A comparison study of various sensitivity analysis methods in building applications

Zhihong Pang, Zheng O'Neill

The University of Alabama, Tuscaloosa, AL, United States

Abstract

This paper aims to present a comparison study on the performances of different SA methods in building energy simulation. Firstly, a literature review on the principles and applications of several popular SA methods was conducted. Secondly, an EnergyPlus based Monte Carlo simulation was performed as the case study. Then, various SA methods were used to compute the sensitivity indexes. Finally, a comprehensive comparison of the SA methods in terms of the time and computational costs, accuracy, convergence rate, and the complexity was provided. This study is expected to provide a specific and detailed guidance on the selection of the SA method to the practitioners in the building simulation field.

Introduction

With the depletion of fossil fuels and the increased focus on the occupant health, building energy efficiency and high-performance buildings are arousing more and more attention these years. To serve this purpose, detailed building performance simulation (BPS) are extensively used in both design and operation stages of high-performance building (Hensen and Lamberts, 2012). However, the increasing model complicity also brings the curse of dimensionality, which makes it hard to understand the model characteristics (Pang et al., 2018). As an effective tool to help understand how sensitive the model outputs are to the various input parameters, sensitivity analysis (SA) is gaining its popularity in both the academia and the industry. Sensitivity analysis is the study of how uncertainty in the model output can be apportioned to the different sources of uncertainty in the model inputs (Saltelli et al., 2008). It is effective to enhance the various model-based studies such as traditional “if-then” analyses, model-based controls, etc. (Hopfe et al., 2006). Building models are highly non-linear and commonly involve a diverse collection of input parameters (Gu et al., 2018; Hensen and Lamberts, 2012; Pang et al., 2017). Thus, building simulation is one of the fields that can significantly benefit from the SA. In practice, it has been widely used to rank the input parameters to facilitate the model simplification, calibration, and optimization.

Last decades have emerged a large amount of SA-based researches and applications in the area of building performance simulation, e.g., Lam and Hui used SA to identify the key parameters for the building energy consumption of an office building in Hong Kong (Lam and Hui, 1996).

Their results suggested that the SA techniques were useful in assessing the variabilities of the building system. O'Neill and Eisenhower leveraged SA to select the most influential parameters on the building energy consumption. These parameters were then refined to calibrate the building energy models automatically with little expert intuition (O'Neill and Eisenhower, 2013). Sun et al. explored using uncertainty analysis (UA) and SA to replace the traditional safety factor with quantified margins based on the quantification of various uncertainties (Sun et al., 2014). Such method was found effective to solve the over-sizing problem and also provide risk management capabilities. Gaetani et al. conducted an SA based simulation to investigate the most influential aspects of occupant behaviour and the appropriate model complexity level for the simulation (Gaetani et al., 2017). The study concluded that increasing the modelling complexity does not necessarily enhance the robustness and accuracy of the simulations. O'Neill and Niu conducted an SA to quantify the impacts of occupant behaviours on the residential building energy usage in three different locations (O'Neill and Niu, 2017). Their simulation result indicated that the influences of the various occupant behaviours tend to be varied with the climate but the thermostat setpoints have the biggest influences in all three cities.

However, even though the last two decades have witnessed a large amount of model-based SA studies of building system and sub-systems, how to select the most appropriate SA method based on given objectives is still a challenging problem for building scientists and engineers. This is mainly due to the relative absence of the expertise in statistics. There have been several review or critical review papers focusing on guiding the selection of the SA methods for the BPS practice. These studies provide some useful information to the practitioners when they implement an SA research. However, they all suffer from some limitations and restrictions.

Tian conducted comprehensive reviews on the SA application for the building energy assessment (Tian, 2013; Tian et al., 2018). The detailed procedures to implement an SA was introduced, including the pros and cons of various sampling strategies, SA methods, software and tools, etc. But how the various SA methods perform in terms of accuracy and efficiency was not investigated. Menberg et al. scrutinized three SA methods, i.e., the Morris method, the linear regression method, and the Sobol method in building energy modelling (Menberg et al., 2016). The results suggested that computationally intensive methods

reveal insights into complex model behaviours. However, many other popular SA methods, such as standard regression coefficient (SRC) were not included in Menberg's case study. Nguyen and Reiter conducted a comprehensive critical review which covered a wide range of SA methods (Nguyen and Reiter, 2015). But the two building models used in the study were relatively simple for the real-world applications, with one being a selected part of a building and the other being a detached residential house. Besides, the numbers of the studied input parameters were small, which may hide some features of the SA methods in terms of accuracy and computational efficiency.

