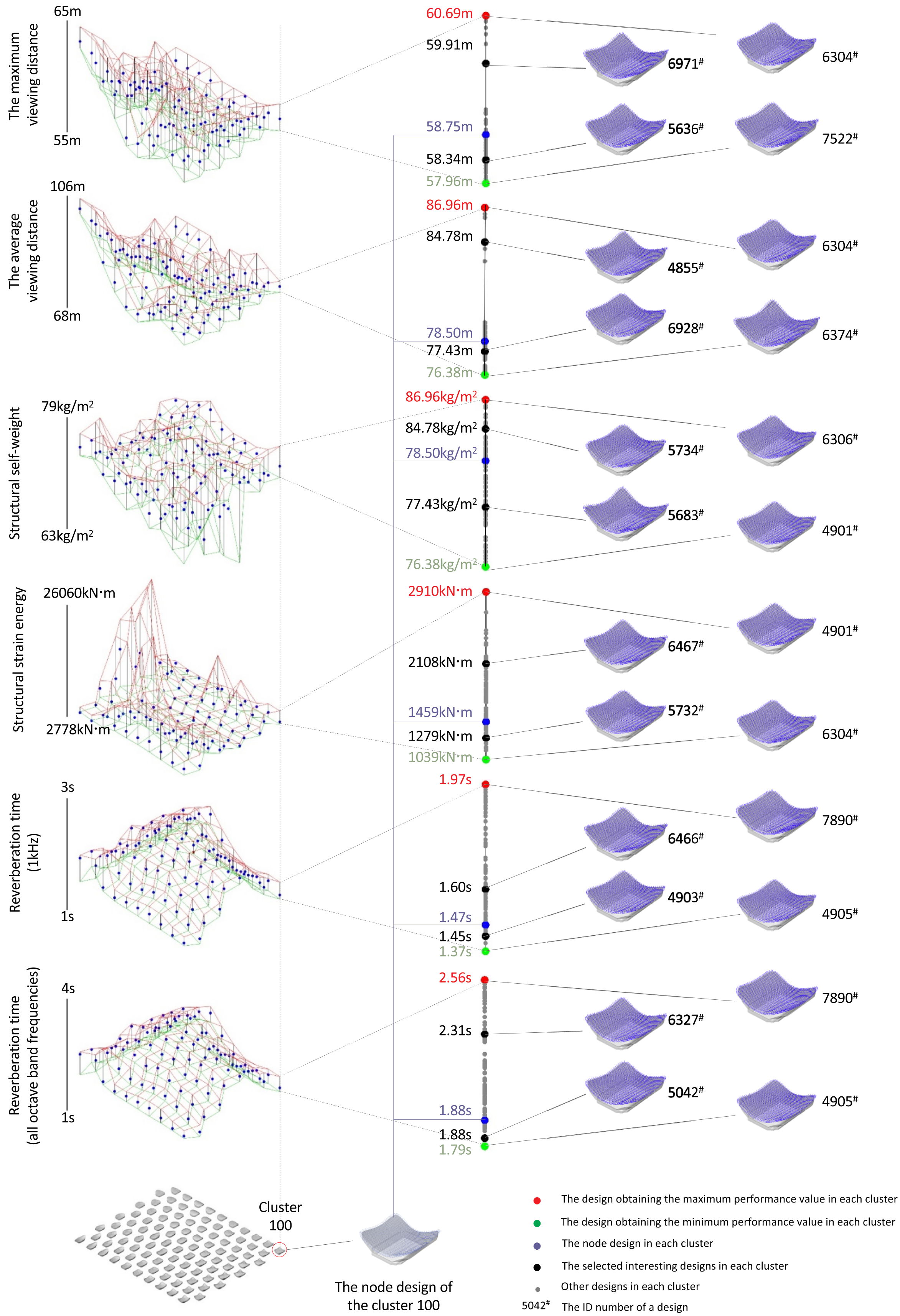



**Fig. 12.** Exploring different types of designs and the related performance values. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

