

# *A comparative analysis of building energy estimation methods in the context of demand response*

Article

Accepted Version

Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Curtis, M., Torriti, J. ORCID: <https://orcid.org/0000-0003-0569-039X> and Smith, S. T. ORCID: <https://orcid.org/0000-0002-5053-4639> (2018) A comparative analysis of building energy estimation methods in the context of demand response. *Energy and Buildings*, 174. pp. 13-25. ISSN 0378-7788 doi: <https://doi.org/10.1016/j.enbuild.2018.06.004> Available at <https://centaur.reading.ac.uk/78070/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <https://www.sciencedirect.com/science/article/pii/S0378778817336393>

To link to this article DOI: <http://dx.doi.org/10.1016/j.enbuild.2018.06.004>

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

1 **Title:** A Comparative Analysis of Building Energy Estimation Methods in the Context of Demand  
2 Response

3  
4 **Target Journal:** Energy and Buildings

5  
6 **Authors:** Mitchell Curtis <sup>1, \*</sup>, Jacopo Torriti <sup>2</sup>, Stefan Thor Smith <sup>2</sup>

7 <sup>1</sup>Technologies for Sustainable Built Environments Centre, University of Reading, Reading, UK

8 <sup>2</sup>School of the Built Environment, University of Reading, Reading, UK

9 \* Corresponding author at: Technologies for Sustainable Built Environments Centre, University of Reading,  
10 Reading, UK. E-mail [m.r.curtis@pgr.reading.ac.uk](mailto:m.r.curtis@pgr.reading.ac.uk)

11

12

13 **Abstract:**

14 A critical element of assessing a building's suitability for Demand Side Response (DSR) is understanding  
15 its turndown potential to ensure that DSR participation will be financially viable. While research has  
16 been undertaken on site level DSR estimation methods, there is currently no research that compares  
17 the outcomes of these methods. This paper compares four non-domestic energy estimation methods  
18 used for understanding the DSR potential of electrical appliances in a building to provide insights about  
19 uncertainty levels based on input requirements. Each method is deployed to estimate the DSR  
20 potential of HVAC chiller assets at two UK hotels over two years. The results show the methods have a  
21 range of error levels from the highest Mean Average Percentage Error (MAPE) of 159% to the lowest  
22 MAPE of 39%. The input requirements followed a general trend of more complex informational inputs  
23 resulting in lower error values. The outcomes of this research enable users to make informed decisions  
24 in selecting DSR estimation methods based on information availability and acceptable estimation error  
25 levels.

26

27

28 **Highlights**

- 29
- Four DSR estimation methods were evaluated using empirical data from two hotels
  - The DSR estimation methods were found to have a wide range of error levels
  - The comparisons of methods allows for informed selection of a DSR estimation method based on available input information
- 30  
31  
32

33

34

35 **Keywords:**

- Demand Side Response
  - Estimation
  - Comparative Analysis
  - Method Review
  - Electricity demand
- 36  
37  
38  
39  
40

## 41 **1 Introduction**

42 Demand Side Response (DSR) programmes generally require a detailed understanding of the turndown  
43 potential of participating buildings to accurately forecast DSR capacity for electricity system balancing.  
44 This detail is needed as DSR programmes will apply penalties if contracted levels of turndown are  
45 missed. As an example, the UK Short Term Operation Reserve (STOR) programme requires participants  
46 to provide a guaranteed turndown kW amount for set periods of time of up to 14 hours per day  
47 (National Grid, 2016). If STOR participants underdeliver by more than 5%, then financial penalties are  
48 applied and progressively increased with the potential for ultimately removing non-performing  
49 participants from the programme if they fail in meeting guaranteed turn down levels too many times.  
50 The severity of penalties will vary by country and DSR programme. For example, the American San  
51 Diego Gas & Electric programme has a low severity based on payments being reduced proportionally  
52 to the contracted amount delivered (SDG&E, 2016). Whereas the Spanish programme is very strict  
53 with exclusion if the site fails to meet their obligations twice (SEDC, 2017). This means that correctly  
54 determining the long-term DSR potential of a building is important for appraising its suitability for DSR.

55  
56 As DSR aggregators play a key role in provide 80% of DSR capacity (SEDC, 2017), this research has  
57 focused on the estimation methods aggregators apply when determining the turndown potential and  
58 suitability of buildings for DSR. DSR aggregators operate by combining small flexible loads from  
59 multiple buildings into a virtual single load and take responsibility for managing the DSR process.  
60 Research into how aggregators decide if a building is suitable shows that the key assessment tasks  
61 focus on determining the long-term DSR potential of a buildings' assets (Curtis, 2017). Therefore, the  
62 ability to correctly analyse a building's DSR potential is a critical element of an aggregator's business  
63 process. This is expressly important when dealing with small to medium enterprises with smaller  
64 overall levels of DSR potential as the ability to lowering the contracted amount of DSR to avoid  
65 penalties due to estimation uncertainty is limited. While an aggregator can perform building surveys to  
66 gain a detailed understanding of a building's DSR potential, surveys have a time and cost impact and  
67 therefore are normally only undertaken once an initial desktop assessment has been completed.  
68 However, performing a desktop assessment to determine a building's potential is often difficult as  
69 detailed usage information (from sub-meters for example) about the individual electrical assets that  
70 are being assessed for DSR is normally unavailable (Merry, 2017). Instead, the only information  
71 normally available is the building's overall electricity usage as recorded at half hourly (UK standard  
72 practice) or similar intervals by the building's utility supplier. Half hourly information will provide a  
73 usage profile that can be used for estimating the building's DSR potential and suitability if all electricity  
74 demand from the grid is reduced by either using backup generators or turning off all assets. For  
75 buildings that can only turndown a limited subset of assets for DSR, a building level profile is unable to  
76 provide the individual assets' usage patterns needed to understand their suitability for DSR. To gain  
77 this necessary level of detail requires additional analysis to try and determine what proportion of the  
78 building's usage is represented by individual assets.

79  
80 Research on understanding energy usage in buildings is extensive, with a review by Borgstein,  
81 Lamberts, & Hensen (2016) identified five categories (Engineering calculations, Simulation, Statistical,  
82 Machine Learning, and Other) that each contained multiple approaches. However, research on  
83 application of these approaches for DSR estimation is limited and is influenced by whether the  
84 estimation is for implicit or instead explicit DSR (SEDC, 2016). Implicit DSR covers price-based  
85 measures, whereby demand might be reduced based on users responding to electricity price signals  
86 (for example, temporarily high electricity prices that encourages reduction in usage to reduce costs).

87 Explicit DSR covers incentive-driven measures, whereby demand reduction is specifically requested  
88 based on an external signal (for example, demand is reduced temporarily based on a site or its  
89 appliance receiving a signal in return for financial compensation for participation). As implicit DSR  
90 relies upon optional participation, research into estimating reduction potential focuses on how groups  
91 with similar DSR assets behave in response to different pricing signals, for the purposes of gaining an  
92 understanding of their combined potential. This is illustrated in research by Shen et al. (2016) where a  
93 genetic algorithm is used to estimate the DSR potential for a group of buildings based on time of use  
94 and dynamic pricing signals. The authors showed that if each building responds independently to  
95 pricing signals, then this can cause higher peak demand usage and therefore recommended that  
96 responses are coordinated across similar groups of buildings to achieve the desired peak reduction.  
97 Similarly, Chassin & Rondeau (2016) utilised the Random Utility Model to understand the potential  
98 provided by groups of fast-acting demand response loads under real-time pricing. In contrast to  
99 implicit DSR's reliance upon optional participation, explicit DSR participation is established by contract  
100 and the application of penalties where sites fail to respond to a specific reduction request or do not  
101 deliver pre-agreed levels of usage reduction. This means that estimation methods for explicit DSR  
102 focus on assessing the likely long-term potential of specific buildings to ensure that contractual  
103 commitments can be met. As 80% of DSR is currently provided by aggregators, who rely upon explicit  
104 DSR, this paper focuses on comparing only energy estimation methods used for explicit DSR (SEDC,  
105 2017).

106  
107 The majority of contributions to the research field of explicit DSR have originated from the Lawrence  
108 Berkeley National Laboratory - Demand Response Research Center (DRRC, 2017). Their research into  
109 DSR has covered several areas including methods for assessing the DSR potential of buildings. To help  
110 improve the assessment process the DRRC developed the Demand Response Quick Assessment Tool  
111 (DRQAT) (Yin & Black, 2015) which uses the EnergyPlus whole building energy simulation program (U.S.  
112 Department of Energy, 2017) to predict DSR potential using predefined building models and a limited  
113 set of user selectable variables. While the DRQAT program helps to make the assessment process  
114 easier, it introduces other limitations, notably that it will only work for predefined building models  
115 which are currently offices and retail buildings based in California. They also recognise that are still  
116 many input uncertainties like operational behaviour and space loads that are hard to capture in the  
117 DRQAT model. To overcome these uncertainties, they use metrics of peak demand (kW), absolute  
118 demand savings (kW), and relative demand savings (%) to compensate for differences in actual and  
119 forecasted load shapes. The DRRC have also looked into understanding the predictors that influence  
120 how well a building will perform when enabled for DSR (Mathieu et al., 2010; Piette et al., 2011). This  
121 research showed that the level of turndown potential could be linked to temperature if the DSR assets  
122 demonstrate varying levels of usage based on external weather conditions with prediction uncertainty  
123 being approximated using the standard error. The limitation of using this approach for assessing a  
124 building is the need for the building to have already been involved in DSR to have access to event  
125 outcomes for analysis. Another assessment approach proposed by Panapakidis et al. (2014) is to  
126 cluster electricity usage of a building into profiles that can then be used to ascertain DSR turndown  
127 opportunities based on variance between the profiles. They try to reduce uncertainty by testing a  
128 range of cluster lengths to find the optimal number to use that minimises the overall sum of squared  
129 errors. This method has the advantage of only needing the building's overall electricity usage records,

130 yet is limited by the assumptions required when deciding what the differences between profiles mean  
131 in terms of individual asset usage. There are other proprietary commercially developed analysis  
132 methods that have not been published. One such method has been provided by an aggregator in  
133 association with this research. They have two approaches when performing building asset assessment  
134 for DSR. The first approach assumes that the asset will work at a set level all year. To help reduce the  
135 uncertainty of this estimation a second approach is used that analyses the building's overall electricity  
136 records for a year to create a baseload usage amount for 95% of the time. The aggregator then takes a  
137 proportion of this 95% to represent the asset usage. Using the baseload value reduces uncertainty by  
138 knowing that at least this amount of electricity is being used 95% of the time and therefore taking a  
139 proportion of it prevents over estimating the assets potential usage. The major limitation of both  
140 approaches is the assumed consistent usage of the asset across the year, which they recognise, but  
141 still use the method to provide an initial understanding of anticipated potential before deciding on  
142 further investigations.