This paper aims at investigating the performances of the widely-used SA methods in whole-building energy simulation. The large hotel prototype building model developed by U.S. DOE (Department of Energy) was used as the case study. 33 parameters were selected as the inputs, based on which 3,000 Monte Carlo (MC) simulations were performed to generate a metamodel of the gas consumption of the Domestic Hot Water (DHW) system. Then, a systematic comparison of ten SA methods including the screening method, correlation analysis, regression analysis, and the variance analysis were conducted. A wide range of metrics including the time and computational costs, the convergence rate, the accuracy of the result, and the complicity of the methods were investigated to facilitate the selection of the SA method in building application. The evaluation results suggested that Morris and SRC demonstrated outstanding performances in terms of convergence rate and accuracy in this case study.

Literature Review

Generally, there are three typical steps to implement an SA for building performance simulation (Tian, 2013):

- Firstly, a baseline energy model should be created as the base case.
 - Secondly, the inputs of interest should be sampled with appropriate distributions, uncertainty ranges, and sampling methods to generate enough cases for Monte Carlo simulation.
 - Thirdly, these models are executed in batch. The inputs and output are extracted for sensitivity analysis.
- The categorization of SA methods varies a lot in different researches based on their scope, applicability, and characteristic. As usual, they are classified as local and global approaches. However, there are some studies insisting that the screening method is neither local nor global, thus should be grouped independently (Heiselberg et al., 2009). In this study, the ten SA methods to be investigated were classified into four groups, i.e., the screening method, the correlation analysis, the regression analysis, and the variance analysis (Pang et al., 2019).

Screening method

The screening method, which is also known as the Morris method or the factorial design, was first proposed by Morris in 1991 (Morris, 1991). Morris method is an efficient method to calculate the sensitivity of the input parameters

by varying all the parameters simultaneously. Specifically, it divides the ranges of the input parameters into k levels, thus the input space of a n -parameter model is transformed into a grid formed by $n*(k+1)$ points, with each point representing a model evaluation. Then, the values of all the input parameter are changed one by one in one iteration, until all the points in the input spaces have been evaluated. Therefore, it is widely considered as a global method. Morris method evaluates the sensitivity of an input with a factor called elementary effect (EE), or μ^* . The definition of EE is presented in Eq. (1). Besides EE, Morris method also outputs the standard deviation, or σ , which evaluates the interaction among input parameters, as defined in Eq. (2) (Morris, 1991). There are some extended usage of the Morris method in recent years (Campolongo et al., 2007; Menberg et al., 2016). However, in this study, only the EE was computed for the sensitivity since we only care about the input ranking.

$$\mu^* = \frac{1}{k} \sum_{i=1}^k \left| \frac{Y(X + \Delta_i) - Y(X)}{\Delta_i} \right| \quad (1)$$

$$\sigma = \sqrt{\frac{1}{k} \sum_{i=1}^k (\mu_k - \mu^*)^2} \quad (2)$$

Morris method has the advantages of better computational efficiencies and easy implementations. Besides, it is effective to discriminate the non-influential input parameters and can take the interactions among inputs into account (Tian, 2013). It was widely used in BPS to remove the insignificant parameters prior to the further SA, e.g., (Brembilla et al., 2018; De Wit and Augenbroe, 2002; Sanchez et al., 2014). But some studies found that the Morris method is not robust enough (Menberg et al., 2016; Yang et al., 2016). It also cannot quantify the specific interaction between two input parameters.

Correlation analysis

The correlation analysis measures the strength of the linear relationship between two input parameters. The most common correlation index is the Pearson correlation coefficient (PEAR), which is also known as the product moment correlation coefficient. Such method is computational efficient, and easy to understand. But it is only suitable to the linear and monotonic models. In the monotonic but non-linear cases, its rank transformation, Spearman's rank correlation coefficient (SPEA) should be used instead. It works by replacing the raw values of the inputs and outputs with their ranks. The definition of such methods is defined in Eq. (3) and Eq. (4) respectively. The value of the correlation indexes varies between -1 and 1, with the sign and the absolute numerical value indicating the direction and the strength of the linear relationship. The application of the correlation analysis methods in building application can be found in (Hester et al., 2017; O'Neill and Niu, 2017; Østergård et al., 2017)

$$\rho = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (3)$$

$$r_s = \frac{\sum(r_{xi} - \bar{r}_x)(r_{yi} - \bar{r}_y)}{\sqrt{\sum(r_{xi} - \bar{r}_x)^2 \sum(r_{yi} - \bar{r}_y)^2}} \quad (4)$$