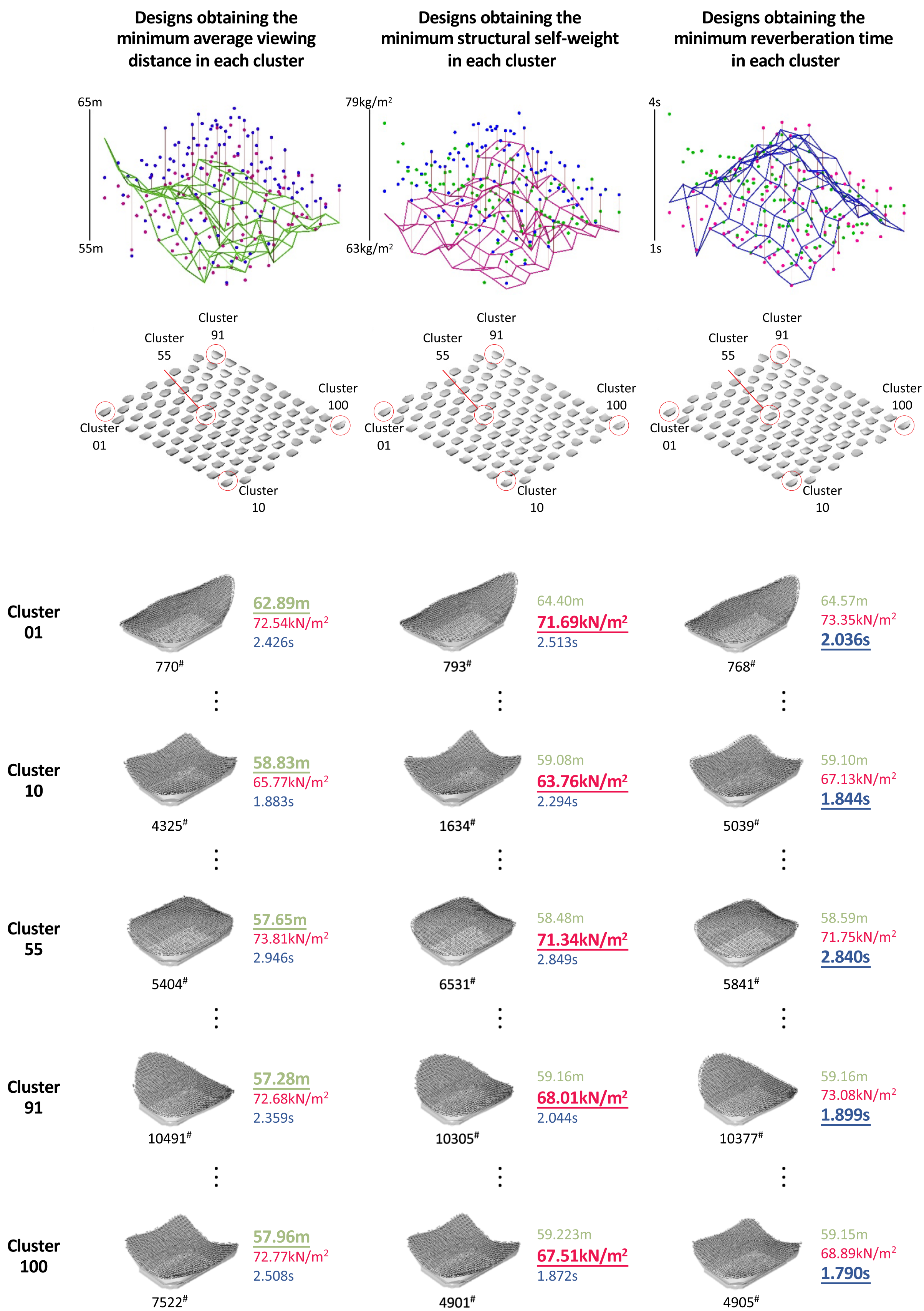



**Fig. 13.** Exploring different types of designs according to the extreme performance values.

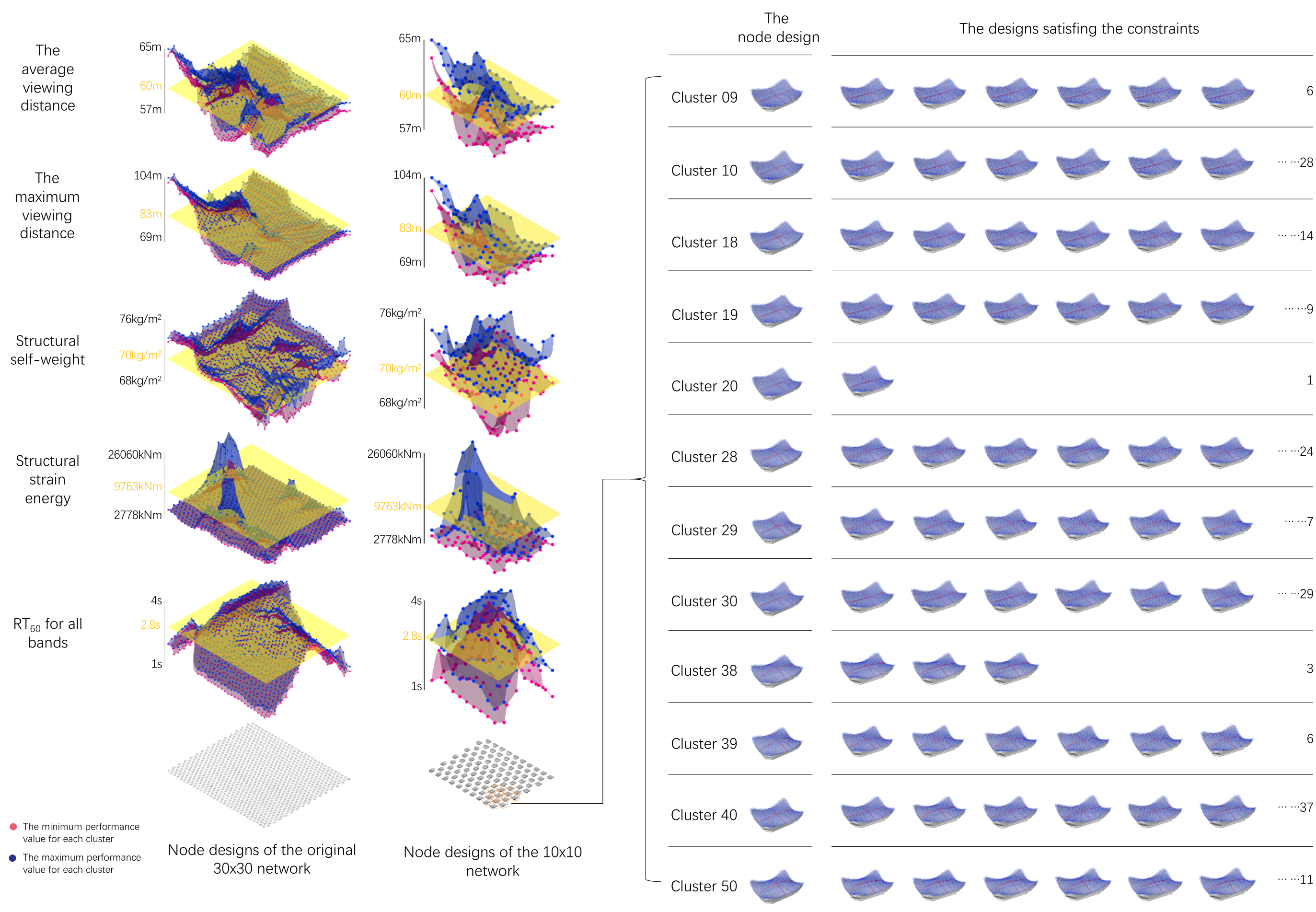



**Fig. 14.** Exploring different types of geometries according to design constraints related to multiple performance indicators.

#### 4.2.4. *Exploring the geometries and performance values of preferred designs*

To support this design exploration, the data visualization aims to 1) highlight the preferred types of designs among all the design alternatives on the SOM network, 2) display the geometries of the preferred designs, 3) demonstrate and compare the performance values of the preferred designs and other designs.

Fig. 15 illustrates an example, in which six clusters of designs are supposed to be preferred by designers and the related node designs are highlighted on the SOM network. Three performance indicators (the average viewing distance, structural self-weight, and reverberation time of all the octave band frequencies) are selected to assess the designs. A three-dimensional scatter chart is used to visualize the data. The x, y, and z axes indicate the three performance indicators, respectively, and the six clusters of designs are represented by the dots in six colours while other designs within the design space are represented by grey dots. Moreover, for each cluster, the geometries of the node designs and the designs obtaining the minimal values for the three performance indicators are presented. This exploration can aid designers to select 'well-performing' designs within various preferred design candidates, which combines quantitative performance and design preference during the early stage of architectural design.

