Uncertainty Analysis of Building Energy Performance Based on Replicated Latin Hypercube Sampling

Wei Tian¹,², Miaorou Jin¹, Pieter de Wilde³, Xing Fu⁴, Guangchen Li¹
¹College of Mechanical Engineering, Tianjin University of Science and Technology, China
²Tianjin International Joint Research and Development Center of Low-Carbon Green Process Equipment, Tianjin, China
³Chair of Building Performance Analysis, Environmental Building Group, University of Plymouth, Plymouth, UK
⁴Tianjin Architecture Design Institute, Tianjin, China

Abstract
Uncertainty analysis has become increasingly important to provide reliable energy efficiency assessment of buildings whilst considering inherent uncertain factors. The one-batch sample method is usually used with the Latin Hypercube sampling (LHS) to evaluate building performance, however it is difficult to assess the stability of uncertainty results using this approach. Hence, the uncertainty of energy performance in buildings may be exaggerated or understated. This paper explores the application of replicated LHS in order to properly evaluate the stability of uncertainty results in building energy assessment. The results indicate that the replicated LHS can quantify the stability of energy analysis by obtaining the confidence intervals of energy use. The similarity of replicated results can be analysed using formal statistical distance measures, including Kolmogorov–Smirnov distance and Dowd test statistics. Moreover, the combined results from several repetitions can improve the convergence of uncertainty results in building energy analysis.

Key Innovations
- Use of the replicated Latin Hypercube sampling method to assess the stability of sampling result in uncertainty analysis of building energy predictions
- Application of statistical distance to quantify the similarity of building energy use from the replicated samples
- Improvement of the convergence of energy use in uncertainty analysis of building energy simulation by combining the repetition results
- Construction of the confidence intervals of building energy from the replicated Latin Hypercube sampling simulation

Practical Implications
The replicated Latin Hypercube sampling can be used to properly assess the stability of uncertainty analysis by quantifying the similarity of individual samples and computing the confidence intervals of energy use in building energy analysis.

Introduction
Uncertainty analysis has been widely used in building energy simulation to provide robust predictions of energy performance of buildings while taking into account inherent uncertain variables (Tian et al., 2018; de Wilde, 2018). The choice of both sampling method and sample size is vital for the sampling-based uncertainty analysis in building energy assessment (Hou et al., 2019; Helton et al., 2005).

Latin Hypercube Sampling (LHS) is commonly used to obtain reliable results in order to reduce computational cost by using the stratified sampling techniques (Kang, 2021; Tian et al., 2018). There are many examples of its application (Liu et al., 2021; Tian et al., 2018). It is used in the seminal PhD dissertations of uncertainty analysis in building performance analysis by de Wit (2001) and Macdonald (2002). More recently, Li and Wang (2020) apply the LHS to evaluate the energy performance of zero/low energy buildings in Hong Kong. Liu et al. (2021) implement the LHS to assess the energy and carbon performance of campus buildings in Tianjin, China. Westermann and Evins (2021) apply the LHS to create uncertainty-aware building energy surrogate models using Bayesian deep learning methods. Zhu et al. (2020) use the LHS to infer the uncertain parameters in building energy models by applying the approximate Bayesian computation. Risch et al. (2021) implement the LHS to reduce parameter uncertainty in urban building energy models. Most of these previous studies apply the LHS only one single time in order to obtain the variation of energy performance of buildings. Yet it is difficult to assess the sufficiency of sample size using only one LHS sample in building energy analysis. As a result, the variation of building performance may be exaggerated or understated due to the insufficient sample size.

The most straightforward method of obtaining the variance of LHS results is the replicated LHS as suggested by Owen (2019) to run simply independent replicates of LHS to analyse the convergence of uncertainty results. Hansen et al. (2012) apply the replicated LHS to assess the uncertainty results of geologic disposal of radioactive waste. More studies have been focused on the extension of LHS size, comparison of LHS and random sampling, and the correlation control of LHS (Schmidt et al., 2019; Pleming and Manteufel, 2005; Janssen, 2013). Note that the augmented LHS methods can only double (or add even multiples) initial sample size due to the stratification feature of LHS. Hence, the replicated LHS can provide more flexibility to add several independent samples.
(whether even or odd multiple of initial sample size) to properly assess the convergence of energy variations of buildings.

