Gulf Organisation for Research and Development

International Journal of Sustainable Built Environment

ScienceDirect
www.sciencedirect.com

CrossMark

Original Article/Research

A comparative study of benchmarking approaches for non-domestic buildings: Part 1 – Top-down approach

Sung-Min Hong [a,*], Greig Paterson [b], Esfandiar Burman [c], Philip Steadman [a], Dejan Mumovic [b]

[a] UCL Energy Institute, University College London, London WC1H 0NN, United Kingdom
[b] The Bartlett School of Graduate Studies, University College London, London WC1H 0NN, United Kingdom
[c] UCL Centre for Urban Sustainability & Resilience, University College London, London WC1E 6BT, United Kingdom

Abstract

Benchmarking plays an important role in improving energy efficiency of non-domestic buildings. A review of energy benchmarks that underpin the UK's Display Energy Certificate (DEC) scheme have prompted necessities to explore the benefits and limitations of using various methods to derive energy benchmarks. The existing methods were reviewed and grouped into top-down and bottom-up approaches based on the granularity of the data used. In the study, two top-down methods, descriptive statistics and artificial neural networks (ANN), were explored for the purpose of benchmarking energy performances of schools. The results were used to understand the benefits of using these benchmarks for assessing energy efficiency of buildings and the limitations that affect the robustness of the derived benchmarks. Compared to the bottom-up approach, top-down approaches were found to be beneficial in gaining insight into how peers perform. The relative rather than absolute feedback on energy efficiency meant that peer pressure was a motivator for improvement. On the other hand, there were limitations with regard to the extent to which the energy efficiency of a building could be accurately assessed using the top-down benchmarks. Moreover, difficulties in acquiring adequate data were identified as a key limitation to using the top-down approach for benchmarking non-domestic buildings. The study suggested that there are benefits in rolling out of DECs to private sector buildings and that there is a need to explore more complex methods to provide more accurate indication of energy efficiency in non-domestic buildings.

© 2014 The Gulf Organisation for Research and Development. Production and hosting by Elsevier B.V.
Open access under CC BY-NC-ND license.

Keywords: Energy benchmark; Energy efficiency; Display Energy Certificate; Non-domestic building; School

* Corresponding author.
E-mail addresses: s.hong@ucl.ac.uk (S.-M. Hong), greig.paterson.10@ucl.ac.uk (G. Paterson), esfandiar.burman.10@ucl.ac.uk (E. Burman), j.p.steadman@ucl.ac.uk (P. Steadman), d.mumovic@ucl.ac.uk (D. Mumovic).

Production and hosting by Elsevier

1. Introduction

There is an imperative to improve energy efficiency of non-domestic buildings owing to global and domestic issues notably energy security and climate change. An approach to achieving this is to operate buildings efficiently and minimise any energy wasted due to inefficiency. In schools, such improvements would lead to cost savings that could be invested for educational purposes.

In the built environment, benchmarking is often employed as part of an energy management practice in existing buildings to assess and improve their energy efficiency. This involves evaluation of the operational energy efficiency of buildings through comparison with standards such as historical energy uses or established energy benchmarks. As a result, establishing and understanding the operational energy efficiency of a building assists and encourages building operators to achieve higher levels of energy efficiency. Benchmarking, therefore, is a technique that is important for achieving higher levels of energy efficiency in non-domestic buildings.

In the UK, benchmarking has gained prominence in recent years when it became part of a mandatory Display Energy Certificate (DEC) scheme under the Energy Performance of Buildings Directive (CIBSE, 2003). Under the directive, public buildings with floor areas greater than 500 m^2 and frequently visited by the public are required to produce a DEC (DCLG, 2012). A key feature of the certificate is the Operational Rating (OR) that indicates how well a building is being operated. The rating is produced based on a comparison of the actual consumption of a building to the benchmarks which represent the typical energy performance of buildings with similar activities (CIBSE, 2009). Having a robust benchmark is, therefore, important in providing building operators with an accurate evaluation of their operational energy efficiency.

In recent years, a study conducted by the Chartered Institution of Building Services Engineers (CIBSE) gave rise to concerns regarding the robustness of the benchmarks that underpin the DEC scheme (Bruhns et al., 2011). The study found that there was a noticeable trend towards higher electricity consumption and lower fossil thermal energy consumption in many benchmark categories. This trend, which was attributed to recent changes in the pattern of energy use of the stock and climate change, suggested that the benchmarks were no longer accurately depicting the pattern of energy use of various types of public buildings in the stock. This, therefore, raised the need to investigate the factors that compromised the robustness of these benchmarks. Moreover, opportunities were presented to explore and compare the benefits and limitations of using various existing methods that are used to benchmark energy performances of non-domestic buildings.

Top-down and bottom-up approaches are two fundamentally different approaches that are used to analyse or design a system in engineering disciplines. A top-down approach refers to ways in which a system is designed by formulating an overview of the system without details of sub-systems. The system would then be refined further subject to availability of more detailed information. A bottom-up approach, on the other hand, would involve specification of lower level system information that would be used to build up a more precise overview. In much the same way, the methods that are used to derive energy benchmarks for buildings can be grouped into these approaches based on the granularity of the information involved in deriving benchmarks.

The top-down approach refers to ways in which energy benchmarks are derived based on building-level energy performance figures. These benchmarks are usually expressed as energy use intensity (EUI)[1] and indicate how other buildings with similar demand use energy. A review has shown that there is a range of methods with varying levels of complexity that are top-down in their nature. A top-down method which is widely used in the UK is to derive energy benchmarks based on descriptive statistics such as 50th and 25th percentiles from a distribution of energy performances of sample buildings (Carbon Trust, 2003; CIBSE, 2012; Hernandez et al., 2008; Jones et al., 2000). Similar methods were also used to assess the energy performance of schools in Argentina and Greece (Filippin, 2000; Santamouris et al., 2007). In recent years, the method was improved through the introduction of procedures to normalise the benchmarks to tailor them to the individual circumstances of buildings in different regions with varying occupancy levels (CIBSE, 2008).