## 5. Conclusion

### 5.1. *Summary of contributions*

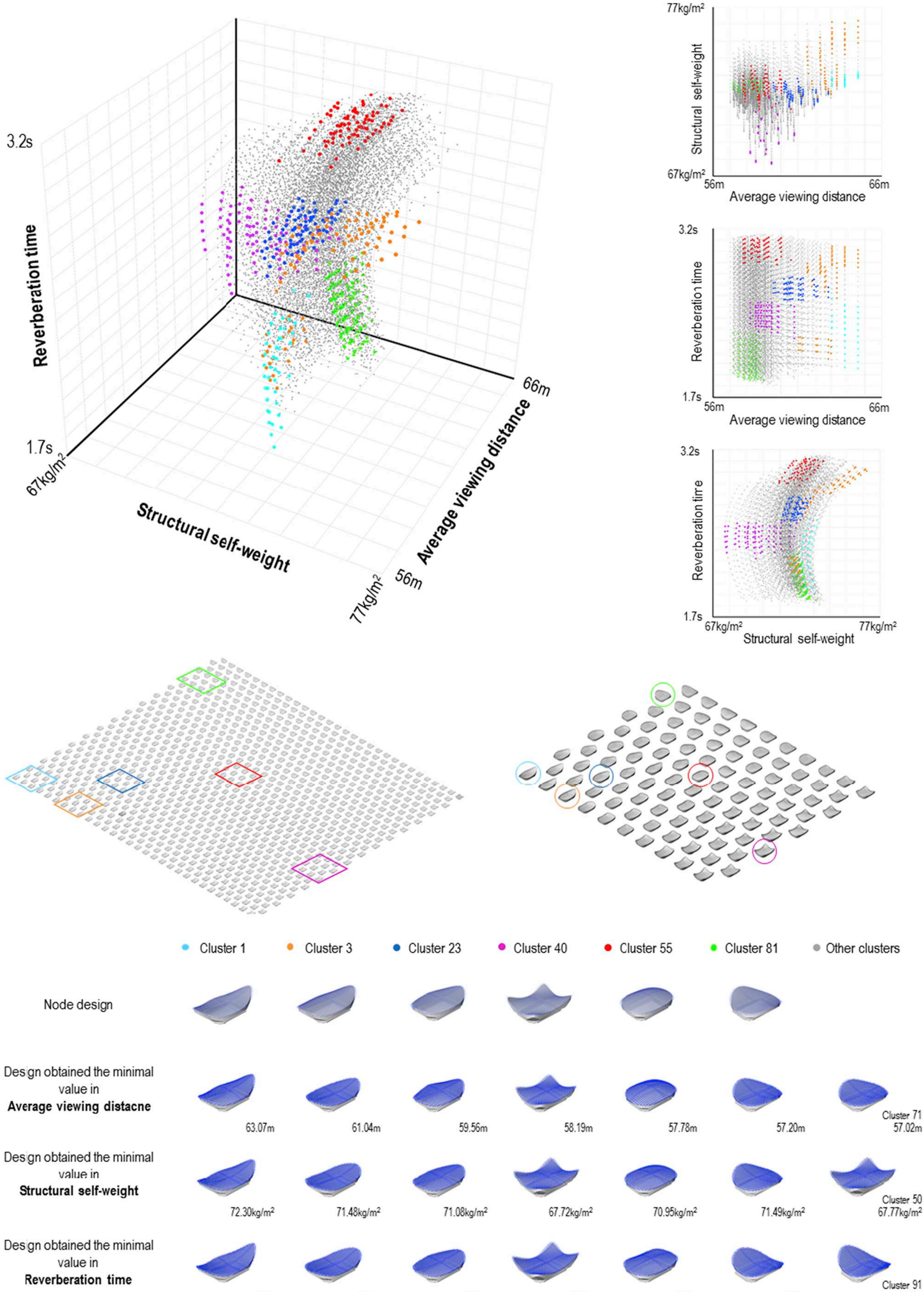
The main contribution of this paper is developing a novel design method based on SOM-LLM or SOM-MLPNN to support the design explorations of indoor arenas, which demands the integration of multi-functional space and long-span roof structure and the investigation of various design alternatives according to multiple performance data (related to quantitative design requirements) and the overall geometries of the buildings (related to qualitative design requirements). In the proposed method, SOM-LLM and SOM-MLPNN are two independent workflows to achieve the goal. SOM-LLM is an existing method for data approximations, but this paper uses it in a different way to perform both data approximation and clustering at the same time. SOM-MLPNN is a new method, formulated based on the inspiration of SOM-LLM, which takes the advantage of the capability of MLPNNs in universal fitting. According to the results of the case studies, the performance of both the workflows in data approximation is acceptable for conceptual design.

Bedsides the main contribution, this paper also studies how the size of the SOM network and the amount of the design inputs (which are used to train the network) influence the performance of SOM-LLM and SOM-MLPNN in data approximations. The results indicate that for both SOM-LLM and SOM-MLPNN, using a larger size of network trained by a small set of design inputs can obtain better performance. This study provides a useful approach for the application of the proposed method to save the computation time and obtain acceptable accuracies of data approximations.

Moreover, a series of data visualizations are proposed to demonstrate the results of the proposed method, which facilitates designers to perform design explorations based on the outcomes, therefore, makes the method more practical.

### 5.2. *Limitations and future work*

There are still some limitations of the proposed method (for both




*(caption on next page)*

**Fig. 15.** Exploring the geometries and the related performance values of the preferred designs. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

SOM-LLM and SOM-MLPNN). First, in this paper, it only focuses on the multi-functionality and structural performance of indoor arena, which impedes the application of the method in the designs of other buildings. In fact, the method is potentially used for other types of buildings, but it is necessary to provide flexible parametric models which can generate diverse designs for these buildings based on several key design parameters. Second, there are constraints on design inputs/variables. The inputs/variables should be the design parameters directly related to the overall geometry, to ensure the direction of the mapping between the inputs and the geometry. The dimensionality of the inputs/variables should be also limited to ensure the effect of the SOM in reflecting the design space (in the example of the case studies, the dimensionality of inputs/variables is four). These constraints for design variables may not influence the early design stage of the building during which designers usually focuses on the overall form of the building and several crucial performance indicators and a small number of key design parameters are emphasized. However, it does not mean it is not necessary to study other design parameters which are not directly related to the overall form but are crucial for building performance.