Regression analysis

Like the correlation analysis, the regression techniques are only suitable to the linear models as well. The most common regression method is the standardized regression coefficient (SRC). The SRC works by using a regression model, as presented in Eq. (5), to fit the input and output data. The terms b and ε in Eq. (5) represent the coefficient of the input and the residual due to the linear approximation respectively. The accuracy and reliability of the SRC are highly dependent on the linearity of the model, i.e., the R-squared value of the regression model. When the interactions of the input parameters are unneglectable, the partial correlation coefficient (PCC), which is able to clean off such coupling effect, is recommended to be used. Similar to the PEAR method, they both have rank transformations, i.e., the standardized rank regression coefficient (SRRC), and the partial rank correlation coefficient (PRCC). Those rank transformations are suitable to the nonlinear models. They are effective to reduce the impacts of the extreme values, but some researchers found that such rank transformations resulted in a hard convergence and unstable outcomes (Nguyen and Reiter, 2015).

$$y_i = a + \sum_j b_j x_{ij} + \varepsilon \quad (5)$$

Tian concluded that SRC was the most popular SA method in the building energy assessment (Tian, 2013). Its wide application in the building science can be found in (de Wilde and Tian, 2009; Menberg et al., 2016; Qiu et al., 2018; Rodríguez et al., 2013).

The regression-based methods are easy to implement and comprehend. Also, they are computationally efficient compared with the variance-based methods. However, such approaches are model-dependent, which means that their usage is restricted to the linear models.

Variance-based methods

The variance-based sensitivity techniques are often used when it comes to the nonmonotonic and nonlinear models. They decompose the variance of the outputs into a sum of contributions by the inputs and the combinations thereof, as defined in Eq. (6), where X is the input parameter and f is the function of the input parameters. For this reason, they are also known as the analysis of variance (ANOVA) methods.

$$Y = f_0 + \sum_{i=1}^d f_i(X_i) + \sum_{i<j}^d f_{i,j}(X_i, X_j) + \dots + f_{1,2,\dots,d}(X_1, X_2, \dots, X_d) \quad (6)$$

The commonly used variance-based methods are Sobol method and the Fourier Amplitude Sensitivity Test (FAST) method. Both methods are model-free, which promises an extensive feasibility. Besides, both methods can output two indexes, i.e., the first order effect which evaluates the impact of the individual parameter with the parameter interaction being eliminated, and the total order effect which shows the contribution of the inputs with the interactions considered. Theoretically, both methods can quantify the high-order interactions among the input parameters, but some researchers found that the FAST is not suitable for high dimensions (Saltelli et al., 1999; Xu and Gertner, 2011). Besides, the FAST was found less efficient when the number of the input variables is large

(larger than ten) (Iooss and Lemaître, 2015). For this reason, the extended-FAST and the FAST with a Random Balance Design (RBD-FAST) methods were proposed to overcome the drawbacks.

Variance-based methods are considered as the most informative method and have established a good reputation as the reference method among. However, they ask for an amount of model evaluations to ensure the convergence and the robustness of the result. Hence, their time and computational costs are much more expensive than other methods aforementioned.

Meta-model-based sensitivity analysis

The Metamodel, which is also known as the surrogate model, or the response surface, refers to a set of techniques which can be used to predict the outcome using the inputs. The Metamodel works by fitting the data generated by the simulation software at the selected points in the sample space. In BPS, it can be used to replace the physical model based on a certain number of samples instead of executing the original model over and over again, which is time consuming and computationally intensive (Iooss and Lemaître, 2015).

There have been a variety of metamodeling techniques available in this era, e.g., polynomials, neural networks (NN), support vector regression (SVR), regression trees, etc. (Iooss and Lemaître, 2015; Iooss et al., 2006; Kleijnen and Sargent, 2000; Song et al., 2015). The top priority in the selection of a metamodel is that such approach can accurately capture the behaviour of the original models using the influential parameters. In this study, the multi-variate linear regression is used to create the metamodel.

Methodology

Model description

The large hotel prototype building model developed by the U.S. DOE was used for the case study. The building is located in Houston, TX, USA. The outlook of the model is as shown in Figure 1. It is a 6-story detached structure with a basement. The total area is 11,345 square meters. The thermos-physics characteristics of the model as well as its architectural plan and the HVAC system design are as presented in Table 1.

Figure 1: Outlook of the large hotel prototype building model.

The domestic hot water (DHW) associated gas energy consumption was specified as the output of interest in this case study. The topology of the DHW system is illustrated in Figure 2. As can be seen from the illustration, the DHW system of this building consists of three sub-systems, i.e.,

the main DHW system (blue in Figure 2), the kitchen system (orange in Figure 2), and the laundry system (green in Figure 2), which supply the hot water to the guestrooms, kitchen, and laundry service respectively. Each system has a dedicated water heater. It should be noted that the booster heater (i.e., the water heater of the kitchen system) is not working independently. Its inlet pipe is connected to the distribution network of the main heater. The main heater and the laundry heater are fuelled by the natural gas, while the booster heater is fuelled by the electricity.

Figure 2: The topology of the DHW system.