There are top-down methods that employ more complex methods in order to evaluate the operational energy efficiency more precisely. The earliest attempts were made using multiple linear regression models to identify significant determinants of the energy use of buildings in the US (Monts and Blissett, 1982; Sharp, 1996, 1998). The approach now forms the basis of the US Environment Protection Agency's (EPA) Energy Star scheme (Environmental Protection Agency, 2011). In addition, the same approach was used to benchmark the energy efficiency of commercial buildings in Hong Kong (Chung et al., 2006). In recent years, the possibility of using artificial neural networks (ANN) to benchmark energy performances of buildings was explored in the US (Yalcintas and Ozturk, 2007; Yalcintas, 2006). There were also studies that used Data Envelopment Analysis techniques to identify frontiers and use them to identify inefficient buildings (Lee and Kung, 2011; Lee and Lee, 2009; Lee, 2008; Zhou et al., 2008).

The bottom-up approach, on the other hand, refers to ways in which whole-building energy benchmarks are built up by aggregating the system level information. For example, benchmarks for schools would be derived by first estimating energy performance of individual systems, such as a ventilation system. These system level consumption figures would then be aggregated together into a single EUI representing a hypothetical performance of a building. The bottom-up methods are described and explored in detail in the following work by Burman et al. (in press). The extensive nature of the work has led it to be divided into two separate papers.

---

[1] Energy Use Intensity (EUI) is a performance metric that is used to express all energy uses in buildings. In general, annual electrical and fossil-thermal energy uses of buildings are normalised by floor area (kWh/m^2).

Recently, Chung (2011) carried out a comprehensive review of various benchmarking methods that were used to derive energy benchmarks for non-domestic buildings. The strengths and weaknesses of various methods discussed in the study, however, focussed on their mathematical properties. The critical review did not provide insights, therefore, into the benefits and limitations of using various methods in deriving robust benchmarks in a real context, such as a policy framework or constraints on resources.

This study aimed to develop a better understanding of the benefits and limitations of using top-down methods to derive energy performance figures for non-domestic buildings. As part of the study, attempts are made at deriving energy benchmarks using descriptive statistics and ANNs. Schools in England were used as a demonstration case.

The study was carried out in the following steps. Initially, stock-level and building characteristics databases were developed, each with size and granularity that are adequate for the methods. Once the data were prepared, a statistical analysis was carried out to assess the current benchmarks and to derive energy benchmarks based on descriptive statistics. Artificial neural networks were used to explore a multivariate approach for benchmarking. The benefits and limitations of using each of the methods were then discussed. These were also compared to the findings from the bottom-up approaches to draw a more general picture of using different approaches.

2. Methodology

2.1. Deriving energy benchmarks using DEC data

2.1.1. Preparation of data

The energy consumption data for schools in England and Wales were acquired from the most recent set of Display Energy Certificate (DEC) data. The data were acquired from Landmark via CIBSE. It comprised 120,253 DEC records with assessment periods ending between October 2008 and June 2012. To avoid duplication of results with previous studies, only the latest records, which were lodged between mid-February 2010 and June 2012, were extracted for analyses (Bruhns et al., 2011; Godoy-shimizu et al., 2011). Statistical Analysis Software (SAS) 9.3[2] was used to prepare and analyse the data.

To ensure that the analysis is solely based on primary and secondary school records, DEC records that were identified as either 'Primary school', 'State primary school', 'Secondary school' or 'State secondary school' within the 'Schools and seasonal public buildings' benchmark category were used for analyses.

The subset of raw data was then cleaned and filtered so that analyses would yield accurate results. Criteria

developed in a previous study by Hong et al. (2013) were used to identify and remove records, which were considered to be uncertain in their nature. For example, records with default OR[3] or where fossil-thermal EUI was equal to zero were removed. In addition, DEC records from schools that are likely to have unusual patterns of energy use such as electrically heated buildings were excluded from the dataset.

2.1.2. Assessing and deriving benchmarks

As described in Section 3.1, the energy performance of primary and secondary schools was analysed separately due to their intrinsically different patterns of energy use (Hong et al., 2013).

First, it was necessary to prepare the data so that the energy performance figures were comparable to the CIBSE TM46 method in that they are adjusted to standard weather and occupancy conditions. It was therefore necessary to make adjustments and filter the cleaned data to the standard conditions that form the basis of the TM46 benchmarks. The actual fossil-thermal EUI of schools were adjusted to the average UK climate (excluding Scotland) of 2021 heating degree days (Bruhns et al., 2011). The adjustment was made by using the method outlined in CIBSE TM41 which underpins adjustments made to the benchmarks in CIBSE TM46 (CIBSE, 2006, 2009). It should be noted that these adjustments were made based on an assumption that 80% of fossil-thermal energy is generally used for space heating (BRECSU, 1996; Carbon Trust, 2012). This was due to insufficient information on the percentage of energy used for space heating in actual individual buildings.

Once the adjustments were made, the dataset was filtered to retain only those schools that were identified as operating at standard occupancy hours. This was achieved by using the variable 'occupancy level', which identifies whether a school is operating at standard or extended occupancy conditions. The dataset prepared to comparable conditions as set out in TM46 was then used as a basis for deriving benchmarks for primary and secondary schools.

The changes in the pattern of energy use of schools, relative to a dataset composed of DECs lodged between 2008 and 2010 and the CIBSE TM46 benchmarks, were illustrated by using a performance rating rather than the actual EUI. These ratings were produced using the Eq. (1) shown below:

$$\text{Performance rating} = \frac{\text{Actual electrical or fossil-thermal EUI}(\text{kWh/m}^2)}{\text{Adjusted electrical or fossil-thermal benchmarks}(\text{kWh/m}^2)} \times 100$$

(1)

---

[3] An operational rating of 200 is a default rating given when there is insufficient information about energy consumption figures. Such cases are therefore not suitable for the analyses. The default rating was later changed to 9999 in 2010.

The ratings generated in relation to the benchmarks, which were adjusted for climate and occupancy hours of individual circumstances, were intended to raise the accuracy of the comparison between actual consumption and the benchmarks.