143

144 The issue that faces aggregators and anyone trying to perform DSR estimations using these methods is  
145 knowing which one to use and how they compare in terms of uncertainty and cost to undertake.

146 Therefore, the aim of this paper is to provide an understanding of uncertainty levels in current non-  
147 domestic DSR potential estimation methods based on the input requirements. By understanding the  
148 uncertainty levels and costs of DSR estimation methods this research hopes to increase usage of DSR  
149 from businesses that are currently excluded due to risk aversion resulting from not knowing the level  
150 of estimation uncertainty. The research is undertaken by examining and applying four DSR estimation  
151 methods to two UK hotels as described in Section 2. Section 3 sets out the research results and  
152 discusses these findings. Section 4 concludes by summarising the implications of this research.

153

## 154 **2 Methods**

155 Four DSR potential estimation methods were applied to two medium-sized UK hotels (~200 rooms) to  
156 evaluate outcome uncertainty against the level of information required for estimation. The four  
157 methods are: asset assessment; baseline comparison; historical event analysis; and building energy  
158 modelling. Figure 1 provides an overview of the explicit DSR estimation methods reviewed in this  
159 paper, including the primary data and parameter inputs and the analytical approaches used. The  
160 methods are to be used as part of an initial desktop assessment to determine the potential DSR of a  
161 building or business. The assessment provides a decision on whether further assessment or inclusion  
162 of the business in a DSR aggregation programme is valid. All methods estimate the half-hourly kW  
163 usage profile of electrical assets over a one-year period to assess if sufficient DSR potential exists. To  
164 explain how the methods were used and compared this section is divided into seven subsections. The  
165 first section describes the comparison of estimation method outcomes, followed by four sections  
166 describing the input requirements and calculation steps for each estimation method. Section six  
167 describes the sensitivity analysis approach used to highlight the influence of input parameter  
168 uncertainties on method outcomes. Finally, section seven describes the approach used to calculate the  
169 cost of using each method.

170

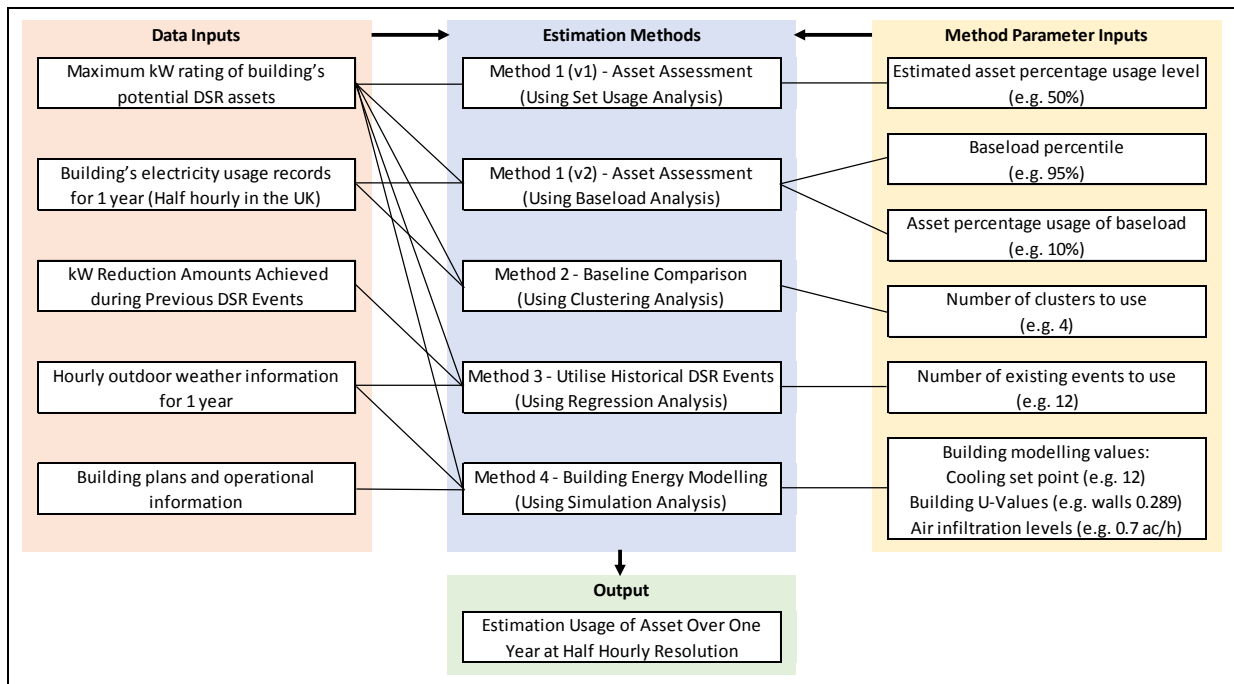


Figure 1 – Overview of DSR Estimation Methods

171  
172  
173  
174

## 2.1 Comparison Approach

175 This comparison of estimation methods was undertaken by using each method to determine the usage  
 176 profiles of HVAC chillers located at two UK hotels. Chillers are large centralised assets that cool water  
 177 for distribution around each hotel's HVAC system to provide space cooling that were identified by  
 178 Curtis et al. (2018) as being suitable for DSR due to the flexibility they offer through their ability to be  
 179 temporally turned off without impacting end-users. The hotel chillers have a maximum rating of 333  
 180 kW for Hotel 1 and 290 kW for Hotel 2. The two hotels have been chosen due to having access to  
 181 detailed information about each building's overall electricity usage as well as high-quality sub-metered  
 182 electricity usage data for the chillers during the years 2013 and 2016 for Hotel 1 and 2015 and 2016 for  
 183 Hotel 2. The sub-metered data enables a direct comparison of the estimation method outcomes  
 184 against actual usage. While chillers are used as an example of an electrical appliance with DSR  
 185 potential in this paper, its purpose is not to assess the suitability of chillers for DSR. Instead, the aim  
 186 and focus of this research is to compare methods for estimating the potential levels of electricity usage  
 187 by assets with potential for explicit DSR programmes, of which chillers are only one example. The  
 188 resulting usage estimates for chillers, as a sample appliance, can then be used as an input for  
 189 determining the specific DSR potential of a building based on the appliances characteristics and  
 190 intended DSR programme requirements. The application of the estimate to a DSR programme is not  
 191 covered in this paper as this is dependent on the ability of an appliance to meet specific programme  
 192 requirements. Therefore, evaluation of the estimations is kept independent by using the Mean  
 193 Absolute Percentage Error (MAPE) and Mean Bias Error (MBE) methods.

194  
195  
196  
197  
198  
199

The MAPE values provide an overall indication of the level of difference between the actual and  
 predicted results while the MBE values indicate the direction of error with positive and negative  
 results indicating over estimation and under estimation respectively. These methods were selected as  
 De Gooijer & Hyndman (2006) define them as the most common measures to use for time series  
 evaluation as they provide an easy to understand percentile value to indicate the level of forecasting

200 error that can be used to compare uncertainty across the four estimation methods. They are also  
201 deemed suitable based on their general usage across the literature on DSR estimation methods (Aman  
202 et al., 2016; Larsen et al, 2015).

203

## 204 **2.2 DSR Estimation Method 1 - Asset Assessment**

205 The asset assessment method is based on a review of current estimation approaches undertaken at a  
206 UK DSR aggregator. This is the simplest of the four methods as it is based on using very limited  
207 information with two variations to the approach as follows (see Appendix A for detailed calculation  
208 steps):

- 209 • *Variation 1 – Minimum Information:* This approach uses only the maximum kW rating of the  
210 asset being assessed. The expected kW usage level of the asset across the year is calculated as  
211 a set percentage of the maximum rating. The set percentage can vary based on the assessor’s  
212 prior knowledge of the asset type and building.
- 213 • *Variation 2 – Utilise Baseload Calculation:* This approach uses the building’s overall electricity  
214 usage records over one year (in the UK this is provided in half-hourly intervals) to calculate its  
215 baseload usage. The baseload amount is calculated for each half-hourly period by taking all  
216 usage records for each period (i.e. 365 usage records for the 00:00 to 00:30 half-hour period),  
217 ordering the records by value, then finding the 5<sup>th</sup> percentile value. This provides half hour  
218 electricity usage values that the building will use at least 95% of the time over the year and is  
219 therefore classified as the baseload. The expected kW usage level of the asset across the year  
220 is then calculated by taking a percentage share of the baseload that is attributed to the asset  
221 to be used in DSR. Again, the percentage will be set according to prior knowledge of this type  
222 of asset and building.

223

## 224 **2.3 DSR Estimation Method 2 - Baseline Comparison**

225 The second estimation method utilises clustering techniques to identify DSR opportunities through  
226 comparison of each building’s different usage profiles over a year. This method works on the basis that  
227 a building has different usage profiles throughout the year, and once profile clusters are identified,  
228 representative profiles of each cluster can be used to ascertain DSR turndown opportunities based on  
229 variance between the profiles. Panapakidis et al. (2014) reviewed a selection of clustering methods for  
230 electricity load curve analysis of buildings and identified that the k-means method offers a balanced  
231 approach for finding appropriate clusters that would be suitable for understanding building energy  
232 efficiency opportunities, including for DSR. However, they did not actually provide specific DSR  
233 estimation outcomes for the test building. Research by Van Wijk et al. (1999) also looked into how to  
234 use clustering techniques to identify patterns and trends on multiple timescales (days, weeks,  
235 seasons). They found that using k-means and then associating the resulting clusters to the different  
236 timescales allowed for identification and exploration of usage profiles. Their technique succeeds in  
237 identifying weekend vs weekday profiles and other significant periods, such as holidays. These  
238 clustering techniques show that building energy usage normally follows a small set of similar profiles.  
239 By identifying these profiles, it is then possible to understand different usage levels, which can then  
240 potentially be used to derive DSR estimations based on the business type.