Table 1: The architectural and HVAC system design of the building model.

Type	Large Hotel
Number of Floors	6 above-ground floors with a basement
Total Floor Area	11,346 m ²
Wall U-Factor	2.882 W/(m ² K)
Roof U-Factor	0.219 W/(m ² K)
Window U-Factor	3.511 W/(m ² K)
Glass SHGC	0.227
Heating source	Gas-fired boiler × 1
Cooling source	Air-cooled chiller × 1
Water Heating Equipment	Main and central gas water heater with storage tank Electric dishwasher booster water heater Gas water heater for laundry with storage

Monte Carlo simulation

33 input parameters associated with the gas consumption of the DHW system were selected for the Monte Carlo analysis, as listed in Table 2. The detail about the sampling is not included here due to the page limit. Readers are recommended to check (Pang and O'Neill, 2018) for more information about the sampling process. It should be noted that the room peak use flowrate was counted nine times since it actually corresponds to the flowrates of nine different rooms.

3,000 EnergyPlus based Monte Carlo simulations were performed for the demonstrations of the various SA methods. 3,000 is considered enough for a model with 33 input parameters (Iooss and Lemaître, 2015). Python Scripts were used to create the building models in batch and extract the inputs and outputs.

Table 2: The uncertainty of the parameters of the large hotel prototype building model.

System	Parameter	Nominal Values	Unit	N O.
Main System	Inlet temp.	22	°C	1 - 8
	Setpoint temp.	60	°C	
	Tank volume	2.27125	m ³	
	Max. temp. limit	82.22	°C	
	Max. capacity	175.8	kW	
	Thermal eff.	0.8028	/	
	Use side effectiveness	1	/	
	Source side effectiveness	1	/	
Booster System	Inlet temp.	60	°C	9 - 15
	Setpoint temp.	82.22	°C	
	Tank volume	0.02271	m ³	
	Max. temp. limit	82.22	°C	
	Max. capacity	8	kW	
	Thermal eff.	1	/	
Laundry System	Peak flow rate	8.4E-5	m ³ /s	16 - 21
	Setpoint temp.	82.22	°C	
	Tank volume	1.1356	m ³	
	Max. temp. limit	82.22	°C	
	Max. capacity	87.9	kW	
	Thermal eff.	0.804	/	
Water Use Equipment	Peak flow rate	1.93E-3	m ³ /s	22 - 33
	Target temp.	55	°C	
	Hot water temp.	60	°C	
	Room peak flowrate	1.31E-6	m ³ /s	
	Kitchen peak use flowrate	1.4E-4	m ³ /s	

Sensitivity analysis

Ten SA methods which were widely used in the BPS community were selected for comparison in this case study. They are Morris method, PEAR, SPEA, SRC, PCC, SRRC, PRCC, Sobol, FAST, and RBD-FAST respectively. The R sensitivity package (Gilles Pujol et al., 2017), and Python SALib package (Herman and Usher, 2017) were used for sampling and the computation of the SA indexes.

A multivariate linear regression metamodel was trained using the input and output data from the MC simulation to save the time and computational cost by avoiding running the original models for too many times. The Python scikit-learn package was used to train the metamodel. The R-squared value of the metamodel was 0.85, which means that the accuracy of the training was satisfactory. There is no universal standard for an acceptable R-squared value, but in general, R² of 0.75 is considered as a reasonable value for an acceptable model accuracy in the building simulation (Efficiency Valuation Organisation, 2007).

The sampling strategies for the various SA methods vary by their requirements: the factorial sampling was used for the Morris method; the Latin Hypercube Sampling (LHS) was used for the correlation and regression methods; the quasi-random sampling methods (e.g., the Sobol

sequence and FAST sampling) was used for the variance-based methods. The R randtoolbox package (Christophe and Petr, 2018) and Python SALib package (Herman and Usher, 2017) were used for the sampling.

To investigate the convergence rate and the accuracy of the various SA methods, a sampling plan was created to generate multi-groups of inputs with an increasing number of model evaluations. It should be noted that the sampling strategies for the Morris method, Sobol method and FAST are fixed (e.g., the Sobol sequence for the Sobol method, and the Factorial design for the Morris method), thus the number of the model evaluations cannot be freely defined for them. In this study, the number of samplings for these three methods should follow $k*(n+1)$, $k*(2n+2)$, and $k*n$ respectively, where n is the number of the model inputs and k is a factor defined by the users. The sampling plan is listed in Table 3.

It should be noted that the maximum number for Morris is 1000 because compared with other methods, the Morris method is more efficient, and this number is big enough for it to reach the convergence (Iooss and Lemaître, 2015).

Table 3: The sampling plan.

