

2018-10-15

Uncertainty and Sensitivity Analysis of Energy Assessment for Office Buildings based on Dempster-Shafer Theory

Tian, W

<http://hdl.handle.net/10026.1/12198>

10.1016/j.enconman.2018.08.086

Energy Conversion and Management

Elsevier

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Uncertainty and Sensitivity Analysis of Energy Assessment for Office Buildings based on Dempster-Shafer Theory

Wei Tian^{1,2*}, Pieter de Wilde³, Zhanyong Li^{1,2}, Jitian Song^{1,2}, Baoquan Yin⁴

¹ Tianjin Key Laboratory of Integrated Design and On-line Monitoring for Light Industry & Food Machinery and Equipment, College of Mechanical Engineering, Tianjin University of Science and Technology, Tianjin 300222, China

² Tianjin International Joint Research and Development Center of Low-Carbon Green Process Equipment, Tianjin 300222, China

³ Chair of Building Performance Analysis, Environmental Building Group, University of Plymouth, Plymouth, Devon PL4 8AA, United Kingdom

⁴ Tianjin Architecture Design Institute, 300074, Tianjin, China

Abstract: Uncertainty and sensitivity analysis of building energy has become an active research area in order to consider variations of input variables and identify key variables influencing building energy. When there is only limited information available for uncertainty of building inputs, a specific probability for a given variable cannot be defined. Then, it is necessary to develop alternative approaches to probabilistic uncertainty and sensitivity analysis for building energy. Therefore, this paper explores the application of the Dempster-Shafer theory (DST) of evidence to conduct uncertainty and sensitivity analysis for buildings. The DST method is one of imprecise probability theories to allow combining uncertainty from different sources in terms of interval-valued probabilities in order to construct the belief and plausibility (two uncertainty measures) of system responses. The results indicate that the DST uncertainty analysis in combination with machine learning methods can provide fast and reliable information on uncertainty of building energy. It is recommended that at least two inherently different learning algorithms should be applied to provide robust simulation results of building energy. A spectrum of distributions should be implemented in global sensitivity analysis with the DST method because there are no specific distributions for intervals of input factors. Moreover, the stability of results from uncertainty and sensitivity analysis should be assessed when applying the DST method in building energy analysis.

Keywords: Uncertainty analysis; building performance; Dempster-Shafer theory; sensitivity analysis; machine learning

*Corresponding author: Wei Tian, Tel: +86 (022) 60600705, Fax: +86 (022) 60600705
Email: tjtianjin@gmail.com

39 **Abbreviations**

BPA	Basic probability assignment
CBF	Cumulative belief function
CCBF	Complementary cumulative belief function
CCPF	Complementary cumulative plausibility function
CDF	Cumulative density function
CHP	Combined heat and power
CL	Cooling set-point temperature (°C)
CPF	Cumulative plausibility function
CSWD	Chinese standard weather data
DST	Dempster-Shafer theory
ED	Equipment peak value (W/m ²)
FT	Infiltration rate (ACH)
HT	Heating set-point temperature (°C)
LD	Lighting power density (W/m ²)
MARS	Multivariate adaptive regression splines
OD	Occupancy density (people/m ²)
RMSE	Root mean square error
SHGC	Solar heat gain coefficient
SVM	Support vector machine
VAV	Variable air volume system

41 1 Introduction

42 Building energy is affected by a number of inherently uncertain variables, including
43 weather conditions, internal heat gains, and occupant behaviours [1, 2]. Therefore, uncertainty
44 analysis of building energy has become an active research field [3-5]. Most previous studies
45 have implemented probabilistic uncertainty methods to consider the influences of these
46 uncertain parameters [6-8]. Urbanucci and Testi [9] use the Monte Carlo risk analysis to
47 estimate the long-term uncertainty of energy demands for a hospital facility in order to
48 optimize the size of CHP (combined heat and power) system. Tian et al. [10] consider the
49 influences of variations of building form on energy performance of buildings located at
50 Harbin (China) based on the Monte Carlo sampling method. Faggianelli et al. [11] implement
51 sampling-based sensitivity analysis by regarding input factors as uniform or normal
52 distributions. Hopfe and Hensen [12] assume the normal distributions for input factors to
53 assess energy performance of an office building using the Latin hypercube sampling method.
54 These examples demonstrate that probabilistic uncertainty and sensitivity analyses have
55 become very popular and widely used in the field of building performance simulation.
56 However, variations of building variables are difficult to obtain and it can be a challenge to
57 gather sufficient information for the definition of a specific probability (such as uniform,
58 normal, triangle, and lognormal) when predicting energy use, especially in the stage of
59 building design [13]. Hence, the alternative approach to probabilistic analysis is needed to
60 handle the imprecise building data in properly estimating energy performance of buildings.

61 The Dempster-Shafer theory (DST) of evidence can be regarded as a generalization of
62 classical probability theory that allows one to deal with the imprecise information on data,
63 often in the form of interval-valued data. The mathematical foundations of DST analysis have
64 been well established [14] and the DST approach has been used in various fields, including
65 studies on reliability of pressure vessels [15], petroleum engineering [16], urban environment
66 [17], and computer voice detection [18]. More recently, the DST analysis is also being
67 applied to the analysis of building energy. Tian et al. [3] implement the DST to assess
68 uncertainty of energy performance for an office building using the EnergyPlus program. Four
69 scenarios are used in their research to represent the level of availability for uncertain inputs
70 from the simple to detailed information. Chaney et al. [19] use the DST to add multiple-
71 sensor data in a house simulation model. They found that the evidence theory is a reasonable
72 approach for providing rich information about occupant interaction with systems in the house.
73 Kim et al. [20] report that the DST can be used to effectively combine uncertainties from five
74 experts into single uncertainty when predicting energy use for a 33-storey office building in
75 Seoul, Korea.

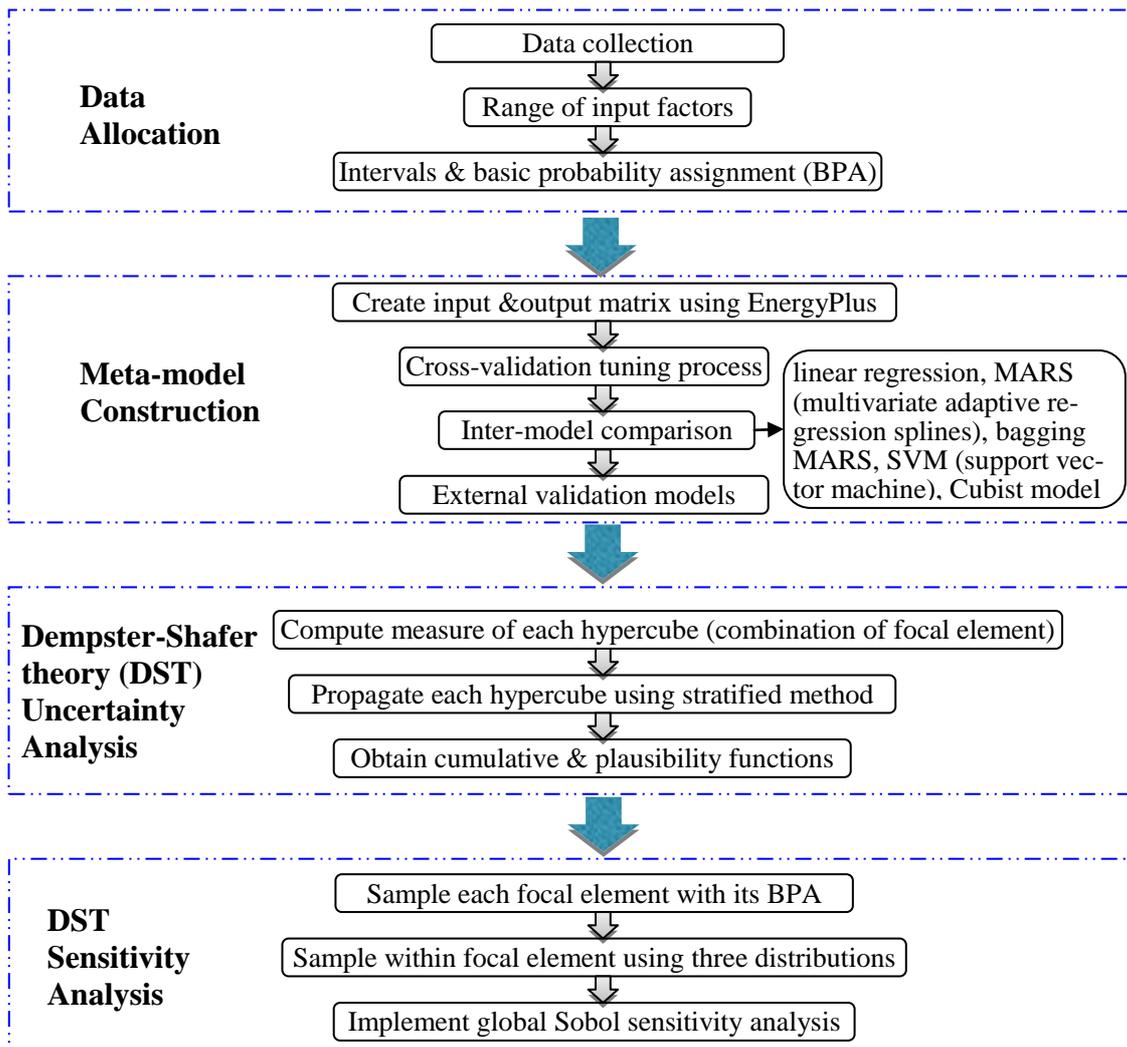
76 These previous studies provide valuable information on the implementation of DST
77 analysis in building energy assessment. However, there are several issues that have not been
78 explored when applying the DST method in building energy assessment. One issue is how to
79 reduce high computational cost of DST analysis in building simulation using engineering-
80 based energy models. A large number of simulation runs are usually required to provide the
81 minimum and maximum output values in order to obtain the stable results of output range for
82 the DST method. Another issue is how to implement sensitivity analysis within the context of
83 DST analysis in assessing building energy performance. The sampling-based sensitivity
84 analysis requires the structured distributions of input variables to obtain a matrix of inputs and
85 outputs. The DST method, however, does not include specific distributions for the data within
86 the intervals.

87 Therefore, this paper explores a systematic approach towards implementation of the DST
88 method in uncertainty and sensitivity analysis of building energy when only limited
89 information on building input variables is available. An office building located in Tianjin
90 (China) is used as a case study to demonstrate the suitability of DST method in assessing

91 building energy performance. The building energy simulation is carried out with the
 92 EnergyPlus program [21]. The originality of this paper is two-fold: (1) implementation of
 93 global sensitivity analysis in conjunction with the DST analysis in assessing building energy
 94 performance; (2) demonstration of using machine learning models to reduce high
 95 computational cost of building energy simulation for both uncertainty and sensitivity analysis
 96 within the DST analysis. Moreover, this research discusses two important issues in the
 97 application of DST analysis: how to choose reliable machine learning models and how to
 98 assess the stability of uncertainty and sensitivity analysis. This provides practical guidance in
 99 applying the DST method into building energy assessment. The combination of DST and
 100 machine learning algorithm can significantly expedite computation, which can make DST
 101 analysis feasible in building energy assessment. However, a number of machine learning
 102 models should be evaluated to choose suitable ones for replacing building energy models on a
 103 case-by-case basis. More discussion on the method used will be presented in section 2.

104 The remaining parts of this paper are structured as follows. Section 2 describes the
 105 statistical methods applied in this research, including the DST analysis, machine learning
 106 models, and sensitivity analysis. Section 3 presents a case study of building energy model to
 107 implement DST analysis. Section 4 discusses the results from these three types of statistical
 108 approaches when assessing the energy performance of an office building. Section 5 presents
 109 the conclusions and further research required in this field.