241



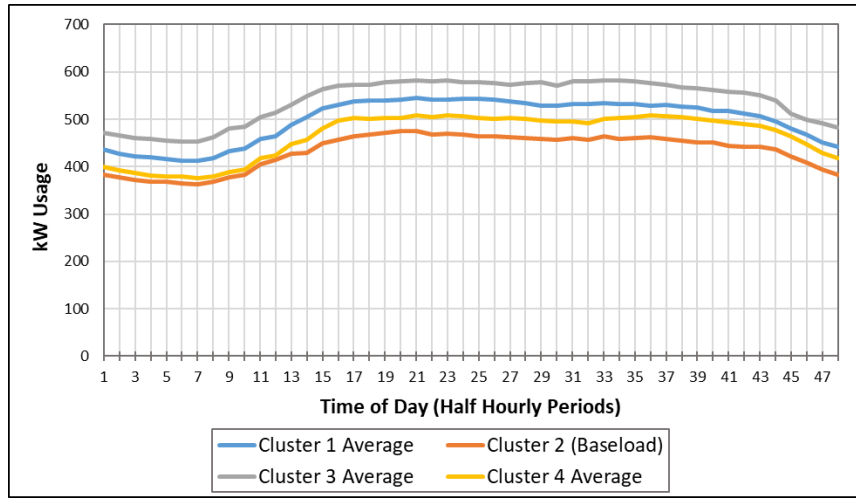
242 The k-means cluster method is used for the baseline comparison (Sayad, 2017). The clustering method  
243 works by first selecting how many groups the usage dataset will be clustered into. For each group, a  
244 random point within the dataset is selected and deemed the centroid value. Each value in the dataset  
245 is assigned to the closest centroid. The mean of the values for each centroid is then calculated. The  
246 centroids are then moved to the mean position and the values are reassigned to the now closest  
247 centroids. This process is repeated until a pre-defined number of iterations is achieved or the level  
248 of centroid position change reaches a set tolerance. The number of clusters for the baseline  
249 comparison will vary for each building. One approach for determining the optimum number of k-  
250 means clusters to use is called 'elbow' method. This method works by repeating the k-means method  
251 using a range of clusters to determine each clusters percentage of variance. The percentage of  
252 variance (dependent variable) is plotted against number of clusters (independent variable) in order to  
253 find the 'elbow' of the curve that signifies the optimum number of clusters, as adding more will have  
254 limited benefit in reducing variance (Ketchen & Shook, 1996). The k-means elbow identification  
255 process is undertaken for each hotel's electricity usage data. The data within each cluster is then  
256 averaged by half-hourly period. The half-hourly averages in each cluster are then used to generate  
257 daily profiles at half-hourly resolution for each cluster of each hotel. Figure 2 provides an example of  
258 the daily profiles developed for the four identified clusters of a hotel.

259  
260 Using the profiles to estimate DSR requires informed assumptions about what the profiles represent  
261 based on available information about the business. For the case of hotels, as in this study, information  
262 on energy sources related to heating and cooling (gas for heating, electricity for cooling), industry  
263 studies/reports on proportional breakdown of electricity use identifies that HVAC demand typically  
264 accounts for 34% of electricity demand in UK hotels (CIBSE, 2012). The consistent daily profiles of  
265 demand across all days of a week, consistent annual occupancy profiles found in hotels, and a high  
266 proportion of HVAC related demand provide the basis for assuming that variation in cluster profiles is a  
267 result of differing HVAC loads. It follows that the profile with the highest demand represents a high  
268 level of chiller usage, whilst the profile of lowest demand represents a baseline level of chiller usage.

269  
270 For a different case, such as an office, where weekday and weekend profiles are likely to be  
271 represented in different clusters, a larger optimum set of clusters is likely. Identifying baseline level  
272 chiller usage would potentially be more difficult in such circumstances where greater variability in  
273 demand related activity is found. Determining what the profiles represent highlights the primary  
274 drawback of this method as it requires assumptions to be made on limited data. Incorrectly assuming  
275 what the profiles represent will result in incorrect DSR estimations and therefore this method needs to  
276 be used with caution.

277  
278 Based on the assumption that profiles represent differences in chiller usage levels, the first step is to  
279 identify days associated with baseline use. In the context of the UK, chillers are not typically in use  
280 during the winter months. The baseline is, therefore, considered as days when the chiller is switched  
281 off during the heating season. The remaining clusters represent days when the chiller is in use. For this  
282 case, the kW usage levels of the chiller on these days is estimated by the difference between the  
283 cluster's usage value and the baseline value. Even in the case where the baseline cluster does not

284 represent chiller switch-off the differences in usage could still be considered as representative of  
285 maximum available turndown. See Appendix B for detailed calculation steps used in this method.  
286



287  
288  
289  
290 **Figure 2 - Example Chart of Clustered Averages**

## 2.4 DSR Estimation Method 3 - Utilise Historical DSR Event Outcomes

291 If a building has previously participated in DSR, then information gained on the kW amount reduction  
292 during each event can be utilised to estimate future performance. Research on this method has  
293 traditionally focused on confirming the DSR performance of a building by calculating the ‘residual  
294 demand’ (referred to as ‘turndown’ in this research), which is deemed as the difference between  
295 normal non-DSR building usage and the actual usage during a DSR event (Mathieu et al., 2010). Further  
296 research into understanding the expected level of residual demand using weather-based regression  
297 analysis was undertaken by Piette et al. (2011). They showed that the level of turndown potential  
298 could be linked to temperature if the DSR assets demonstrate varying levels of usage based on  
299 external weather conditions. This DSR estimation method utilises these concepts to identify a  
300 predictor that determines the expected turndown amount of historical DSR events. The predictor can  
301 then be utilised to determine the expected turndown amount at any time over a one-year period.

302  
303 This method relies on access to historical DSR event outcomes for the building. To provide consistency  
304 for testing this method with both hotels, a set of 24 DSR events were randomly created. The DSR  
305 events were then matched to each hotel’s actual chiller sub-metered data to provide real kW events  
306 outcome for each year of analysis (on the basis that during the event the chillers would have been  
307 temporarily turned off). Secondary data sources include any values that can be used for regression  
308 analysis to find a suitable predictor of the DSR event outcomes. For this research, the predictors  
309 selected for analysis were Outside Air Temperature, Building’s Electricity Usage Level, Half Hour Period  
310 of Day, and Day of Week. The first step in this method is to calculate the R-squared value of each  
311 predictor against the historical DSR event outcomes to decide which predictor to use. The regression  
312 calculation results of Table 1 show that the Outside Air Temperature predictor achieved the highest r-  
313 squared score and therefore this predictor is selected for the next step. The second step then uses the  
314 Outside Air Temperature values for each half-hourly period of the year in conjunction with the  
315 predictors slope and y-intercept to calculate the DSR estimation potential for the buildings. See  
316 Appendix C for detailed calculation steps used in this method.

317

Table 1 - Method 3's R-squared Regression Results

Hotel / Year	Time of Day	Day of Week	Buildings Electricity Usage Level	Outside Air Temperature
Hotel 1 - 2013	0.003	0.036	0.273	0.722
Hotel 1 - 2016	0.003	0.040	0.087	0.636
Hotel 2 - 2015	0.007	0.017	0.046	0.434
Hotel 2 - 2016	0.019	0.028	0.066	0.447

318

## 319 2.5 DSR Estimation Method 4 - Building Energy Modelling

320 Building energy modelling provides insight into DSR potential by modelling the energy usage of  
 321 building assets under different operational and environmental scenarios. Modelling gives insight into  
 322 flexibility of asset usage that can then be used for DSR estimation. However, this is very time  
 323 consuming in comparison to the previous estimation methods, and requires a very high level of  
 324 information and specialised skills to complete. Utilising a database of archetypal building models for a  
 325 building stock can help reduce the modelling burden for DSR, as demonstrated by Yin & Black (2015).  
 326 The predefined model archetypes can be modified as necessary, but its success is dependent on the  
 327 maturity of the database of archetypes and level of modification needed to provide results deemed of  
 328 value to DSR estimation. Another issue with energy building models is the 'performance gaps' between  
 329 model designs and actual performance of completed buildings, which can result in high levels of  
 330 output uncertainty (Menezes, Cripps, Bouchlaghem, & Buswell, 2012). For this research, the building  
 331 energy model DSR estimation method utilises the Yin & Black (2015) methodology by creating a  
 332 building energy model of the test hotels using EnergyPlus. The outcome of the simulation includes the  
 333 expected level of cooling in kW per half hour that will be used for DSR estimation.

334

335 To undertake this energy modelling approach, the building plans for each hotel were used to provide  
 336 both accurate building dimensions as well as the fabric structure of the building (outlined in Table 2).  
 337 The building plans are used to create a representative model of the building using the software  
 338 package 'DesignBuilder' v5.0.2 (DesignBuilder, 2017b). The DesignBuilder program then utilises the  
 339 EnergyPlus simulation program (U.S. Department of Energy, 2017) to estimate the buildings energy  
 340 usage over one year at half hourly intervals. The simulated energy usage results of the modelled chiller  
 341 units were then exported from DesignBuilder to provide the DSR estimation potential for each  
 342 building. See Appendix D for detailed calculation steps used in this method.

343

344

Table 2 - Build Energy Model Components

Component	Hotel 1 Description	Hotel 2 Description
External Walls	400mm thick wall (formed of stone masonry, brick, glass wool insulation, and plasterboard) total U-Value of 0.289	300mm thick wall (formed of brick, polystyrene insulation, concrete, and plasterboard) total U-Value of 0.351
External Windows	Double glazed (formed of two 3mm panes with a 6mm air gap) total U-Value of 3.365	Double glazed (formed of two 3mm panes with a 6mm air gap) total U-Value of 3.365
Roof	400mm flat roof (formed of asphalt, glass wool insulation, air gap, plasterboard) total U-Value of 0.322	320mm Flat roof (formed of asphalt, glass wool insulation, air gap, plasterboard) total U-Value of 0.346

HVAC System	Fan Coil Unit (4-Pipe), 333kW air-cooled chiller with a cooling set point of 23°C	Fan Coil Unit (4-Pipe), 290kW air-cooled chiller with a cooling set point of 23°C
Property Details	7 stories, ~21,000 m <sup>2</sup> isolated building located in Bristol, UK.	6 stories, ~15,000 m <sup>2</sup> isolated building located in London, UK.
Weather File	Custom DesignBuilder weather data file created for each year of analysis (DesignBuilder, 2017a).	

