

**Impacts of problem scale and sampling strategy on surrogate model accuracy
An application of surrogate-based optimization in building design**

Yang, Ding; Sun, Y; Di Stefano, D.; Turrin, Michela; Sariyildiz, Sevil

DOI

[10.1109/CEC.2016.7744323](https://doi.org/10.1109/CEC.2016.7744323)

Publication date

2016

Document Version

Accepted author manuscript

Published in

2016 IEEE Congress on Evolutionary Computation (CEC)

Citation (APA)

Yang, D., Sun, Y., Di Stefano, D., Turrin, M., & Sariyildiz, S. (2016). Impacts of problem scale and sampling strategy on surrogate model accuracy: An application of surrogate-based optimization in building design. In *2016 IEEE Congress on Evolutionary Computation (CEC)* (pp. 4199-4207). IEEE.
<https://doi.org/10.1109/CEC.2016.7744323>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Impacts of Problem Scale and Sampling Strategy on Surrogate Model Accuracy

An Application of Surrogate-based Optimization in Building Design

Ding Yang^{1,2}, Yimin Sun^{1,3}, Danilo di Stefano⁴, Michela Turrin^{2,3}, Sevil Sariyildiz^{2,5}

¹School of Architecture, South China University of Technology, Guangzhou, China

²Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, The Netherlands

³State Key Laboratory of Subtropical Building Science, South China University of Technology, Guangzhou, China

⁴ESTECO, Trieste, Italy

⁵Faculty of Architecture, Yaşar University, Izmir, Turkey

D.Yang-2@tudelft.nl, arymsun@scut.edu.cn, distefano@esteco.com, M.Turrin@tudelft.nl, sevil.sariyildiz@yasar.edu.tr

Abstract—Surrogate-based Optimization is a useful approach when the objective function is computationally expensive to evaluate, compared to Simulation-based Optimization. In the surrogate-based method, analytically tractable "surrogate models" (also known as "Response Surface Models - RSMs" or "metamodels"), are constructed and validated for each optimization objective and constraint at relatively low computational cost. They are useful for replacing the time-consuming simulations during the optimization; quickly locating the area where the optimum is expected to be for further search; and gaining insight into the global behavior of the system. Nevertheless, there are still concerns about the surrogate model accuracy and the number of simulations necessary to get a reasonably accurate surrogate model. This paper aims to unveil: 1) the possible impacts of problem scale and sampling strategy on the surrogate model accuracy; and 2) the potential of Surrogate-based Optimization in finding high quality solutions for building envelope design optimization problems. For this purpose, a series of multi-objective optimization test cases that mainly consider daylight and energy performance were conducted within the same time frame. Then, the results were compared, in pair, based on which discussions were made. Finally, the corresponding conclusions were obtained after the comparative study.

Keywords—surrogate-based optimization; problem scale; sampling strategy; response surface model; design of experiments; multi-objective optimization

I. INTRODUCTION

Computational Design Optimization (CDO) is a rising field of research in sustainable building design. It has been applied to many aspects including building envelope design, building service system, and renewable energy generation, etc. [1]. In this context, simulation-based optimization is frequently employed by architects and engineers to assist the early design decision-making. However, building performance simulations are usually time-consuming, for instance, annual hourly daylight and energy simulations or computational fluid dynamics simulations. This hinders the efficient application of simulation-based optimization in the building design practice, within the feasible time frames of projects.

Instead, surrogate-based optimization is a promising solution to this problem, in which surrogate models (also known as Response Surface Models - RSMs or metamodels) are utilized during the optimization. Basically, the surrogate model method is an approximation approach that mimic the behavior of the original simulation model at a reduced computational cost. It contains a group of mathematical and statistical techniques used to explore the functional relationship between input and output variables. In this method, the time-consuming simulations are replaced by the surrogate models during the optimization, which helps to improve the optimization efficiency. But, a reasonably accurate and valid surrogate model should be ensured before using it. This is the most challenging and crucial aspect of the method.

In the literature of sustainable building design, various types of surrogate models have been applied in the prediction of energy performance [2] and indoor environmental quality, such as thermal, daylighting [3] and ventilation performance [4]. In the applications, computationally expensive simulations (e.g. annual dynamic energy and daylight simulations, or CFD simulations) were replaced by validated surrogates during the optimization. Nevertheless, there are still concerns regarding the surrogate model accuracy and thus the applicability of the surrogate-based optimization. Because, in some cases, the number of simulations necessary to get a reasonably accurate surrogate may be approaching the number of simulations needed for the simulation-based optimization, as reported in [5].

In general, the paper aims at evaluating the applicability of the surrogate-based optimization to the building envelope design optimization problem that mainly considers daylight and energy performance. Specifically, the work investigates how, and to what extent, 1) different problem scales and sampling strategies may affect the accuracy of RSMs, and 2) the potential of RSMs in finding high quality solutions for the multi-objective optimization problems in question.

This paper is organized as follows. Section 2 introduces the method used in the study and relevant background knowledge; Section 3 describes the test cases, including the optimization problem formulation and the computational setup; Section 4

This research is sponsored by the Key Project of National Natural Science Foundation of China (Grant No.51138004).

The first author is sponsored through a scholarship by China Scholarships Council (CSC) for his joint Ph.D. research at Delft University of Technology, Delft, The Netherlands.

reports and discusses the results obtained; Finally, Section 5 presents conclusions and future research.

II. METHODOLOGY

A. Comparative Study

In order to achieve the research goal, a series of multi-objective optimization tests are arranged, based on two different problem scales, sampling strategies and optimization approaches respectively, as shown in Table I. By conducting comparative studies in pair (i.e. 1 vs 2; 2 vs 3; 2 vs 4), possible impacts of the each factor are investigated.