SA method	Sampling numbers
Morris	340, 680, 1700, 3400, 6800 ($k = 10, 20, 50, 100, 200$)
Correlation & regression	50, 100, 200, 500, 1000, 2000, 5000, 10000
Sobol	6800, 13600, 34000, 68000, 136000, 340000, 680000 ($k = 100, 200, 500, 1000, 2000, 5000, 10000$)
FAST	3300, 6600, 16500, 33000, 66000, 165000, 330000 ($k = 100, 200, 500, 1000, 2000, 5000, 10000$)
FAST-RBD	3300, 6600, 16500, 33000, 66000, 165000, 330000 ($k = 100, 200, 500, 1000, 2000, 5000, 10000$)

Comparison metrics

Multiple metrics were used to compare the performances of various SA methods, including the time for the computation, the convergence rate, the accuracy, and the complicity.

The Kendall Correlation Coefficient (KCC) was computed to quantify the differences between the two rankings with different SA methods or different sampling numbers. The definition of KCC was presented in Eq. (7), where n , N_{con} , and N_{dis} represent the numbers of the input parameters, concordant pairs, and discordant pairs respectively. The KCC considers the two rankings as a set of ordered pairs (S_i^1, S_i^2) , where S^x represents the different methods and i represents the number of the input parameter. For two different input parameters, if $S_i^1 < S_j^1$, and $S_i^2 < S_j^2$, this pair is considered as concordant, if not, then it was considered as discordant. Therefore, a larger KCC index represents that the rankings of two sequences are more concordant.

$$\tau = \frac{N_{con} - N_{dis}}{n*(n-1)/2} \quad (7)$$

Results

Computational cost

As listed in Table 3, a number of model evaluations were performed for each SA method to reach the convergence. The computational time of each method with the maximum number of the model evaluation was presented in Table 4.

It is obvious that the PEAR and SRC are the most computational efficient methods in this case study, with each taking less than 0.05 seconds. The result also confirmed that the Sobol method is the most computationally intensive SA method. The time use for the Sobol method is nearly one minute and a half. It's worth noting that even if FAST has been considered as a computationally intensive method in the past, its computational time is actually much less than its counterpart, the Sobol method. In this case, the time cost of the Sobol method is approximately 45 times longer than the FAST method.

It should be noted that the SA index calculation in this study was based on the metamodel instead of the whole-building energy models. Considering that the minimum number of the model evaluations for the convergence vary significantly for different methods, this difference in time and computational costs will be amplified in practice. For example, in this case study, a single EnergyPlus simulation took approximately one minutes with a Dell Precision T7610 workstation (Xeon E5-2630 12-logical-core, and 32 GB random-access memory).

Table 4: Computational time for the various SA methods.

SA	Morris	PEAR	SPEA	SRC	SRRCC
Time	0.13 s	0.02 s	0.11 s	0.03 s	0.12 s
SA	PCC	PRCC	Sobol	FAST	RBD
Time	2.52 s	2.48 s	82 s	1.8 s	1.9 s

Accuracy

The accuracy of the input ranking is one of the most important metrics for the evaluation of an SA method. In this study, a large number of model evaluations were executed to make sure that all the results were convergent for the comparison. For instance, as usual, the k value for the Sobol and FAST methods is usually 1000 in most researches. However, in this case study, the k value was 10,000 to avoid the mis-convergence.

The ranking result of the top ten parameters were presented in Figure 3. The correspondence of the parameters and their numbers is listed in Table 2. For instance, the variable 10 in Figure 3 is the setpoint temperature of the booster water heater in Table 2.

It's obvious from Figure 3 that the investigated SA methods had a good agreement when it comes to the identification of the most significant parameters (i.e., the top 10 parameters), except for the PRCC. All of the methods identified the thermal efficiency and maximum capacity of the laundry water heater (variable 20 and variable 19) as the most important parameters. Besides, the temperature setpoint and the thermal efficiency of the main water heater (variable 2 and variable 6), as well as the peak use

flowrate of the guest room (variable 24) were also identified as influential parameters. These conclusions made sense because the laundry heater and main water heater are the major end users of the DHW. However, compared with other methods, PRCC misidentified five input parameters, which made it not preferred in this case.

Figure 3: The comparison of the top ten parameters by all the ten SA methods.

Figure 4: The comparison of the rankings by all the ten SA methods.

Figure 5: The carpet plot for all the input parameters.

Figure 6: The Kendall correlation coefficients of all the SA methods with reference to the Sobol method.

The rankings of all the input parameters were presented in Figure 4. Compared with the rankings only for the most significant inputs, the rankings for the whole set of inputs varied a lot, especially for the parameters which is not significant, e.g., PEAR and SPEA identified variable 24, the peak flowrate of a guestroom on the first floor, as unimportant, while other methods considered it with moderate influence. Besides, Sobol ranked variable 4, the maximum temperature limit of the main water heater as the least significant, while PEAR and SPEA rated it as the top 15 input parameters.