345  
346  
347

## 2.6 Sensitivity Analysis

348 The accuracy of estimation method is an important factor in creating credible/robust DSR portfolios  
349 that can meet grid-operator needs. Appropriate interpretation of uncertainty in inputs to the proposed  
350 methods is, therefore, critical to DSR estimation. To understand the impact of each estimation  
351 method's input uncertainty on the DSR estimation, and so give insight as to where more accurate  
352 information should be sought, a one-at-a-time local sensitivity analysis test was carried out, as in  
353 Saltelli, Chan, & Scott (2008). The sensitivity results are compared using the HVAC chillers yearly MWh  
354 usage estimation output as generated by of the four methods, as this provides scale context to the test  
355 outcomes. In performing the sensitivity tests, each method was first run using base values for each  
356 input parameter, as described in Figure 1 and sections 2.2 to 2.5. Completing this step provides  
357 baseline outcomes for comparison against. Each input parameter was then adjusted from the base  
358 values, as outlined in Table 3, and the sensitivity test for each method re-run using the adjusted input  
359 parameter, generating the sensitivity comparison results. As estimation methods 1-3 only have one or  
360 two input variables, all inputs for each method are tested during the analysis. The detailed modelling  
361 approach of Method 4, however, has a wide range of input variables ranging from building form and  
362 structure, to operational schedules of appliances and occupancy profiles. In this instance, it is assumed  
363 that the availability of building plans and detailed information of HVAC and lighting infrastructure  
364 reduces uncertainty in many of the structural aspects of the model. Menberg, Heo, & Choudhary  
365 (2016) identified temperature set points, thermal conductivity, and air infiltration as having a  
366 significant impact on building energy model results. These three variables are the focus of our analysis  
367 for Method 4.

368  
369

**Table 3 - Summary of Estimation Method Sensitivity Analysis Input Parameters**

Method	Base Values	Input Adjustment
1 (1)	50%	Adjust asset usage percentage by +/- 5 and 10 points
1 (2)	10%	Adjust asset percentage usage of baseload value by +/- 2.5 and 5 points
	5%	Adjust baseload percentile by +/- 1 and 2 points
2	4	Adjust number of clusters used by +/- 1 cluster
3	12	Adjust number of available existing events by -50%, +50%, +100%
4	23 °C	Adjust cooling set point by +/- 1 and 2 °C
	0.289 to 3.365	Adjust U-Values of External Walls, Windows, and Roof by +/- 10% and 20%
	0.7	Adjust air infiltration levels by +/- 0.1 and 0.2 ac/h

370  
371

372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394

## 2.7 Determining the Cost of each Estimation Method

The final output of the review of DSR estimation methods is a comparison of each method’s estimation errors in relation to its cost to run. This comparison is performed to provide context on the usage of each method in a business setting. It enables consideration of the cost/benefit selection of a higher error method that is cheaper or vice-versa. To calculate each method’s cost to run in a business setting required estimating the time it would take an experienced user to perform the tasks needed to run the estimation method and the cost of any external data input requirements. Table 4 provides a summary of the expected time required and external cost (if any) for each informational input. The time estimations used in this table are necessarily subjective, as the actual time and cost required will depend on and vary by individuals and organisations. Given the potential for variability, creating a cost factor provides a means of understanding the representative scale of effort required to undertake each method. The figures used in this table provide a point of reference, comprising estimations based on experience gained through application of these methods within a UK aggregator for this research and observations of users. The time value includes both the time taken to obtain information about the building (this covers talking to the building representative to obtain the sites half hourly electricity usage data and information about the DSR assets) and the time required to format, analyse and interpret the data. Most external information has no direct cost, as it is obtained for free from the building users or other sources. The only externally sourced information incurring cost is historical weather observations (ECMWF, 2017), which has a fixed yearly fee of £5,000 and has been split into individual usage costs on the assumption of performing 500 assessments per year (£10 per usage).

**Table 4 - Summary of Estimation Methods Information Input Costs**

Information Input	Time to obtain / use (minutes)	Usage Cost (@ £20 per hour)
Maximum kW rating of building’s DSR assets	30	£10
Building’s electricity usage records for 1 year	60	£20
Previous DSR Event Outcomes	120	£40
Hourly outdoor weather information for 1 year	60	£20 + £10 (data)
Building plans and operational information	420	£140

395  
396  
397  
398  
399  
400  
401  
402

To calculate the total cost of performing each method, the individual costs of gaining data for each input from Table 4 are associated with each method as per Table 5. This table shows the cumulative total running cost of each method, based on the information required. This information combined with the MAPE results from section **Error! Reference source not found.** enables a comparison of estimation error against method cost to be performed, as shown in section 3.3.

**Table 5 – Summary of Costs to Perform Each Estimation Method**

Information Input & Cost		Information Usage and Cost per Method				
		1 (1)	1 (2)	2	3	4
Maximum kW rating of building’s DSR assets	£10	£10	£10	£10	£10	£10
Building’s electricity usage records for 1 year	£20		£20	£20		
Previous DSR Event Outcomes	£40				£40	
Hourly outdoor weather information for 1 year	£30				£30	£30
Building plans and operational information	£140					£140
<b>Total Cost per Method</b>		<b>£10</b>	<b>£30</b>	<b>£30</b>	<b>£80</b>	<b>£180</b>

403

404  
405

### 3 Results and Discussion

406 The results of applying the four DSR estimation methods to two hotels is reviewed and discussed over  
407 three sections. The first section reviews the initial outputs of each method by applying 'base case'  
408 values to the input variables, and comparing the estimation error between methods. The second  
409 section then reviews the sensitivity analysis results to understand the impact of input variables on the  
410 estimation error levels. Finally, the error levels are compared against the estimated cost of  
411 undertaking each method, to gain an understanding of how cost and error levels correspond.

412  
413

#### 3.1 Estimation Method MAPE and MBE Outcomes

414 The estimation errors of MAPE and MBE for each estimation method, when using default (base) values  
415 for input variables, are given in Table 6. The methods were applied to each hotel over two years to  
416 generate a predicted half hourly kW usage value for their HVAC chillers. The predicted kW values were  
417 then compared to the actual kW usage values (as recorded by sub-meters), and MAPE and MBE were  
418 calculated for annual estimation errors. The average, minimum, and maximum MAPE and MBE values  
419 were then calculated, as shown in Figure 3. The MAPE values provide an overall indication of the level  
420 of difference between the actual and predicted results. Figure 3 and Table 6 show a range of MAPE  
421 estimation errors across the methods, with M1-V1 'Asset Assessment' having the worst average level  
422 of error at 159%. In contrast, M3 'Utilise Historical DSR Event Outcomes' had the lowest average level  
423 of error at 39%.

424  
425  
426  
427  
428  
429  
430  
431  
432

The MBE values indicate the direction of error between the actual and prediction values, with positive  
and negative results indicating over estimation and under estimation respectively. Figure 3 shows that  
all methods, except M1-V1, under predict usage levels. As seen with the MAPE result, the M1-V1  
outcome also has the highest average MBE value at 150%, which indicates that this method  
dramatically over predicted the expected usage of the HVAC chiller. In contrast, with an average MBE  
of -10%, M4 provides the closest prediction to actual usage.

Table 6 – Individual hotel summary of Estimation Method error levels

Method	Hotel 1 - 2013		Hotel 1 - 2016		Hotel 2 - 2015		Hotel 2 - 2016	
	MAPE	MBE	MAPE	MBE	MAPE	MBE	MAPE	MBE
M1-V1	193%	122%	250%	136%	98%	236%	96%	104%
M1-V2	35%	-46%	59%	-50%	71%	1%	75%	-38%
M2	57%	-41%	59%	-29%	40%	16%	70%	-12%
M3	33%	-15%	40%	-7%	36%	-6%	46%	-18%
M4	58%	-1%	63%	5%	39%	2%	45%	-31%

433  
434  
435  
436  
437  
438  
439

**Abbreviation Key:**

- M1-V1 = Method 1- Variation 1 - Minimum information using set percentage of asset usage
- M1-V2 = Method 1- Variation 2 - Utilise baseload calculation with set usage percentage
- M2 = Method 2 - Baseline comparison using cluster analysis
- M3 = Method 3 - Regression analysis utilising historical DSR event outcomes
- M4 = Method 4 - Building energy modelling

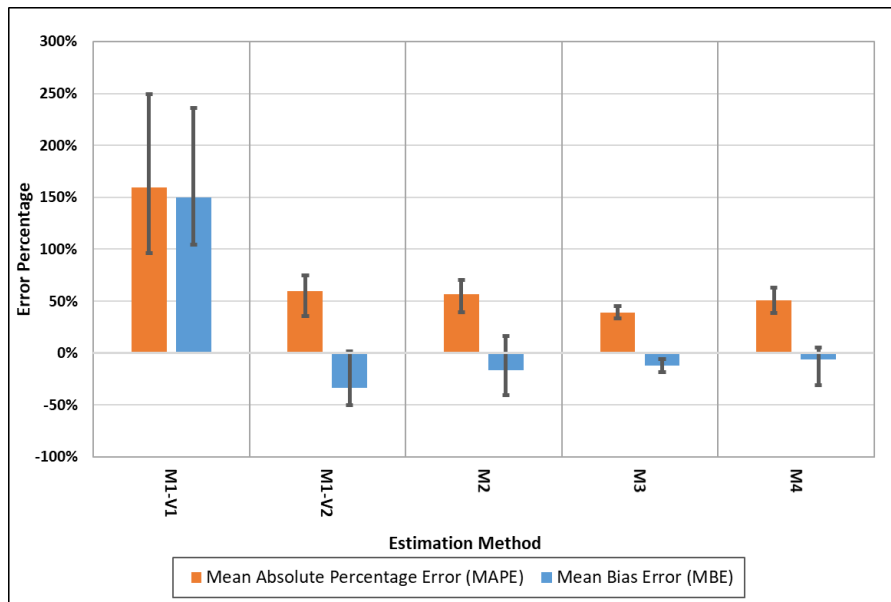


Figure 3 – Summary of Each Estimation Methods Error Levels

440  
441  
442

443 Considering the outcomes of each method: the two sub-variations of M1 had contrasting results with  
 444 M1-V1 having the highest overall average error level at 159%, while M1-V2 had a considerably lower  
 445 error level of 60%. The high uncertainty level of the M1-V1 method could be a result of it assuming a  
 446 fixed usage level of a chiller when the actual sub-meter data shows a highly variable pattern based on  
 447 a usage percentage mean of 20.8% with a variance of 252.5%. In contrast, M1-V2 uses the more  
 448 variable input of the building’s overall electricity usage levels for a year to first calculate the buildings  
 449 baseload usage. A percentage (in this case 10%) of the baseload is then deemed to be used by the DSR  
 450 asset, producing a much lower average MAPE value of 60%. This result is unexpectedly low considering  
 451 the method still uses a fixed proportion of buildings usage, which only considers time of day variation  
 452 and results in the same half hour prediction values being used for the entire year. The error level is still  
 453 high due to this method only taking time of day variation into account and does not consider day of  
 454 year variation which will impact the estimation results of a chiller that is highly influenced by  
 455 seasonality