As mentioned in Section 1, MOOs focus on the 'well-performing' designs. Even through there is interactive optimization which allow designers to explore the geometries of designs during the iterations and select preferred designs to change the direction of optimization [5,10], it is still limited in supporting designers to investigate various designs in the design space. However, in a standard MOO, there are no aforementioned constraints about design variables, which makes it can be used in various fields and design process. Moreover, since MOOs are based on simulations, in which the performance data are high-fidelity.

In this light, the combination of the proposed method and a MOO is a potential way to overcome the limitations for the proposed method, which can be studied in the future work. First, the proposed method supports designers perform design exploration to define several proper designs. Then, based on the selected designs, more design variables can be considered, and MOOs can be used to further search for 'well-performing' designs based on the high-fidelity performance data obtained by simulations. Moreover, to generalize the method, more building types (e.g. residence building, office buildings) can be considered.

## Declaration of competing interest

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

## Acknowledgement

This research is related to the project (201707020041) supported by the Plan of Science and Technology of Guangzhou and the project (51761135025) supported by National Natural Science Foundation of China as well as the project of independent subject supported by the State Key Lab of Subtropical Building Science in South China University of Technology. The research is also supported by the Urban System and Environment Joint Research Centre (USE) between South China University of Technology (SCUT) and Delft University of Technology (TUD). The first author is supported by the scholarship of China Scholarship Council (CSC) and SCUT for his joint Ph.D. research at Delft University of Technology (TUD), Delft, the Netherlands.