This mis-ranking of the non-significant input parameters is probably due to the fact that the sensitivity indexes of the input parameters with smaller influences are too small, so that the absolute differences are negligible. The carpet plot of the sensitivity indexes obtained by various SA methods were presented in Figure 5. The colour code on the right suggests the influence of the input variable on the output with larger indices being denoted in blue and the smaller indices being denoted in pink. This validates that for most of the input parameters, their impacts are very similar since their influences are neglectable.

To evaluate the accuracy of each method quantitatively, the KCC indexes were calculated to compare all the methods with reference to the Sobol method, since Sobol is one of the most accurate approach (Nguyen and Reiter, 2015). The result was plotted in Figure 6. Figure 6 shows that Morris and FAST had the closest accuracy to the Sobol method, followed by RBD-FAST and the four regression-based methods. The two correlation analysis methods had the worst performances.

Convergence rate

How fast can an SA method reach the convergence is also one of the most concerned topics for its evaluation. In this case study, a sampling strategy with a set of gradient sample numbers was created to investigate the SA method's capability of achieving a fast convergence.

The rankings of the input parameters obtained by the same SA method but with different sampling times were presented in Figure 7. The results of the top three sampling times were bolded and coloured in red, blue, and green respectively to facilitate the view. Besides, a black bold auxiliary line was plotted in each subplot to mark the ranking of the top ten parameters.

It was apparent that the Morris method, SRC, and FAST archive the best performance in terms of the convergence rate. Their rankings are nearly stable from the very beginning of the computation, which suggests that they may need less model evaluations for the sensitivity analysis in practice. Except for these three methods, the Sobol method performed relatively well too, with its ranking fluctuating in an acceptable range. However, it should be noted that the actual times of model evaluations for Sobol and FAST are not the k values used for the caption. The true values are of the model evaluations were listed in Table 3.

The KCC index were computed to evaluate the convergence rate of each method quantitatively. For each method, the case with the maximum sampling time was set as the reference. The result was as presented in Figure 8. The result of Morris was not included since it was exactly convergent from the very beginning, which means that it has the best performance in terms of the convergence rate. This figure further verifies our previous conclusion that the SRC and FAST are the fastest to output the stable results (other than Morris), followed by the Sobol method. The other three regression-based methods and RBD-FAST perform a bit worse but are still acceptable. The correlation-based methods have the worst convergence rates.

Figure 7. The rankings for all the input parameters with different methods and sampling strategy.

Conclusions and future work

Conclusions

This paper used an EnergyPlus based MC simulation to compare ten SA methods based on three metrics: computational/time cost, accuracy, and convergence rate. The large hotel prototype building model developed by US DOE was used as the case study. 33 input parameters associated with the DHW system were selected and sampled using the Latin Hypercube sampling method. A total of 3,000 simulation models with different uncertainties were created and analysed.

The DHW gas energy consumption was specified as the output of interest. A multivariate linear regression metamodel of the building DHW system was trained to facilitate the further analysis. The R-squared value of the regression was 0.85. The evaluations of these methods based on the case study were summarized in Table 5.

- The PEAR and SPEA are not recommended due to their bad accuracies and slow convergence rates. Though they both can rank the significant input parameters correctly, they perform poorly in ranking the non-significant parameters.
- Generally, the SRC and Morris method were the most recommended methods for the linear model, for their outstanding performances of less computational intensiveness, accuracy, and fast convergence. According to the result of the comparison, Morris and SRC have the lowest demand for the computation

except for the correlation analysis-based methods (i.e., PEAR and SPEA). Both approaches have an outstanding convergence rate, which means that less model evaluations are required to produce stable results. Besides, compared with SRC, the accuracy of Morris is slightly higher, but the sampling strategy of Morris (i.e., factorial design) will need more efforts.

- The rank transformation based methods (i.e., SPEA, SRRC, and PRCC) do not have obvious benefits over their prototypes in this cases study.
- FAST has a similar performance of the Morris method and SRC in terms of computation and accuracy. Besides, since it is a variance analysis based method, its result can be directly used for uncertainty analysis, unlike the screening and regression based methods. However, this method requires more model evaluations in practice, which would greatly increase the time for simulation if the study is not based on a metamodel.
- Sobol method has long been considered as the most accurate SA method, and is recommended by many researchers from various fields to be used as a benchmark. Moreover, its ability to decompose uncertainty in the output makes it an outstanding tool for uncertainty analysis as well. But compared with SRC, Morris, and FAST, a much larger amount of model evaluations is required by the Sobol to reach convergence. This would greatly increase the time consumed by the simulation.

Figure 8. The KCC indexes for the same SA method with different number of samples (model evaluations).