456

457 An average MAPE value of 56% placed M2 as the method with the second highest level of absolute  
 458 error. Comparatively, however, the average MAPE is similar to the M1-V2 and M4 results. This  
 459 outcome, which is based on the method outlined by Panapakidis et al. (2014), helps support usage of  
 460 their profile clustering technique based on the DSR estimation results being comparable to the other  
 461 methods. Caution however needs to be taken on assuming this method is comparable to M1-V2 and  
 462 M4 due to its assumptions around the differences between profiles indicating usage of a particular  
 463 electrical asset, which may be difficult to determine in different businesses.

464

465 The lowest MAPE of all the methods was M3 at 39%. The ranking of method suitability by MAPE  
 466 supports research by Piette et al. (2011) where the inclusion of temperature dependency of DSR assets  
 467 in predictors improves prediction. For non-weather impacted assets other potential regression  
 468 parameters could be used including time of day, occupancy levels, or operational schedules. The  
 469 drawback to this method is access to historical DSR events and obtaining suitable predictor data, which  
 470 could be hard to come by.

471 An average MAPE value of 51% placed M4 as the method with the second lowest level of absolute  
472 error. It is possible to achieve lower levels of error as demonstrated by the researchers at the  
473 Lawrence Berkeley National Laboratory (Dudley, 2010) who used calibrated Energy Building Models for  
474 accurate DSR forecasting. However, the calibration methods require obtaining sub-metered data of  
475 key electrical assets which, if available, could be used directly for predicting the building's DSR usage,  
476 limiting the need for using an Energy Building Model. While this method achieves comparatively good  
477 error estimation levels even without calibration, it does have the drawbacks of requiring access to  
478 detailed plans of a building and the skill and time needed to construct the model.

479

480

### 481 **3.2 Sensitivity Analysis of Estimation Methods**

482 The previous review of the error in the estimation methods provides a comparative analysis of  
483 methods without accounting for the uncertainty in their input values. The error range in DSR  
484 estimation depends not only on the estimation methodology, however, but also on these input  
485 uncertainties and the sensitivity of method outcome to these uncertainties. Figure 4 summarises the  
486 sensitivity profiles for each method's inputs, as determined by re-running each method with adjusted  
487 inputs. To facilitate comparison of sensitivity between methods, the charts shown in Figure 4 have  
488 been normalised. Plotting change in input variable as a percentage of the base case value against the  
489 percentage difference in estimated energy use (MWh), Figure 4 shows varying sensitivity to inputs  
490 within and across the four methods. This section examines each method's sensitivity profiles to gain  
491 further insights into how they are influenced by input variation.

492

493

494

495

496

497

498

499

500

501

502

503

504



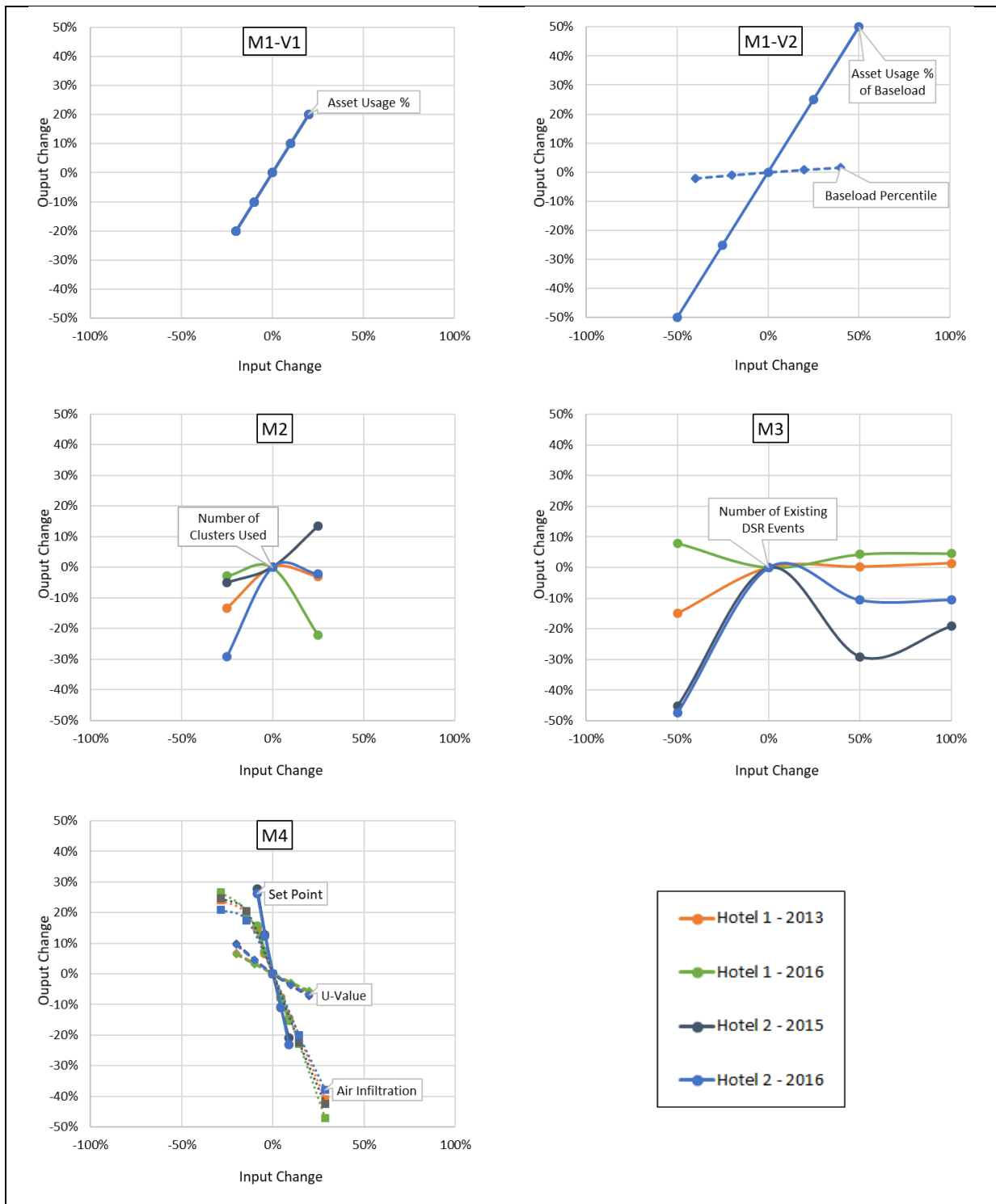


Figure 4 - Estimation Method Sensitivity Analysis Results

505  
506

507 The asset usage percentage input gradients of M1-V1 (1:1) and M1-V2 (1:1) shows they are both  
 508 sensitive to changes, whilst adjustments in the percentile value used for baseload estimation in M1-V2  
 509 has little effect (0.04:1). Altering the asset usage percentage input values for M1-V1 and M1-V2  
 510 however had different impacts on the resulting MAPE outcomes across both hotels and years. The M1-  
 511 V1 MAPE outcomes varied from -28.2% to 29.5% with a consistent pattern of the MAPE value  
 512 decreasing as the percentage of asset usage value lowered. This indicates that the base usage value of  
 513 50% is too high and a lower value should be used to better represent actual usage of the chillers. The  
 514 M1-V2 MAPE outcomes had a greater variance level of -11.5% to 82.8% and in contrast to M1-V1,  
 515 when the asset usage percentage of the baseload value is lowered the MAPE values increased.

516 However, when the usage percentage is increased MAPE values for Hotel 1 initially lower before  
517 increasing - indicating that the base value is close to optimal. MAPE values for Hotel 2 continue to  
518 decrease as the usage percentage increases indicating that a higher base value would be more  
519 appropriate. The other input for M1-V2, percentile baseload value, has negligible effect on the MAPE  
520 outcomes with a variance range of -1.0% to 1.6% across both hotels and years and therefore the base  
521 value of 10% is deemed appropriate.

522  
523 M2 has a non-linear sensitivity profile, with each hotel and data year being impacted differently with  
524 no clear pattern. The percentage change in MAPE values resulting from the input changes has a  
525 variance range of -6% to 7% across both hotels and years. This level of MAPE variance implies that  
526 changing the number of clusters has only a small impact, and that the base value is appropriate for this  
527 application of the estimation method. The limited output variance could be the result of this method  
528 calculating the chiller usage values based on differences between cluster profiles that means adding or  
529 removing a single cluster will only cause the redistribution of input values into other similar clusters  
530 without causing major changes in the generated profiles.

531  
532 M3 also has a non-linear sensitivity profile that varies differently between the two hotels. The general  
533 pattern of the profile shows that when the number of historical events is lowered by 50% from 12 to 6,  
534 this has the greatest impact on estimation outputs, with MAPE values increasing by 4% and 23% for  
535 Hotel 1, and 28% and 75% for Hotel 2. When the number of events is increased to 18 and 24, the  
536 profile shows a more consistent change, except for Hotel 2 -2015. When excluding Hotel 2 -2015, the  
537 MAPE values had a minimal change range of -2% to 5%. However, Hotel 2 - 2015 showed far greater  
538 changes, with the MAPE value increasing by 45% and 28%. A potential cause of this difference could be  
539 due to the facilities manager of Hotel-2 deciding when to turn the chiller system on and off during the  
540 year. In 2015 it was turned on in April and off in October, whereas in 2016 it was turned on in May but  
541 not turned off again. In contrast the Hotel-1 system is left running all year with output adjusted  
542 automatically as required to meet the set point conditions. Based on the overall results of this method  
543 it is clear that reducing the number of historical events has a negative impact on the outcomes.  
544 Whereas the impact on increasing the number of events used is unclear due to the outcomes of Hotel  
545 2 – 2015.

