Using machine learning and mathematical programming to benchmark energy efficiency of buildings

Alexander Severinsen
Data scientist, Storekeeper AS
Thunes vei 2, Skoyen
0021 Oslo
Norway
as@storekeeper.no

Helen Marita Sorensen Holst
Phd scholar/lecturer, The Arctic University of Norway
9019 Tromso
Norway
helen.m.holst@uit.no

Keywords

energy analysis, energy saving methodology, energy saving potential, benchmarking, energy savings calculation, machine learning, mathematical programming

Abstract

In this paper, we demonstrate a novel benchmarking technique to identify buildings with a potential to reduce energy consumption, taking into account both weather effects and building characteristics. The proposed method will quantify if a building's energy usage is dependent on outside temperature, and the degree of dependency over time. Our method prepares the data for analysis in a way that can improve the results from benchmarking techniques. A demonstration is performed with the use of two different data envelopment analysis (DEA) models. The first DEA model without taking into account temperature, and a second where temperature is included as a non-controllable variable. Results suggest that the that the average store is 28 % less efficient than the most efficient stores. Further, our analysis show that using our suggested analytical framework will improve the accuracy of the efficiency scores compared to more standard methods. This is an important finding suggesting that our proposed methodology has advantages over existing benchmarking methods.

Introduction

Somewhere between 30 and 40 % of the global energy consumption occurs in buildings (United Nations Environment Programme, 2007). Thus, buildings represent an important opportunity to reduce energy consumption, and further help mitigate global warming, one of the world's most important problems. In recent years, researchers have proposed a number of methods to analyze and improve energy efficiency in buildings. A number of data driven approaches have been developed (Wang et al. (2012), Chung (2011), Lee and Kung (2011)). While simulation approaches have high accuracy, the data driven approaches have the advantage of being able to analyze a large number of buildings, and at the same time consider multiple parameters; both weather conditions and buildings characteristics (Lee and Kung, 2011). Some data driven studies use regression models to to evaluate energy efficiency of both buildings (Lee (2010), Lee and Lin (2011a), Lee and Lin (2011b), Chung (2012), Kavousian and Rajagopal (2014)), schools (Hong et al (2014)) and guest rooms (AlFaris et al (2016)). In this paper we use data envelopment analysis (DEA) to study the most energy efficient buildings. Nonparametric DEA is a mathematical programming technique used to find an efficiency frontier which consists of the most efficient buildings. The DEA methodology has been applied to benchmarking energy efficiency in buildings (Lee (2008), Lee and Lee (2009), Lee and Kung (2011), Wang et al (2015)), in hotels (Önüt and Soner (2006)), in university departments (Tu (2015)), of technical equipment (Du, Jin and Fan (2015), Blum (2015)), on an industrial level (Blomberg et al (2012), Al-Refaie et al (2016)) and on a more regional level (Wang et al (2013)). DEA is emerging as one of the leading energy benchmarking methods (Lee and Lee (2009), Grösche (2009), Önüt and Soner (2006), Lu et al. (2014), Chung(2011), Borgstein et al (2016)). Further, for example, Lee and Lee (2009) use DEA together with climate adjusted energy consumption. They analyze 47 government buildings in Taiwan, and find the energy performance to vary between 31 % and 100 %, where the average was 65% energy efficiency. However, while an important contribution to further the understanding of what drives energy usage in government buildings; because of the lack of detailed daily temperature data, they had to incorporate monthly average temperatures in the model. Using daily temperature data could possibly improve the accuracy of the result. Next, in a more recent publication, Wang et al (2015) use data envelopment analysis with a degree-day based simple normalization approach. To normalize the climate effect, the degree-day method has typically been used (Ferrier and Hirschberg, 1992). For example, the cumulative variation of temperatures in a period can be expressed as number of degree-days through eq 1. (Eto, 1988).