1) Problem Scale

To investigate the possible impacts of the problem scale on the accuracy of RSMs (by comparing test 1 and 2), two cases with different numbers of design variables are shaped based on a similar building envelope design optimization problem. Case 1 has two design variables, while Case 2 includes forty-one design variables. The increase of design variables usually represents the increase of complexity and design freedom; while, a large number of design variables is common, especially in the conceptual design stage. (more details in Section III)

2) Sampling Strategy

To investigate the possible impacts of the sampling strategy on the accuracy of RSMs (by comparing test 2 and 3), two sampling strategies are used, namely, "normal" and "adaptive" sampling strategies. Actually, they are differentiated from each other by the ways of generating sample points within the design space. The former selects all the sample points at once; while the later generates sample points iteratively by considering the information accumulated from the previous iteration. (more details in Section II - B)

3) Optimization Approach

To investigate the potential of RSMs in finding high quality solutions, RSM-based and simulation-based optimization approaches are compared (i.e. comparing test 2 and 4). In fact, the main difference between these two approaches lies in whether or not "shifting" (instead of eliminating) the computational efforts for simulations from within an optimization loop to a prior time. It is worth noting that running a certain number of simulations is required no matter whether RSMs are used or not. Simulations are used either for training response surfaces (in the RSM-based optimization), or for coupling directly with the optimization algorithm (in the simulation-based optimization).

Moreover, in this comparison, the "normal" sampling strategy is used in the RSM-based optimization, while the random selection of design points for the initial generation is used in the simulation-based optimization.

4) Time Frame for Comparison

For the sake of comparison, the same time frame of implementing each test should be ensured, which is assumed within 24 hours. Considering that simulations account for a major portion of time spent in all the tests, the number of simulations to be run for each test is an important monitoring factor. Additionally, also due to the comparability, the selection of optimization algorithms should be the same as well, namely, NSGA-II [6] in this study.

B. Normal and Adaptive Sampling Strategies

The sampling strategy (or sampling plan) in design variable space is known as Design of Experiments (DoE) [7]; thus it is sometimes called DoE strategy. The experiments in this paper are referred to as "computer experiments" or "simulations".

Basically, there are two types of DoE strategies or sampling strategies: space filling and adaptive sampling [8]. Hereby, we call them "normal DoE" and "adaptive DoE" respectively. The former selects a fixed number of samples uniformly covering the space or avoiding clusters, such as Latin Hyper Cube; while the later iteratively adjusts or leads the samplings to the complexity of the design space (i.e. the most irregular regions of the design space). In some cases, certain regions are much more complex than remaining regions, thus a more intensive exploration on those specific regions might be more efficient than a uniform space exploration [9]. By using a proper DoE strategy, a good sampling of design space can be obtained. This means that we could extract the most relevant information possible by making the smallest number of simulations.

1) Application of Normal Design of Experiments

Normal DoE is widely used in the surrogate-base optimization for training surrogate models [7, 10, 11, 12]. The traditional procedure of the surrogate-base optimization is a two-stage process [13], as shown in Fig. 1. The first stage is to generate surrogate models with the initial dataset obtained from running simulations. While in the second stage, the surrogate-base (or virtual) optimizations are run for searching promising areas where the optimum is expected to be globally; and decisions need to be made regarding the local accuracy of surrogates until valid optimal solution(s) are obtained. During the whole process, normal DoE is used in the first stage to generate the initial sample points for training surrogate models. Specifically, the two-stage process is described as follows:

TABLE I. TEST ARRANGEMENT AND COMPARATIVE STUDIES

Test No.	Problem Scales	Sampling Strategies	Optimization approaches	Impacts of Problem Scales	Impacts of Sampling Strategies	Impacts of Optimization approaches
1	Case 2 (41 vars)	Normal DoE	-	□		
2	Case 1 (2 vars)	Normal DoE	RSM-based	□	○	△
3	Case 1 (2 vars)	Adaptive DoE	-		○	
4.	Case 1 (2 vars)	Random	Simulation-based			△

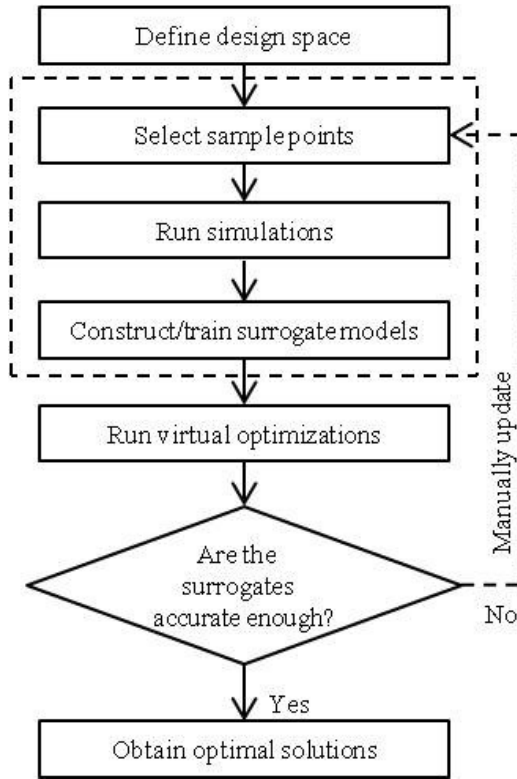


Fig. 1. Traditional surrogate-base optimization procedure

Step 1: Define design space and parameterize geometry;

Step 2: Select sample points from the design space by using "normal DoE";

Step 3: Run simulations for the selected sample points to obtain a database;

Step 4: Construct or train surrogate models based on the database, by selecting a proper type of surrogate model (e.g. Kriging);

Step 5: Run surrogate-base optimizations to quickly locate the promising areas where the optimum is expected to be globally (i.e. optimal regions of the design space);

Step 6: Check and verify the local accuracy of the surrogate models by comparing the simulation results and approximation results (to see if objective functions have been adequately approximated by the surrogate models)

Step 7: If no, then go back to step 2 (to manually select the sample points from the promising areas, run update simulations and update the surrogate models), and iterate until accurate enough optimal solution(s) are obtained.

Although this process has been widely used, there are still limitations. For instance: how many simulation runs will be needed for generating good initial surrogate models, and/or for updating them? Actually, it cannot be predicted in advance. In case of heavy simulations, this will be a very time-consuming procedure. More than this, it heavily relies on the experience of engineers in iteratively and manually updating the surrogate models.