546  
547 M4 has three different input variables of Cooling Setpoint, U-Value, and Air Infiltration. The Set Point  
548 Temperature and U-Value inputs have linear sensitivity profiles with gradients of (1:0.32) and (1:0.6)  
549 respectively. The Air Infiltration input range of 0.6 to 0.9 ac/h had a linear profile of (1:0.7), however  
550 the lowest input value of 0.5 ac/h was not linear with a smaller change in output compared to the  
551 linear values. Air Infiltration changes displayed the biggest impact on output and resulting MAPE  
552 values. This is shown with the MAPE values for Air Infiltration having a variance range of -18% to 54%.  
553 In contrast, the MAPE values range for the U-Value input was -8% to 8% and the Cooling Setpoint input  
554 range was -18% to 27%. The results show how changing the Set Point temperature and Air Infiltration  
555 rates have significant impacts on the chiller usage compared to only a minor impact from changing U-  
556 Values. This could reflect the usage of mechanical space cooling, which actively responds to  
557 temperature requirements and causes pressurised losses through Air Infiltration. The Air Infiltration  
558 input having the biggest impact does raise concern for this type of estimation method, as this is one of

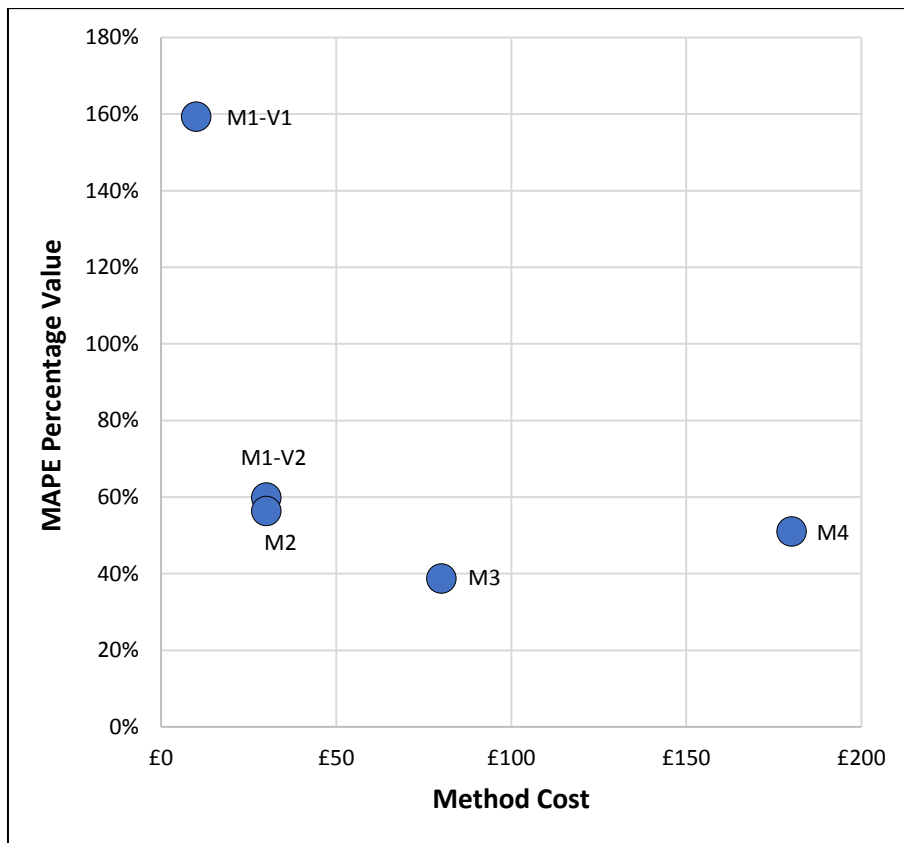
559 the hardest parameters to determine when constructing the energy building model. The other inputs  
560 can be obtained with relatively high accuracy by obtaining the Set Point directly from the building's  
561 current setup and the U-Values from visual inspections of the existing construction and building plans.  
562 In contrast, the Air Infiltration rate can only be accurately obtained through a building pressure test  
563 which would be infeasible for a building of this size. Therefore, the default building model Air  
564 Infiltration rates will need to be used, and caution taken on the final outputs.

565

### 566 3.3 Cost versus Method Estimation Errors

567 The final set of results compares the cost of running each method against the expected level of  
568 estimation error. This comparison helps provide context to usage of the methods when balancing cost  
569 against acceptable error levels. Figure 5 maps out the links between each method's average MAPE  
570 results as per Table 6 and the estimated cost to run as per Table 5. The figure shows a rough trending  
571 direction of a higher method cost resulting in lower estimation errors. This is reflected in the lowest  
572 cost method M1-V1 having the highest error level while the lowest error level M3 has the second  
573 highest cost. Each method will be further examined to understand the implications of method costs  
574 and input requirements on error outcomes.

575



576

577

Figure 5 - Comparison of Estimation Method Error versus Cost

578

579 M1-V1 has the distinction of being the cheapest estimation method with the worst error level. This can  
580 be directly related to the input requirement of only needing to know the asset's maximum kW rating,  
581 and then using a percentage of this for the estimation. This requires minimal time for a person to  
582 undertake, both in collecting the required information and using it to calculate the estimation.  
583 Unfortunately, the high error level means that this method can only be used for a very rough and quick  
584 estimation before proceeding with a lower error method. In comparison, M1-V2 reduces the error

585 level by two thirds compared to M1-V1 while costing 3 times more to run. While M1-V2 is more  
586 expensive than M1-V1, it is still comparatively cheap compared to all the methods tested. This method  
587 also uses relatively accessible data of the building's electricity usage records, which in the UK is  
588 available in half hourly format for any business with peak electricity usage of 100 kW or greater.

589  
590 M2 is the third equally cheapest method to run due to the primary input requirement being the  
591 building's half hourly electricity usage records. It also has the fourth lowest error level and therefore,  
592 of the methods analysed, provides a representatively balanced error to cost ratio, which makes it a  
593 potentially suitable approach. However, as discussed previously, this method's usage of clustering  
594 means that care needs to be taken on its application to suitable buildings and assets.

595  
596 M3 achieved the lowest error level of all methods tested at 39%. However, it also has the second  
597 highest cost at £80, which is a result of requiring two expensive input requirements. Firstly, it uses  
598 detailed historical air temperature readings over a year for the building's location, which requires  
599 paying for access to the necessary weather archive. Secondly, it uses previous DSR event outcomes  
600 which require time to obtain from the building users, and then formatting and verifying before using. It  
601 is also anticipated that obtaining previous DSR event outcomes could be difficult, due to the limited  
602 current uptake of DSR and even if the client has participated, then it could be difficult for them to  
603 provide the necessary information based on how it has been provided from their current aggregator.

604  
605 M4 had the highest cost at £180 with the second lowest error level of 51%. The high cost is primarily  
606 due to the time required to model the building in the building energy modelling tool. As the resulting  
607 error level is similar to M1-V1, M2 and M3 methods, which are significantly cheaper to run, this  
608 method is not recommended. Although a potential justification for using this method would be if  
609 multiple assets within one building were being estimated, thereby reducing the individual assessment  
610 costs while providing a combined view of the building's potential.

611

## 612 **4 Conclusion**

613 This paper has undertaken an examination and comparison of four non-domestic DSR estimation  
614 methods to provide insights into uncertainty levels based on the input requirements. The examination  
615 was performed by using each method to estimate the DSR potential of HVAC chiller assets at two  
616 hotels over two years. The estimation outcomes were then compared against the chiller's actual sub-  
617 metered usage records by calculating MAPE and MBE values to understand each method's level of  
618 estimation error. The results showed a wide range of estimation errors. Method 1 - Sub-variation 1  
619 yields the highest error level MAPE of 159%, while the lowest error level MAPE of 39% was achieved  
620 with method 3. While method 3 could be a recommended approach based on its low error level alone,  
621 its usage is restricted by information input considerations. The primary limitations of this research  
622 were a reliance on usage of one electrical appliance (HVAC chillers) and business (hotels) type,  
623 uncertainty of the method usage time and cost input variables due to the subjectivity of how each  
624 organisation could apply them, and being restricted to using only known estimation methods that  
625 excludes unpublished proprietary approaches. Based on this paper's findings, each method requires  
626 review to understand the implications of input requirements on outcome uncertainty. These findings  
627 can be summarised as follows:

628

- 629
- 630
- 631
- 632
- 633
- 634
- 635
- 636
- 637
- 638
- 639 • **Method 1** sub-variation 1 has the lowest informational requirement and cost of £10 to use  
640 based on only needing to know the maximum kW rating of the asset being assessed to apply  
641 this method. However, the penalty of this low informational requirement is the highest error  
642 level of all methods at 159%. Sub-variation 2 achieved a much lower error level of 60% by  
643 using the building's half-hourly electricity records that increases the usage costs to £30. The  
644 sensitivity results for this method showed a high impact on the outcomes based on variations  
645 of the inputs. This means that the error results might differ substantially when used in other  
646 scenarios. Therefore, the error levels reported in this research for method 1 need to be used  
647 with care when deciding on suitable assessment approaches.
  - 648
  - 649 • **Method 2** had the second worst error level of 56% while being the third cheapest to run at £30  
650 through clustering of the building's half hourly electricity usage data. The sensitivity analysis of  
651 this method showed a medium to low impact on error levels arising from changes in the  
652 primary input of how many clusters are used. These results indicate that baseline comparison  
653 is a suitable method for assessment though it has two limitations that need to be fully  
654 understood by users to ensure valid results. Firstly, it requires the user to select the  
655 appropriate number of clusters, which is open to individual interpretation. Secondly, this  
656 method will only work on electrical assets that have enough variation within the building's  
657 overall usage to be identified by the clustering.
  - 658
  - 659 • **Method 3** had the lowest overall error level of 39% with the second highest cost of £80. The  
660 low error level makes its utilisation of historical DSR event outcomes an attractive method.  
661 However, its practical usage is limited as it requires the building to have previously undertaken  
662 DSR and have access to historical DSR events outcomes. The sensitivity analysis also showed a  
663 significant increase in error if less than 12 historical event records over a year are available for  
664 analysis. In new DSR markets these limitations may restrict usage of this method. Even in  
665 established markets it could be difficult or time consuming to obtain any adequate historical  
666 information from the existing DSR aggregator.
  - 667
  - 668 • **Method 4** had the second lowest error level at 51% but had the highest cost of £180, which is  
669 over twice that of method 3, the next most expensive, as a consequence of the amount of  
670 time required to develop a building energy model. While this method had the second lowest  
671 error level, it is only slightly lower than many other cheaper options and method 2, for  
672 example, costs 6 times less with only a slightly higher error level of 56%. The usage  
673 requirements of this method also restrict its practical application given its reliance on detailed  
674 building plans and the skills to develop building models. The importance of having the right  
675 information and skills is highlighted by the sensitivity analysis, which showed major impacts  
676 from variations in temperature set-points and air infiltration model values.