Table 5: Comparison of the SA methods.

SA	Time	Accuracy	Convergence
Morris	Very Short	Good	Fast
PEAR	Very Short	Bad	Slow
SPEA	Very Short	Bad	Slow
SRC	Very Short	Moderate	Fast
SRRC	Very Short	Moderate	Moderate
PCC	Short	Moderate	Moderate
PRCC	Short	Moderate	Moderate
Sobol	Long	Good	Moderate
FAST	Short	Good	Fast
RBD	Long	Moderate	Moderate

Future work

- The DHW model used in this case study is a highly linear system. Thus, the conclusions in this paper may not be directly extended to a nonlinear system. It is necessary to include more case studies in the future to investigate the performances of the various SA methods for nonlinear models.
- More SA methods should be investigated in the future to include more options for the practitioners, e.g., the regionalized SA method, the local one-at-a-time (OAT) SA method, some optimized regression methods like Lasso regression, etc.

Abbreviations

Abbreviations	
AHU	Air handling unit
ANOVA	analysis of variance
BPS	Building performance simulation
DHW	Domestic hot water
DOAS	Dedicated outdoor air system
FAST	Fourier amplitude sensitivity test
FCU	Fan-coil unit
KCC	Kendall Correlation Coefficient
MC	Monte Carlo
NN	Neural network
OAT	One-at-a-time
PCC	Partial correlation coefficient
PEAR	Pearson correlation coefficient
PRCC	Partial rank correlation coefficient

RBD	Random Balance Design
SA	Sensitivity analysis
SPEA	Spearman correlation coefficient
SRC	Standard regression coefficient
SRRC	Standard rank regression coefficient
SVR	Support vector machine
UA	Uncertainty analysis
VAV	Variant air volume

References

- Brembilla, E., Hopfe, C. J., & Mardaljevic, J. (2018). Influence of input reflectance values on climate-based daylight metrics using sensitivity analysis. *Journal of Building Performance Simulation*, 11(3), 333-349.
- Campolongo, F., Cariboni, J., & Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*, 22(10), 1509-1518.
- Christophe, D., & Petr, S. (2018). randtoolbox: Generating and Testing Random Numbers. *R package version 1.17.1*. Retrieved from <https://CRAN.R-project.org/package=randtoolbox>
- de Wilde, P., & Tian, W. (2009). *Identification of key factors for uncertainty in the prediction of the thermal performance of an office building under climate change*. Paper presented at the Building Simulation.
- De Wit, S., & Augenbroe, G. (2002). Analysis of uncertainty in building design evaluations and its implications. *Energy and Buildings*, 34(9), 951-958.
- Efficiency Valuation Organisation. (2007). International performance measurement and verification protocol. *Efficiency Valuation Organisation, Oakridge, Tennessee*.
- Gaetani, I., Hoes, P.-J., & Hensen, J. L. (2017). On the sensitivity to different aspects of occupant behaviour for selecting the appropriate modelling complexity in building performance predictions. *Journal of Building Performance Simulation*, 10(5-6), 601-611.
- Gilles Pujol, Bertrand Iooss, Alexandre Janon, Boumhaout, w. c. f. K., Veiga, S. D., Delage, T., Fruth, J., Gilquin, L., Guillaume, J., Gratiet, L. L., Lemaitre, P., Nelson, B. L., Monari, F., Oomen, R., Ramos, B., Roustant, O., Song, E., Staum, J., Touati, T., & Weber, F. (2017). sensitivity: Global Sensitivity Analysis of Model Outputs. *R package version 1.15.0*. Retrieved from <https://CRAN.R-project.org/package=sensitivity>
- Gu, J., Xu, P., Pang, Z., Chen, Y., Ji, Y., & Chen, Z. (2018). Extracting Typical Occupancy Data of Different Buildings from Mobile Positioning Data. *Energy and Buildings*.
- Heiselberg, P., Brohus, H., Hesselholt, A., Rasmussen, H., Seinre, E., & Thomas, S. (2009). Application of sensitivity analysis in design of sustainable buildings. *Renewable Energy*, 34(9), 2030-2036.
- Hensen, J. L., & Lamberts, R. (2012). *Building performance simulation for design and operation*: Routledge.