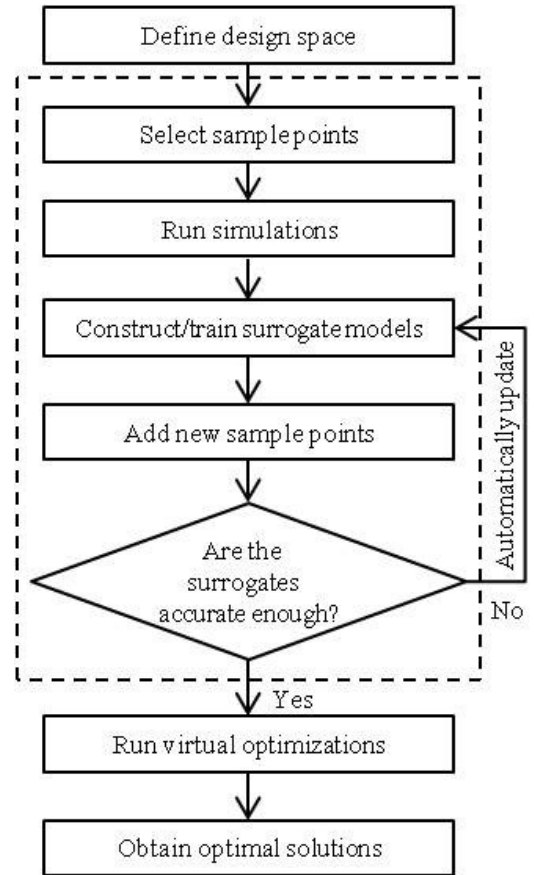


Fig. 2. Surrogate-base optimization procedure by using adaptive DoE

2) Application of Adaptive Design of Experiments

The strategy of adaptive DoE is slightly different, but the goal is to have an automated procedure. It is meant to minimize the number of simulation runs needed to generate a good surrogate model that can be used for the virtual optimizations. Specifically, the process is shown in Fig. 2.

The main difference between normal DoE and adaptive DoE lies in the incorporation of the Adaptive Space Filler (ASF) [14]. It is an adaptive DoE algorithm which iteratively adds new points to an existing database, and iteratively validates and trains the surrogate models until the required accuracy level (or the maximum number of iterations) is reached. The stopping criterion is defined by users, e.g. mean and max absolute error, relative error, normalized error, or R-squared (i.e. the coefficient of determination).

III. CASE STUDY

In order to perform the comparative study, two test cases are shaped based on a similar building envelope design optimization problem that mainly considers daylight and energy performance.

A. Test Case Description

The two hypothetical test cases are assumed to be one-story sports halls with rectangular floor plans and spherical roofs, located in Guangzhou, in south China, as shown in Fig. 3 and 4.

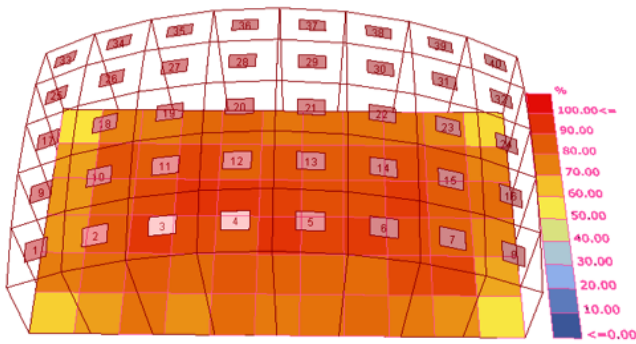


Fig. 3. Case 1 with two design variables

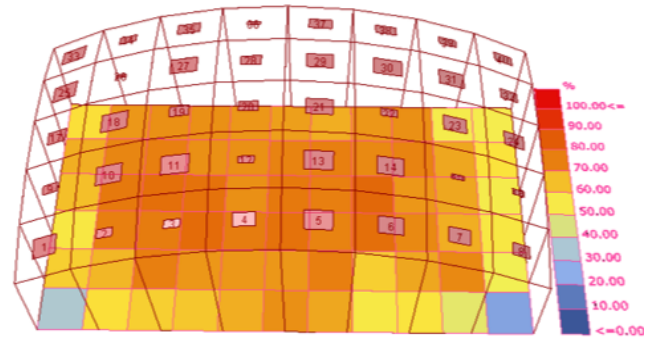


Fig. 4. Case 2 with forty-one design variables

The floor plan sizes of the both cases are the same and fixed (i.e. 40m*70m); while both roofs are changeable in height within certain ranges. The skylights are allocated to the "cells" of the roofs (i.e. 40 cells in total in each case); and the window-to-roof ratio(s) are also changeable in both cases.

Basically, the optimization problems of these two cases are similar, namely, figuring out the optimal roof configurations (i.e. roof heights and window-to-roof ratios) for both daylight and energy performance.

B. Optimization Problem Formulation

1) Design Variables

Case 1 and Case 2 share one common design variable - "Roof Height (RH)" which is defined as the vertical height from the bottom of the roof to the top. But, they are differentiated by their numbers of design variables needed for characterizing "Window-to-Roof Ratios (WRR)". In both cases, WRR(s) are defined as the ratio(s) between glazing area and roof area in each cell. Specifically, in Case 1, only one design variable is needed for characterizing it, because each cell has the same WRR. While in Case 2, forty design variables are needed in total, because all the cells are free to have an independent WRR respectively. This increases the degree of design freedom and enlarges the design space as well. In addition, considering that the focus here is the application of surrogate models, possible design variables regarding shading devices and/or constructions are not discussed in this paper for the sake of simplicity.

As for the ranges of design variables, it is important to make sure that they are meaningful for the specific cases. The range of Roof Height is between 1m to 15m; and the interval is 0.1m. By adding the fixed height below the roof (i.e. 15m, which is also the minimum height requirement for many sports activities), the overall variation of the total height of the building is between 16m to 30m. This matches the normal volume of a 40m*70m sports hall. For the range of Window-to-Roof Ratio(s), they vary from 0.01 to 0.50 and the interval is 0.01. The reason for this is that they are sensitive to the objective functions and meet the accuracy requirement of the research.

2) Objectives

A modification of Useful Daylight Illuminance (UDI) and Energy Use Intensity (EUI) are selected as optimization objectives for daylight and energy performance in this study, respectively, forming a multi-objective optimization problem

Useful Daylight Illuminance (UDI) is an annual, hourly metric for assessing daylight provision in buildings, devised by Mardaljevic and Nabil [15]. It uses hourly sun and sky condition data, providing more complete and accurate information than the traditional daylight factor approach. It is considered useful also because of the attempt to integrate the evaluation of daylight level and glare risk simultaneously.