110 **2 Method**



111 **Figure 1. Flow chart used in this research**
 112

113 The computational procedure used in this study is shown in Figure 1. The first step is to
 114 collect the data for uncertainty and sensitivity analysis of building energy assessment. The
 115 data should be processed based on the requirement of the DST analysis (as described in
 116 sections 3). The second step is to create fast-computing machine learning models based on the
 117 engineering-based energy models (sections 2.2). These computationally cheap models then
 118 will be used for the uncertainty and sensitivity analysis of DST method. The third step is to
 119 implement DST uncertainty analysis with the data obtained from the first step using the
 120 learning models from the second step (section 2.1). The fourth step is to apply global
 121 sensitivity analysis in conjunction with the DST analysis based on the machine learning
 122 models obtained from the second step (section 2.3).

123 2.1 Dempster-Shafer Theory

124 For most engineering problems, the relationships among inputs and outputs can be written
 125 as,

$$y = f(x) \quad (1)$$

126 where x is a vector of system inputs, y is a vector of system outputs, and f is a function to
 127 describe the relationships among x and y . In this case study, x represents input variables in
 128 building energy assessment, such as occupant density, equipment heat gains, and heating set-
 129 point temperature, while y is energy performance, such as annual heating energy, annual
 130 cooling energy, annual total electricity, or carbon emissions. The f functions in this research
 131 are to represent complex relationships among building input variables and energy
 132 performance that can be computed using building simulation programs, such as EnergyPlus,
 133 ESP-r, and DOE-2 [22].

134 It is common practice to define probabilities for building input variables in uncertainty
 135 analysis of building performance. The sampling-based probabilistic uncertainty propagation
 136 could be used to obtain uncertain performance of buildings by running building energy
 137 models with a large number of times. However, in the case of limited information for these
 138 input data, it may be difficult to justify the choices of specific probabilities. As a result, a less
 139 structured representation of uncertainty for building input factors is needed instead of specific
 140 probabilities. The Dempster-Shafer theory (DST) of evidence is a generalization of classical
 141 probability theory to handle the imprecise information on input data [14].

142 For the DST analysis, an evidence space for input data x is specified as a triple $(X, X_E,$
 143 $m_{EX})$, where X is the set of all possible values (i.e. sample space), X_E is a set of subsets of X
 144 (i.e. focal elements), and m_{EX} is a function for a subset U of X (i.e. basic probability
 145 assignment, BPA).

$$146 m_{EX}(U) > 0 \text{ if } U \subset X \text{ and } U \in X_E \quad (2)$$

$$147 m_{EX}(U) = 0 \text{ if } U \subset X \text{ and } U \notin X_E \quad (3)$$

148 The $m_{EX}(U)$ function denotes the amount of information assigned to U . Similar to the
 149 probability theory, the sum of basic probability assignment equals one. However, in the DST
 150 method, there are two uncertainty measures: belief and plausibility defined as follows,

$$Bel_X(U) = \sum_{V \subset U} m_{EX}(V) \quad (4)$$

$$Pl_X(U) = \sum_{V \cap U \neq \emptyset} m_{EX}(V) \quad (5)$$

151 $Bel_X(U)$ can be interpreted as a measure of the amount of information to support U
 152 containing true values, while $Pl_X(U)$ represents the absence of information to support U
 153 containing false values. In a similar way, the uncertainty of system output defined in Eq.(1)
 154 can be obtained with an evidence space of triple (Y, Y_E, m_{EY}) . The resultant uncertainty of
 155 outputs can be summarized with a cumulative belief function (CBF) and a cumulative

156 plausibility function (CPF) by the corresponding beliefs and plausibilities. More detailed
157 descriptions on the DST can be found in [14, 23].

158 As shown in Figure 1, the computation procedure for the DST uncertainty analysis can be
159 divided into three steps. The first step is to define a hypercube that is the combination of focal
160 elements for input variables and then calculate the composite evidential measure (i.e. BPA) of
161 each hypercube. For the case study building as will be described in section 3, the numbers of
162 hypercubes in case A and case B are 32 and 512, respectively. These two values can be
163 calculated based on the fundamental of combination from the numbers of intervals listed in
164 Table 1 and Table 2.

165 The second step is to compute the minimum and maximum values of the system response
166 in each hypercube, which are the most computationally intensive procedure in the DST
167 analysis. The computational methods available are optimization, sampling, and vertex
168 techniques [14, 24]. To reduce computation cost, machine learning algorithms are used to
169 provide reliable results with fast computing as will be described in section 2.2. The sampling
170 method is chosen in this paper since the whole output space can be sufficiently explored by
171 using machine learning energy models. The conventional probability sampling method [25,
172 26] is used to obtain the uncertainty of outputs for each hypercube cell. The uniform
173 distribution for all input variables is assumed to propagate input uncertainty to output
174 uncertainty and then find the minimum and maximum output values for this hypercube space.
175 Note that this does not mean the distribution within intervals is uniform and the purpose of
176 assuming uniform distribution is only to obtain the minimum and maximum values within the
177 intervals. Latin Hypercube sampling with a sampling size of 10,000 is used due to its high
178 stratification. Discussion of convergence of outputs with sample size will be presented in
179 section 4.2.1.

180 The third step is to form the cumulative belief (CBF) and plausibility functions (CPF) by
181 aggregating the minimum and maximum values of system response obtained from the second
182 step. The uncertainty results with the DST method are bounded between the CBF and CPF.
183 The CBF is a lower bound on a probability value consistent with the evidence space, whereas
184 the CPF is an upper bound on a probability value consistent with the evidence space. Hence,
185 uncertainty results can be interpreted by the CBF (the smallest probability) and CPF (the
186 largest probability) that are combined together to have the complete information of all
187 possible output values (i.e. energy use in this case study). Similar to the probability theory,
188 the CCBF (complementary cumulative belief function) and CCPF (complementary
189 cumulative plausibility function) may be more useful in the field of risk analysis or reliability
190 analysis.

191 **2.2 Machine learning algorithms**

192 Machine learning is used to create reliable and fast-computing models (also called meta-
193 models or surrogate models) based on the inputs and outputs computed from the EnergyPlus
194 program. In this study, 400 EnergyPlus models are used to construct a matrix containing
195 inputs variables and outputs for this office building using the Sobol sequence. This Sobol
196 sequence is a quasi-random low-discrepancy sequence with a better performance in
197 comparison with the Monte Carlo sampling [27, 28]. The 400 simulation runs of EnergyPlus
198 models are used in this case study based on preliminary studies to create reliable meta-models
199 of building energy use. This simulation number is higher than ten times of input variables
200 used in most of building energy simulation studies [13]. The simulation number of energy
201 models required for creating accurate energy models can be evaluated using the RMSE (root
202 mean square error) of energy models. In this study, the RMSE for three performance indicator
203 has become stable after around 300 times and the extra 100 times of simulation models (total
204 number 400) are used to ascertain better performance of energy analysis. The determination
205 of simulation number for constructing reliable meta-models is likely to be problem dependent,
206 depending on number of input variables, simulation output, complexity among inputs and

207 outputs, and accuracy required by building projects. The R caret package [29] is used to
208 create these machine learning models in this study. The R caret package combines more than
209 100 machine learning models to provide a streamlined process for creating predictive models.

210 Five machine learning algorithms have been selected since they have better performance
211 in terms of predictive capability and are also widely used in the field of building energy
212 analysis [30, 31]. These five models are linear regression, MARS (multivariate adaptive
213 regression splines), bagging MARS, SVM (support vector machine), and Cubist model. The
214 reason for exploring these five options is as follows. (1) The linear model is still used here
215 because the linear model has good performance with better interpretation [32]. (2) MARS
216 creates a piecewise linear model to replace original predictors with new surrogate features to
217 account for non-linear effects [33]. If necessary, the interaction terms of these new features
218 can be also considered in MARS models to further improve predictive performance. The
219 number of new features and the number of degrees of interactions can be determined using an
220 automatic pruning procedure. (3) The bagging MARS approach implements the bagging
221 (bootstrap aggregating) technique to stabilize the predictive results from MARS models. The
222 bagging technique, an ensemble learning method, simply creates a number of new data set
223 using the bootstrap method (i.e. randomly sampling with replacement) to create a number of
224 corresponding models instead of only one regression model [33]. Then the prediction results
225 are averaged from these regression models to reduce the variance of outcomes. A
226 disadvantage of the bagging technique is high computational cost because more
227 computational time is required with an increase in the number of bootstrap samples. A multi-
228 core workstation is used here to expedite the calculation using parallel computing. (4) The
229 SVM is similar to robust regression that tries to mitigate the influence of influential
230 observations. Several kernel functions (polynomial, radial basis, hyperbolic tangent) are
231 available in SVM to encompass nonlinear functions of inputs. The radial basis function is
232 chosen in this study based on the suggestion from Kuhn and Johnson [33]. (5) The Cubist
233 belongs to the rule-based models with the boosting technique. The boosting technique is one
234 of ensemble methods to provide the unequally weights for different models in terms of model
235 errors [3]. More detailed information on these machine leaning techniques is available in [33,
236 34].

237 In the five models above, model variables need to be tuned except for the linear model.
238 The cross-validation method is used to find the optimal values for these models. The cross
239 validation in this study is based in randomly dividing the original data set into ten sets of
240 roughly equal size (also named ten-fold) [34]. Then one new data set is used as a test set to
241 assess the performance of regression models obtained from the remaining nine data sets as
242 training data. This process repeats ten times until all the ten data sets are used as test sets. To
243 further test the predictive performance of the optimal model for five algorithms, an extra 200
244 EnergyPlus models are simulated except for the 400 simulation that are used for regression.
245 Two measures are used to assess predictive performance of regression models: RMSE (root
246 mean square error) and R^2 (coefficient of determination). RMSE is widely used in the field of
247 machine learning and is the absolute fit measure how the regression model predicts the
248 outcomes. The lower RMSE, the better regression model is. R^2 is the relative measure to
249 account for the proportion of total variance explained by the model. The higher R^2 indicate a
250 better regression model.

251 The choice of suitable machine learning models involves a lot of efforts, which is related
252 to the prior knowledge of both building physics and machine learning algorithm. For instance,
253 if there are interactions among input variables on building energy use, it is necessary to
254 choose machine learning models that can consider interactions. For the MARS approach,
255 second or higher degree terms should be added to tune optimal models. This is the case for
256 London domestic gas use influenced by a number of factors, including dwelling type,
257 household composition, and building area [30]. However, if there are no strong interactions,

258 the MARS model without second degree terms should be used since more terms actually
 259 deteriorate model predictive performance. This is the case for assessing annual heating and
 260 cooling energy of an office building located in London [35]. When there exists highly
 261 nonlinear relationships between inputs and output in building energy performance, non-
 262 parametric machine learning models usually perform better than linear models. However, if
 263 there are approximately linear relationships in building energy analysis, linear models would
 264 have more robust performance in comparison with most of complex non-parametric
 265 relationships. This has been confirmed in [36] to assess energy performance of campus
 266 buildings at Georgia Institute of Technology, USA. If using support vector machine, linear
 267 kernel function should be used instead of non-linear polynomial or radial basis functions.
 268 Tian et al [36] also discuss another important issue on correlation of input variables, which
 269 usually leads to unstable meta-models. For instance, equipment heat gains are usually
 270 associated with lighting use in office buildings. Then, the principal component approach or
 271 partial least square method can be used to reduce the number of correlated variables to
 272 increase the stability of meta-models. More research is required to choose suitable machine
 273 learning algorithms based on building features, building type, and thermal performance of
 274 buildings,

275 **2.3 Sensitivity analysis for Dempster-Shafer Theory**

276 The DST method causes two issues in implementing sensitivity analysis in building
 277 energy assessment. The first issue is that the DST analysis does not assume any distribution
 278 for the intervals of input variables. However, the sampling-based global sensitivity analysis
 279 requires to have the specific distributions for a variable, which may have significant influence
 280 on sensitivity results [37]. The second issue is the high computational cost for sensitivity
 281 analysis due to the nature of DST analysis as discussed in section 2.1. For the first issue, the
 282 sensitivity method used here is based on the recommendation from Helton et al. [38] to
 283 specify a spectrum of distributions to represent possible variations within the intervals of
 284 focal elements. Three types of distributions are considered to cover the larger values, middle
 285 value, and lower values with the left quadratic, uniform, and right quadratic distributions,
 286 respectively. For the second issue, similar to the DST uncertainty analysis, reliable machine
 287 learning models as described in section 2.2 are used for running global sensitivity analysis
 288 instead of the engineering-based EnergyPlus models. The fast-computing machine learning
 289 models can assure the convergence of the global sensitivity analysis by running a large
 290 number of times of simulation models.