In addition, a bootstrapping analysis was carried out to examine the robustness of the benchmarks derived from varying sample sizes. The 7455 electrical EUI values for primary schools were used as a sample. Initially, a normality test, which was conducted using the Kolmogorov–Smirnov method, indicated that the distribution was not normal. This suggested that a percentile interval should be used to assess the confidence of the medians rather than using standard errors, which are typically used for normally distributed samples. The varying sample sizes explored were 5, 10, 20, 40, 80, 160, 320, 640 and 1280. For each sample size, a sampling distribution of medians was plotted from 100 randomly selected bootstrap samples. To examine the accuracy of EUI from 10 records, for example, medians from 100 samples of 10 randomly selected records were used. From the sampling distribution, the median was selected as a central measure and the 2.5th and 97.5th percentiles were used as lower and upper limits of the 95% confidence intervals.

2.2. Exploring a multivariate approach to benchmarking

This section describes how artificial neural networks (ANNs) were used to examine the feasibility of using a multivariate approach to energy benchmarking. An ANN method proposes the potential to account for the influence of intrinsic features of buildings such as shape and age. In this way, buildings with similar characteristics may be more closely comparable.

2.2.1. Building characteristics database

In order to conduct a multivariate analysis, a database of building characteristics is required. Currently, databases that provide information on the built form and architectural features of schools do not exist in the UK. Therefore, an initial step was taken to collect information on built form using a desktop-based approach. The data collection process is outlined in detail by Hong et al. (2013).

Tables 1–3 outline the input and output parametres for the ANN models respectively. The schools selected for analysis were sourced from the DEC database. The annual electricity and heating fuel use (kWh/m^2/annum) figures were used as the output in this study. The following criteria were used to select the school buildings for analysis, ensuring the buildings are comparable with each other:

- The school has a valid DEC
- The school has one main building

- Age and materials of construction use are consistent

The database has data on 502 schools across England. To prevent extreme values from affecting results during the training process, output values that were more than 1.5 interquartile ranges away from the upper and lower quartile figures were removed. The final dataset contained 465 schools.

2.2.2. Artificial Neural Network

Artificial neural networks (ANNs) are machine learning techniques which are a subset of artificial intelligence. They are inspired by the biological neural processes that take place within the brain. Often they are used to learn complex and nonlinear patterns between inputs (e.g. age of building) and outputs (e.g. electricity energy consumption) in a database. Samarasinghe (2007) provides an explanation of neural network concepts and their roles in applied sciences and engineering problems.

The ANN models were constructed using Matlab's Neural Network Toolbox.[4] A feedforward multilayer perceptron method was used for the study – Fig. 1 shows the conceptual structure of this ANN. Each ANN consisted of an input layer, an output layer and a hidden layer. The input layer accepts the building characteristic, such as building age, number of pupils and floor area. The output layer shows the energy consumption prediction associated with a pattern of building characteristics. The hidden layer enables the system to generate nonlinear and complex relationships by intervening between the input and output neurons (Haykin, 1999). Each neuron in the input and output layer took continuous, categorical or binary values as outlined in Tables 1 and 2. Prior to the training of the network, all continuous inputs were normalised to values between −1 and 1 to generalise the calculation process. Synaptic weights connect neurons in adjacent layers. ANN training involves the modification of these weights until the predicted outputs are as close as possible to the actual outputs collected during the data collection process. Two ANN models were constructed, one with heating energy consumption as an output and one with electrical energy consumption as an output.

The accuracy of the ANN method was assessed on the basis of the mean absolute percentage error (MAPE), that is the mean difference between the ANN predictions and the actual energy consumption values.

In order to understand the influence each input parametre has on the predicted output, a study was conducted that tested the change in output as the inputs were altered.

---

[4] For Matlab Neural Network Toolbox, please see: http://www.mathworks.co.uk/products/neural-network/.

Table 1
List of building characteristics collected using the desktop approach.

Input parametre	Input neuron type	Data range/activation criteria	Description
Construction year	Continuous	1860–2010	Year the school was built
Phase of education	Binary	(−1) Primary/elementary, (1) secondary/high school	Primary schools or secondary schools/sixth form colleges
Number of pupils	Continuous	44–2013	Part-time pupils divided by 2, plus the number of full-time pupils
Internal environmental conditioning	Categorical	(−1) Nat. vent, (0) mixed mode, (1) mech. vent	Primary internal environmental conditioning strategy
Site exposure	Categorical	(−1) Exposed, (0) semi-sheltered, (1) sheltered	'Exposed': no obstructions present (4 × the height of the school away); 'semi-exposed': obsts. lower than the school; 'sheltered': obsts. taller than the school
Orientation	Continuous	−45° to +45°	Angle at which the external walls differ from absolute north, south, east and west. Positive angle for clockwise orientations
North façade adjacency	Binary	(−1) Open, (1) obstructed	Obstructed if a building or tree is within 1 × the height of the building from the majority of the façade orientation
South façade adjacency	Binary	(−1) Open, (1) obstructed	See North façade adjacency
East façade adjacency	Binary	(−1) Open, (1) obstructed	See North façade adjacency
West façade adjacency	Binary	(−1) Open, (1) obstructed	See North façade adjacency
Floor area	Continuous	861 m^2–15396 m^2	Gross internal area (GIA)
Building depth ratio	Continuous	2.50–16.60	Building volume/exposed external wall area
Compactness ratio	Continuous	1.01–4.59	Perimetre of the building footprint/perimetre of a circle with the same area as the building footprint
Surface exposure ratio	Continuous	1.71–5.67	Building volume/exposed surface area

Table 2
Collected building characteristics (ANN inputs).

Input parametre	Input neuron type	Data range/activation criteria	Description
North glazing ratio	Continuous	0.00–0.13	Glazed area on the north façade/total floor area
South glazing ratio	Continuous	0.00–0.15	Glazed area on the south façade/total floor area
East glazing ratio	Continuous	0.00–0.11	Glazed area on the east façade/total floor area
West glazing ratio	Continuous	0.00–0.14	Glazed area on the west façade/total floor area
Glazing type	Binary	(−1) Single, (1) double	Single or double/secondary glazing
Roof shape	Binary	(−1) Pitched, (1) flat	Pitched or flat roof
Roof glazing	Binary	(−1) None, (1) glazing	Existence of any roof glazing
Heating degree days	Continuous	1635.6–2843.3	Heating degree days during the DEC monitoring period
Cooling degree days	Continuous	73.9–425.2	Cooling degree days during the DEC monitoring period

Table 3
Energy consumption outputs.