In order to adapt to the large-volume cases in question, a modification based on the original UDI metric is needed. The original UDI is defined as the annual occurrence of illuminances that are within a "useful" range considered by occupants across the work plane. To be specific, it refers to the percentage of time during the occupancy hours that a test point receives hourly illuminances between 100 and 2000 lux. However, in order to know the overall daylighting condition throughout the entire space (i.e. 40m*70m), more than one test points (i.e. 66, in this study) is needed and should be evenly spread at the work plane. Thus, a modification based on UDI is necessary for characterizing the overall condition of the whole space by using one single value. In this case, we define UDI_{mod} as the percent of test points across the whole analysis area with the original UDI value larger than 50%. To make it clear, it represents the percentage of floor area that receives "useful" illuminance (i.e. 100 - 2000 lux) for at least 50% of the occupancy hours. UDI_{mod} should be maximized in the multi-objective optimization.

Energy Use Intensity (EUI) is a basic approach to benchmark a building's energy efficiency or performance. It is defined as the energy consumption per unit of floor area (kWh/m^2) of a building over one year, thus facilitates direct comparison with other buildings, giving us a general idea of how energy efficient the building is. In this study, we only consider "site energy" instead of primary or secondary energy; and a Coefficient of Performance (CoP) value of 3 is set, assuming that only electricity was used. EUI should be minimized in the multi-objective optimization.

C. Computational Setup

Computational setup includes the integration of several tools used by the case studies; each of these tools is mainly responsible for a different task during the whole process. It also includes the setup of the parametric model, the coupled simulation and the automatic process (or mathematical optimization); while the coupled simulation setup is the focus of this section, due to its importance for the evaluations of objectives.

1) Tool Integration

Programs involved in this study include: Rhinoceros [16] and Grasshopper [17]; Daysim [18] and EnergyPlus [19]; modeFRONTIER [20]. They are responsible for parametric modeling, numeric simulation and mathematical optimization (or process automation), respectively. The connection or data exchange between them mainly relies on the related plug-ins in Grasshopper, for instance, Honeybee and Ladybug [21] for connecting Daysim and EnergyPlus; and customized plug-ins for connecting modeFRONTIER. Generally, the integration between these tools is illustrated in Fig. 5, and a more detail description can be found in [22].

2) Coupled Simulation

UDI_{mod} and EUI can be evaluated or simulated by the daylight simulation in Daysim and the energy simulation in EnergyPlus, "separately". But, this will lead to the result that energy saving by using daylight cannot contribute to or affect the total energy use. Thus, it requires a coupled daylight and energy simulation.

Specifically, for the evaluation of UDI_{mod} , the annual, hourly dynamic daylight simulation is required; and Daysim is competent to manage this, which has been proved by many other works [23, 24, 25]. While, for the evaluation of EUI, a coupled simulation by integrating Daysim and EnergyPlus programs is recommended, because of the significant limitation regarding the daylighting module in EnergyPlus [26]. In this case, Daysim generates a report that describes the control of the electric lighting depending on the admission of daylight (i.e. electric lighting schedule), which is then used by EnergyPlus for calculating the final energy consumption. Thus, there is a coupling (i.e. data transfer) between Daysim and EnergyPlus; and they run in a sequential way. In short, the coupled simulation is required and applied in this study.

IV. RESULTS AND DISCUSSION

A. Impacts of Problem Scale on RSM Accuracy

In order to investigate possible impacts of the problem scale on the surrogate model accuracy, test 1 and test 2 are carried out by using two cases with different problem scales (i.e. number of design variables). Some common settings shared by the both cases are described below.

First, Uniform Latin Hypercube [27] sampling strategy - a frequently used normal DoE, is applied here for the both cases to guarantee a relatively uniform distribution of design points over each dimension. Then, 360 sample points are selected at once to run the simulations (i.e. 21.5 hours). Moreover, five types of surrogate models are trained for UDI_{mod} and EUI respectively by using 330 sample data; and the remaining 30 sample data are used for surrogate model validation.

The five surrogate model types considered include Polynomial Singular Value Decomposition (SVD); Stepwise Regression (STEP); Kriging (KR); Shepard K-Nearest (KN) and Radial Basis Functions (RBF). Detailed descriptions of these types can be found in [27]. Moreover, the surrogate model accuracy is expressed by means of the Mean Absolute Error (MAE), Mean Normalized Error (MNE) and R-squared

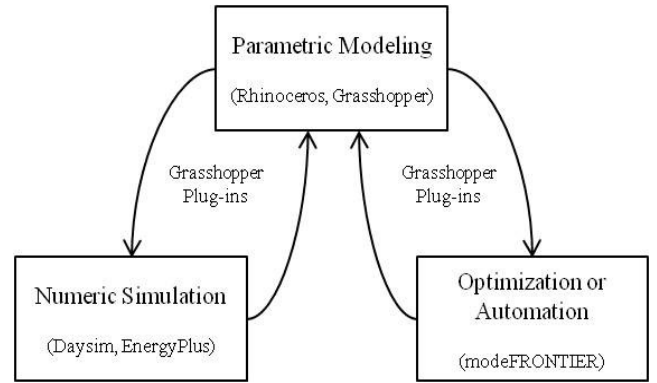


Fig. 5. Tool integration

(R^2) [20]. In general, the lower the value of MAE and MNE, the better the surrogate accuracy; while R^2 should be as close as possible to 1, indicating the model fits the data well.