$$\text{Degree day} = \sum(A-B) \quad (1)$$

Where A = average daily outdoor air temperature, and B = changing point temperature (CPT); the temperature at which no heating is required. Earlier studies (including Wang et al. 2015) use standard published values for the CPT; in the U.S., they use 18 °C. However, each store might have a separate CPT, and to increase the accuracy of degree-days it is an advantage to calculate each building's individual CPT. Historically, lack of data has made this very difficult. However, today, detailed energy and climate data is readily available. In Norway, weather data can easily be accessed online through met.no and setting up an API that downloads data at request (sometimes even with hourly temperature data) from a building's closest weather station is relatively easy. In addition, increased use of smart meters have made energy data both more available (often through online solutions at request) and at a more detailed interval, normally kWh. Our approach takes advantage of these data, and to the best of our knowledge, we are not aware of any studies that uses a similar approach.

The remainder of the paper is structured as follows. In the methodology section we explain in more detail how regression analysis is used as a technique to prepare data for data envelopment analysis (DEA), and further the idea behind DEA is explained in more detail. We then present the results, discuss the implications of the results and give some further research suggestions, at last the conclusion.

Methodology

The analysis starts with standard linear regression to analyze the relationship between daily average outside temperatures and daily energy use for 132 Norwegian retail stores. If temperature is not found statistically significant, no adjustment for weather effects needs to be undertaken. However, if temperature is found to be an important variable to explain energy use we need to quantify this. The quantification is done with piecewise linear regression to estimate the changing point temperature (CPT) for each of the buildings. Once the CPT is established we automatically collect the average daily temperature from each store's closest weather station, and finally we can use the CPT and the outside temperature to calculate degree-days. This is in contrast to using a fixed CPT for all stores, which is the traditional approach. At this point we have an accurate understanding of the impact temperature has on energy usage. This, together with data about building characteristics can increase the accuracy of the results from energy benchmarking techniques like data envelopment analysis (DEA). More specific, our approach consist of the following four steps to benchmark energy efficiency of buildings:

1. Collect meter data for all the 132 buildings in the analysis (done automatically through a web-service directly to a database). The meter data is on hourly basis, but we aggregate them into daily energy usage. The reason for this is that temperature data is on a daily basis. We further use each stores longitude and latitude together with the corresponding longitude and latitude for every Norwegian weather station, and map the closest weather station to the store; then download the actual daily average temperature. Data is collected for the year 2015.
2. We further perform a linear regression with the data (at a daily level). One regression model for each store. Energy consumption is used as a dependent variable, and temperature as the independent variable. We then collect R^2 (a measure of how many percent of energy usage can be explained by temperature) and the corresponding p-value (for testing if temperature is statistical significant, we use $p >= 0.01$ to reject any relationship between energy usage and temperature).
3. In this step we find each buildings individual changing point temperature (CPT), following the approach of Day et al. (2003). Traditionally, because of lack of data, a fixed value has been used (in the U.S. 18 °C). However, buildings vary in terms of heating schedule, thermal properties, and solar gains. To find the CPT we apply piecewise linear regression (PLR), a method that allow us to detect breakpoints in a time series (Hansen, 2000) – in this case the breakpoint indicates the outside temperature when a building start to use more energy as the temperature decrease. The PLR model is an iterative routine that search through all possible breakpoints for a particular building, and the breakpoint that "wins" is the model with the lowest residual MSE (mean square error). It is an efficient machine learning technique. This routine is then implemented for all the buildings, resulting in 132 individual CPT; at last we use eq.1 (above) with the appropriate daily temperature (from the closest weather station) to calculate the degree-days. This is done programmatically through a program that loops through all the data and stores the results in a data file.
4. The data is now aggregated from a daily to a yearly figure for benchmarking purposes. We now have yearly energy usage and yearly heating degree-days (based on the results from a regression model and a PLR model) At last, we join this with building characteristics such as size ($m^{2)}$ and run-time (number of hours the store is open). Then data envelopment analysis (DEA) is conducted and the results are presented and discussed.

In Figure 1 our conceptual model of the benchmarking framework is presented.