291 The computational procedure for sensitivity analysis is illustrated in Figure 1. The first
 292 step is to sample each focal element with its BPA using random sampling. The next step is to
 293 sample within the corresponding focal element using three types of distributions (left
 294 quadratic, uniform, and right quadratic), respectively. The reasons for choosing these three
 295 distributions are to cover as many as possible situations within intervals defined from the first
 296 step. The left quadratic distribution can emphasize the smaller values with each focal element,
 297 whereas the right quadratic distribution can emphasize the larger values with each focal
 298 element. The uniform distribution can cover the whole range of each focal element. The
 299 density functions for left, uniform, and right distributions, respectively, are

$$300 \quad f_{left}(x) = 3(b - x) / (b - a)^3 \quad (6)$$

$$301 \quad f_{uniform}(x) = 1 / (b - a) \quad (7)$$

$$302 \quad f_{right}(x) = 3(x - a) / (b - a)^3 \quad (8)$$

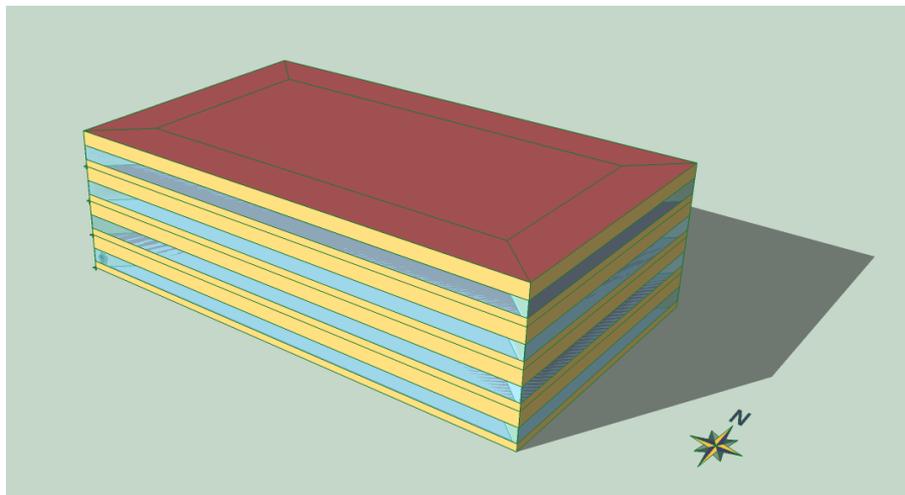
303 where a and b are minimum and maximum values, respectively, within an interval defined in
 304 the first step. More detailed descriptions on these distributions are available in [38].

305 The final step is to implement global Sobol sensitivity analysis with machine learning
 306 models to provide importance ranking of input variables [39]. The Sobol sensitivity method is

307 one of variance-based approach to decompose the variance of output to the corresponding
308 input variable. The detailed procedure for the Sobol sensitivity analysis is available in [37].
309 Two sensitivity indicators are often used in the variance-based method: main effect and total
310 effect. The main effect represents the effects of one individual variable without considering
311 other variables, whereas the total effect is due to the effects of this specific variable and
312 interactions with the other variables. R sensitivity package [40] is used here to implement the
313 Sobol sensitivity analysis.

314 Note that the sensitivity analysis used here is different from quantitative risk assessment
315 in which the probability of input variables need to be specified in the first place. For the DST
316 sensitivity analysis, there is no assumption on the probability for input variables. The results
317 from quantitative risk assessment are similar to the probabilistic sensitivity analysis by
318 considering only one specified probability. In contrast, the results from the DST sensitivity
319 analysis are the combined ranking importance to consider all the possibilities within the
320 intervals. Hence, the results from the DST sensitivity analysis depend on a number of factors,
321 including the BPA and the relationships among inputs and outputs, but not on the
322 specification of probabilities within these intervals of input variables.

323 3 A case study of building energy model



324
325 **Figure 2. An office building used for this research**

326 Figure 2 illustrates an office building studied in this paper. It is a four-storey building
327 with a total floor area of 6,000 m². The window-wall ratio is 40%. The thermal properties of
328 building envelope are commensurate with the requirements of energy efficiency for office
329 buildings in China [41]. The U-values for wall and roof are taken as 0.45 and 0.23 W/m²K,
330 respectively. The U-value and SHGC (solar heat gain coefficient) for windows are 2.40
331 W/m²K and 0.35, respectively. A VAV (variable air volume) air system with perimeter
332 hydronic baseboard heaters is used to provide ventilation, heating, and cooling to maintain
333 indoor thermal comfort. A gas boiler is used to supply hot water, and a centrifugal chiller with
334 air cooling is used to supply chilled water for the VAV system.

335 Internal heat gains for occupants, lighting, and equipment are derived from two Chinese
336 standards [41, 42] and expert opinions as listed in Table 1. The interpretation of values in
337 Table 1 will be discussed in the end of this subsection. The office building is located in
338 Tianjin, China and the typical year weather data (CSWD, Chinese standard weather data) is
339 obtained from the EnergyPlus website [21]. The climate in Tianjin has a cold, windy winter
340 and hot, humid summer, which requires heating in winter and cooling in summer.

341 The EnergyPlus V8.8 program is used to simulate the thermal behaviours of the building
342 [21]. EnergyPlus is widely used in the field of building energy analysis and has been
343 validated extensively. Typical one-core and four-perimeter zones are used for zoning this

344 building when creating an energy model. Three performance measures are annual heating
 345 energy, annual cooling energy, and annual carbon emissions normalized by the floor area.
 346 Heating and cooling energy values are directly obtained from the simulation results of
 347 EnergyPlus models. Carbon emissions are calculated by multiplying the carbon emission
 348 factors of electricity use (1.00 kgCO₂/kWh) [43] and natural gas (0.20 kgCO₂/kWh) [44],
 349 respectively, with annual electricity and gas use from the results of EnergyPlus models. Note
 350 that other performance measures can be also used for this method, such as overheating risk in
 351 natural ventilation buildings.

352 **Table 1. Intervals and BPA (basic probability assignment) from Expert I and Expert II**

Variable	Short names	Expert I (case A)		Expert II	
		<i>Intervals</i>	<i>BPA</i>	<i>Intervals</i>	<i>BPA</i>
Infiltration rate (ACH)	FT	[0.3, 0.4], [0.4, 0.5]	0.6, 0.4	[0.3, 0.4]	1
Equipment peak value (W/m ²)	ED	[13, 15], [16, 17]	0.6, 0.4	[14, 15], [15, 16]	0.8, 0.2
Lighting power density (W/m ²)	LD	[6, 8], [7, 9]	0.3, 0.7	[6, 7], [7, 8]	0.4, 0.6
Occupancy density (people/m ²)	OD	[8, 10], [11, 12]	0.5, 0.5	[9,10], [10, 11]	0.8, 0.2
Heating set-point temperature (°C)	HT	[19, 21]	1	[20, 21]	1
Cooling set-point temperature (°C)	CL	[24, 25], [25, 26]	0.6, 0.4	[24, 25]	1

353
 354 **Table 2. The combined intervals and BPA (basic probability assignment) for case B**
 355 **from two experts**

Variable	<i>Intervals</i>	<i>BPA</i>
Infiltration rate (ACH)	[0.3, 0.4], [0.4, 0.5]	0.8, 0.2
Equipment peak value (W/m ²)	[13, 15], [14, 15], [15, 16], [16, 17]	0.3, 0.4, 0.1, 0.2
Lighting power density (W/m ²)	[6, 7], [6, 8], [7, 8], [7, 9]	0.2, 0.15, 0.3, 0.35
Occupancy density (people/m ²)	[8, 10], [9, 10], [10, 11], [11, 12]	0.25, 0.4, 0.1, 0.25
Heating set-point temperature (°C)	[19, 21], [20, 21]	0.5, 0.5
Cooling set-point temperature (°C)	[24, 25], [25, 26]	0.8, 0.2

357 The uncertain input factors considered in this study are listed in Table 1. The purpose of
358 the case study is to explore how the DST can help to provide more reliable simulation outputs
359 by considering uncertainty of new buildings in the preliminary design stage. These variables
360 are closely related to occupant behaviour, including infiltration rate, equipment heat gains,
361 lighting heat gains, heating & cooling set-point temperatures. The information on these input
362 variables is obtained from two experts in the area of building energy engineering, Expert I and
363 Expert II as summarized in Table 1. For the infiltration rate, Expert I states that the actual
364 infiltration rate is in one of two contiguous intervals: in the interval [0.3, 0.4] with a 60%
365 level of subjective belief (named as basic probability assignment in evidence theory), or in the
366 interval [0.4, 0.5] with a 40% level of subjective belief. In contrast, in Expert II's opinion, the
367 infiltration rate lies in the interval [0.3, 0.4], with a 100% level of subjective belief. The other
368 values in Table 1 can be interpreted in the same way. Based on the suggestion from these two
369 experts and previous studies [8, 13], the infiltration rate is treated as a constant ACH (air
370 exchange per hour) value in a whole year in this research since the infiltration rate is very
371 uncertain, depending on building age, construction quality, building use, and weather
372 conditions [13].

373 Two cases (named as case A and case B) are considered in this paper to represent two
374 different uncertain situations. Case A is directly obtained by the opinion of the Expert I as
375 listed in Table 1. Case B is derived by combining the opinions from Expert I and Expert II as
376 summarized in Table 2 in which the input variables can be explained in the same way as the
377 values in Table 1. It is assumed that the two sources are weighted equally since both Expert I
378 and Expert II are senior building engineers. The detailed calculation procedure is available in
379 a book chapter written by Oberkampf and Helton [23]. A number of methods are available to
380 combine the evidence from different sources; please refer to [45, 46].

381 In order to compare the results from DST and probability-based analysis, the simulation
382 results from uniform distributions are used as a special case for probability-based analysis to
383 represent the results for conventional probabilistic method. The corresponding uncertainty
384 results are named as cumulative distribution function (CDF) in this research. Note that the
385 results from the DST method are interval-based for a specific probability, whereas the results
386 from the probabilistic method are specific values for a given probability.