This paper compares one time LHS with replicated LHS to obtain more stable energy predictions for buildings. The replicated LHS method generates the energy data of a sample size with different sequences of random numbers. The energy results from these independent LHS datasets are analysed to determine whether there are consistent uncertainty results of building energy performance. A four-storey office building is used as a case study to demonstrate the application of the replicated LHS technique. The energy simulation of buildings is carried out with the EnergyPlus simulation program (DOE, 2020). The statistical analysis and result visualization are conducted using the R computational environment (R Core Team, 2020).

**Building energy models**

![Figure 1: An office building used in the uncertainty analysis.](image)

A four-storey office campus building is used in this research as shown in Figure 1. The total floor area is 2590 m² with the window-wall-ratio of 0.45. The building thermal parameters are in line with the China building code released in 2005 (MOC, 2005). The hourly schedules for occupants, lights, and equipment, are derived from previous research (Tian et al., 2020; Fu et al., 2019) and the site survey. District heating and window air-conditioners are used to provide heating and cooling in order to maintain indoor thermal comfort. The building is located in Tianjin, China, and the typical weather data is the Chinese Standard Weather Data (CSWD) of Tianjin developed by the China Meteorological Administration (2005). Annual heating energy intensity normalized by the total floor area (kWh/m²) is used as the energy performance measure in this study.

**Table 1: Input ranges of building energy models.**

<table>
<thead>
<tr>
<th>Input</th>
<th>Unit</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lighting power density</td>
<td>W/m²</td>
<td>5-13</td>
</tr>
<tr>
<td>Occupancy density</td>
<td>m²/person</td>
<td>8-12</td>
</tr>
<tr>
<td>Equipment density</td>
<td>W/m²</td>
<td>11-19</td>
</tr>
<tr>
<td>Wall U-value</td>
<td>W/m² K</td>
<td>0.3-0.5</td>
</tr>
<tr>
<td>Window U-value</td>
<td>W/m² K</td>
<td>1.5-2.5</td>
</tr>
<tr>
<td>Infiltration rate</td>
<td>ACH</td>
<td>0.3-0.6</td>
</tr>
</tbody>
</table>

The uncertain parameters are summarized in Table 1. Uniform distributions are assumed in this study since the design uncertainty used in this study can be considered as equal distributions (Tian et al., 2018). Normal distributions is more applied in the case of inherent random uncertainty in building energy analysis (Tian et al., 2018). These values in Table 1 are derived from previous studies (Tian et al., 2020; Liu et al., 2021) and the building codes (MOC, 2005; MOC, 2019; MOC, 2015).

The energy simulation is carried out with the EnergyPlus Program (DOE, 2020). The IDFs (Input Data File) of EnergyPlus models are created using the R computational environment (R Core Team, 2020) since a large number of energy models are required for uncertainty analysis.

**Methods**

The replicated LHS method is used to obtain the energy results for uncertainty analysis of building energy consumption. Three types of statistical analysis are used to quantify the results from the replicated LHS: statistical distance, overlapping index, and confidence interval. The statistical distance analysis is used to compare the density distributions from several LHS samples in the same sample size. The overlapping analysis can be used to compare the density from the individual distributions and the reference distributions (i.e. the combination of all samples). Lastly, the confidence interval analysis is used to obtain the intervals of energy use to assess the variability among the replicated samples.

**Replicated Latin Hyper-cube Sampling approach**

The replicated LHS approach is based on repeating the LHS process several times in order to properly assess the stability of uncertainty results (Hansen et al., 2012; Alban et al., 2017).

The replicated LHS method includes the six steps. The first step is to choose the distributions of building input factors. The second step is to obtain the combinations of input factors using the LHS. The third step is to create and run the EnergyPlus models. The fourth step is to collect and visualize the energy results using the probability density functions and the cumulative density functions. The fifth step is to repeat the first four steps to obtain the results using the different random seeds (i.e. different repetitions). The final step is to statistically analyse the stability of uncertainty results as will be described in the following three subsections.