Output	Output neuron type	Data range	Description
Heating energy consumption	Continuous	7–272 kWh/m^2/annum	Annual heating fuel use
Electricity energy consumption	Continuous	7–95 kWh/m^2/annum	Annual electricity fuel use

The process was as follows:

– Across the input patterns, the mean values of all continuous inputs and the mode values for all binary or categorical inputs were set to form a base-case ANN configuration.
– For one input at a time, the normalised values of the input were set to their extremes across their range, -1 then 1, and the two outputs calculated. All other inputs remained in their base-case condition as each individual input was altered.
– The change in output across the range in each input was recorded and compared against the base-case ANN outputs.

The ANN training, testing and analysis process used in this study is outlined in detail in a study by Hong et al. (2013).
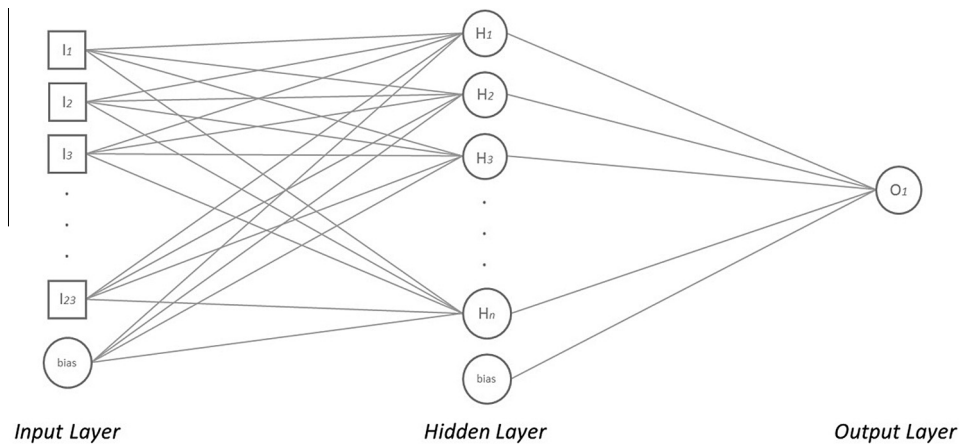
Figure 1. Conceptual structure of the ANN used in this study.

3. Results

3.1. Robustness of statistical energy benchmarks for schools

Table 4 shows 10th, 25th and 50th percentiles of the latest energy performance figures for schools in England presented in the form of EUI.

The 25th and 50th percentiles indicate the performance of primary and secondary schools that are commonly considered to exemplify 'good practice' and 'typical' performance, respectively. The robustness of using these descriptive statistics as benchmarks comes from the simplicity and effectiveness of the method in describing the distribution of the latest pattern of energy use in the sample of schools in England. Comparing the annual energy performance of a primary school to the corresponding statistics would provide a good indication of how energy-intensive it is relative to its peers that have similar demands. For example, an electrical EUI of a primary school building that is less than the 50th percentile would indicate that the building is less intensive in electricity consumption than at least half of the school buildings in the sample, or more broadly the school stock. Similarly, achieving energy performances similar to the 25th percentile would indicate that a building is likely to be more energy-efficient than the majority of school buildings in England given its low EUI.

There is, on the other hand, a limitation in using these simple indicators to assess the operational energy efficiency of buildings. This is that feedback from using these benchmarks is not likely to indicate precisely whether a building is being operated efficiently or not. This is largely due to the low granularity of the data that were used to derive these benchmarks, which are usually based on minimal information about buildings such as the annual energy performance, floor area and regional weather conditions. This means that varying implications of intrinsic features of buildings such as the built form or age on the pattern of energy demand are not accounted for (Hong et al., 2013). A comparison without a way of incorporating these characteristics, therefore, is more likely to indicate buildings that are more or less intensive in energy use but not necessarily with regard to their efficiency. A newly built school whose EUI is close to the 10th percentile, for example, would suggest that it is likely to be very energy-efficient based on the fact that the EUI is lower than 90% of the buildings in the stock. This, however, would ignore the fact that the building would have been built to much higher standards, therefore have intrinsically less demand for energy use than Victorian buildings built more than a hundred years ago.

Table 4 also shows a comparison of the statistics from the latest data to the previous CIBSE benchmarks. This

Table 4

Energy consumption statistics by type of fuel for primary and secondary schools (Feb 2010–Jun 2012).

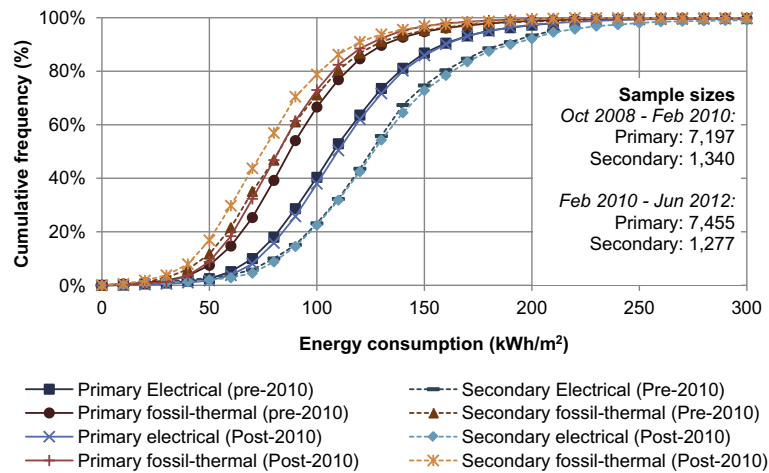| Phase of education	N	%	Electrical EUI (kWh/m^2)			Weather-corrected fossil-thermal EUI (kWh/m^2)		
			10th%	25th%	50th%	10th%	25th%	50th%
Latest data (2010–2012)								
– Primary	7455	85	29	35	43	76	96	121
– Secondary	1277	15	32	41	50	64	83	111
CIBSE TM46	–	–	–	–	40	–	–	150
CIBSE Guide F								
– Primary	–	–	–	22	32	–	113	164
– Secondary	–	–	–	25	33	–	108	144