- Herman, J., & Usher, W. (2017). SALib: An open-source Python library for sensitivity analysis. *Journal of Open Source Software*, 2(9). doi:doi:10.21105/joss.00097
- Hester, J., Gregory, J., & Kirchain, R. (2017). Sequential early-design guidance for residential single-family buildings using a probabilistic metamodel of energy consumption. *Energy and Buildings*, 134, 202-211.
- Hopfe, Hensen, J., & Plokker, W. (2006). *Introducing uncertainty and sensitivity analysis in non-modifiable building performance software*. Paper presented at the Proceedings of the 1st Int. IBPSA Germany/Austria Conf. BauSIM.
- Iooss, B., & Lemaître, P. (2015). A review on global sensitivity analysis methods. In *Uncertainty management in simulation-optimization of complex systems* (pp. 101-122): Springer.
- Iooss, B., Van Dorpe, F., & Devictor, N. (2006). Response surfaces and sensitivity analyses for an environmental model of dose calculations. *Reliability Engineering & System Safety*, 91(10-11), 1241-1251.
- Kleijnen, J. P., & Sargent, R. G. (2000). A methodology for fitting and validating metamodels in simulation1. *European Journal of Operational Research*, 120(1), 14-29.
- Lam, J. C., & Hui, S. C. (1996). Sensitivity analysis of energy performance of office buildings. *Building and Environment*, 31(1), 27-39.
- Menberg, K., Heo, Y., & Choudhary, R. (2016). Sensitivity analysis methods for building energy models: Comparing computational costs and extractable information. *Energy and Buildings*, 133, 433-445.
- Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2), 161-174.
- Nguyen, A.-T., & Reiter, S. (2015). A performance comparison of sensitivity analysis methods for building energy models. *Building Simulation*, 8(6), 651-664.
- O'Neill, Z., & Niu, F. (2017). Uncertainty and sensitivity analysis of spatio-temporal occupant behaviors on residential building energy usage utilizing Karhunen-Loève expansion. *Building and Environment*, 115, 157-172.
- O'Neill, Z., & Eisenhower, B. (2013). Leveraging the analysis of parametric uncertainty for building energy model calibration. *Building simulation*, 6(4), 365-377.
- Østergård, T., Jensen, R. L., & Maagaard, S. E. (2017). Early Building Design: Informed decision-making by exploring multidimensional design space using sensitivity analysis. *Energy and Buildings*, 142, 8-22.
- Pang, Z., O'Neill, Z., Li, Y., & Niu, F. (2019). The Role of Sensitivity Analysis in the Building Performance Analysis: A Critical Review. *Energy and Buildings*, 109659.
- Pang, Z., & O'Neill, Z. J. A. e. (2018). Uncertainty quantification and sensitivity analysis of the domestic hot water usage in hotels. 232, 424-442.
- Pang, Z., Xu, P., Lu, X., Qiu, S., Chen, L., & Hou, J. (2017). Evaluation of the performance of a new solar ventilated window: Modeling and experimental verification. *Journal of Renewable and Sustainable Energy*, 9(6), 065101.
- Pang, Z., Xu, P., O'Neill, Z., Gu, J., Qiu, S., Lu, X., & Li, X. (2018). Application of mobile positioning occupancy data for building energy simulation: An engineering case study. *Building and Environment*, 141, 1-15.
- Qiu, S., Li, Z., Pang, Z., Zhang, W., & Li, Z. (2018). A quick auto-calibration approach based on normative energy models. *Energy and Buildings*, 172, 35-46.
- Rodríguez, G. C., Andrés, A. C., Muñoz, F. D., López, J. M. C., & Zhang, Y. (2013). Uncertainties and sensitivity analysis in building energy simulation using macroparameters. *Energy and Buildings*, 67, 79-87.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Global sensitivity analysis: the primer*: John Wiley & Sons.
- Saltelli, A., Tarantola, S., & Chan, K.-S. (1999). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1), 39-56.
- Sanchez, D. G., Lacarrière, B., Musy, M., & Bourges, B. (2014). Application of sensitivity analysis in building energy simulations: Combining first-and second-order elementary effects methods. *Energy and Buildings*, 68, 741-750.
- Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M., & Xu, C. (2015). Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications. *Journal of hydrology*, 523, 739-757.
- Sun, Y., Gu, L., Wu, C. J., & Augenbroe, G. (2014). Exploring HVAC system sizing under uncertainty. *Energy and Buildings*, 81, 243-252.
- Tian, W. (2013). A review of sensitivity analysis methods in building energy analysis. *Renewable and Sustainable Energy Reviews*, 20, 411-419.
- Tian, W., Heo, Y., de Wilde, P., Li, Z., Yan, D., Park, C. S., Feng, X., & Augenbroe, G. (2018). A review of uncertainty analysis in building energy assessment. *Renewable and Sustainable Energy Reviews*, 93, 285-301.
- Xu, C., & Gertner, G. (2011). Understanding and comparisons of different sampling approaches for the Fourier Amplitudes Sensitivity Test (FAST). *Computational statistics & data analysis*, 55(1), 184-198.
- Yang, S., Tian, W., Cubi, E., Meng, Q., Liu, Y., & Wei, L. (2016). Comparison of sensitivity analysis methods in building energy assessment. *Procedia Engineering*, 146, 174-181.