The efficiency scores are calculated based on data envelopment analysis (DEA). DEA is a non-parametric method first introduced by Charnes, Cooper and Rhodes (CCR), (1978). DEA measures the relative efficiency between homogeneous units and estimates a composite score for each unit under consideration. Efficient units will obtain a score of 1 (100 %), while inefficient units will receive a score less than one but greater than zero. The objective is to minimize input (maximize output) holding output (input) fixed. The method calculates efficient utilization of resources by applying mathematical programming and has gained popularity as a well-suited tool for benchmarking, within both operations research and business economics. One of the advantages of the method is the ability to incorporate multiple inputs and multiple outputs. DEA, unlike parametric methods, do not require an a priori functional form. The main disadvantage is its sensitivity to outliers and a thorough investigation of possible outliers is necessary for a reliable result.

The CCR model assumes constant returns to scale, allowing possible scaling of units in the analysis. This implies allowing units of different size to be compared. Banker, Charnes and Cooper (BCC), (1984) further developed the model to account for variable returns to scale. Assuming variable return to scale, the BCC model ensures that units in the analysis will be compared to other units of similar size. Banker and Morey (1986) developed a model for incorporating exogenously fixed or non-controllable variables as some variables affecting the production process may be outside the control of individual units.

Investigating energy efficiency, the term decision-making unit (DMU) represents the different retail stores. Considering a set of n observations of DMUs, each $DMU_j (j \in n)$, uses m_C controllable inputs $x_{ij} (i \in m_{NC})$ to produce s outputs $y_{rj} (r \in s)$, affected by m_{NC} non-controllable inputs. By creating a piecewise linear approximation DEA determines an efficient frontier or best practice frontier by these n observations and returns an efficiency estimate, θ_0, for each DMU_0. The model is specified for DMU_0 as follows:

$$\theta^* = Min\ \theta_0 \qquad (2)$$

$$\text{s.t.}\ \ \theta_0 x_{i0} \geq \sum_{j=1}^{n} \lambda_j x_{ij}\,, \quad i \in m_C \qquad (3)$$

$$x_{i0} \geq \sum_{j=1}^{n} \lambda_j x_{ij}\,, \quad i \in m_{NC}$$

$$y_{r0} \leq \sum_{j=1}^{n} \lambda_j x_{rj}\,, \quad r \in s$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geq 0, \qquad j \in n$$

The LP problem in eq.2 is an input-oriented variable returns to scale (VRS) minimizing controllable inputs whilst holding non-controllable inputs and output at a constant level. Using the VRS model the additional restriction, $\sum_{j=1}^{n} \lambda_j = 1$, is included compared to the constant returns to scale (CRS) model in Charnes, Cooper et al. (1978). This allows for comparing the DMU under consideration (DMU_0) to a similar DMU(s), of approximately the same size.

Figure 2 shows a conceptual illustration of the DEA model compared to a standard ordinary least regression (OLS) model.

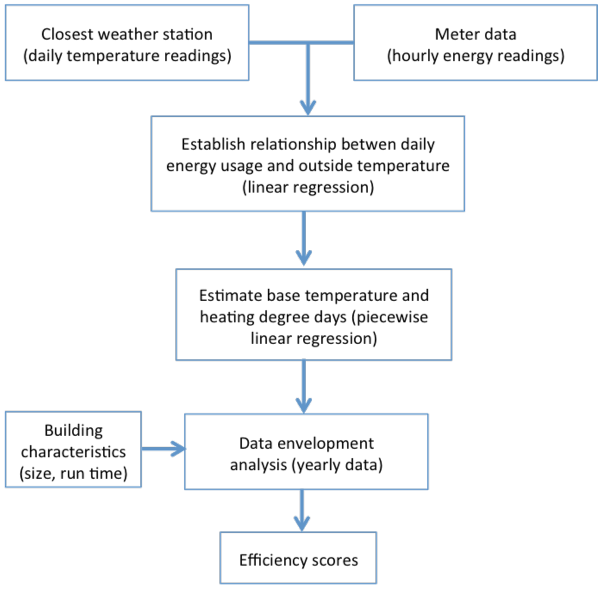

Figure 1. Conceptual model of benchmarking framework.