A total of 18 different sample sizes has been used in this study to thoroughly understand the effects of variation in sample size: 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, and 300. The sample intervals increase from 5 to 25 with an increase of sample size since the convergence trend from a large sample size shows an insensitivity to the variations of sample size. The maximum sample size 300 is chosen based on the preliminary simulation to present the small percentile variations of building energy in this case study. The repetition number (i.e. random seed) is taken as five in every sample size in this study. Hence, the total number of energy models employed in this study is 10,500. The R LHS package is used to construct the LHS (Carnell, 2020).
Statistical distance between probability functions

Statistical distances (or divergences) can quantify the distance among probability distributions to determine the similarity or difference of these distributions. Hence, the distance metrics can be used to assess the stability of results from the replicated LHS. If the statistical distance from several distributions is small, then the results from replicated LHS are sufficient. There are a number of statistical distance measures, such as Kolmogorov–Smirnov distance, Bhattacharyya distance, Kullback–Leibler divergence, Jensen–Shannon divergence and Lepage test. Two distances are used in this research: KS and DTS (Dowd Test Statistics) (Dowd, 2020) to provide robust statistical distance analysis.

The Kolmogorov-Smirnov test has been widely used to compare the empirical distributions of two samples. The KS test can be written as,

\[ KS = \max_x |F(x) - E(x)| \]  

where \( F(x) \) and \( E(x) \) are the empirical cumulative distribution functions from two samples, respectively. Hence, the KS statistic is the maximum distance between two empirical distributions. A larger KS value indicates a large difference of two distributions. The KS test is only focused on the maximum difference of two cumulative distributions although this KS test is one of non-parametric approaches, not based on a specific distribution assumption. The ks.test function in the R environment is used to compute the KS distance (R Core Team, 2020).

The DTS is a new distance metric proposed by Dowd to detect the differences of two distributions by using the reweighted integral of the distance between two density distributions (Dowd, 2020). More weights would be specified for the extreme observations. The large DTS means there are large differences between two cumulative density functions. The DTS can be written as,

\[ DTS = \int \frac{|F(x) - E(x)|}{D(x)(1-D(x))} \, dx \]  

where \( D(x) \) is the empirical cumulative distribution of a combined sample (E and F), the other symbols are the same in Eq. (1). The DTS has the same advantage as one of non-parametric methods, which do not rely on the underlying distributions. The DTS has more power in comparison with the KS test (Dowd, 2020). The R two samples package is used to compute the DTS distances (Dowd, 2020).

The purpose of using the two statistical distance methods is to obtain robust results on the probability distributions for energy use of buildings. The KS distance is more widely used to only focus on the maximum difference, while the DTS is a new statistics to integrate the Wasserstein and Anderson-Darling notions together to provide more thorough analysis. More detailed discussion on this topic is available in (Markatou et al., 2021).

Overlapping computation

Overlapping is usually used to describe the similarity of two empirical distributions. A distribution-free overlapping index is selected to compute the similarity of distributions for the individual sample and the reference sample (Pastore and Calcagni, 2019). The overlapping index \( OI \) can be written as,

\[ OI = \int \min(E(x), F(x)) \, dx \]  

where \( E(x) \) and \( F(x) \) are the density functions from two samples, respectively. The overlapping index is normalized between 0 and 1. In the case of overlapping value 0, the two distributions are completely separated without any overlap. In contrast, the two distributions are exactly the same in the case of the overlapping index being 1. The overlapping index is computed using the R overlapping package (Pastore, 2018).

Confidence intervals from replicated samplings

The confidence intervals can be obtained from the replicated LHS. The sample size is assumed as \( n \) and the repetition number is assumed as \( m \). The estimate of mean energy use \( D \) can be computed from the following,

\[ D = \frac{1}{m} \sum_{r=1}^{m} D_r \]  

The estimate of standard error of the estimate \( D \) can be written as,

\[ SE(D) = \sqrt{\frac{m}{m(m-1)} \sum_{r=1}^{m} (D - D_r)^2} \]  

The 1-\( \alpha \) confidence interval is computed by,

\[ D \pm t_{(1-\alpha/2)} SE(D) \]  

where \( t_{(1-\alpha/2)} \) is the 1-\( \alpha/2 \) quantile of the \( t \)-distribution with \( m \) degrees of freedom for the energy use. More detailed information on the choice of the \( t \)-distribution in the replicated LHS can be found in Hansen et al. (2012).

Results and discussion

Comparison of replicated distributions

Figure 2 illustrates the change of density plots for annual heating energy using different repetitions (i.e. sampling random seeds) for three sample sizes. In the sample size of 10, the distribution differences from the violin plots are visually significant. In the first and third repetitions, the distributions are much flatter than the other three repetitions. As the sample size gradually increases, the differences among different repetitions becomes less significant. In the case of 300 sample size, the distributions of heating energy from five repetitions are visually almost the same.