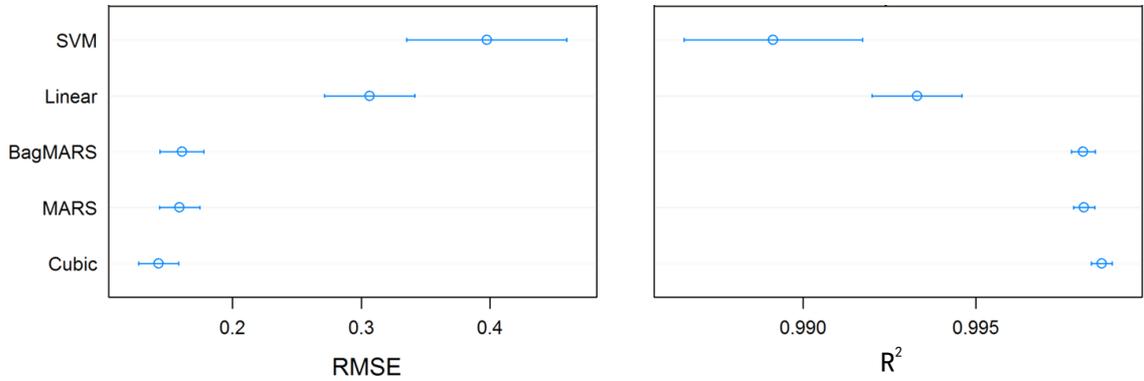
387 **4 Results and discussion**

388 **4.1 Performance comparison of machine learning models**

389 Figure 3 shows the comparisons of predictive performance of five machine learning
390 models (as described in section 2.2) for heating, cooling, and carbon emissions in order to
391 choose reliable models for uncertainty and sensitivity analysis based on the Dempster-Shafer
392 theory of evidence. The variations of RMSE (root mean square error) and R^2 are expressed as
393 95% confidence interval of two statistics in this study. A reliable machine learning model
394 should have low RMSE values and high R^2 . Moreover, the variations for these two measures
395 should be also low to provide stable estimation of simulation outputs for buildings in this case
396 study.

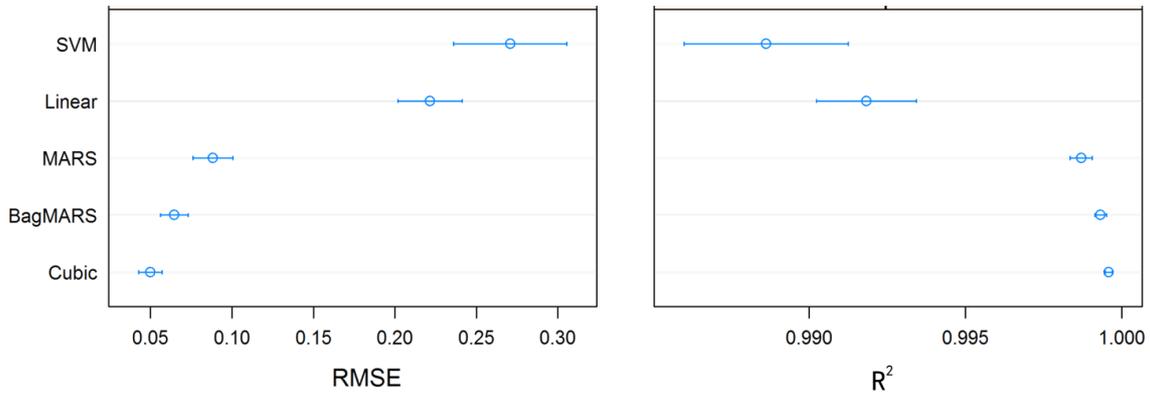
397 Figure 3a indicates that the Cubist model performs the best amongst the five models for
398 estimating heating energy. The mean value of corresponding RMSE for heating energy is
399 only 0.14 kWh/m² and the variation of RMSE is also small in terms of 95% confidence
400 interval. The next two models are the MARS and bagging MARS models with similar
401 accuracy, which indicates that the bagging technique does not significantly improve
402 predictive accuracy in this case study. The linear and SVM models do not perform well for
403 predicting annual heating energy use. To further validate regression models, the extra 200
404 simulation runs that are not used to obtain regression models are applied to the external
405 validation of models. The statistics from this external validation are listed in Table 3 in which
406 the Cubist model has the best performance in terms of both R^2 and RMSE. The MARS and

407 bagging MARS have similar predictive capability for heating energy. These conclusions are
 408 the same as those obtained in the internal cross-validation method, which indicates that the
 409 Cubist and MARS regression can be used to provide reliable heating energy use for this office
 410 building.



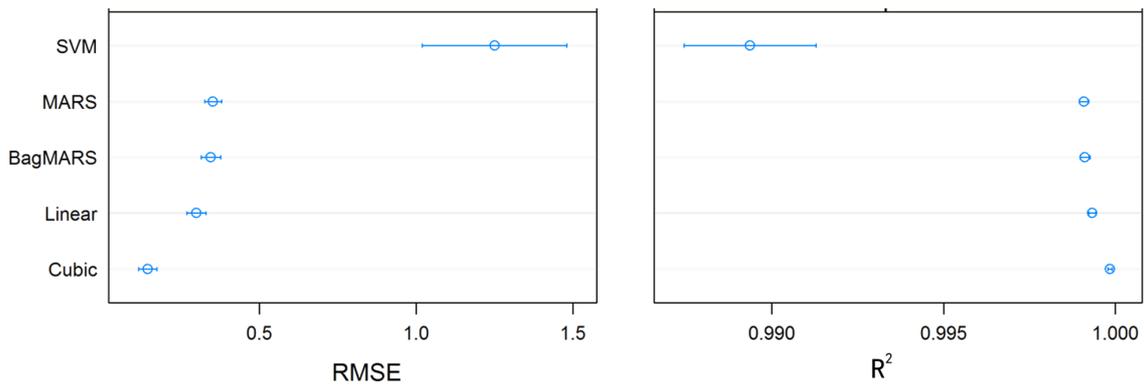
411
412

Confidence Level: 0.95
(a) Heating energy use



413
414

Confidence Level: 0.95
(b) Cooling energy use



415
416

Confidence Level: 0.95
(c) Carbon emissions

Figure 3. Comparison of five machine learning models for estimating performance of building

417
418
419

420
421

Table 3. Comparison of predictive performance from five machine learning models using external validation

Model	Heating energy		Cooling energy		Carbon emission	
	R ²	RMSE	R ²	RMSE	R ²	RMSE
Linear	0.992	0.326	0.991	0.223	0.999	0.300
MARS	0.998	0.176	0.998	0.085	0.999	0.343
Bagging MARS	0.998	0.162	0.999	0.065	0.999	0.343
SVM	0.990	0.365	0.990	0.245	0.987	1.315
Cubist	0.999	0.136	0.999	0.041	0.999	0.204

422

423 Figure 3b shows the comparison of five machine learning models for cooling energy use.
424 The Cubist model is still the best performer in this case study, similar to heating energy use.
425 The next model is the bagging MARS model, which is better than the MARS models in terms
426 of both R² and RMSE. Therefore, the bagging technique has more influence for cooling than
427 heating energy in this office building by providing more stable predictions. The linear and
428 SVM models do not perform as well as the other three learning models. The corresponding
429 statistics of R² and RMSE from external validation are summarized in Table 3 to indicate that
430 the Cubist and MARS models are two best performers for predicting cooling energy in the
431 office building. In terms of RMSE, the Cubist model is approximately 5 times better than the
432 linear model.

433 Figure 3c shows the predictive performance of five machine learning models for carbon
434 emissions using cross validation. The best learning model is from the Cubist method. The
435 next model is from linear regression, which may be unexpected. This suggests the linear
436 model may have better performance compared to non-parametric models when the
437 relationship between inputs and outputs is approximately linear. The linear model shows
438 slightly better performance than the MARS and bagging MARS models. As a result, the
439 Cubist and linear models are selected to validate the model performance using external
440 EnergyPlus simulation runs. The corresponding RMSE and R² from external validation are
441 listed in Table 3 to show that two best models are the Cubist and linear model in this case
442 study to estimate carbon emissions.

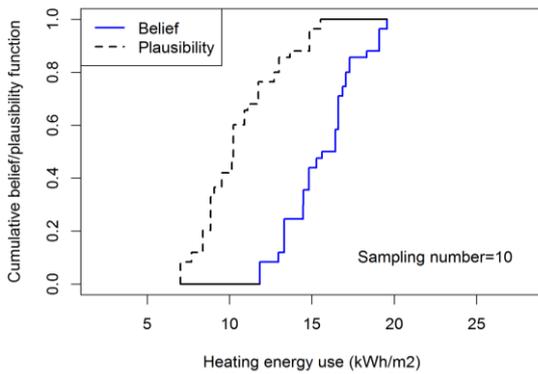
443 The bootstrap approach is used as an alternative model selection method to compare the
444 results from the cross-validation method. A bootstrap sample is a random sample of the
445 original data set taken with replacement, which has the same size as the original data. Hence,
446 the samples that are not selected in a bootstrap sample can be used as out-of-bag samples to
447 validate models obtained from a bootstrap sample. The results indicate that most of results are
448 similar to those obtained from the cross-validation method as shown in Figure 3 although the
449 variations from the bootstrap approach are smaller than those from the cross-validation
450 method.

451 It is recommended to determine the accuracy of machine learning models required for
452 building projects in order to decide when to stop choosing suitable machine learning models.
453 This is because a large number of machine learning models are available [33] and it is
454 unnecessary to try a large number of machine learning models for a specific building project.
455 The acceptance criterion for model performance can refer to the values set out by ASHRAE
456 Guideline 14 in which the coefficient variations of RMSE for energy models should be lower
457 than 5% [47]. This threshold value has been widely used in building energy analysis [48]. The

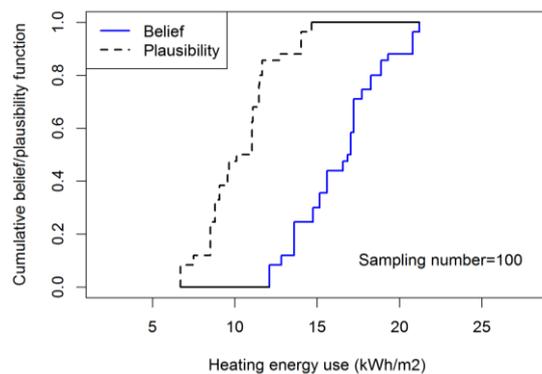
458 coefficient variations of RMSE for all three performance output (heating, cooling, and carbon
 459 emissions) are lower than this criterion in this case study. The mean coefficient variations of
 460 RMSE for Cubist models are 1.02%, 0.13%, and 0.09% for heating, cooling, and carbon
 461 emission, respectively. Therefore, the Cubist models have good performance in terms of
 462 ASHRAE Guideline 12 [47]. This criterion from ASHRAE can be regarded as the minimum
 463 requirement for performance of energy models. This is because buildings typically are
 464 bespoke, one-off products that are designed in response to a unique client brief. The
 465 translation of client brief to technical requirements is mostly conducted by expert consultants
 466 who have a considerable freedom in setting accuracy targets and thresholds [49], which is
 467 similar to the processes observed in Systems Engineering [50, 51].