## References

[1] M. Turrin, Performance Assessment Strategies: A Computational Framework for the Conceptual Design of Large Roofs, Ph.D. Thesis Department of Architectural Engineering and Technology, Delft University of Technology, 2014, https://books.bk.tudelft.nl/index.php/press/catalog/book/269 [Accessed: 7-April-2019].
[2] Gul E. Okudan, S. Tauhid, Concept selection methods - a literature review from 1980 to 2008, International Journal of Design Engineering 1 (3) (2008) 243–277, https://doi.org/10.1504/IJDE.2008.023764.
[3] W. Pan, M. Turrin, C. Louter, S. Sariyildiz, Y. Sun, Integrating multi-functional space and long-span structure in the early design stage of indoor sports arenas by using parametric modelling and multi-objective optimization, Journal of Building Engineering 22 (2019) 464–485, https://doi.org/10.1016/j.jobe.2019.01.006.
[4] 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 (2011) 656–675, https://doi.org/10.1016/j.aei.2011.07.009.
[5] P. von Buelow, ParaGen: performative exploration of generative systems, Journal of International Association of Shell and Spatial Structure 53 (2012) 271–284 https://www.ingentaconnect.com/content/iass/jiass/2012/00000053/00000004/art00010 [Accessed: 12-January-2020].
[6] M. Turrin, P. Von Buelow, A. Kilian, R. Stouffs, Performative skins for passive climatic comfort: a parametric design process, Autom. Constr. 22 (2012) 36–50, https://doi.org/10.1016/j.autcon.2011.08.001.
[7] D.J. Gerber, S.E. Lin, B.P. Pan, A.S. Solmaz, Design optioneering: multidisciplinary design optimization through parameterization, domain integration and automation of a genetic algorithm, Proc. SIMAUD '12, Simulations for Architectural and Urban Design (SIMAUD), Orlando, 2012, pp. 23–30 http://www.simaud.org/proceedings/index.php#2012 [Accessed: 12-January-2020].
[8] S. Sariyildiz, Performative computaitonal design, Proc. ICONARCH '12, Selcuk University, Konya, 2012, pp. 313–344 http://resolver.tudelft.nl/uuid:b637bb68-7e90-44f7-9c45-74a05e29a69d [Accessed: 27-February-2020].
[9] S.H.E. Lin, D.J. Gerber, Designing-in performance: a framework for evolutionary energy performance feedback in early stage design, Autom. Constr. 38 (2014) 59–73, https://doi.org/10.1016/j.autcon.2013.10.007.
[10] C.T. Mueller, J.A. Ochsendorf, Combining structural performance and designer preferences in evolutionary design space exploration, Autom. Constr. 52 (2015) 70–82, https://doi.org/10.1016/j.autcon.2015.02.011.
[11] N.C. Brown, C.T. Mueller, Design for structural and energy performance of long span buildings using geometric multi-objective optimization, Energy and Buildings 127 (2016) 748–761, https://doi.org/10.1016/j.enbuild.2016.05.090.
[12] D. Yang, S. Ren, M. Turrin, S. Sariyildiz, Y. Sun, Multi-disciplinary and multi-objective optimization problem re-formulation in computational design exploration: a case of conceptual sports building design, Autom. Constr. 92 (2018) 242–269, https://doi.org/10.1016/j.autcon.2018.03.023.
[13] H. Hajela, L. Berke, Neural networks in structural analysis and design: an overview, Comput. Syst. Eng. 3 (1992), https://www.sciencedirect.com/science/article/pii/0956052192901389 [Accessed: 12-January-2020].
[14] T. Wortmann, A. Costa, G. Nannicini, T. Schroepfer, Advantages of surrogate models for architectural design optimization, Analysis and Intelligence for Engineering Design 29 (4) (2015) 471–481, https://doi.org/10.1017/S0890060415000451.
[15] S. Tseranidis, N.C. Brown, C.T. Mueller, Data-driven approximation algorithms for rapid performance evaluation and optimization of civil structures, Autom. Constr. 72 (2016) 279–293, https://doi.org/10.1016/j.autcon.2016.02.002.
[16] D. Yang, Y. Sun, D. Di Stefano, M. Turrin, S. Sariyildiz, Impacts of problem scale and sampling strategy on surrogate model accuracy, IEEE Congr. Evol. Comput. ' 16 (2016) 4199–4207, https://doi.org/10.1109/CEC.2016.7744323 IEEE.
[17] J. Harding, Dimensionality reduction for parametric design exploration, in: S. Adriaenssens, F. Gramazio, M. Kohler, A. Menges, M. Pauly (Eds.), Advances in Architectural Geometry, vdf Hochschulverlag AG, Zürich, 2016, pp. 274–287, https://doi.org/10.3218/3778-4 2016.
[18] J. Harding, C. Brandt-Olsen, Biomorpher: interactive evolution for parametric design, Int. J. Archit. Comput. 16 (2018) 144–163, https://doi.org/10.1177/1478077118778579.
[19] W. Pan, Y. Sun, M. Turrin, C. Louter, S. Sariyildiz, Integrating the multi-functional space and long-span structure for sports arena design: a design exploration process based on design optimization and self-organizing map, in: C.T. Mueller, S. Adriaenssens (Eds.), Proceeding of the IASS Symposium 2018: Creativity in Structural Design, 2018 International Association for Shell and Spatial (IASS), Boston, 2018, pp. 1–8 (8), https://www.ingentaconnect.com/content/iass/piass/2018/00002018/00000004/art00015# [Accessed: 7-April-2019].
[20] IEEE CIS task force on many-objective optimisation – IEEE Computational Intelligence Society [Online], Available https://www.cs.bham.ac.uk/~limx/MaOP.html , Accessed date: 4 July 2019.
[21] D. Shepard, A two-dimensional interpolation for irregularly-spaced data function, in: R.B. Blue, A. Rosenberg (Eds.), 23rd ACM National Conference, ACM, New York, 1968, pp. 517–523 https://dl.acm.org/citation.cfm?id=800186 [Accessed: 7-April-2019].
[22] J. Goppert, W. Rosenstiel, Topology-preserving interpolation in self-organizing maps, Proc. NeuroNimes '93, International Conference Neural Networks and Their Industrial and Cognitive Applications, Nanterre, 1993, pp. 425–434 https://www.