668 These findings have three key implications on the selection of DSR estimation methods. Firstly, the  
669 wide range of error levels means the outputs of these methods will need to be carefully considered  
670 when being used to make decisions about the suitability of buildings for DSR. Secondly, care needs to  
671 be taken in ensuring accurate input selection as sensitivity analysis demonstrates that adjusting the

672 inputs on most methods will result in large variations to the outputs. Thirdly, this research tested four  
673 methods using HVAC chillers in hotels only. Therefore, other assets and businesses may result in  
674 different error outcomes and caution needs to be taken before this research is used to select  
675 estimation methods outside of this scope. This final implication highlights a potential future area for  
676 research which would entail re-running the method comparisons on different DSR assets and  
677 businesses to understand the different impacts on estimation outcomes.

678

## 679 **5 Funding**

680 This work has been supported and funded by the Technologies for Sustainable Built Environments  
681 (TSBE) centre, Reading University, in conjunction with the Engineering and Physical Sciences  
682 Research Council (EPSRC) [grant number EP/G037787/1].

683

## 684 **6 Appendix A**

685 The following steps outline the calculations performed for Method 1 - Asset Assessment:

686

### 687 **1. Variation 1 – Minimum Information**

- 688 1.1. An anticipated set percentage usage amount of the asset is selected based on either a  
689 default 50%, or another amount if the assessor has prior knowledge of the type of asset  
690 and site.
- 691 1.2. The expected kW usage level of the asset is calculated for each half-hour of a year by  
692 multiplying the anticipated percentage usage amount by the maximum rating of the asset,  
693 with the resulting values being saved into a DSR asset usage estimation dataset.

694

### 695 **2. Variation 2 – Utilise Baseload Calculation**

- 696 2.1. Using the site's Metered Electricity Usage Records, a baseload value is calculated by  
697 obtaining the 5<sup>th</sup> percentile kW value for each half-hour period of the day based on one  
698 year's worth of data as per formula (1) (e.g. for each half-hour period of a day, the 365  
699 daily values for the year are obtained and then ranked before determining the 5<sup>th</sup>  
700 Percentile value).

$$n_{HH} = \left[ \frac{P}{100} \times N_{HH} \right] \quad (1)$$

Where:

$n$  = kW value of percentile for selected half-hour

$P$  = Percentile

$N$  = Ordered list of kW values for selected half-hour (sorted from least to  
greatest)

$HH$  = Selected half-hour

- 701 2.2. A percentage value is then selected that represents how much of the baseload is expected  
702 to be used by the asset. This can either be a default 10%, or another amount if the  
703 assessor has prior knowledge of the asset type and site.

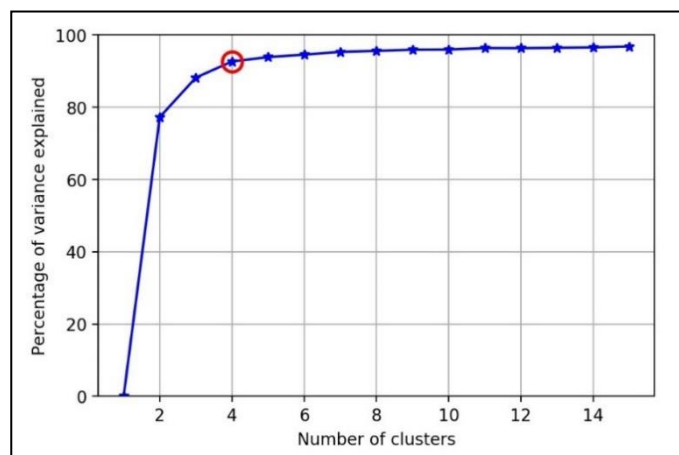
- 704 2.3. The expected kW usage level of the asset is calculated for every half-hour period in a year  
705 by multiplying the anticipated percentage usage amount against the baseload kW value,  
706 with the resulting values being saved into a DSR asset usage estimation dataset.

707 2.4. If the usage outcome is higher than the maximum usage rating of the assets, then the  
708 previous step is re-run with a lower percentage.  
709

## 710 7 Appendix B

711 The following steps outline the calculations performed for Method 2 - Baseline Comparison:

- 712 1. The k-means cluster method is used for the baseline comparison (Sayad, 2017). This clustering  
713 method works by first selecting how many groups the usage dataset will be clustered into. For  
714 each group, a random point within the dataset is selected and deemed the centroid value.  
715 Each value in the dataset is assigned to the closest centroid. The mean of the values for each  
716 centroid is then calculated. The centroids are then moved to the mean position and the values  
717 are reassigned to the now closest centroids. This process is repeated until a pre-defined  
718 number of interactions is achieved or the level of centroid position change reaches a set  
719 tolerance.
- 720 2. The number of clusters for the baseline comparison will vary for each site. For this analysis the  
721 'elbow' method for determining the optimum number of k-means clusters is used. This  
722 method works by repeating the k-means method using a range of clusters to determine each  
723 cluster's percentage of variance. The percentage of variance (dependent variable) is plotted  
724 against the number of clusters (independent variable) in order to find the 'elbow' of the curve,  
725 which signifies the optimum number of clusters as adding more will have limited benefit in  
726 reducing variance (Ketchen & Shook, 1996). Figure provides an example of identified 'elbow'  
727 for clustering of one hotel's daily electricity usage profiles over one year. The main recognised  
728 limitations of the elbow method is its reliance on a manual decision-making process to  
729 determine where the elbow sits, and that the chart might not have a recognisable elbow if the  
730 line is consistent across the clusters (Ketchen & Shook, 1996). The elbow method calculation is  
731 performed by:
  - 732 i. Calculating the percentage of variance explained for a range of clusters (normally 1-15)  
733 using the equation (Imran, 2015).
  - 734 ii. Create a line chart with markers that shows each cluster's percentage of variance as  
735 shown in Figure for Hotel 1 in 2016
  - 736 iii. Determine the elbow based on the chart and record the cluster number.



739  
740 **Figure 6 - Example of Cluster Identification using the Elbow Method**  
741 **(with the Elbow being indicated by the red circle)**  
742

743 3. Once the number of clusters to be used has been decided, then the k-means method as shown  
 744 in equation (2) (Sayad, 2017) can be used to group the Site’s Half Hourly Electricity dataset into  
 745 similar days. The dataset is then updated with a new column 49 containing a value that  
 746 represents which cluster each day belongs to.

$$J_n = \sum_{j=1}^K \sum_{i=1}^n (x_i - c_j)^2 \quad (2)$$

Where:

$n$  = Objects being clustered

$J_n$  = Cluster outcome for  $n$  value

$K$  = Clusters

$c_j$  = Centroid for cluster  $j$

$x_i$  = Object  $i$

- 747 4. The half-hourly averages in each cluster are then used to generate daily profiles at half-hourly  
 748 resolution for each cluster of each hotel. **Error! Reference source not found.** provides an  
 749 example of the daily profiles developed for the four identified clusters of a hotel.
- 750 5. The baseline profile is then identified based on the assumption that the profiles represent  
 751 differences in chiller usage levels. In the context of the UK, chillers are not typically in use  
 752 during the winter months. Therefore, the baseline is considered as days when the chiller is  
 753 switched off during the heating season and, as a result, profile cluster 2 in **Error! Reference**  
 754 **source not found.** comprises the baseline profile as it has the lowest usage values. The  
 755 remaining cluster profiles then represent days when the chiller is in use.
- 756 6. A new dataset is created that covers all half-hourly periods for one year, and has an additional  
 757 column identifying which cluster profile is associated with each day of the year. For each day in  
 758 the dataset, the kW usage levels of the chiller is estimated by the calculating the difference  
 759 between that day’s cluster profile usage value and the baseline value. If a day in the new  
 760 dataset is associated with the baseline cluster, then the chiller is deemed to be off during this  
 761 day, so the expected usage is set to 0.
- 762 7. The dataset now represents the DSR asset usage estimation dataset of the chiller. The results  
 763 are then checked to verify that no values are greater than the maximum usage rating of the  
 764 chiller asset. If there are, then the values are adjusted down to the maximum rating or, if the  
 765 values are consistently too high, then this method is rejected if the assessor believes the  
 766 method is providing unrealistic results based on the assessor’s (or their colleagues’) prior  
 767 knowledge of customary usage for this type of asset.

## 769 8 Appendix C

770 The following steps outline the calculations performed for Method 3 - Utilise Historical DSR Event  
 771 Outcomes:

- 772 1. The first step is to determine what variables are available for predicting the event turndown  
 773 amount. For this example, the variables of Outside Air Temperature, Site Electricity Usage, Half  
 774 Hour Period of Day, and Day of Week are used.
- 775 2. For each variable, a two-column dataset is created for each year of data with the first column  
 776 containing the event turndown results, and the second column containing the predicting  
 777 variable value.
- 778 3. Using equation (3) the R-squared / coefficient of determination for each dataset is calculated.