468 4.2 Results of uncertainty analysis

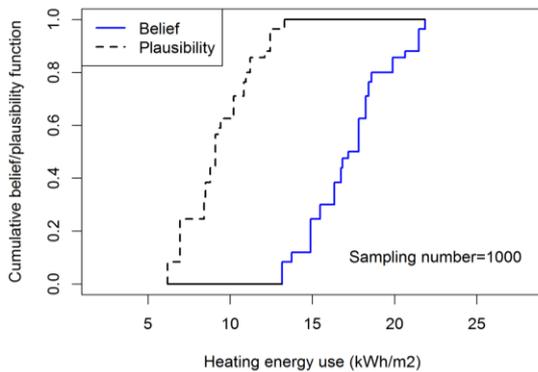
469 4.2.1 Annual heating energy



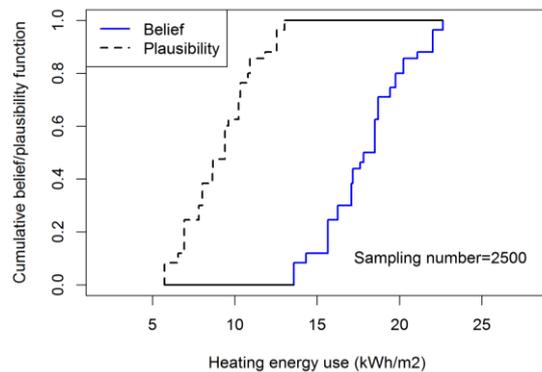
470 (a) sampling number 10



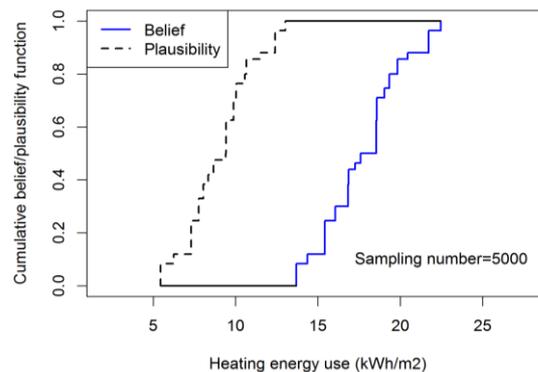
471 (b) sampling number 100



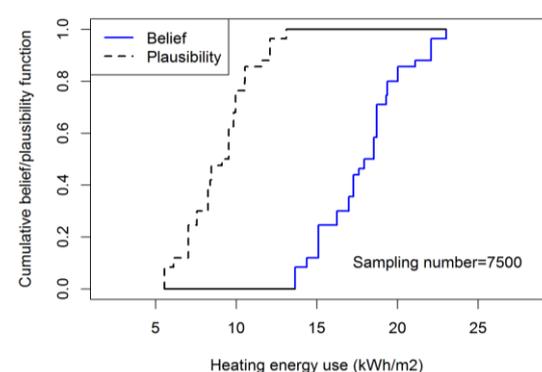
472 (c) sampling number 1,000



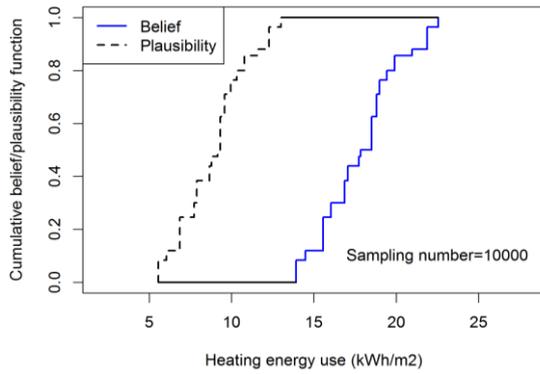
473 (d) sampling number 2,500



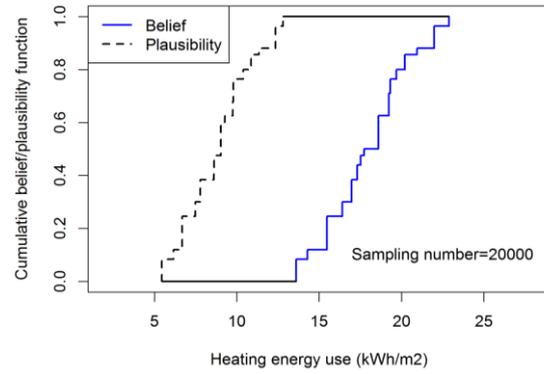
474 (e) sampling number 5,000



475 (f) sampling number 7,500



(g) sampling number 10,000

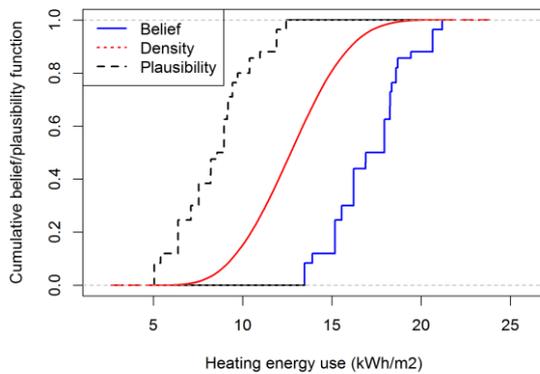


(h) sampling number 20,000

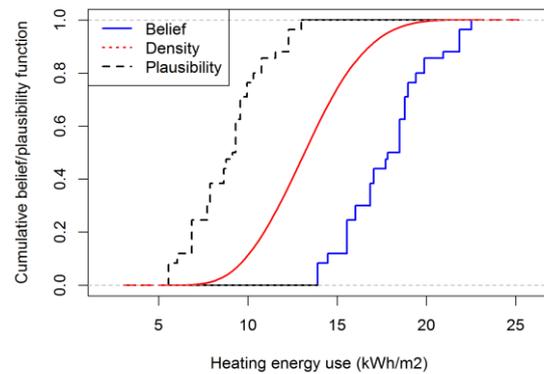
Figure 4. Comparison of stability of uncertainty results with sampling number for heating energy using the MARS model in case A

For uncertainty analysis, it is necessary to assess the convergence of results. Figure 4 compares change of cumulative belief function (CBF) and cumulative plausibility functions (CPF) for annual heating energy using the MARS model as the sample size increases. The area between the CBF and the CPF becomes larger with an increase in sample size because more resamples are required to find the minimum and maximum values for every cell defined in the DST analysis. For the sampling size of 10, the results are apparently inadequate and not converged by comparing the shapes of CBF and CPF for the case of sampling size 10 and 100. As the sample size increases to 2,500, the uncertainty for the CBF and CPF tends to become stable. Beyond sample size of 5,000, the shapes of CBF and CPF only change slightly. To obtain fully converged results, sample size of 10,000 are used in this research for all three outputs: heating, cooling, and carbon emissions.

Figure 5 shows the uncertainty results for annual heating energy of the office building based on the Dempster-Shafer theory with a sample size 10,000. The shapes of cumulative belief function (CBF) and cumulative plausibility functions (CPF) from the Cubist and MARS models are similar for case A although the predicted values from the MARS model are slightly larger than those from the Cubist model. The same conclusion can be also obtained for the case B as illustrated in Figure 5c and Figure 5d. Hence, both the Cubist or MARS models can produce reliable results instead of using the computationally expensive EnergyPlus models.



(a) Cubist model for case A



(b) MARS model for case A

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

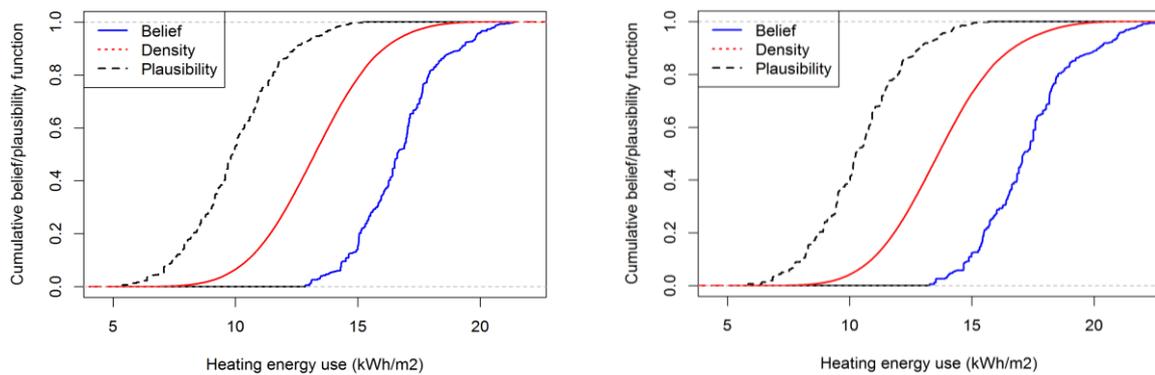
496

497

498

499

500



(c) Cubist model for case B (d) MARS model for case B

Figure 5. Uncertainty analysis of annual heating energy from Dempster-Shafer theory

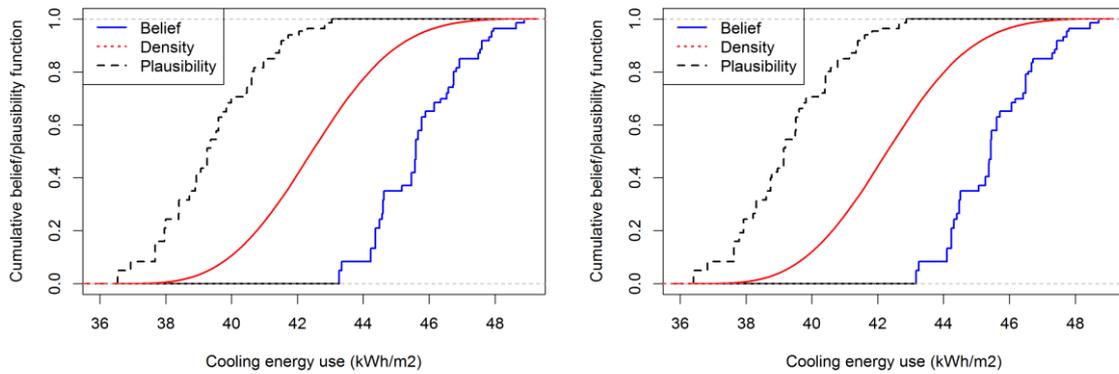
501
502
503

504 Based on the DST analysis, annual heating energy in this office building should be within
505 the ranges of solid blue (CBF) and dashed black (CPF) lines in Figure 5. As might be
506 expected, the red density plot (cumulative density function, CDF) falls between the CBF and
507 the CPF associated with the evidence space as defined in Table 1 and Table 2. This is because
508 the CDF with uniform distribution (as described in section 3) is a special case of the DST
509 results. For this office building, annual heating energy is unlikely to be more than 23 kWh/m²
510 and less than 5 kWh/m². If the annual heating energy quota for office buildings (i.e. upper
511 limit of energy use recommended or required by government) is 20 kWh/m², then the lowest
512 and highest probabilities above this quota in case A with the Cubist model are 0 and 12%,
513 respectively, as shown in Figure 5a. If using the MARS model, the corresponding lowest
514 probability is the same as the Cubist model (0%) and the highest probability is slightly higher,
515 around 11 %.

516 As also can be seen from Figure 5, the CBF and CPF are more smooth for case B than for
517 case A. This is in line with the input uncertainty defined in Table 2. A large number of
518 hypercube for input variables in the DST approach means more discretization and smaller
519 intervals, which usually results in smoother output. The number of hypercubes in case B is
520 512, whereas the number of hypercube in case A is only 32 as described in section 2.1. As a
521 result, a smoother CBF and CPF is observed for case B in Figure 5.

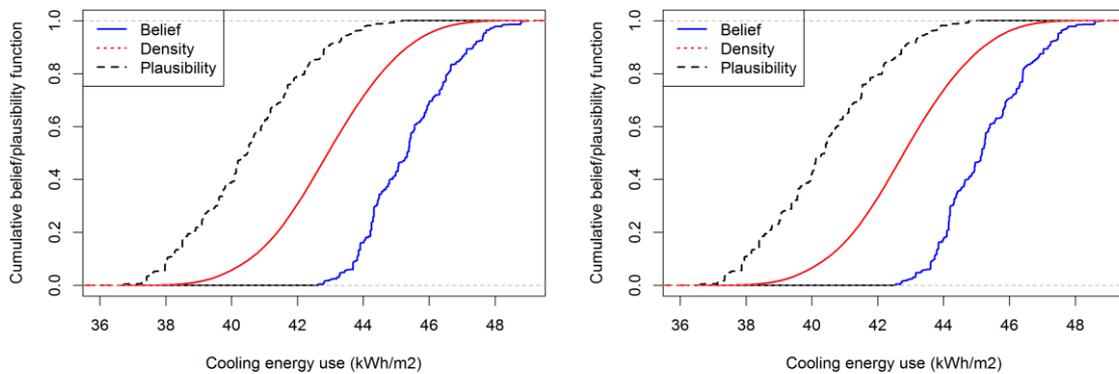
522 4.2.2 Annual cooling energy

523 Figure 6 shows the uncertainty results for cooling energy using the Cubist and bagging
524 MARS models in case A and case B for the office building. The shapes of CBF and CPF are
525 similar from the Cubist and bagging MARS models in case A. This statement also holds true
526 in the case B. Hence, the uncertainty of output is reliable using machine learning models
527 instead of engineering-based EnergyPlus models.