TABLE II. RESULTS ON THE ACCURACY OF GENERATED SURROGATES IN TEST 1, 2 AND 3

Objectives	Type of Surrogate Model	Mean Absolute Error (*)	Mean Normalized Error	R-squared
Test 1 (Case 2, forty-one design variables, normal DoE)				
UDI_{mod}	SVD	4.70E0	7.39E-2	8.48E-1
	RBF	4.83E0	7.59E-2	8.65E-1
	STEP	4.85E0	7.63E-2	8.57E-1
	KR	5.79E0	9.10E-2	8.30E-1
	KN	1.28E1	2.02E-1	3.59E-1
EUI	SVD	2.11E-1	1.47E-2	9.88E-1
	STEP	2.23E-1	1.55E-2	9.88E-1
	RBF	2.73E-1	1.90E-2	9.84E-1
	KR	3.32E-1	2.32E-2	9.79E-1
	KN	1.97E0	1.37E-1	3.57E-1
Test 2 (Case 1, two design variables, normal DoE)				
UDI_{mod}	RBF	3.56E0	3.56E-2	9.59E-1
	KN	5.11E0	5.11E-2	9.22E-1
	STEP	1.86E1	1.86E-1	6.52E-1
	SVD	1.89E1	1.89E-1	4.76E-1
	KR	3.79E1	3.79E-1	-1.86E0
EUI	RBF	7.53E-1	1.99E-2	9.76E-1
	KN	1.30E0	3.43E-2	9.49E-1
	STEP	3.43E0	9.07E-2	7.85E-1
	SVD	4.25E0	1.12E-1	5.84E-1
	KR	3.11E1	8.21E-1	-3.14E1
Test 3 (Case 1, two design variables, adaptive DoE)				
UDI_{mod}	KR	3.96E0	3.96E-2	9.65E-1
EUI	KR	9.36E0	2.47E-1	1.95E-1

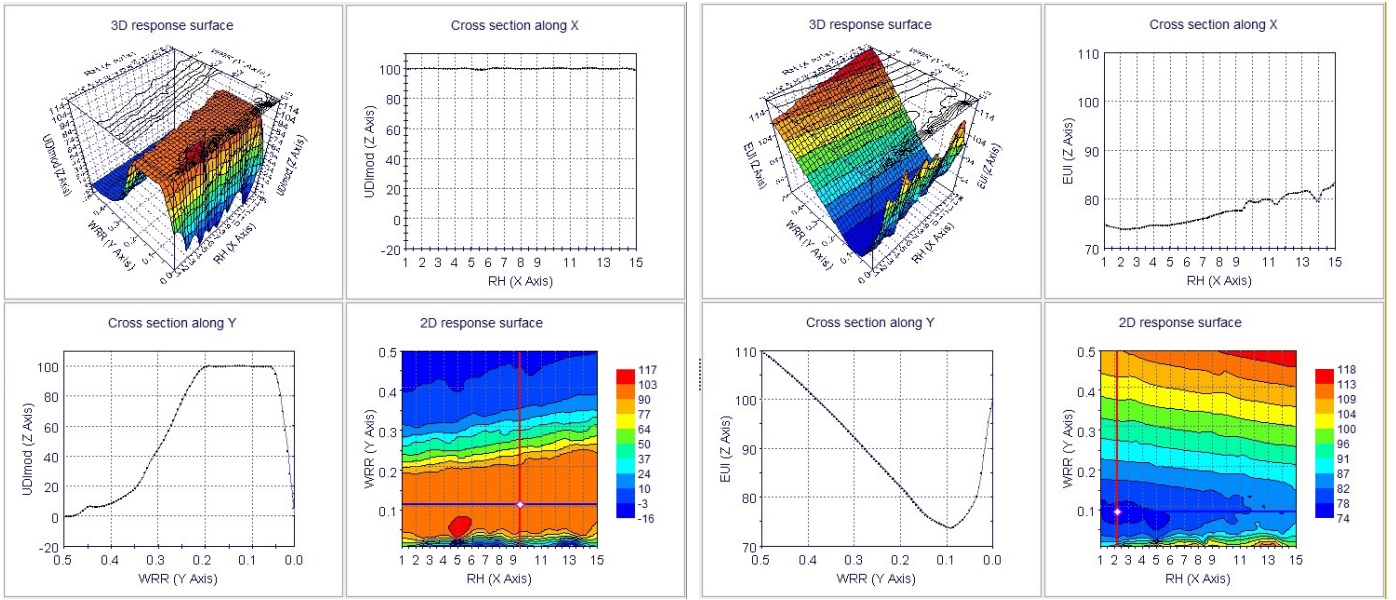


Fig. 6. Response surface for UDI_{mod} in test 2 (left); Response surface for EUI in test 2 (right)

The results on the surrogate model accuracy in test 1 and 2 (shown in Table II) are compared within each case and between cases. In order to facilitate the comparison, all the results are sorted by Mean Absolute Error in ascending order. Within each case, we found that the approximation abilities of the five surrogate models are different in varying degrees. Considering the performance of all the listed accuracy indices, SVD (marked in Blue) and RBF (marked in Yellow) are eventually selected out of the five surrogate models for Case 2 and Case 1, respectively. By further comparing the selected surrogate models within each case, we also found that the surrogate model for EUI is more accurate than that for UDI_{mod} . Moreover, by the comparison between cases, UDI_{mod} is approximated relatively better in Case 1 than that in Case 2 (due to the relatively large difference in R^2); while EUI is approximated equally well in both cases (because of similar values for all accuracy indices).

In summation, the above results indicate that (1) simultaneously comparing multiple types of surrogate models for a specific objective is helpful for obtaining a more accurate surrogate model; (2) EUI is relatively easier to be approximated compared to UDI_{mod} when using the number of training data; and (3) the problem scale has a larger impact on UDI_{mod} than EUI in this specific case study (i.e. the surrogate model accuracy for UDI_{mod} decreases obviously along with the increase of the number of design variables). The latter two indications may be associated with the complex input-output behavior of the original model.

In order to have more informed knowledge of the input-output behavior of the original model, visualization of the generated response surfaces is helpful. As examples, the response surfaces (i.e. RBF) for UDI_{mod} and EUI in test 2 are illustrated in Fig. 6. Specifically, as we can observe from the 3D and 2D response surfaces for both UDI_{mod} and EUI, WRR (i.e. Y axis) is more sensitive than RH (i.e. X axis), because the slope along Y axis is much more steep than the slope along X

axis. And, there is an obvious "flat area" in the response surface for UDI_{mod} , while the response surface for EUI is relatively smooth, which confirms that the input-output behavior of UDI_{mod} is more complex than that of EUI. Moreover, by observing the cross sections along X and Y axes, we can identify the most sensitive ranges of WRR, and have a general idea of how UDI_{mod} and EUI change along with WRR and RH and the corresponding variation ranges. The input-output behavior obtained from Fig. 6 is also consistent with the educated guess in terms of building performance of this specific case, which in turn proves the correctness of the response surfaces.