$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (3)$$

Where:

$R^2$  = R-squared / coefficient of determination

$y_i$  = Current value from event data set

$\bar{y}$  = Mean of event data set values

$f_i$  = Predicted value for  $y_i$

- 779 4. The R-squared values of each variable used as shown in Table 1 are compared, and the highest  
 780 value selected as the predictor variable to be used for estimating DSR asset usage. In this case  
 781 the Outside Air Temperature has the highest values.  
 782 5. The Outside Air Temperature values for each half-hourly period of the year in conjunction with  
 783 the predictor's slope and y-intercept are used to calculate the DSR estimation potential for the  
 784 hotels.  
 785

## 786 9 Appendix D

787 The following steps outline the calculations performed for Method 4 - Building Energy Modelling:

- 788 1. The building plans for each hotel were used to provide both accurate building dimensions as  
 789 well as the fabric structure of the building (outlined in Table 2). The building plans are used to  
 790 create a representative model of the building using the software package 'DesignBuilder'  
 791 v5.0.2 (DesignBuilder, 2017b). The DesignBuilder program then utilises the EnergyPlus  
 792 simulation program (U.S. Department of Energy, 2017) to estimate the building's energy usage  
 793 over one year at half-hourly intervals.  
 794 2. Customised weather files were generated for each hotel for the years 2013 and 2016 and  
 795 loaded into DesignBuilder. These were created using MIDAS weather data (UK Met Office,  
 796 2017) that was then converted into an EnergyPlus formatted hourly weather data.epw file  
 797 using the process outlined on the DesignBuilder online help (DesignBuilder, 2017a)  
 798 3. Each model's energy usage was then simulated at half-hour intervals for one year using  
 799 DesignBuilder/EnergyPlus, with the results of the chiller assets electricity usage being  
 800 extracted to provide the DSR estimation potential for each hotel.  
 801  
 802

## 803 10 References

- 804 Aman, S., Frincu, M., Chelmis, C., Noor, M., Simmhan, Y., & Prasanna, V. K. (2016). Prediction models for dynamic demand  
 805 response: Requirements, challenges, and insights. *2015 IEEE International Conference on Smart Grid Communications,*  
 806 *SmartGridComm 2015*, 338–343. <https://doi.org/10.1109/SmartGridComm.2015.7436323>  
 807 Borgstein, E. H., Lamberts, R., & Hensen, J. L. M. (2016). Evaluating energy performance in non-domestic buildings: A review.  
 808 *Energy and Buildings*, 128, 734–755. <https://doi.org/10.1016/j.enbuild.2016.07.018>  
 809 Chassin, D. P., & Rondeau, D. (2016). Aggregate modeling of fast-acting demand response and control under real-time pricing.  
 810 *Applied Energy*, 181, 288–298. <https://doi.org/10.1016/j.apenergy.2016.08.071>  
 811 CIBSE. (2012). *CIBSE Guide F*. London: CIBSE Publications.  
 812 Curtis, M. (2017). Demand side response aggregators: How they decide customer suitability. In *2017 14th International*  
 813 *Conference on the European Energy Market (EEM)* (pp. 1–6). IEEE. <https://doi.org/10.1109/EEM.2017.7981909>  
 814 Curtis, M., Torriti, J., & Smith, S. T. (2018). Demand Side Flexibility and Responsiveness: Moving Demand in Time Through  
 815 Technology. In A. Hui, R. Day, & G. Walker (Eds.), *Demanding Energy: Space, Time and Change* (pp. 283–312). Cham:  
 816 Springer International Publishing. [https://doi.org/10.1007/978-3-319-61991-0\\_13](https://doi.org/10.1007/978-3-319-61991-0_13)  
 817 De Gooijer, J. G., & Hyndman, R. J. (2006). 25 Years of Time Series Forecasting. *International Journal of Forecasting*, 22(3),  
 818 443–473. <https://doi.org/10.1016/j.ijforecast.2006.01.001>  
 819 DesignBuilder. (2017a). DesignBuilder - Edit / Translate Hourly Weather Data. Retrieved February 3, 2017, from

820 [http://www.designbuilder.co.uk/helpv5.0/Content/\\_Edit\\_hourly\\_weather\\_data.htm](http://www.designbuilder.co.uk/helpv5.0/Content/_Edit_hourly_weather_data.htm)

821 DesignBuilder. (2017b). DesignBuilder V5.0.2. Retrieved from <https://designbuilder.co.uk/software/product-overview>

822 DRRC. (2017). Demand Response Research Center. Retrieved August 5, 2016, from <https://drcc.lbl.gov/about-us>

823 Dudley, J. H. (2010). Comparison of Demand Response Performance with an EnergyPlus Model in a Low Energy Campus

824 Building. Retrieved from <http://www.escholarship.org/uc/item/9zw4m12g>

825 ECMWF. (2017). European Centre for Medium-Range Weather Forecasts - Meteorological archive. Retrieved May 2, 2017,

826 from <https://www.ecmwf.int/en/forecasts/accessing-forecasts/order-historical-datasets>

827 Imran. (2015). What is "Within cluster sum of squares by cluster" in K-means. Retrieved March 2, 2017, from

828 <https://discuss.analyticsvidhya.com/t/what-is-within-cluster-sum-of-squares-by-cluster-in-k-means/2706/2>

829 Ketchen, D., & Shook, C. (1996). The application of cluster analysis in strategic management research: An analysis and

830 critique. *Strategic Management Journal*, 17(6), 441–458. [https://doi.org/10.1002/\(SICI\)1097-](https://doi.org/10.1002/(SICI)1097-0266(199606)17:6<441::AID-SMJ819>3.0.CO;2-G)

831 [0266\(199606\)17:6<441::AID-SMJ819>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1097-0266(199606)17:6<441::AID-SMJ819>3.0.CO;2-G)

832 Larsen, E. M., Pinson, P., Leimgruber, F., & Judex, F. (2015). From demand response evaluation to forecasting - Methods and

833 results from the EcoGrid EU experiment. *Submitted to IEEE Transactions on Power Systems*, 10, 75–83.

834 <https://doi.org/10.1016/j.segan.2017.03.001>

835 Mathieu, J. L., Gadgil, A. J., Callaway, D. S., Price, P. N., & Kiliccote, S. (2010). Characterizing the Response of Commercial and

836 Industrial Facilities to Dynamic Pricing Signals From the Utility. *ASME 2010 4th International Conference on Energy*

837 *Sustainability, Volume 1*, 1019–1028. <https://doi.org/10.1115/ES2010-90266>

838 Menberg, K., Heo, Y., & Choudhary, R. (2016). Sensitivity analysis methods for building energy models: Comparing

839 computational costs and extractable information. *Energy and Buildings*, 133, 433–445.

840 <https://doi.org/10.1016/j.enbuild.2016.10.005>

841 Menezes, A. C., Cripps, A., Bouchlaghem, D., & Buswell, R. (2012). Predicted vs. actual energy performance of non-domestic

842 buildings: Using post-occupancy evaluation data to reduce the performance gap. *Applied Energy*, 97, 355–364.

843 <https://doi.org/10.1016/j.apenergy.2011.11.075>

844 Merry, J. (2017). Inferior submetering is destroying performance. Retrieved April 8, 2017, from

845 <https://www.cibsejournal.com/opinion/inferior-submetering-means-substandard-performance/>

846 National Grid. (2016). Short Term Operating Reserve Information. Retrieved January 30, 2017, from

847 [http://www2.nationalgrid.com/UK/Services/Balancing-services/Reserve-services/Short-Term-Operating-](http://www2.nationalgrid.com/UK/Services/Balancing-services/Reserve-services/Short-Term-Operating-Reserve/Short-Term-Operating-Reserve-Information/)

848 [Reserve/Short-Term-Operating-Reserve-Information/](http://www2.nationalgrid.com/UK/Services/Balancing-services/Reserve-services/Short-Term-Operating-Reserve/Short-Term-Operating-Reserve-Information/)

849 Panapakidis, I. P., Papadopoulos, T. A., Christoforidis, G. C., & Papagiannis, G. K. (2014). Pattern recognition algorithms for

850 electricity load curve analysis of buildings. *Energy and Buildings*, 73, 137–145.

851 <https://doi.org/10.1016/j.enbuild.2014.01.002>

852 Piette, M. A., Mathieu, J. L., Price, P. N., & Kiliccote, S. (2011). Quantifying changes in building electricity use, with application

853 to demand response. *IEEE Transactions on Smart Grid*, 2(3), 507–518. <https://doi.org/10.1109/TSG.2011.2145010>

854 Saltelli, A., Chan, K., & Scott, E. (2008). *Sensitivity Analysis*. Chichester: Wiley.

855 Sayad, S. (2017). K-Means Clustering. Retrieved March 2, 2017, from [http://www.saedsayad.com/clustering\\_kmeans.htm](http://www.saedsayad.com/clustering_kmeans.htm)

856 SDG&E. (2016). *San Diego Gas & Electric - Capacity Bidding Program*. Retrieved from

857 [http://regarchive.sdge.com/tm2/pdf/ELEC\\_ELEC-SCHEDS\\_CBP.pdf](http://regarchive.sdge.com/tm2/pdf/ELEC_ELEC-SCHEDS_CBP.pdf)

858 SEDC. (2016). *Explicit and Implicit Demand-Side Flexibility Complementary Approaches for an Efficient Energy System*.

859 SEDC. (2017). *Explicit Demand Response in Europe Mapping the Markets 2017*.

860 Shen, L., Li, Z., & Sun, Y. (2016). Performance evaluation of conventional demand response at building-group-level under

861 different electricity pricings. *Energy and Buildings*, 128, 143–154. <https://doi.org/10.1016/j.enbuild.2016.06.082>

862 U.S. Department of Energy. (2017). EnergyPlus. Retrieved April 3, 2017, from <https://energyplus.net/>

863 UK Met Office. (2017). MIDAS Data Guide. Retrieved February 15, 2017, from [https://badc.nerc.ac.uk/data/ukmo-](https://badc.nerc.ac.uk/data/ukmo-midas/ukmo_guide.html)

864 [midas/ukmo\\_guide.html](https://badc.nerc.ac.uk/data/ukmo-midas/ukmo_guide.html)

865 Van Wijk, J. J., Van Selow, E. R., Wijk, J. J. Van, & Selow, E. R. Van. (1999). Cluster and calendar based visualization of time

866 series data. *Proceedings 1999 IEEE Symposium on Information Visualization (InfoVis'99)*, 4.

867 <https://doi.org/10.1109/INFVIS.1999.801851>

868 Yin, R., & Black, D. (2015). *Improvement of Demand Response Quick Assessment Tool ( Drqat ) And Tool Validation Case*

869 *Studies*.