semanticscholar.org/paper/Topology-preserving-Interpolation-in-Maps-RosenstielLehrstuhl/65a372386a205ec2dac42978875dffa338556268#citing-papers [Accessed: 7-April-2019].
[23] M. Aupetit, P. Couturier, P. Massotte, Function approximation with continuous self-organizing maps using neighboring influence interpolation, in: H.-H. Bothe, R. Rojas, C. Fyfe (Eds.), Proceeding of Neural Computation (NC2000), International Computer Science Conventions, Berlin, 2000 https://www.semanticscholar.org/paper/Function-Approximation-with-Continuous-Maps-using-.-Aupetit/d37291982e5757db3b89a09636efdacf15ef99b7 [Accessed: 7-April-2019].
[24] G. Cybenko, Approximation by Superpositions of a Sigmoidal Function, Mathematics of Control, Signals and Systems, 2 (1989), pp. 303–314 https://link.springer.com/article/10.1007/BF02551274 [Accessed: 7-April-2019].
[25] S. Koziel, D.E. Ciaurri, L. Leifsson, Surrogate-based methods, in: S. Koziel, X.S. Yang (Eds.), Computational Optimization, Methods and Algorithms. Studies in Computational Intelligence, 356 Springer, Berlin, 2011, pp. 33–59, https://doi.org/10.1007/978-3-642-20859-1_3.
[26] K.P. Murphy, Machine Learning: A Probabilistic Perspective, The MIT press, Cambridge, 2012https://mitpress.mit.edu/books/machine-learning-1 [Accessed: 7-April-2019].
[27] K. Hiyama, L. Wen, Rapid response surface creation method to optimize window geometry using dynamic daylighting simulation and energy simulation, Energy and Buildings 107 (2015) 417–423, https://doi.org/10.1016/j.enbuild.2015.08.035.
[28] S. Zhang, Y. Sun, Y. Cheng, P. Huang, M. Olaide, Z. Lin, Response-surface-model-based system sizing for nearly/net zero energy buildings under uncertainty, Appl. Energy 228 (2018) 1020–1031, https://doi.org/10.1016/j.apenergy.2018.06.156.
[29] S. Reza, X. Zhang, A. Mahdiyar, A comprehensive review on the application of artificial neural networks in building energy analysis, Neurocomputing 340 (7) (2019) 55–75, https://doi.org/10.1016/j.neucom.2019.02.040.
[30] F. Bre, J.M. Gimenez, V.D. Fachinotti, Prediction of wind pressure coefficients on building surfaces using artificial neural networks, Energy and Buildings 158 (2018) 1429–1441, https://doi.org/10.1016/j.enbuild.2017.11.045.
[31] B. Si, J. Wang, X. Yao, X. Shi, X. Jin, X. Zhou, Multi-objective optimization design of a complex building based on an artificial neural network and performance evaluation of algorithms, Adv. Eng. Inform. 40 (2019) 93–109, https://doi.org/10.1016/j.aei.2019.03.006.
[32] Z. Wang, Y. Wang, R. Zeng, R.S. Srinivasan, S. Ahrentzen, Random Forest based hourly building energy prediction, Energy and Buildings 171 (2018) 11–25, https://doi.org/10.1016/j.enbuild.2018.04.008.
[33] F. Smarra, A. Jain, T. De Rubeis, D. Ambrosini, A.D. Innocenzo, R. Mangharam, Data-driven model predictive control using random forests for building energy optimization and climate control, Appl. Energy 226 (2018) 1252–1272, https://doi.org/10.1016/j.apenergy.2018.02.126.
[34] Y.K. Yi, Adaptation of Kriging in daylight modeling for energy simulation, Energy and Buildings 111 (2016) 479–496, https://doi.org/10.1016/j.enbuild.2015.11.036.
[35] Y.K. Yi, Dynamic coupling between a Kriging-based daylight model and building energy model, Energy and Buildings 128 (2016) 798–808, https://doi.org/10.1016/j.enbuild.2016.05.081.
[36] S. Shanmuganathan, Artificial neural network modelling: an introduction, in: S. Shanmuganathan, S. Samarasinghe (Eds.), Artificial Neural Network Modelling, 628 Springer, Berlin, 2016, pp. 1–14, https://doi.org/10.1007/978-3-319-28495-8.
[37] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning internal representations by error propagation, in: D.E. Rumelhart, J.L. McClelland (Eds.), Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations, MIT Press, Cambridge, 1987, pp. 318–362 https://ieeexplore.ieee.org/document/6302929 [Accessed: 7-April-2019].
[38] P. Werbos, Backpropagation through time: what it does and how to do it, in: Proc. IEEE '90, vol. 78, IEEE, 1990, pp. 1550–1560. doi:https://doi.org/10.1109/5.58337.
[39] A.S. Ahmad, M.Y. Hassan, M.P. Abdullah, H.A. Rahman, F. Hussin, H. Abdullah, R. Saidur, A review on applications of ANN and SVM for building electrical energy consumption forecasting, Renew. Sust. Energ. Rev. 33 (2014) 102–109, https://doi.org/10.1016/j.rser.2014.01.069.
[40] C. Rubio-bellido, R. Pino-mejías, P. Alexis, J.A. Pulido-arcas, Comparison of linear regression and artificial neural networks models to predict heating and cooling energy demand, energy consumption and $CO_2$ emissions, Energy 118 (2017) 24–36, https://doi.org/10.1016/j.energy.2016.12.022.