The OLS fits through the data points, A-F, illustrated by the dotted line, whereas the DEA model assuming variable return to scale technology envelops the data points creating a piecewise linear frontier (solid line). This DEA frontier, created by the best practice units, serves as the benchmark technology for inefficient DMUs. In short, DEA optimizes each company individually (by benchmarking it against its closest peers), whereas traditional statistical methodologies rely on averages or single optimization approaches.

Results

LINEAR REGRESSION – ENERGY VERSUS OUTSIDE TEMPERATURE

In the first stage of the analysis, we run 132 regression models with daily energy as the dependent variable and temperature as independent. In 91 of these models we get a p-value < 0.01, thus in 68 % of the stores we find that temperature from a statistical standpoint affects daily energy usage. However, we find large variation among these stores in terms of R^2 (the percentage variation of energy usage explained by temperature). The minimum R^2 between the 91 stores is 0.019, while maximum is 0.83, and the average is 0.16. One important reason that we see such variation is that many stores have systems to take advantage of waste heat from the refrigeration. In Figure 2, we see a histogram of the distribution of the R^2 values, showing that most of the stores have an R^2 lower than 0.4 – seven stores have a larger R^2. These results enable us to filter out the stores where we need to calculate heating degree-days, performed in the next step of the analysis.

OPTIMIZED CHANGING POINT TEMPERATURES (CPT)

In this step we present the results from the piecewise linear regression (PLR). The method detects CPT for each store by partitioning the independent variable (temperature) into intervals, and a separate regression line is fit to each interval. Then, the model with the lowest residual MSE (mean square error) is used to find the CPT-value. Thus, PLR runs a series of mod-

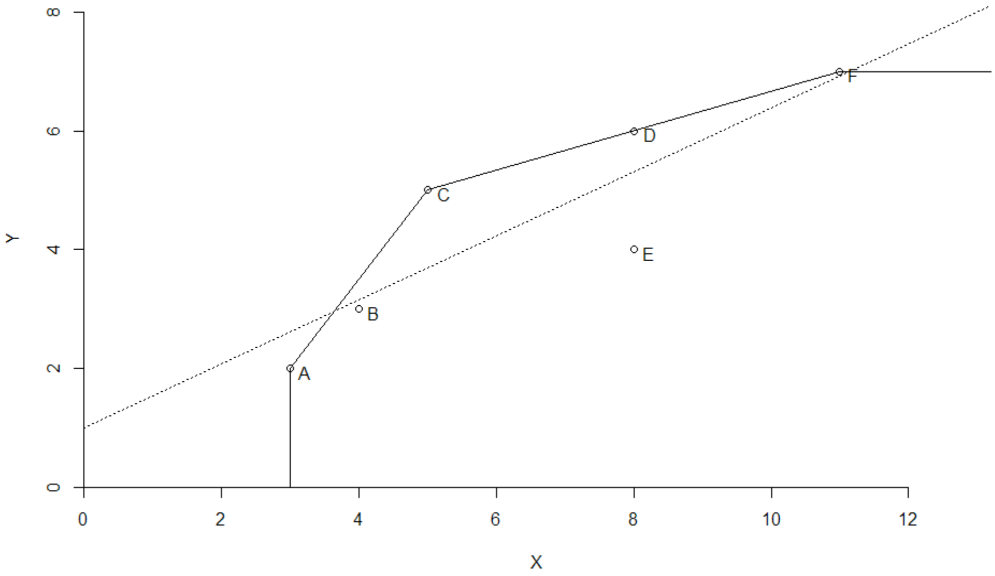

Figure 2. Conceptual visualization of DEA methodology.