In short, apart from using response surfaces for the surrogate-based optimization, visualizing response surfaces gives us an easy way to understand the overall behavior of the original model. With this information, we could reformulate the design space in a more informed manner when it is needed.

B. Impacts of Sampling Strategy on RSM Accuracy

In order to investigate possible impacts of the sampling strategy on the surrogate model accuracy, test 2 and test 3 are carried out by using different DoE strategies (i.e. normal DoE and adaptive DoE). Different from the normal DoE in test 2, an adaptive DoE algorithm called "Adaptive Space Filler (ASF)" [14] is applied in test 3. The total simulation time for both daylight and energy is around 22.3 hours (similar with test 2); other settings for test 3 are described here below in Table III.

TABLE III. SETTINGS FOR ADAPTIVE SPACE FILLER IN TEST 3

	Iterations	Sample Size	Space Filling Algorithm	Type of Surrogates	Termination Criterion
UDI_{mod}	43	10	Incremental Space Filler	Kriging	Mean Absolute Error < 0.5
EUI	21				

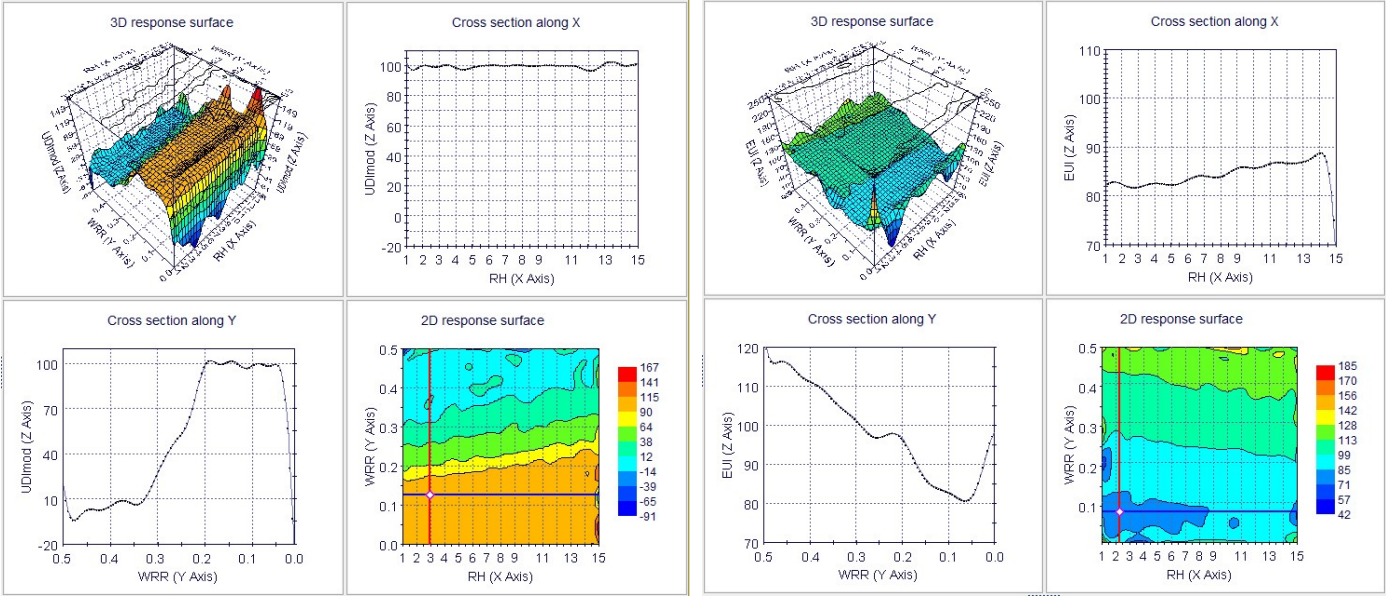


Fig. 7. Response surface for UDI_{mod} in test 3 (left); Response surface for EUI in test 3 (right)

The results on the surrogate model accuracy in test 3 (shown in Table II) are compared with that in test 2. As we can observe from the results, the surrogate model for UDI_{mod} in test 3 has similar accuracy level with the one in test 2. However, we are aware that some small variations or fluctuations occur compared with the original smooth response surface in test 2. This could be interpreted as the ability of the adaptive DoE in exploring complex regions. Moreover, regarding the surrogate model for EUI in test 3, its approximation ability is much lower than the one in test 2 according to the current results, although the general trend of the overall behavior is similar. The low accuracy level achieved is far away from the expectation; possible reasons might be associated with some detail settings for ASF, which needs further research.

In summation, the above results indicate that the adaptive DoE has potential to explore complex regions of design space, apart from the overall behavior of the original model. But, some settings for ASF need to be fine-tuned in further research.

C. Impacts of RSM-based Approach on Optimization Results

In order to investigate possible impacts of the RSM-based approach on the optimization results, both the RSM-based and simulation-based approaches are applied on the same case (i.e. Case 1) within the similar time frame. The settings for the both approaches are described here below in Table IV.

TABLE IV. SETTINGS FOR RSM-BASED AND SIMULATION-BASED OPTIMIZATION IN TEST 2 AND 4

	Iterations	Sample Size	Optimization Algorithm	Simulation Runs	Time (hrs)
RSM-based	500	30	NSGA-II	360 (for RSMs)	21.5
Simulation-based	14			420	22.3

As shown in the scatter plot charts of Fig. 8, it should be first noted that UDI_{mod} and EUI show a highly negative correlation between them. This means that, in general, both objectives change in the same direction, namely, EUI decreases along with the increase of UDI_{mod} as expected. It just matches the purpose of UDI_{mod} , as a metric that is meant to utilize daylight and prevent overheating simultaneously for energy saving. In this case, the optimal result(s) obtained from this multi-objective optimization can be either a unique solution or a very "short" Pareto Front consisted of a few solutions; plus, the optimal result(s) will appear on the top-left corner of the scatter plot charts.