[41] K. Amasyali, N.M. El-gohary, A review of data-driven building energy consumption prediction studies, Renew. Sust. Energ. Rev. 81 (2018) 1192–1205, https://doi.org/10.1016/j.rser.2017.04.095.
[42] S. Singaravel, J. Suykens, P. Geyer, Deep-learning neural-network architectures and methods: using component-based models in building-design energy prediction, Adv. Eng. Inform. 38 (2018) 81–90, https://doi.org/10.1016/j.aei.2018.06.004.
[43] C. Series, A methodology for daylight optimisation of high-rise buildings in the dense urban district using overhang length and glazing type variables with surrogate modelling, J. Phys. Conf. Ser. 1343 (2019), https://doi.org/10.1088/1742-6596/1343/1/012133.
[44] H. Salehi, R. Burgueño, Emerging artificial intelligence methods in structural engineering, Eng. Struct. 171 (2018) 170–189, https://doi.org/10.1016/j.engstruct.2018.05.084.
[45] S. Hashemi, S., K. Sadeghi, A. Fazeli, M. Zarei, Predicting the weight of the steel moment-resisting frame structure, International Journal of Steel Structures 12 (2019) 168–180. doi:https://doi.org/10.1007/s13296-018-0105-z.
[46] S. Samarasinghe, Order in the black box: consistency and robustness of hidden neuron activation of feed forward neural networks and its use in efficient optimization of network structure, in: S. Shanmuganathan, S. Samarasinghe (Eds.), Artificial Neural Network Modelling, 628 Springer, Berlin, 2016, pp. 15–43, https://doi.org/10.1007/978-3-319-28495-8.
[47] K.S. Kasiviswanathan, K.P. Sudheer, J. He, Quantification of prediction uncertainty in artificial neural network models, in: S. Shanmuganathan, S. Samarasinghe (Eds.), Artificial Neural Network Modelling, 628 Springer, Berlin, 2016, pp. 145–159, https://doi.org/10.1007/978-3-319-28495-8.
[48] D. Moshou, H. Ramon, K.U. Leuven, Extended self-organizing maps with local linear mappings for function approximation and system identification, Proc. WSOM'97, Helsinki University of Technology, Helsinki, 1997, https://www.semanticscholar.org/paper/Extended-Self-Organizing-Maps-with-Local-Linear-for-Moshou-Ramon/b8ac9be7cc102b70d06e2b2e5fcb147963174960 [Accessed: 7-April-2019].
[49] M. Aupetit, P. Massotte, P. Couturier, C-SOM: a continuous self-organizing map for function approximation, Proc. ISC '99, International Association of Science and Technology Development (IASTED), Santa-Barbara, 1999 https://www.semanticscholar.org/paper/C-SOM%3A-A-CONTINUOUS-SELF-ORGANIZING-MAP-FOR-Aupetit-Massotte/62aeb7ebcd21006d4e759609a9010d924084c52f [Accessed: 7-April-2019].
[50] T. Kohonen, Self-organized formation of topologically correct feature maps, Biol. Cybern. 43 (1) (1982) 59–69 https://link.springer.com/article/10.1007/BF00337288 [Accessed: 7-April-2019].
[51] Kohonen T., MATLAB Implementations and Applications of the Self-organizing Map, Unigrafia, Helsinki, http://docs.unigrafia.fi/publications/kohonen_teuvo/, (2014) [Accessed: 7-April-2019].
[52] M.M. Van Hulle, L. Neurofysiologie, K.U. Leuven, Self-organizing maps, in: G. Rozenberg (Ed.), Handbook of Natural Computing, Springer, Berlin, 2012, pp. 586–618, https://doi.org/10.1007/978-3-540-92910-9_19.
[53] J. Zhang, Kohonen self-organizing map: an artificial neural network, in: J. Zhang (Ed.), Visualization for Information Retrieval. The Information Retrieval Series, 23 Springer, Berlin, Heidelberg, 2008, pp. 107–125, https://doi.org/10.1007/978-3-540-75148-9_5.
[54] Rhinoceros 3D, [Online]. Available https://www.rhino3d.com/ [Accessed: 7-January-2019].
[55] Grasshopper, [Online]. Available http://www.grasshopper3d.com [Accessed: 7-January-2019].
[56] C. Preisinger, M. Heimrath, Karamba - a toolkit for parametric structural design, Struct. Eng. Int. 24 (2) (2014) 217–221, https://doi.org/10.2749/101686614X13830790993483.
[57] Self-organizing map – Matlab selforgmap, [Online]. Available https://nl.mathworks.com/help/deeplearning/ref/selforgmap.html [Accessed: 11-January-2019].
[58] MATLAB, [Online]. Available https://nl.mathworks.com/products/matlab.html [Accessed: 11- January-2019].
[59] Feedforward neural network – Matlab feedforwardnet, [Online]. Available https://nl.mathworks.com/help/deeplearning/ref/feedforwardnet.html [Accessed: 11-January-2019].
[60] M. Barron, Auditorium Acoustics and Architectural design, 2nd ed, Spon Press, London and New York, 2009https://researchportal.bath.ac.uk/en/publications/auditorium-acoustics-and-architectural-design-2nd-edition , Accessed date: 27 February 2020.