Figure 2. Cumulative frequency distribution curves of electricity and fossil-thermal performance ratings based on data collected from 2008 to 2010 and 2010 to 2012.

indicates that schools are now much more intensive in electricity use and less intensive in heating consumption than they were more than a decade ago. This is likely due to the fact that the energy data that formed the basis of these benchmarks were mainly acquired from earlier research conducted in the late 1990's (CIBSE, 2008). As discussed in previous studies, schools are likely to have had different demands for energy owing to changes in various factors such as developments in technologies and their rate of uptake, and changes to the building regulations (Global Action Plan, 2006; Godoy-shimizu et al., 2011; Hong et al., 2013). Recent changes in the pattern of energy use are more clearly shown in Fig. 2, which plots the distributions of performance ratings for electricity and fossil-thermal energy consumption of primary and secondary schools.

Taking the performance rating of 100 as a basis of comparison with the CIBSE TM46 benchmarks, the figure shows that the trends of higher electricity consumption and lower fossil-thermal energy use, which were observed previously, persist (Bruhns et al., 2011). This means that evaluating the energy performance of schools using these benchmarks is not likely to provide a useful feedback to building operators, as it is not an accurate indication of how schools use energy today. This, therefore, highlights that having robust data, which accurately depicts the most recent trends in energy use, is a key to delivering robust energy benchmarks based on descriptive statistics.

Fig. 3 shows the medians derived from sampling distributions of electrical EUI of primary schools with varying sample sizes. The vertical bars indicate the 95% confidence intervals of each statistic based on the 2.5th and 97.5th percentiles.

It can be seen that the fluctuation in the line connecting the medians from varying sample sizes becomes almost flat as the sample size exceeds 300. What is more, there is a dramatic decrease in the confidence intervals as the sample size increases, which suggests significant improvements in
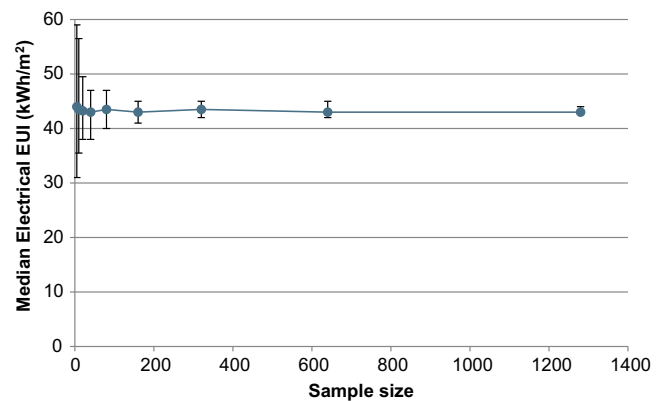

Figure 3. Changes in median electrical EUI of primary schools and the corresponding confidence intervals derived from varying sample sizes.

accuracy. The figure however also shows that the rate at which the confidence intervals decrease becomes very small as the sample size becomes larger. This shows that the accuracy of energy benchmarks derived using descriptive statistics is highly dependent on the size of the sample from which they were derived. Moreover, it is shown that the median electrical EUI of 43 kWh/m^2 presented in Table 4 is highly likely to be an accurate description of the typical performance of a primary school relative to the school population.

Lastly, an aspect of the data that should also be noted is that the sample used to produce the statistics in this section is mostly from schools with floor areas greater than 1000 m^2. As mentioned previously, this is due to the threshold for DECs that has only recently been reduced to 500 m^2. This, therefore, means that the statistics presented in Table 4 do not account for the patterns of energy use of school buildings that are smaller than the threshold, which may be different. This highlights the importance of having a sample that is representative of all the buildings in the stock. It raises the need to investigate the pattern of energy

use of buildings that are less than 1000 m^2 but greater than 500 m^2 once the DECs from these buildings become available.

3.2. Artificial Neural Network approach

The best performing ANN models produced mean absolute percentage errors (MAPE) of 22.0% and 20.6% for the prediction of heating and electricity energy consumption, respectively. That is, on average the prediction of the ANN was 22.0% and 20.6% from the actual heating and electrical energy consumption values when tested on a dataset not used in the ANN training process. A study by Hawkins et al. (2012), which used ANN to identify determinants of energy use of UK higher education buildings, showed MAPEs of 25% and 34% for heating and electricity, respectively. The errors in this paper are therefore 3% and 13.4% lower for heating and electricity respectively than the aforementioned study.

When benchmarking using this method, the performance of each school is compared with the ANN outputs.

Therefore the robustness of this approach is dependent on the accuracy of the model. Currently, there is a lack of building services data, such as boiler efficiencies. Building services are likely to have a significant influence on the energy performance of schools and therefore without such data, the prediction accuracy of such a method may be restricted. Unlike the US, which has the CBECS database (Energy Information Administration (EIA), n.d.), acquiring such detailed information for a large number of school buildings to identify the pattern of energy use and, therefore, identify determinants of energy use requires considerable time and resources. Furthermore, without a central database, the ANN benchmarking method is unable to adapt over time. That is, the benchmarking algorithm is unable to renew itself as new building data are entered into a central database as was done in a study by Yalcintas (2006).