528
529

(a) Cubist model for case A (b) Baging MARS model for case A



530
531

(c) Cubist model for case B (d) Baging MARS model for case B

532

Figure 6. Uncertainty analysis of annual cooling energy from Dempster-Shafer theory

533

534

535

536

537

538

539

540

The CDF results lie between the CBF and CPF since the uniform distribution used for the CDF is one of possible choices in evidence space of input variables using the DST method. Cooling energy is between 36 and 49 kWh/m² in this case study. If the quota of annual cooling energy is 46 kWh/m² for the office building, then the highest probability for the cooling energy above this quota value is around 38% using two machine learning models (Figure 6a and Figure 6b). Compared to the case A, there are less jumps for the CBF and CPF plots in the case B. As discussed in section 4.2.1, this is due to the increase of hypercube number of inputs that leads to smoother outputs.

541

4.2.3 Annual carbon emissions

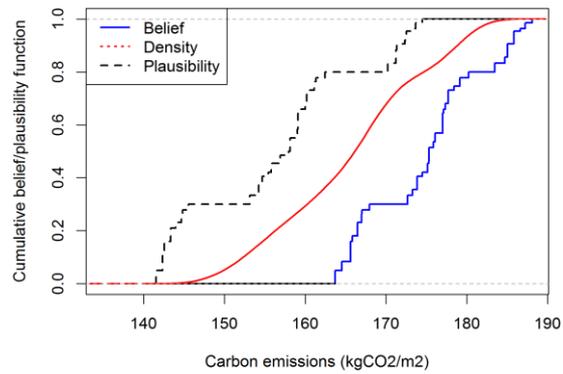
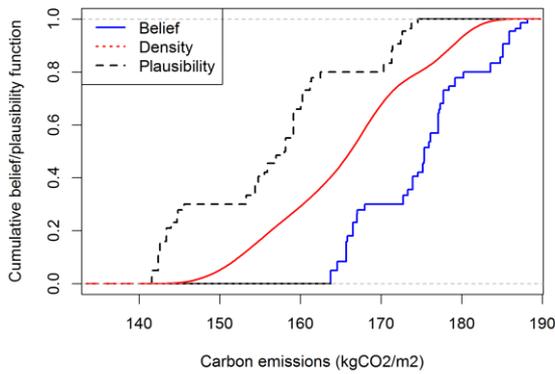
542

543

544

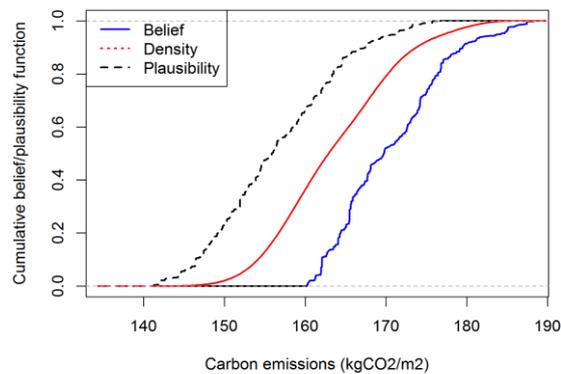
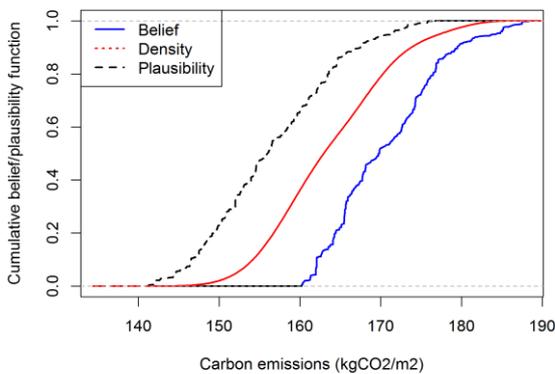
545

Figure 7 shows uncertainty results of annual carbon emissions using the Cubist and linear models in case A and case B. The results from these two machine learning models are very close for both cases. Hence, the results are robust for showing the variation of carbon emissions using the fast-computing learning models.



546
547

(a) Cubist model for case A (b) Linear model for case A



548
549

(c) Cubist model for case B (d) Linear model for case B

550
551

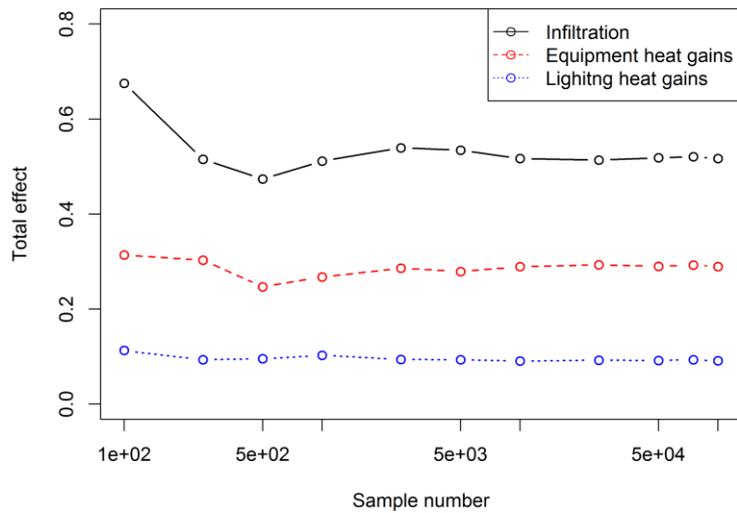
Figure 7. Uncertainty analysis of annual carbon emissions based on Dempster-Shafer theory

552
553
554
555
556
557
558
559
560
561
562
563

It is apparent that in case A there are more obvious jumps in CBF and CPF for carbon emissions in comparison with heating and cooling energy use (Figure 5 and Figure 6). This can be explained by the sensitivity analysis as will be presented in section 4.3. The two dominant variables for carbon emissions are equipment and lighting heat gains that are listed in Table 1. As a result, the trends of carbon emissions are affected substantially by the specification of these two input variables. The equipment peak value has two discontinues intervals [13, 15] and [16, 17], which leads to significant jumps of CBF and CPF in Figure 7a and Figure 7b. The overlapping intervals from lighting peak values ([6, 8] and [7, 9]) also have important influence on carbon emissions. In case B (Table 2), the intervals for equipment and lighting equipment gains become more continuous by combining the opinions from two experts. Therefore, the CBF and CPF in case B become much smoother than those in case A.

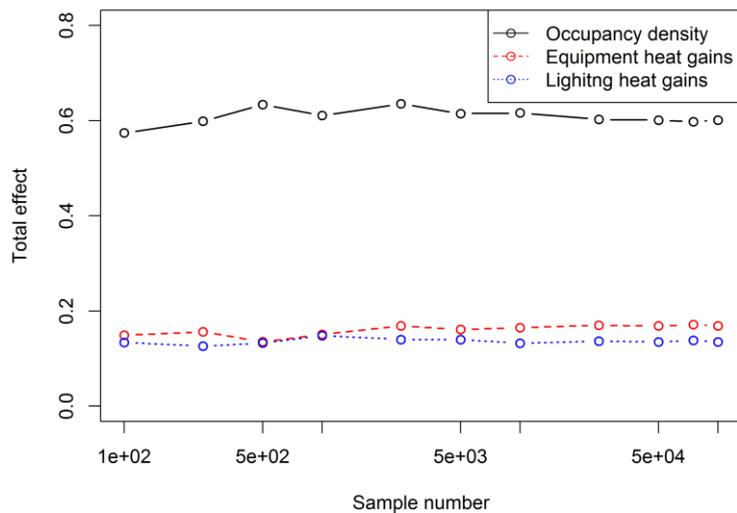
564
565
566
567
568
569
570

Based on the analysis in this subsection, the results from DST analysis are different from the results of assuming uniform distributions for input variables in building energy assessment. The uncertainty from uniform distributions is significantly less than possible variations for building energy performance. Hence, when there is only limited information on input variables, the uniform distributions cannot be regarded as good choices for uncertainty analysis in building energy analysis. Instead, the DST analysis should be implemented to properly estimate uncertainty of building performance.



572
573

(a) Heating energy with the MARS model



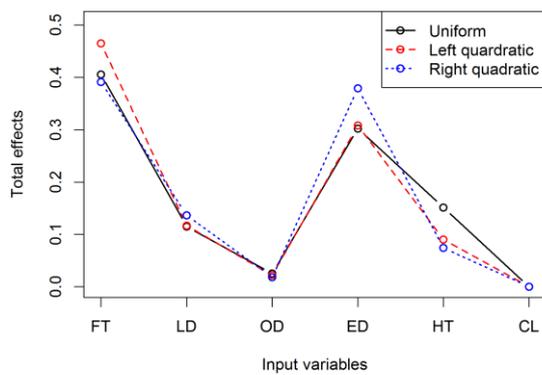
574
575

(b) Cooling energy with the Cubist model

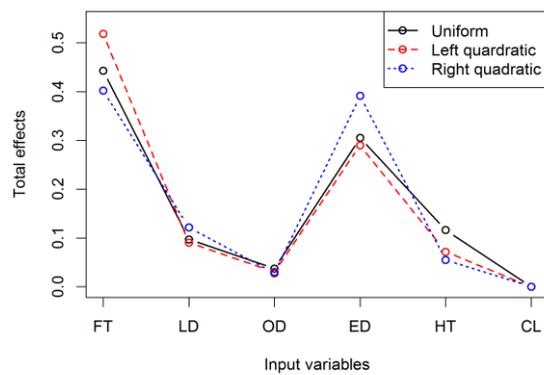
Figure 8. Stability of total effects as a function of sample size from global sensitivity analysis

576
577

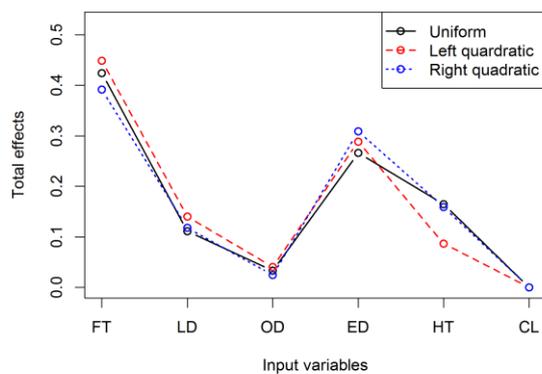
578 The two machine learning models with better predictive performance are used to provide
 579 robust results of sensitivity analysis for three performance measures: heating, cooling, and
 580 carbon emissions. The sensitivity analysis is implemented with a sample size of 100,000 from
 581 uncertain variables and the associated three distribution possibilities (uniform, left quadratic,
 582 and right quadratic) as discussed in section 2.3. Figure 8 demonstrates the stability of total
 583 effects from the Sobol global sensitivity analysis as a function of sample size. The results
 584 vary a lot at the sample sized below 2,500. In Figure 8b, there are intersections of ranking
 585 importance for equipment and lighting heat gains for cooling energy use at the sample size
 586 less than 1,000. After the sample size of 10,000, the total effects become stable in both Figure
 587 8a and Figure 8b. The number of samples is chosen as 100,000 to confirm the convergence of
 588 sensitivity analysis.