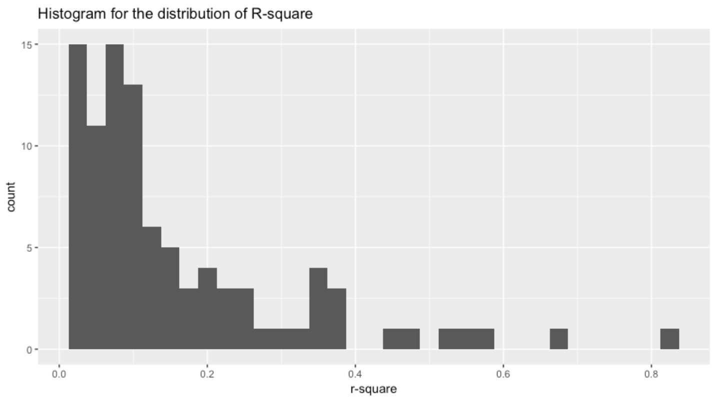

Figure 3. Distribution of R^2 between the 132 different stores where outside temperature was found to be statistically significant.

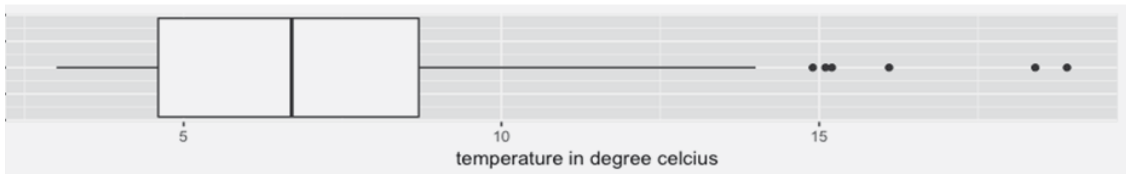

Figure 4. CPT temperature for 95 stores (with statistical significant relationship between energy usage and outside temperature).

els for every store to find the optimized CPT (the algorithm was set up to search within the entire observed temperature range, resulting in a more extensive search for stores with a large difference between the maximum and the minimum observed temperature). The average CPT for the 95 stores where we found a significant relationship between energy usage and temperature is 7.3 °C, median equals 6.7 °C, with minimum equal to 3 °C and maximum 18.9 °C. The distribution of the CPT temperatures can be seen in the boxplot below.

The default changing point temperatures (CPT) in Norway (and Denmark) is 17 °C. While these are undocumented values, they still are often in use as reference temperatures when calculating degree-days (NVE report, 2014). The average CPT we find is almost 10 °C below this default temperature. Previous studies has found CPT values in houses, office buildings, educational buildings, hospitals and hotels ranging from 9.7 °C till 13.2 °C, also with variations between weekdays and weekends (Pedersen, 2007). Other studies have investigated CPT for both heating and cooling demand (Lindberg & Doorman (2013), Lindberg et al. (2015)). Our findings demonstrate larger variation than previous studies. One important reason that the CPT are considerably lower in some of the buildings that we analyze is that many of these stores have heat recovery systems, and the building structures of retail stores might be very different from for example an office building. The variation in CPT and the deviation from standard default temperatures demonstrates that it is important to calculate individual CPT, as the following degree-day calculation will be very different if using standard values (e.g. 17°C). For example, earlier studies (Wang et al. 2015) use standard published values for the CPT; in the U.S., they use 18 °C as their starting point for degree-days calculations before inclusion in a DEA benchmarking study. Our approach improves this, and increases the accuracy of efficiency benchmarking of buildings.