Figs. 4 and 5 show the changes in output values across the range of each input when compared to the base-case output values. Larger changes in output indicate greater influence of the input on the output.
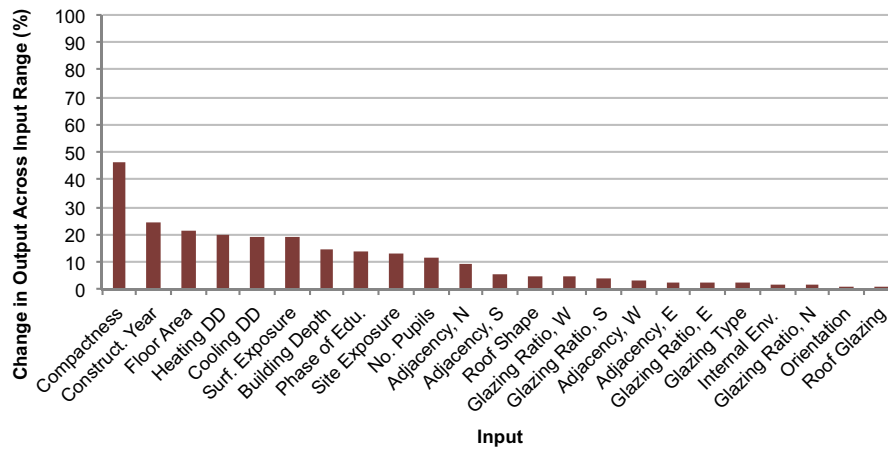

Figure 4. Change in output across input range for the heating energy consumption output (Hong et al., 2013).
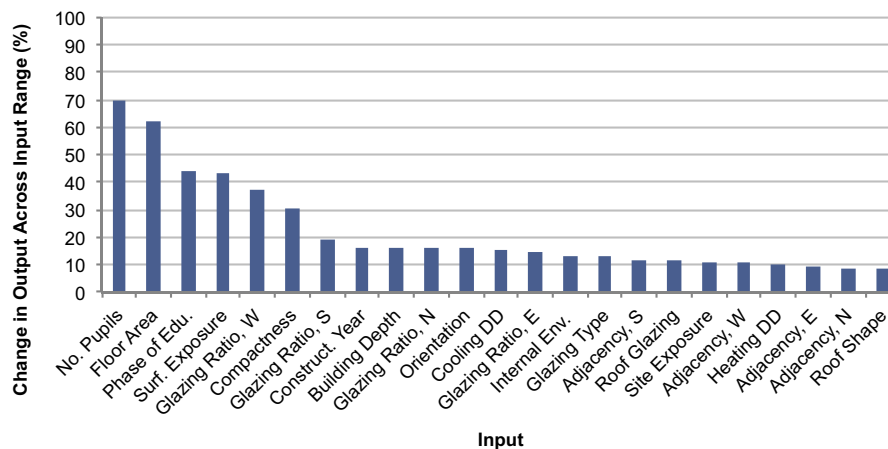

Figure 5. Change in output across the input range for the electricity energy consumption output (Hong et al., 2013).

The above analysis gives an indication of the influence different building characteristics have on heating and electricity energy consumption. For example, how compact a building is and how many pupils a school has are found to have the greatest impact on heating and electricity energy consumption respectively. This shows that schools with a longer perimetre relative to their floor area have greater heat loss through external walls, and therefore use more fossil-thermal energy. Conversely, the number of pupils (occupant density) is likely to have a significant impact on electricity energy consumption, due to the resulting increased use of electrical equipment. It should be noted, however, that ANNs are a machine learning technique that recognise patterns in data. For example, 414 buildings of the 452 school buildings in the ANN dataset were naturally ventilated, which may have affected the influence of the Internal Environment variable on the outputs. The dataset and results are discussed in further detail in the study by Hong et al. (2013).

When the energy performance of a building is benchmarked using this method, the energy performance estimated by the ANNs would be used as benchmarks. These performance figures would be representative of the energy performance of a typical building that has similar characteristics to the building of interest. This would mean that the varying influences of building characteristics recognised by the ANNs are taken into account. Therefore comparing the annual energy performance of a school to the ANN prediction would provide an indication of how energy-efficient the building is likely to be relative to peers with similar building characteristics.

4. Discussion

A summary of the benefits and limitations of applying the respective top-down methods to derive energy benchmarks is shown below (Table 5).

4.1. Type of feedback

A fundamental difference between the top-down and bottom-up approaches was identified as the different perspectives that could be acquired through each approach.

As shown in Table 5, both top-down methods were found to provide assessments of energy efficiency relative to the energy performance of other buildings with similar operating conditions, which in this case were schools in England. These benchmarks therefore presented opportunities for building operators or managers to put their buildings' performance into a broader context. These benchmarks, for example, would be effective in identifying how efficient a given group of buildings is in relation to the stock. With regard to schools, such feedback would be beneficial for local authorities or county councils who have energy efficiency as part of their agenda. Moreover, such feedback would provide motives for improving energy efficiency of buildings based on peer pressure rather than absolute levels of energy efficiency. For other building types such as commercial offices, where reputation is of crucial value, such peer-driven feedback may generate stronger motives to improve energy efficiency.

The drawback of the top-down methods is, however, that there is a limitation as to how accurately one can evaluate the operational energy efficiency of buildings. This is largely due to the low granularity of the information involved in benchmarking energy performance. Unlike the bottom-up approach, which accounts for detailed information on a building's fabric, services and occupancy, comparisons of simple headline figures are more likely to indicate that a building is using more or less energy than others rather than indicating how efficient it is. As shown in Table 5, a more complex approach, which was demonstrated through ANNs, has shown possibilities for obtaining a more accurate indication of energy efficiency through inclusion of intrinsic features of buildings, such as their shape and age, which determine the demand for energy. The forthcoming study by Burman et al. (in press), however, suggests that the like-for-like comparison of performances through a bottom-up approach would provide a more precise indication of energy-efficiency.

4.2. Availability of adequate data

The study has shown that availability of adequate data is crucial in adopting top-down methods to derive robust benchmarks. Compared to the bottom-up approaches,

Table 5
Benefits and limitations of using the two top-down methods to derive robust energy benchmarks.

Method	Benefits	Limitations
Descriptive statistics	• Establish energy performance relative to other primary and secondary schools in England • Minimum data required per building (energy consumption and floor area) • Simplicity of the method and effective description of the energy performance of the stock	• Insufficient in accurately evaluating how efficiently a building is being managed • Large, well-distributed and up-to-date data are required to derive robust benchmarks
Artificial neural networks (ANN)	• Comparison to energy performances of a typical building with the same features • Opportunities to identify and take into account sector specific determinants for benchmarking • Improved accuracy in evaluating operational efficiency	• Complexity associated with training and optimising the model • No existing central database of building characteristics

which require much finer granularity of data, low granularity data such as the annual energy consumption of a building is more likely to be obtainable through utility bills or regular metre readings. Similarly, obtaining more detailed information on a building for more complex methods would be relatively less intensive in resources than the bottom-up approaches, which often require resource-intensive activities such as post occupancy evaluation (Burman et al., in press).