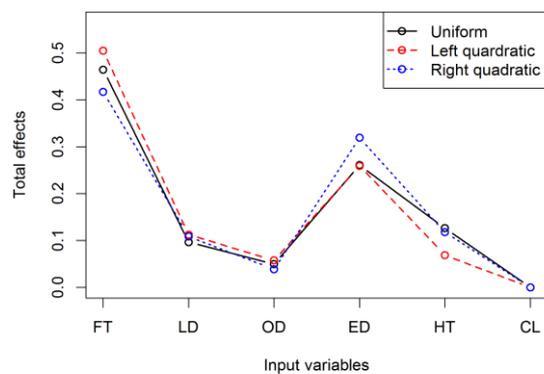
(a) Cubist model for case A



(b) MARS model for case A



(c) Cubist model for case B



(d) MARS model for case B

Figure 9. Results of sensitivity analysis for heating energy from Dempster-Shafer theory (refer to Table 1 for full names of input variables)

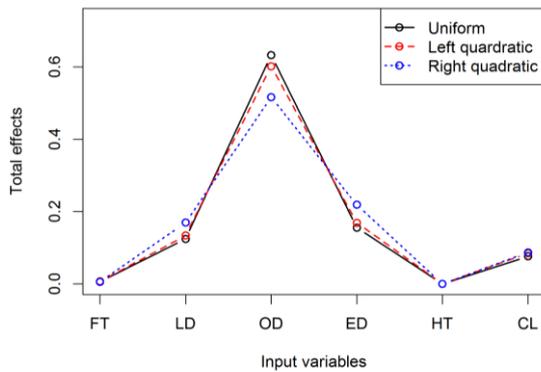
589
590

591
592
593
594

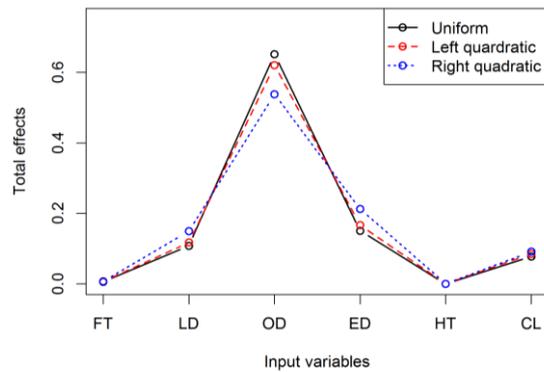
595 Figure 9 shows the total effects of six variables with three types of distributions using the
 596 global Sobol sensitivity analysis for annual heating energy use in the office building. The
 597 ranking results are similar for three distribution possibilities as can be seen from Figure 9a.
 598 The ranking order is also similar from two machine learning models (Cubist and MARS) as
 599 illustrated in Figure 9a and Figure 9b. The most important variable identified here is the
 600 infiltration rate (FT), which accounts for approximate 40% of output variation. Hence, it is
 601 necessary to obtain reliable information on infiltration rate for accurately predicting annual
 602 heating energy. From the perspective of energy saving, it is important to try to reduce
 603 infiltration rate in order to reduce heating energy use. The next important variable is the
 604 equipment heat gains (ED) that also has important influences on annual heating energy use.
 605 For the right quadratic distributions, the importance from equipment heating gains becomes
 606 more evident and its important is in the same level as infiltration rate in the office building.
 607 The lighting heat gains (LD) and heating set-point temperatures (HT) have medium effects on
 608 heating energy use. The occupancy density (OD) and cooling heat-point (CL) have almost no
 609 influence on output variable in this case study. There are apparent similarities for ranking
 610 results in case A and case B. Hence, the ranking results of sensitivity analysis are not
 611 influenced by opinions from different experts although the results of uncertainty analysis are
 612 quite different in two cases as discussed in section 4.2.

613 Figure 10 shows the results of sensitivity analysis for annual cooling energy from two
 614 machine learning models in two cases. The Cubist and bagging MARS models present the
 615 similar results for ranking the importance of six variables influencing cooling energy use. The

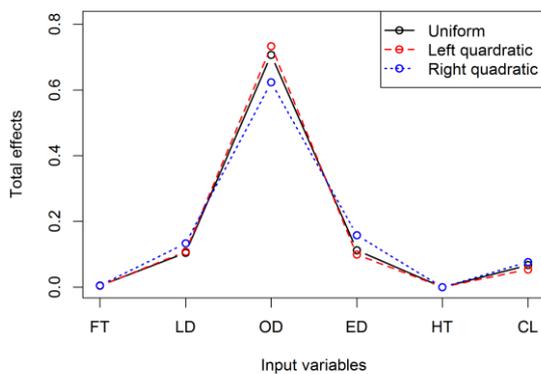
616 sensitivity rankings from case A are also similar to those from case B. In four subplots of
 617 Figure 10, the dominant variable is the occupancy density (OD), which accounts for around
 618 60% of variations of annual cooling energy in the office building. Then, it is necessary to
 619 obtain the reliable data on occupancy density in order to provide accurate estimation on
 620 cooling energy use in this building. The next three variables have similar importance,
 621 equipment heat gains (ED), lighting heat gains (LD), and cooling set-point temperatures (CL).
 622 The remaining two variables (infiltration rate and heating set-point temperatures) have almost
 623 no effect on output variable. As also can be seen from Figure 10, the sensitivity results from
 624 three distributions are similar for cooling energy use in the office building. Hence, the
 625 assumption of various distributions does not influence the validity of ranking results in this
 626 case study.



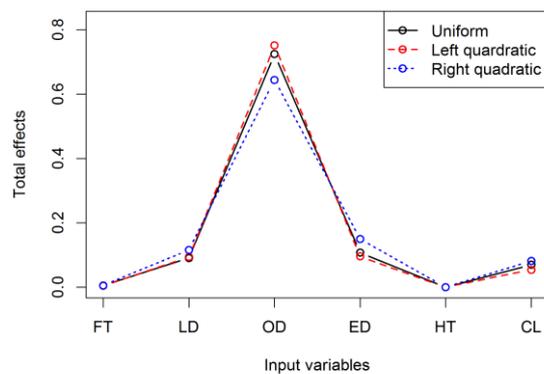
627 (a) Cubist model for case A



628 (b) Bagging MARS model for case A



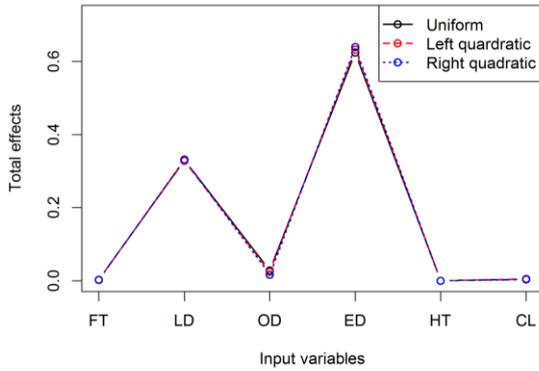
629 (c) Cubist model for case B



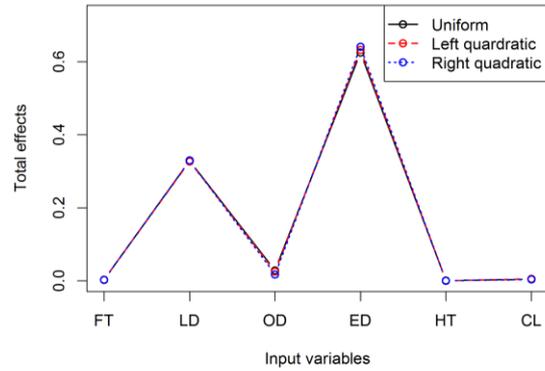
630 (d) Bagging MARS model for case B

631 **Figure 10. Results of sensitivity analysis for cooling energy from Dempster-Shafer**
 632 **theory (refer to Table 1 for full names of input variables)**

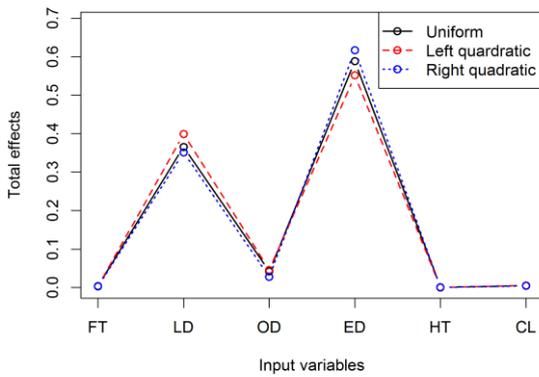
633 Figure 11 shows the ranking results from sensitivity analysis for annual carbon emissions
 634 from two machine learning models in two cases for this office building. As might be
 635 expected, equipment (ED) and lighting heat gains (LD) have significant influence on carbon
 636 emissions. This is because most of electricity use in office buildings is due to office
 637 equipment (such as computers, printing machine, projectors) and lighting. The variations of
 638 carbon emissions are almost not influenced by four remaining variables. It is also observed
 639 that the ranking results from the Cubist and linear models are very similar, which indicates
 640 that the sensitivity results obtained from this study are robust. In two cases, the trend of
 641 important variables is similar although three types of distributions lead to more disperse
 642 results in case B than those in case A.



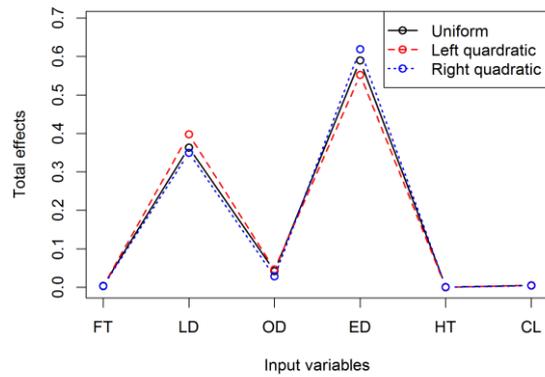
(a) Cubist model for case A



(b) Linear model for case A



(c) Cubist model for case B



(d) Linear model for case B

Figure 11. Results of sensitivity analysis for carbon emissions from Dempster-Shafer theory (refer to Table 1 for full names of input variables)

643
644
645
646
647
648
649

650 Sensitivity analysis applied in this subsection would be useful to make informed
651 decisions, depending on project purposes. For instance, energy saving measures can be
652 determined even in the case of the availability of limited information by using DST sensitivity
653 analysis since a spectrum of distributions for these input factors have been considered. If the
654 aim of project is to reduce variations of energy performance, then more efforts should be
655 made to collect more information on these key variables identified by the DST sensitivity
656 method.

657 5 CONCLUSIONS

658 This research implements uncertainty and sensitivity analysis for assessing energy
659 performance of an office building based on the DST (Dempster-Shafer theory) approach.
660 Machine learning methods are used to expedite the computation since a large number of
661 energy models needs to be run. The following conclusions can be drawn from this study.

662 (1) The DST analysis is applicable to provide informative uncertainty results of energy
663 performance in buildings when only limited information on input variables is available. Note
664 that the uncertainty results for energy use bounded between the CBF and CPF cannot be
665 interpreted as being equally possible (similar to uniform distributions in probability theory).
666 The energy performance may be any possible values within the intervals between the CBF
667 and CPF based on the DST analysis. When the information on uncertainty of inputs is
668 sufficient to specify distributions, the Monte-Carlo based sampling method is preferred in
669 building energy assessment.

670 (2) Machine learning algorithms can be used to reduce high computational cost in
671 implementing DST analysis in building energy analysis, instead of directly using engineering-
672 based energy models. It is recommended to compare several machine learning methods and
673 then choose at least two learning methods that are inherently different in nature in order to
674 provide robust analysis.

675 (3) The DST analysis does not assume any distribution within the intervals of input
676 factors. Hence, a spectrum of distributions should be used in implementing sampling-based
677 sensitivity analysis to provide reliable sensitivity results for building energy analysis.

678 (4) It is necessary to assess the stability of results as a function of sample size from
679 uncertainty and sensitivity analysis in applying the DST into building energy assessment. This
680 is often ignored in applying uncertainty and sensitivity analysis in the field of building energy
681 analysis.