In Figure 5 we show a scatter plot of average daily temperatures against daily energy usage in kWh for three of the stores, two buildings were the outside temperature is highly correlated with energy usage. The figure clearly visualizes a strong linear relationship between kWh and temperature for store number "1810", together with the estimated CPT of 7.1 °C. Store number "1810" has an R^2 of 0.83, meaning that 83 % of the variation in energy usage can be explained by variation in temperatures. Similarly, the R^2 for store number "2003" is 0.67. The vertical line shows the CPT as found through piecewise linear regression. Visually inspecting the plots in figure 5 it could be argued that the CPT looks to be more to the right for building no. 1810, however the "breakpoint" has been set automatically from the piecewise regression model based on the lowest residual MSE. Further, for store "1409" we find no similar relationship, and conclude that this is a building where there is no need to find the CPT and to calculate degree-days. Still, for store "1409" the vertical line shows the proposed CPT, as the piecewise linear regression runs through all stores, not taking into account if the relationship between energy and temperature is statistically significant. Obviously, this needs to be taken into account. The reason we find no relationship between electricity and temperature for building no. 1409 is that it is a food store and has a lot of waste heat from the refrigeration system. We are also investigating if there might be other energy carriers for this particular building. In the DEA analysis in the next section we will run DEA-models with degree days included for all stores, and another DEA-model including degree-days for stores where there is a significant relationship between energy usage and outside temperature.

RESULTS FROM DEA ANALYSIS

In the DEA-analysis, the input variable selected is annual energy consumption (kWh). As outputs the area of the building (m^2) and the number of opening hours are chosen ("runtime"). As we have seen in the previous section of the paper, for some buildings the energy consumption is heavily dependent on the outside temperature. We include the non-controllable input variable yearly degree-days for each store into the model to account for this. Still, we conduct to different DEA models:

Model 1: Modelling with run-time and size as output variables and degree-days as exogenous variables *without* taking into consideration if temperature is statistically significant.

Model 2: Modelling with run-time and size as output variables and degree-days as exogenous variable *taking* into account if temperature is statistically significant. In particular we set yearly degree-days = 0 for all the stores where we find no climate effect.

This approach enables us to look at the efficiency scores between these two modelling choices, and in particular compare the DEA efficiency scores for the stores with and without the climate effect. As indicated, the result of the DEA is highly sensitive to individual units. This potential problem has been resolved by removing potential outliers from the original sample. The distribution of the efficiency scores are shown in Figure 6.

The average efficiency from model 1 is 0.72 or 72 %. Given an optimal store in terms of energy efficiency this means that, we can potentially reduce the electricity consumption with a considerable 28 % (on average). Of the 132 stores in model 1 we find that 32 of them are 100 % efficient. We get exactly the same average efficiency score for model 2 (setting yearly degree-days = 0 in the DEA output variable where no climate effect is found). Moreover, as we can see from Figure 6 the there is little visible change in efficiency between the models. However, only 30 of the stores are found to be 100 % efficient in model 2, and more importantly, the efficiency scores changes between many of the stores. For example, looking at the 37 stores where we found no climate effect and comparing the efficiency scores between model 1 and model 2 we find an average change in efficiency scores of 7 %, ranging from a reduction of 34 % to an increase of 53 %. This is an important finding, and we conclude