The challenge in using top-down methods, however, lies with the fact that it is often difficult to obtain such data in sufficient quantity, that are up-to-date and for a group of buildings that reasonably represent the stock (Table 5). This is often determined by policies or frameworks that enable monitoring and collection of energy performance data from non-domestic buildings. Taking the UK as an example, it is only in recent years that the energy performance of such a large number of schools has become available. This is due to the implementation of the DEC scheme, which enforced the monitoring of the actual energy performance of non-domestic buildings and established a central database to collect data in a systematic manner.

Currently, however, the DEC scheme is applicable only to buildings that are occupied by public authorities and frequently visited by the public. This means that data from building types that are found in the private sector such as commercial offices, retail and industrial buildings, are not in general collected through the scheme. Moreover, the low granularity of information involved in producing DECs would mean that the information in the database is not likely to be sufficient to adopt more complex methods, which could improve the accuracy beyond descriptive statistics. To supply sufficient data for complex methods would require a more comprehensive framework comparable to the US Energy Information Administration's (EIA) Commercial Building Energy Consumption Survey (CBECS) (EIA, n.d.). In the UK, CarbonBuzz,[5] an online platform for collection of detailed building information, has the potential to collect and deliver such data on a large scale. However, the platform is still in its early development and is unlikely to provide sufficient data in the near future. This highlights the need to maximise the potential of the existing schemes such as DECs and CarbonBuzz. One such way would be to roll out the DECs to buildings in the private sector. The possibility of extending DECs to commercial buildings was first mooted by the Department for Business, Innovation and Skills (BIS), and has since gathered momentum with a commitment by HM Government in The Carbon Plan to 'extend Display Energy Certificates to commercial buildings'. This commitment has subsequently been reversed although lobbying continues (BIS, 2010; Gardiner and Lane, 2013; HM Government, 2011).

5. Conclusions

This study explored various approaches to deriving energy benchmarks for non-domestic buildings with the aim of gaining insights into the benefits of using each approach for deriving benchmarks as well as limitations that affect the robustness of those benchmarks. Schools in England were used as a demonstration. Existing benchmarking methods were grouped into top-down and bottom-up approaches, based on the ways in which the benchmarks are derived. This involved using descriptive statistics to derive typical and good practice benchmarks from a distribution of energy performance values, and artificial neural networks (ANNs) based on an analysis of the factors that contribute towards energy consumption. These methods were based on two different types of data including a stock-level DEC dataset of approximately 14,000 primary and secondary schools and a building characteristics dataset containing approximately 500 schools.

The first set of analyses described the latest pattern of energy use in primary and secondary schools in England and Wales using descriptive statistics. 10th, 25th and 50th percentiles were derived from the distribution of electrical and weather-corrected fossil-thermal EUIs. Although these figures are an effective way to describe how the wider population of schools perform, it was found that they offer limited insight as to how energy-efficient a building is. Comparisons of the distribution of the latest data against the previous benchmarks reaffirmed the findings from previous analyses made in 2011 that the current UK benchmarks for schools no longer accurately reflect the pattern of energy use. This was more clearly shown from a comparison of the distribution of energy performance values for schools based on the latest and the previously analysed datasets. Moreover, it also showed that the disparity between electricity consumption and heating consumption relative to the benchmarks, which was found in the previous analyses, persists in both primary and secondary schools. These therefore showed that quality of data, in particular how up-to-date it is and how well it represents the school population plays a crucial role in deriving robust energy benchmarks using descriptive statistics.

ANNs were trained in order to create a model that can predict the energy consumption of a typical school given a set of specific building characteristics. For energy consumption predictions, the ANN mean absolute percentage error (MAPE) was 22% for heating and 20.6% for electricity. It is desirable to increase the predictive accuracy of this method. In order to achieve this, data on building services are likely to be required. Additionally, an analysis was carried out looking at the influence a range of building characteristics had on heating and electricity energy consumption. This showed that some factors have a significant impact, such as compactness of plan for heating energy consumption and number of pupils for electricity energy consumption. The predictions from ANNs therefore account for the varying influences of building

---

[5] For CarbonBuzz, please use: http://www.carbonbuzz.org/.

characteristics on energy consumption, and when compared to actual performance, give an indication of how energy-efficient the building is relative to a typical building with similar characteristics.

An assessment of the results from the two separate analyses has highlighted that a key benefit of using the top-down approach is that it provides an opportunity to put the energy performance of a building into a broader context. Using a more complex approach through ANNs showed the potential for assessing the operational energy efficiency of buildings more accurately than the simple top-down method. Comparison with the bottom-up methods, however, suggested that the top-down methods were less precise. In addition, the availability of robust data was found to be an important factor in deriving robust benchmarks using top-down approaches. This highlighted the importance of developing, continuously refining and endorsing the policies and frameworks that enable the systematic monitoring and collection of robust data of varying granularity.

In summary, the comparison of the benefits and limitations of using top-down and bottom-up approaches has shown that different methods should not be treated separately but rather in combination to maximise the benefits in identifying and improving the energy efficiency of non-domestic buildings. This raises the possibility of conducting further studies to explore ways in which the top-down and bottom-up approaches could be combined.

Acknowledgements

References

Department for Business Innovation and Skills (BIS), 2010. Low Carbon Construction Executive Summary. Available from: <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/31774/10-1266es-low-carbon-construction-igt-executive-summary.pdf> (retrieved August 14, 2011).

BRECSU, 1996. Energy Consumption Guide 73: Saving energy in schools. Building Research Energy Conservation Support Unit.

Bruhns, H., Jones, P., Cohen, R., 2011. CIBSE Review of Energy Benchmarks for Display Energy Certificates – Analysis of DEC results to date. Available from: <http://www.cibse.org/content/Technical_Resources/Technical_Reports/Technical Report_CIBSE Report on 45000 DECs.pdf> (retrieved June 11, 2011).

Burman, E., Hong, S.-M., Paterson, G., Kimpian, J., Mumovic, D., in press. A comparative study of benchmarking approaches for non-domestic buildings: part II – bottom-up methods.