682 The conclusions obtained above and the methods proposed in this paper can be applied to
683 other buildings in various climate zones. However, there are several issues that still need to be
684 addressed in further studies. One is to compare the performance of more machine learning
685 algorithms (such as deep learning and ensemble learning methods) when applying them in
686 DST analysis of building energy performance. Another issue is to investigate whether it is
687 possible to determine the rule-of-thumb sample size (for instance, in terms of variable
688 number) on applying the DST method in building energy analysis in order to make this
689 method more readily available.

690 ACKNOWLEDGEMENT

691 This research was supported by the National Natural Science Foundation of China (No.
692 51778416) and the Key Projects of Philosophy and Social Sciences Research, Ministry of Ed-
693 ucation (China) "Research on Green Design in Sustainable Development" (contract No.
694 16JZDH014, approval No. 16JZD014).

696 REFERENCES

- 697
698 [1] S. Kiluk. Diagnostic information system dynamics in the evaluation of machine learning
699 algorithms for the supervision of energy efficiency of districtheating-supplied buildings.
700 Energy Conversion and Management. 150 (2017) 904-13.
- 701 [2] M. Heidarinejad, J.G. Cedeño-Laurent, J.R. Wentz, N.M. Rekstad, J.D. Spengler, J.
702 Srebric. Actual building energy use patterns and their implications for predictive modeling.
703 Energy Conversion and Management. 144 (2017) 164-80.
- 704 [3] W. Tian, X. Meng, B. Yin, Y. Sun, Y. Liu, X. Fu. Design of Robust Green Buildings
705 Using a Non-probabilistic Uncertainty Analysis Method. Procedia Engineering. 205 (2017)
706 1049-55.
- 707 [4] F.R. Cecconi, M. Manfren, L.C. Tagliabue, A.L.C. Ciribini, E. De Angelis. Probabilistic
708 behavioral modeling in building performance simulation: A Monte Carlo approach. Energy
709 and Buildings. 148 (2017) 128-41.
- 710 [5] R. Choudhary, W. Tian. Influence of district features on energy consumption in non-
711 domestic buildings. Building Research & Information. 42 (2014) 32-46.
- 712 [6] J. Chen, G. Augenbroe, X. Song. Evaluating the potential of hybrid ventilation for small to
713 medium sized office buildings with different intelligent controls and uncertainties in US
714 climates. Energy and Buildings. 158 (2018) 1648-61.
- 715 [7] W. Tian, X. Fu, Y. Sun, B. Yin, X. Meng, Y. Liu. Sustainable Building Design Based on
716 the Second Order Probability Approach. Procedia Engineering. 205 (2017) 1056-63.
- 717 [8] Y. Heo, G. Augenbroe, D. Graziano, R.T. Muehleisen, L. Guzowski. Scalable
718 methodology for large scale building energy improvement: Relevance of calibration in model-
719 based retrofit analysis. Building and Environment. 87 (2015) 342-50.

- 720 [9] L. Urbanucci, D. Testi. Optimal integrated sizing and operation of a CHP system with
721 Monte Carlo risk analysis for long-term uncertainty in energy demands. *Energy Conversion*
722 *and Management*. 157 (2018) 307-16.
- 723 [10] W. Tian, S. Yang, J. Zuo, Z. Li, Y. Liu. Relationship between built form and energy
724 performance of office buildings in a severe cold Chinese region. *Building simulation*. 10
725 (2017) 11-24.
- 726 [11] G.A. Faggianelli, L. Mora, R. Merheb. Uncertainty quantification for Energy Savings
727 Performance Contracting: Application to an office building. *Energy and Buildings*. 152
728 (2017) 61-72.
- 729 [12] C.J. Hopfe, J.L.M. Hensen. Uncertainty analysis in building performance simulation for
730 design support. *Energy and Buildings*. 43 (2011) 2798-805.
- 731 [13] W. Tian, Y. Heo, P. de Wilde, Z. Li, D. Yan, C.S. Park, et al. A review of uncertainty
732 analysis in building energy assessment. *Renewable and Sustainable Energy Reviews*. 93
733 (2018) 285-301.
- 734 [14] J.C. Helton, J.D. Johnson, W.L. Oberkampf, C.B. Storlie. A sampling-based
735 computational strategy for the representation of epistemic uncertainty in model predictions
736 with evidence theory. *Computer Methods in Applied Mechanics and Engineering*. 196 (2007)
737 3980-98.
- 738 [15] C. Xie, G. Li, F. Wei. An integrated QMU approach to structural reliability assessment
739 based on evidence theory and kriging model with adaptive sampling. *Reliability Engineering*
740 *& System Safety*. 171 (2018) 112-22.
- 741 [16] P. Masoudi, T. Aifa, H. Memarian, B. Tokhmechi. Uncertainty assessment of volumes of
742 investigation to enhance the vertical resolution of well-logs. *Journal of Petroleum Science and*
743 *Engineering*. 154 (2017) 252-76.
- 744 [17] X. Zhao, K. Mu, F. Hui, C. Prehofer. A cooperative vehicle-infrastructure based urban
745 driving environment perception method using a D-S theory-based credibility map. *Optik -*
746 *International Journal for Light and Electron Optics*. 138 (2017) 407-15.
- 747 [18] T.-J. Park, J.-H. Chang. Dempster-Shafer theory for enhanced statistical model-based
748 voice activity detection. *Computer Speech & Language*. 47 (2018) 47-58.
- 749 [19] J. Chaney, E. Hugh Owens, A.D. Peacock. An evidence based approach to determining
750 residential occupancy and its role in demand response management. *Energy and Buildings*.
751 125 (2016) 254-66.
- 752 [20] Y. Kim, K. Ahn, C. Park. Aggregated epistemic uncertainty for building energy
753 prediction using Dempster-Shafer evidence theory. *Proceedings of 9th IAQVEC (Indoor Air*
754 *Quality Ventilation & Energy Conservation In Buildings) Conference October 23-26, 2016,*
755 *Incheon, Korea. , 2016.*
- 756 [21] DOE. EnergyPlus V8.8, September 2017, Department of Energy, USA. 2017.
- 757 [22] Y. Kwak, J.-H. Huh. Development of a method of real-time building energy simulation
758 for efficient predictive control. *Energy Conversion and Management*. 113 (2016) 220-9.
- 759 [23] W. Oberkampf, J.C. Helton. Chapter 10: Evidence Theory for Engineering Applications.
760 in: E. Nikolaidis, D.M. Ghiocel, S. Singhal, (Eds.), *Engineering Design Reliability Handbook*.
761 *CRC Press*2004. pp. 269-85.
- 762 [24] H.-R. Bae, R.V. Grandhi, R.A. Canfield. Epistemic uncertainty quantification techniques
763 including evidence theory for large-scale structures. *Computers & Structures*. 82 (2004) 1101-
764 12.
- 765 [25] W. Tian, P.D. Wilde. Impact of global warming on thermal performance of domestic
766 buildings using probabilistic climate data. *International Journal of Global Warming*. 10
767 (2016) 514-35.
- 768 [26] J.C. Helton, J.D. Johnson, C. Sallaberry, C.B. Storlie. Survey of sampling-based methods
769 for uncertainty and sensitivity analysis. *Reliability Engineering & System Safety*. 91 (2006)
770 1175-209.

771 [27] W. Tian, J. Song, Z. Li, P. de Wilde. Bootstrap techniques for sensitivity analysis and
772 model selection in building thermal performance analysis. *Applied Energy*. 135 (2014) 320-8.
773 [28] S. Heidenreich, H. Gross, M. Bär, L. Wright. Uncertainty propagation in computationally
774 expensive models: A survey of sampling methods and application to scatterometry.
775 *Measurement*. 97 (2017) 79-87.
776 [29] M. Kuhn. Building predictive models in R using the caret package. *Journal of Statistical*
777 *Software*. 28 (2008) 1-26.
778 [30] L. Wei, W. Tian, E.A. Silva, R. Choudhary, Q. Meng, S. Yang. Comparative Study on
779 Machine Learning for Urban Building Energy Analysis. *Procedia Engineering*. 121 (2015)
780 285-92.
781 [31] W. Tian, P. de Wilde. Uncertainty and sensitivity analysis of building performance using
782 probabilistic climate projections: A UK case study. *Automation in Construction*. 20 (2011)
783 1096-109.
784 [32] W. Tian, S. Yang, Z. Li, S. Wei, W. Pan, Y. Liu. Identifying informative energy data in
785 Bayesian calibration of building energy models. *Energy and Buildings*. 119 (2016) 363-76.
786 [33] M. Kuhn, K. Johnson. *Applied predictive modeling*. Springer2013.
787 [34] T. Hastie, R. Tibshirani, J. Friedman. *The Elements of Statistical Learning: Data Mining,*
788 *Inference, and Prediction*. Springer2009.
789 [35] J. Song, L. Wei, Y. Sun, W. Tian. Implementation of Meta-modelling for Sensitivity
790 Analysis in Building Energy Analysis. The eSim 2014 conference, , May 7 to 10, Ottawa,
791 Canada, 2014.
792 [36] W. Tian, R. Choudhary, G. Augenbroe, S.H. Lee. Importance analysis and meta-model
793 construction with correlated variables in evaluation of thermal performance of campus
794 buildings. *Building and Environment*. 92 (2015) 61-74.
795 [37] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, et al. *Global*
796 *sensitivity analysis: the primer*. Wiley-Interscience2008.
797 [38] J.C. Helton, J.D. Johnson, W.L. Oberkampf, C.J. Sallaberry. Sensitivity analysis in
798 conjunction with evidence theory representations of epistemic uncertainty. *Reliability*
799 *Engineering & System Safety*. 91 (2006) 1414-34.
800 [39] A. Saltelli. Making best use of model evaluations to compute sensitivity indices.
801 *Computer Physics Communications*. 145 (2002) 280-97.
802 [40] B.I. Gilles Pujol, Alexandre Janon. R package sensitivity V1.12.2: sensitivity: Global
803 Sensitivity Analysis of Model Outputs. <https://CRAN.R-project.org/package=sensitivity>
804 accessed on 2016-11-20. 2016.
805 [41] MOC. GB50189-2015. Design standard for energy efficiency of public buildings,
806 Ministry of Construction (MOC) of P.R.China China Planning Press (2015) (in Chinese)2015.
807 [42] MOC. GB50189-2005. Energy Conservation Design Regulation for Public Buildings.
808 Ministry of Construction (MOC) of P.R.China China Planning Press (2005) (in Chinese)2005.
809 [43] NDRC. China regional grid based line emission factor in 2016 (in Chinese), by NDRC
810 (National Development and Reform Commission). 2017.
811 [44] S. Yang, W. Tian, L. Wei, Z. Wang. Uncertainty analysis of carbon emissions in food
812 industry. *Food Research and Development*. 37 (2016) 215-8, in Chinese.
813 [45] K. Sentz, S. Ferson. Combination of evidence in Dempster-Shafer Theory, Report No.
814 SAND2002-0835, Sandia National Laboratories. 2002.
815 [46] W. Jiang, J. Zhan. A modified combination rule in generalized evidence theory. *Applied*
816 *Intelligence*. 46 (2017) 630-40.
817 [47] ASHRAE. Guideline 14-2002 Measurement of Energy and Demand Savings. ASHRAE,
818 Atlanta, GA. (2002).
819 [48] D. Coakley, P. Raftery, M. Keane. A review of methods to match building energy
820 simulation models to measured data. *Renewable and Sustainable Energy Reviews*. 37 (2014)
821 123-41.

- 822 [49] P. de Wilde. Building performance analysis. . Wiley-Blackwell. 2018.
- 823 [50] INCOSE. INCOSE (International Council on Systems Engineering) Systems
- 824 Engineering Handbook: A Guide for System Life Cycle Processes and Activities. . Wiley.
- 825 2015.
- 826 [51] T. Gilb. Competitive engineering: a handbook for systems engineering, requirements
- 827 engineering, and software engineering using planguage. . Butterworth-Heinemann. 2005.
- 828