As shown in Fig. 9 (left), all the solutions in Fig. 8 are plotted in the same chart. We can find that a much larger number of solutions (i.e. 15,000 in this case) are obtained in very short optimization time (i.e. a few seconds) by using the RSM-based approach. This is the major advantage of the RSM-based optimization over the simulation-based optimization. As we can see, in general, the overall distribution of the solutions from the RSM-based and simulation-based approaches matches. This proves the general correctness of the surrogate model used, to a certain extent. Thus, the RSM-based approach has the potential to search the overall design space sufficiently and locate the promising regions. Nevertheless, the surrogate models used may not be accurate enough locally. As shown in the blue ellipse, some approximate solutions exceed the valid range of UDI_{mod} (i.e. between 0% and 100%), due to the limited local accuracy of the surrogate models. Therefore, it is necessary to verify the local accuracy of the surrogate models at prefer regions and update the surrogate models locally. As shown in Fig. 9 (right), 20 design points are randomly selected to run the simulations. The verification result shows that most of the selected design points are approximated well, except the ones marked (by a boxed point and a circled point).

In summation, the above results indicate that the RSM-based optimization has the potential to search the overall

design space sufficiently and locate the promising regions where the optimum is expected to be. However, due to the limitation on the local accuracy of the surrogate models, further local verification at prefer regions and surrogate model update are suggested.

V. CONCLUSIONS

Due to the time-consuming simulations in building design, the effectiveness of the simulation-based optimization tends to be limited. To conquer this obstacle, the surrogate-based optimization is suggested by many researchers. In this paper, a series of tests were designated to investigate the possible impacts of problem scale and sampling strategy on the surrogate model accuracy, and the potential of surrogate-based optimization in finding high quality solutions for multi-objective optimization problems. Based on the observation of current results, some conclusions are summarized below.

The problem scale has a larger impact on UDI_{mod} than EUI in this specific case study. The accuracy of the surrogate model for UDI_{mod} decreases obviously, along with the increase of the number of design variables; while the accuracy of the surrogate model for EUI remains similar.

The adaptive DoE has potential to explore complex regions of design space, apart from the overall behavior of the original model. But, the accuracy of the generated surrogate model for EUI is low in this case.

The RSM-based optimization has the potential to search the overall design space sufficiently and locate the promising regions where the optimum is expected to be. However, due to the limitation on the local accuracy of the surrogate models, further local verification at prefer regions and surrogate model update are suggested. Additionally, apart from using response surfaces for the surrogate-based optimization, they are also useful for gaining insight into the global behavior of the system.

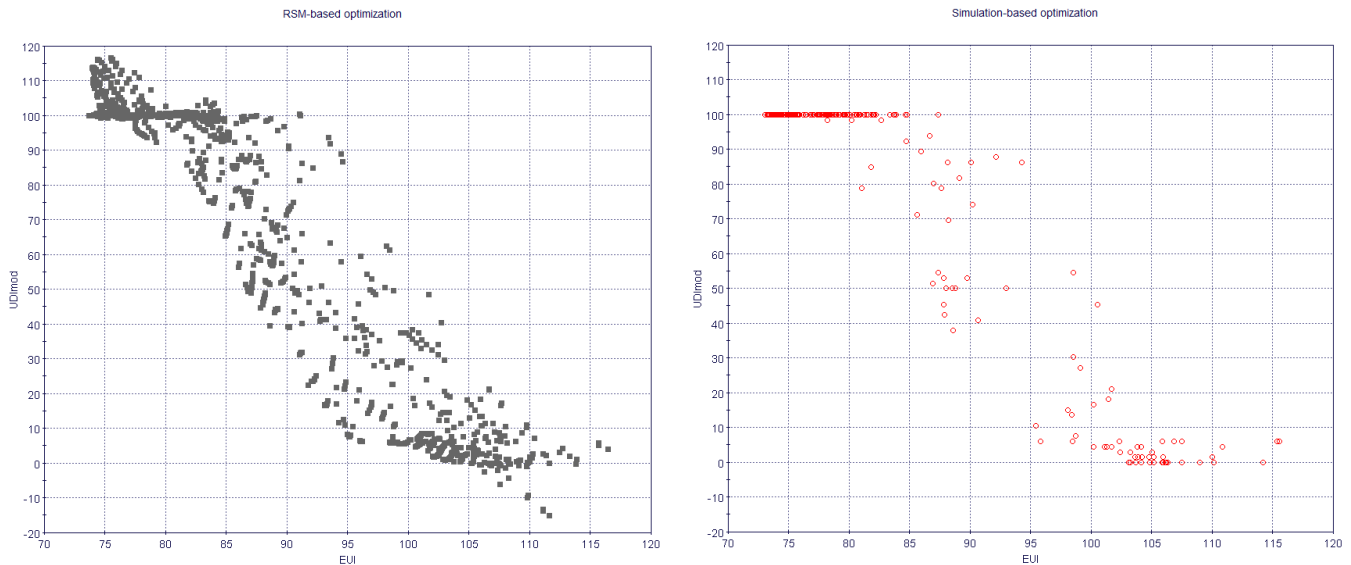


Fig. 8. RSM-based optimization (left); Simulation-based optimization (right)