Table 2 lists the five percentiles (0.10, 0.25, 0.50, 0.75, 0.90) of heating energy in five sample sizes with five repetitions. The coefficient of variation for the 50th percentile would be decreased significantly around 65% from sample size 10 to sample size 300. For the 10th and 90th percentile of heating energy, the coefficient of variation can be up to 86% from the sample size 10 to sample size 300. Hence, the differences of distributions
for the replicated LHS are very significant in the lower sample sizes.

Figure 2: Violin plots of heating energy with five repetitions (different random seeds) and three different sample sizes

Figure 3 shows the quantitative change of probability density for heating energy use using the Kolmogorov-Smirnov (KS) distance and Dowd test statistics (DTS) due to LHS repetitions. There are five repetitions in this research and the two samplings are required to compute these statistical distances. Hence, all the combinations are generated by taking two samplings at a time from all the five elements (i.e. repetitions). As a result, there are ten combinations for every sample size to compute the KS and DTS distances.

Figure 3(a) shows the change of the KS distances for annual heating energy as a function of the sample size. The median of the KS distances is 0.3 in the 10 sample size, while the median is decreased to 0.2 in the 15 sample size. With the increase of sample size, the median is reduced below 0.15 after the sample size of 40. Hence, the similarity from five repetitions has increased significantly from the sample size 10 to 40. Beyond a sample size of 100, most of the KS distances are below 0.1, which indicates that the divergences of distributions among five repetitions has been further reduced. When the sample size increases up to 300, the median of the KS distance is around 0.05. Hence, the similarity of probability distributions for heating energy would increase with an increase of sample size in terms of the KS distance.

Figure 3(b) shows the change of the DTS for heating energy with an increase of sample size. The overall trend of the DTS is the same as the KS distance. The median of DTS would be reduced approximately by 40% from the sample size 10 to 20. Beyond a sample size of 40, all the DTS values are less than 1.0. If the sample size increases up to 100, all the DTS values are below 0.5, which indicates that the divergence of distributions of heating energy among the samples is very small.

Table 2: Summary of percentiles of district heating energy in five sample sizes and five repetitions (unit: kWh/m²)

<table>
<thead>
<tr>
<th>Sample number</th>
<th>Prob.</th>
<th>Repetition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>0.10</td>
<td>60.04</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>67.29</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>78.34</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>83.14</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>84.28</td>
</tr>
<tr>
<td>50</td>
<td>0.10</td>
<td>62.81</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>68.78</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>76.57</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>82.35</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>86.98</td>
</tr>
<tr>
<td>100</td>
<td>0.10</td>
<td>65.47</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>69.73</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>75.17</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>80.55</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>86.78</td>
</tr>
<tr>
<td>200</td>
<td>0.10</td>
<td>63.39</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>68.93</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>75.71</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>81.64</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>87.84</td>
</tr>
<tr>
<td>300</td>
<td>0.10</td>
<td>63.42</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>69.67</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>75.51</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>81.58</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>86.34</td>
</tr>
</tbody>
</table>
Overlapping analysis from individual and all samples

This section is focused on the comparison of density functions from the individual sample and the reference case. All the samples are combined together, which is used as the benchmark due to a large number of sample size 10500.

Figure 4(a) shows the overlapping index as a function of sample size from the individual sample. The blue line is a prediction mean value from the LOESS (Locally Estimated Scatterplot Smoothing) smoother and the grey area around the blue line is the 95% confidence intervals. The overlapping index would increase with an increase of sample size although there are large variations from individual samples. When the sample size is over 100, most overlapping index values exceed 0.85. The variations of overlapping index reduces as the sample sizes increases.

Figure 4(b) shows the overlapping index by combining the five repetitions in a specific sample size. The overlapping results would be much more stable in comparison with the results as shown in Figure 4(a). In the case of sample size 10, the overlapping index of five repetitions is 0.83, which is even higher than the individual overlapping in the sample size up to 75. When the sample size increases to 25, the overlapping index of the combined five repetitions exceeds 0.9, higher than most individual samples. Therefore, the results from the replicated samplings can be combined to provide stable results in comparison with the single time LHS.