Carbon Trust, 2003. Energy Consumption Guide 19. Carbon Trust.

Carbon Trust, 2012. Schools Sector Overview (CTV019): Learning to Improve Energy Efficiency. United Kingdom. Available from: <http://www.carbontrust.com/media/39232/ctv019_schools.pdf> (retrieved January 17, 2013).

Chung, W., 2011. Review of building energy-use performance benchmarking methodologies. Appl. Energy 88 (5), 1470–1479.

Chung, W., Hui, Y., Lam, Y., 2006. Benchmarking the energy efficiency of commercial buildings. Appl. Energy 83 (1), 1–14.

CIBSE, 2003. CIBSE Briefing 6: The Energy Performance of Buildings Directive. Chartered Institution of Building Services Engineers.

CIBSE, 2006. CIBSE TM41: Degree-days: Theory and Application. CIBSE Publications.

CIBSE, 2008. CIBSE TM46 Energy Benchmarks. CIBSE Publications.

CIBSE, 2009. CIBSE TM47 Operational Ratings and Display Energy Certificates. CIBSE Publications.

CIBSE, 2012. CIBSE Guide F: Energy EFFICIENCY in BUILDINGS, third ed. CIBSE Publications.

DCLG, 2012. Improving the Energy Efficiency of Our Buildings: A Guide to Display Energy Certificates and Advisory Reports for Public Buildings. Available from: <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/51164/A_guide_to_display_energy_certificates_and_advisory_reports_for_public_buildings.pdf> (retrieved July 24, 2013).

Energy Information Administration (EIA), n.d. Energy Information Administration (EIA)-Commercial Buildings Energy Consumption Survey (CBECS). Available from: <http://www.eia.gov/consumption/commercial/> (retrieved June 10, 2012).

Environmental Protection Agency, 2011. ENERGY STAR® Performance Ratings Technical Methodology, Methodology. Available from: <http://www.energystar.gov/ia/business/evaluate_performance/General_Overview_tech_methodology.pdf> (retrieved June 08, 2011).

Filippin, C., 2000. Benchmarking the energy efficiency and greenhouse gases emissions of school buildings in central Argentina. Build. Environ. 35, 407–414.

Gardiner, J., Lane, T. 2013. Green for Growth. Building Magazine. Available from: <http://www.building.co.uk/story.aspx?storyCode=5048252> (retrieved February 17, 2013).

Global Action Plan, 2006. UK Schools Carbon Footprint Scoping Study. Stockholm Environment Institute. Available from: <http://www.sd-commission.org.uk/data/files/publications/GAP-Final-Report.pdf> (retrieved July 31, 2013).

Godoy-shimizu, D., Armitage, P., Steemers, K., Chenvidyakarn, T., 2011. Using Display Energy Certificates to quantify schools energy consumption. Build. Res. Inf 39 (6), 37–41.

Hawkins, D., Hong, S.M., Raslan, R., Mumovic, D., Hanna, S., 2012. Determinants of energy use in UK higher education buildings using statistical and artificial neural network methods. Int. J. Sustainable Built Environ. 1 (1), 50–63.

Haykin, S., 1999. Neural networks: a comprehensive foundation. In: Griffin, J. (Ed.), The Knowledge Engineering Review, 13. Prentice Hall, pp. 409–412.

Hernandez, P., Burke, K., Lewis, J.O., 2008. Development of energy performance benchmarks and building energy ratings for non-domestic buildings: an example for Irish primary schools. Energy and Buildings 40 (3), 249–254.

HM Government, 2011. The Carbon Plan. Change, pp. 1–83. Available from: <http://www.decc.gov.uk/en/content/cms/what_we_do/lc_uk/carbon_plan/carbon_plan.aspx>.

Hong, S.-M., Paterson, G., Mumovic, D., Steadman, P., 2013. Improved benchmarking comparability for energy consumption in schools. Build. Res. Inf 42 (01), 47–61.

Jones, P., Turner, R., Browne, D., Illingworth, P., 2000. Energy benchmarks for public sector buildings in Northern Ireland. In: Proceedings of CIBSE National Conference, Dublin, pp. 1–8.

Lee, W., 2008. Benchmarking the energy efficiency of government buildings with data envelopment analysis. Energy Build. 40 (5), 891–895.

Lee, W.-S., Kung, C.-K., 2011. Using climate classification to evaluate building energy performance. Energy 36 (3), 1797–1801.

Lee, W.-S., Lee, K.-P., 2009. Benchmarking the performance of building energy management using data envelopment analysis. Appl. Therm. Eng. 29 (16), 3269–3273.

Monts, J., Blissett, M., 1982. Assessing energy efficiency and energy conservation potential among commercial buildings: a statistical approach. Energy 7 (10), 861–869.

Samarasinghe, S., 2007. Neural Networks for Applied Sciences and Engineering: From Fundamentals to Complex Pattern Recognition. Auerbach Publications, Boca Raton.

Santamouris, M., Mihalakakou, G., Patargias, P., Gaitani, N., Sfakia-naki, K., Papaglastra, M., Zerefos, S., 2007. Using intelligent clustering techniques to classify the energy performance of school buildings. Energy Build. 39 (1), 45–51.

Sharp, T., 1996. Energy benchmarking in commercial office buildings. In: Proceedings of the ACEEE 1996 Summer Study on Energy Efficiency in Buildings, vol. 4, pp. 321–29.

Sharp, T., 1998. Benchmarking energy use in schools. In: ACEEE 1998 Summer Study on Energy Efficiency in Buildings. ACEE, USA, pp. 305–316.

Yalcintas, M., 2006. An energy benchmarking model based on artificial neural network method with a case example for tropical climates. Int. J. Energy Res. 30 (14), 1158–1174.

Yalcintas, M., Ozturk, U.A., 2007. An energy benchmarking model based on artificial neural network method utilizing US Commercial Buildings Energy Consumption Survey (CBECS) database. Int. J. Energy Res. 31, 412–421 (April 2005).

Zhou, P., Ang, B.W., Poh, K.L., 2008. A survey of data envelopment analysis in energy and environmental studies. Eur. J. Oper. Res. 189 (1), 1–18.