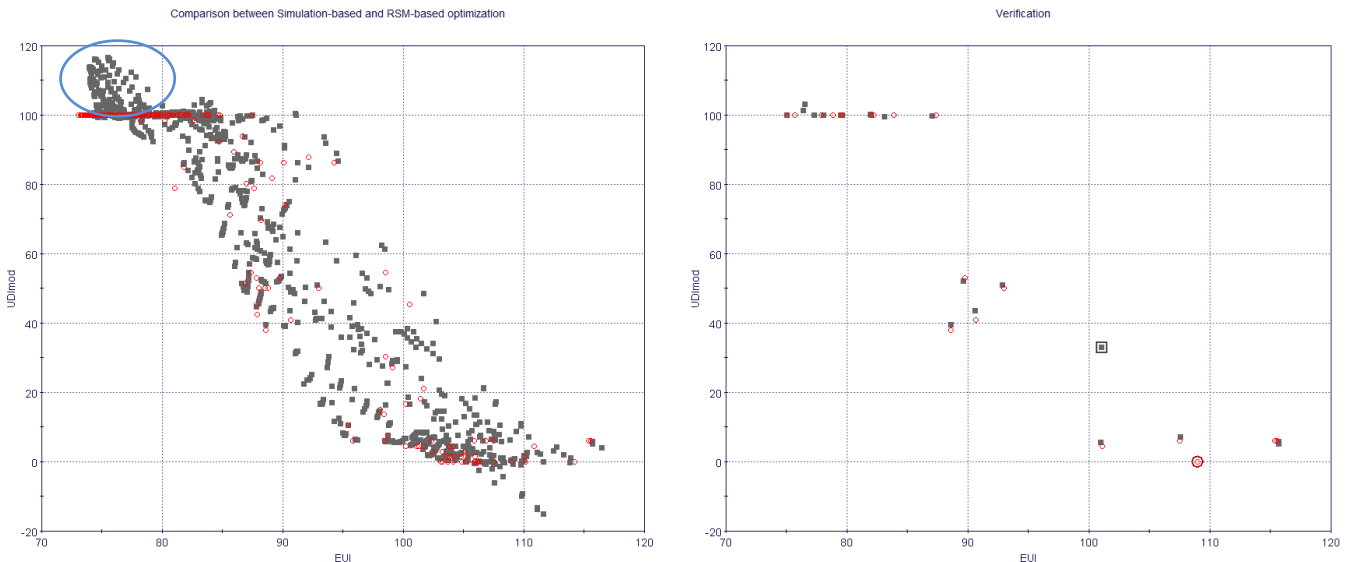


Fig. 9. Comparison between RSM-based and simulation-based optimization (left); Verification of randomly selected 20 design points (right)

In future research, the surrogate-based optimization is expected to be applied in a more complex case, and detail settings for ASF need to be fine-tuned in order to improve the approximation accuracy of the generated surrogate models.

ACKNOWLEDGMENT

We would like to thank for the technical support from ESTECO, and also thank our colleagues, Antonio D'Aquilio and Rusne Sileryte for many useful discussions.

REFERENCES

- [1] R. Evins, "A review of computational optimisation methods applied to sustainable building design," *Renewable and Sustainable Energy Reviews*, vol. 22, pp. 230-245, June 2013.
- [2] F. Pernodet, H. Lahmidi, and P. Michel, "Use of genetic algorithms for multicriteria optimization of building refurbishment," in 11th International IBPSA Conference, 2009.
- [3] K. Hiyama and L. Wen, "Rapid response surface creation method to optimize window geometry using dynamic daylighting simulation and energy simulation," *Energy and Buildings*, vol. 107, pp. 417-423, November 2015.
- [4] P. Sofotasiou, H. Ben, and K. John, Calautit, "Towards an Integrated Computational Method to Optimise Design Strategies for the Built Environment," in 14th International Conference on Sustainable Energy Technologies (SET), Nottingham, UK, 2015.
- [5] L. Dietsche, P. Lee, and J. Dooley, "CFD-Based Optimization Methods Applied to Polymer Die Design," 2011.
- [6] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *Evolutionary Computation*, *IEEE Transactions on*, vol. 6, pp. 182-197, 2002.
- [7] N. V. Queipo, R. T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P. Kevin Tucker, "Surrogate-based analysis and optimization," *Progress in Aerospace Sciences*, vol. 41, pp. 1-28, January 2005.
- [8] P. de Oliveira Castro, E. Petit, J. C. Beyler, and W. Jalby, "ASK: Adaptive Sampling Kit for performance characterization," in Euro-Par 2012 Parallel Processing, ed: Springer, 2012, pp. 89-101.
- [9] A. Lovison and E. Rigoni, "Adaptive sampling with a Lipschitz criterion for accurate metamodeling," *Communications in Applied and Industrial Mathematics*, vol. 1, pp. 110-126, 2011.
- [10] Q. Xu, E. Wehrle, and H. Baier, "Knowledge-Based Surrogate Modeling in Engineering Design Optimization," in *Surrogate-Based Modeling and Optimization*, ed: Springer, 2013, pp. 313-336.
- [11] E. Asadi, M. G. d. Silva, C. H. Antunes, L. Dias, and L. Glicksman, "Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application," *Energy and Buildings*, vol. 81, pp. 444-456, October 2014.
- [12] E. Tresidder, Y. Zhang, and A. I. Forrester, "Acceleration of building design optimisation through the use of kriging surrogate models," *Proceedings of building simulation and optimization*, pp. 1-8, 2012.
- [13] A. I. Forrester, N. W. Bressloff, and A. J. Keane, "Optimization using surrogate models and partially converged computational fluid dynamics simulations," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 2006, pp. 2177-2204.
- [14] T. Montrone, S. Silvetti, and M. Engel, "Adaptive Space Filler: General Description," ESTECO Technical Report 2015-0022015.
- [15] A. Nabil and J. Mardaljevic, "Useful daylight illuminances: A replacement for daylight factors," *Energy and Buildings*, vol. 38, pp. 905-913, July 2006.
- [16] Rhinoceros, <https://www.rhino3d.com/>
- [17] Grasshopper, <http://www.grasshopper3d.com/>
- [18] Daysim, <http://daysim.ning.com/>
- [19] EnergyPlus, <https://energyplus.net/>
- [20] modeFRONTIER, <http://www.esteco.com/modefrontier/>
- [21] Honeybee and Ladybug, <http://www.grasshopper3d.com/group/ladybug/>
- [22] D. Yang, Y. Sun, M. Turrin, P. von Buelow, and J. Paul, "Multi-objective and multidisciplinary design optimization of large sports building envelopes: a case study," in *Proceedings of the International Association for Shell and Spatial Structures (IASS) Symposium - Future Visions*, Amsterdam, 17-20 August 2015.
- [23] C. F. Reinhart and D. A. Weissman, "The daylit area – Correlating architectural student assessments with current and emerging daylight availability metrics," *Building and Environment*, vol. 50, pp. 155-164, April 2012.
- [24] M. R. Fontenelle and L. E. G. Bastos, "The multicriteria approach in the architecture conception: Defining windows for an office building in Rio de Janeiro," *Building and Environment*, vol. 74, pp. 96-105, April 2014.
- [25] M. Manzan, "Genetic optimization of external fixed shading devices," *Energy and Buildings*, vol. 72, pp. 431-440, April 2014.
- [26] E. L. Didoné and F. O. R. Pereira, "Integrated computer simulation for considering daylight when assessing energy efficiency in buildings," *International Building Performance Simulation Association*, Sydney, 14-16 November 2011.
- [27] ESTECO, "modeFRONTIER 4.1 User Manual," Trieste, Italy.