It is interesting to note that the overlapping value for the equivalent number of a single large number is almost the same as the overlapping value for the several independent replicates with a small sample size. For instance, the average overlapping in the case of sample size 150 is 0.917 (Figure 4a), while the overlapping value in the case of all five replicates with a sample size 50 is 0.920 (Figure 4b). The decreasing effect due to the application of replicated LHS is not significant even though the

Figure 3: Box plots of Kolmogorov-Smirnov distance and Dowd test statistics for heating energy with 5 random seeds and 18 different sample sizes
stratification feature has been diminished compared to a single large number of LHS. This indicates that the replicated LHS is an effective method to both produce reliable uncertainty energy results and provide variance estimation for the LHS samples. Note that the original purpose of using the replicated LHS is to obtain the variance of LHS samples.

The replication number can be determined in a progressive manner to gradually add more replications by checking the similarity of density plots for building energy. The change of percentile values for energy use from the combined samples of several replications can be also used as an indicator to check the convergence of building energy. The annual heating energy simulation results of buildings from the three repetitions would be very stable in this study when the sample size is above 30. Moreover, there is a special type of LHS, named as the sliced LHS, in which a large LHS design can be divided into smaller LHS design (Qian, 2012). As a result, these smaller LHS can be used as the replicated LHS to assess the stability of uncertainty results. The final results can be obtained from the combined small LHS to present better convergence performance since this large sample still has the advantage of good stratification. Further analysis is required to assess the suitability of the sliced LHS in the replicated uncertainty analysis of building energy.

Figure 4: Overlapping index of heating energy as a function of sample size

Figure 5: Cumulative density plots of heating energy with 95% confidence intervals in six sample sizes
Confidence intervals from replicated samples

Figure 5 shows the 95% confidence intervals of heating energy in six samples based on the method described in the section of “Confidence intervals from replicated samplings”. The confidence intervals are very large in the case of 20 sample size, which suggests that further samplings are required to obtain stable results for heating energy in buildings. The confidence intervals close to the 50th percentile are less than the intervals in other percentiles. This indicates that the median of heating energy converges faster than the larger or smaller percentiles, such as 90th or 10th percentiles. Hence, more samples are required if interest focus on energy use with a smaller probability. This may occur when concentrating on a strict energy benchmark or building codes. This trend can be observed in other sample sizes, such as 50 and 100. As the sample size increases, the differences in various percentiles for heating energy become less significant. The confidence intervals become smaller with an increase of sample size as also can be seen from Figure 5. Hence, sufficient sample size is required to obtain stable cumulative density functions of heating energy. For building projects, the confidence intervals may be determined firstly based on the project requirements. Then, the confidence interval can be computed as described in the section of “Confidence intervals from replicated samplings” in the specific sample size.

Another way of applying this method is to compute confidence intervals after visual inspection of cumulative distribution functions from replicated samples. If the confidence interval can meet the pre-determined requirements, then the uncertainty results from the combination of all replicated samples can be used to properly estimate the building energy performance.

Conclusion

This paper investigates the uncertainty analysis of building energy using the replicated Latin Hypercube sampling (LHS) method. The following conclusions can be obtained from this research.

1) The replicated LHS can provide the reliable energy use estimation when applying the uncertainty analysis in building energy assessment.

2) The statistical distances (such as Kolmogorov–Smirnov distance, Dowd test statistics) can be used to quantify the difference of probability density for building energy use between the replicated samples.

3) The uncertainty performance from the replicated LHS does not differ significantly from the results with an equivalent large number of LHS.

4) The confidence intervals of building energy use can be obtained from the replicated LHS to quantify the sample-to-sample variability.

This paper presents a case study to demonstrate the suitability of replicated LHS method. The procedures used in this study can be applied to the other studies of uncertainty analysis of building energy use based on the Latin Hypercube sampling. Further research is needed to apply the sliced LHS method to assess its performance when using the replicated sampling technique in uncertainty analysis of building energy. More different LHS methods (such as maximin criteria, genetic algorithm) can be combined with the replicated sampling to improve the convergence of uncertainty analysis in buildings in future research.

Acknowledgement

This research was supported by the National Natural Science Foundation of China (No. 51778416) and the Key Projects of Philosophy and Social Sciences Research, Ministry of Education (China) “Research on Green Design in Sustainable Development” (contract No. 16JZD014).

References


Renewable and Sustainable Energy Reviews 93: 285-301.