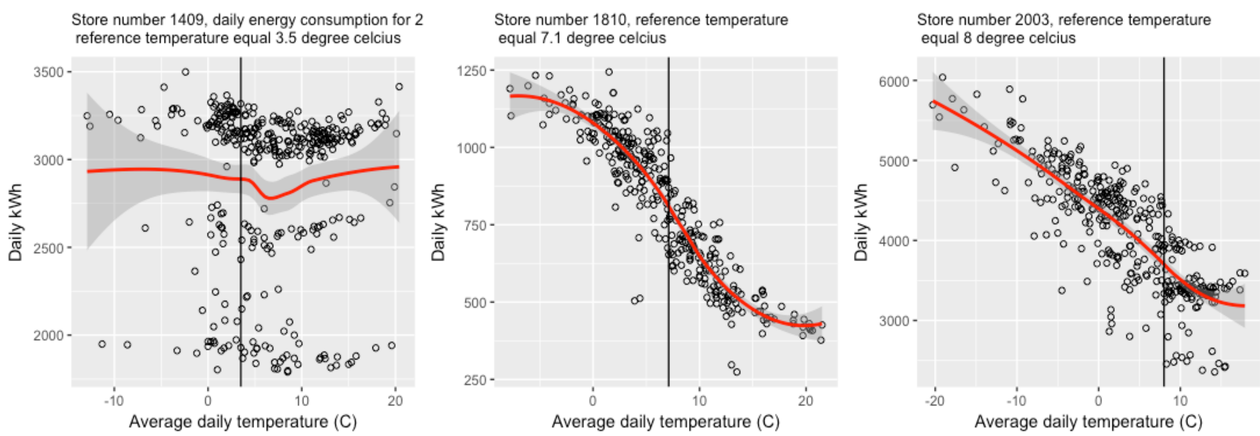

Figure 5. Temperature versus daily electricity usage for three example stores for year 2015, vertical line indicating the CPT from the piecewise linear regression models.

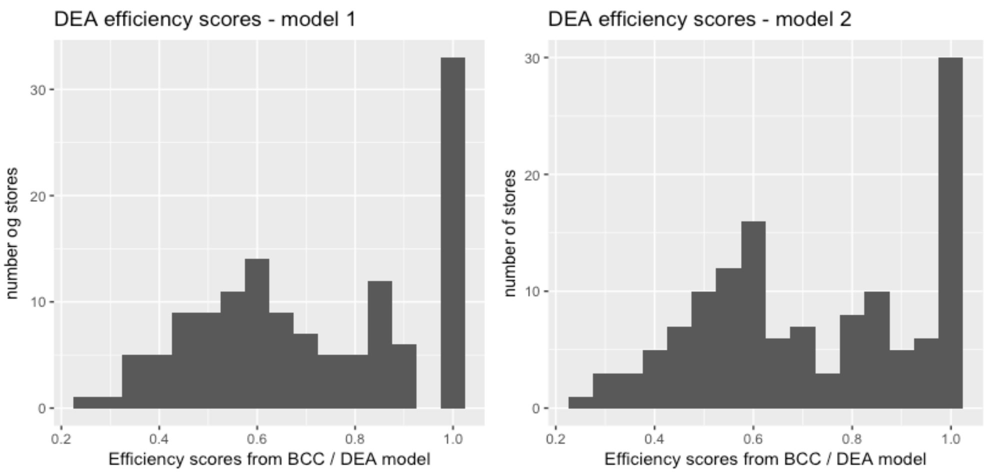

Figure 6. Efficiency scores from DEA model – with and without climate effect.

Figure 7. Change in efficiency scores as a consequence of including climate effects in the DEA model.

that if we do not include climate effect in the analysis the efficiency scores might turn out very differently. The change in DEA efficiency scores between model 1 and model 2 for the 37 stores can be seen in Figure 7. For the tallest bar (10 stores) there is practically no change in efficiency, however we see that for the other 27 stores, including climate effects has an impact on the scores.

Discussion and outlook

We have planned a number of actions based on the results in this paper. First, the stores with the lowest efficiency scores will be visited and analyzed carefully to better understand the factors behind the low efficiency score. Could there be equipment in these stores that work under non-optimal conditions? Further, based on this work (using hourly data from the least efficient stores) we are planning on developing predictive models to be used to indicate stores in need of maintenance (typically service request and maintenance are very expensive for the store owners). A sample of the stores with efficiency scores of 1 will also be visited to better understand the underlying factors. In future work more detailed information about the stores will be included in the DEA model, for example more details about building materials, lightning, building age, heat recovery system, and amount of cooling/freezing equipment. Future work will also consider more detailed estimation of the CPT, as the changing point could vary between weekdays and weekends. For example, Pedersen (2007) found the CPT for an office building to be 11 degree Celsius during weekdays, and 11.6 degrees during weekends. Taking into account this variation could further improve the accuracy of the presented efficiency scores. At last, further studies would benefit from outlining differences between the International Performance Measurement and Verification Protocol (IPMVP) (Energy, 2001).

Conclusions

We found that the presented modeling framework is a good approach to rank the energy performance between retail stores in Norway. We have also demonstrated the importance of a building's individual climate effects in a DEA model. We saw that not taking into account climate will affect the efficiency scores and changes them by, on average, 7 %. The largest change for an individual store was an efficiency increase of 53 %. We have also seen that the availability of detailed climate and energy data (often on hourly level), improves the methods of ranking building from an energy usage perspective. Using our methods to rank a large portfolio gives quick insights into the least and most efficient buildings, and can be used as a foundation for further investigation into retail buildings with low efficiency scores, or to learn from best performers.

References

AlFaris, F., B. Abu-Hijleh, et al. (2016). *"Using integrated control methodology to optimize energy performance for the guest rooms in UAE hospitality sector."* Applied Thermal Engineering 100: 1085–1094.

Al-Refaie, A., Hammad, M. and Li, MH. (2016) *"DEA window analysis and Malmquist index to assess energy efficiency and productivity in Jordanian industrial sector."* Energy Efficiency, Volume 9, Issue 6, pp 1299–1313.

Borgstein, E. H., R. Lamberts, et al. (2016). "*Evaluating energy performance in non-domestic buildings: A review*." Energy and Buildings 128: 734–755.

Banker R. D., A. Charnes and W. W. Cooper (1984) *"Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis",* Management Science 30: 9, 1078–1092.

Banker, R.D. and R.C. Morey (1986) "*Efficiency Analysis for Exogenously Fixed Inputs and Outputs*" Operations Research 34: 4, 513–521.

Charnes, A., W.W. Cooper and E. Rhodes (1978), "*Measuring the efficiency of decision making units*" European Journal of Operational Research, vol. 2, no. 6, pp. 429–444.

Blomberg, J., E. Henriksson, et al. (2012). *"Energy efficiency and policy in Swedish pulp and paper mills: A data envelopment analysis approach."* Energy Policy 42: 569–579.

Blum, H. (2015) *"The economic efficiency of energy-consuming equipment: a DEA approach"* Energy Efficiency, Volume 8, Issue 2, pp 281–298.

Chung (2011), *"Review of building energy-use performance benchmarking methodologies",* Applied energy, vol. 88, no. 5, pp. 1470–1479.

Chung, W. (2012). *"Using the fuzzy linear regression method to benchmark the energy efficiency of commercial buildings."* Applied Energy 95: 45–49.

Day, A.R., Knight, I., Dunn, G., Gaddas, R. (2003). "*Improved methods for evaluating base temperature for use in building energy performance lines*". Building Service Engineering Research and Technology, vol. 24, no. 4, p. 221–228.

Du, Z., X. Jin, et al. (2015). *"Evaluation of operation and control in HVAC (heating, ventilation and air conditioning) system using exergy analysis method."* Energy 89: 372–381.

Energy, D. (2001). International Performance Measurement & Verification Protocol. *Handbook of financing energy projects*, 249.

Eto JH. (1988) *"On using degree-days to account for the effects of weather on annual energy use in office buildings"*. Energy Build; 12: 113–127.

Ferrier GD., and Hirschberg JG. (1992) "*Climate control efficiency"*. Energy J.; 13: 37–54.

Grösche P. Measuring residential energy efficiency improvements with DEA. JProd Anal 2009; 31: 87–94.

Hansen (2000) "*Sample Splitting and Threshold Estimation*". Econometrica, Vol. 68, No. 3. (May 2000), pp. 575603.

Hong, S.-M., G. Paterson, et al. (2014). *"Improved benchmarking comparability for energy consumption in schools."* Building Research & Information 42 (1): 47–61.

Kavousian, A. and Rajagopal, R. (2014). *"Data-Driven Benchmarking of Building Energy Efficiency Utilizing Statistical Frontier Models."* Journal of Computing in Civil Engineering, Volume 28, Issue 1.

Lee, W.-S (2008), *"Benchmarking the energy efficiency of government buildings with data envelopment analysis",* Energy and Buildings, vol. 40, no. 5, pp. 891–895.

Lee, W.-S. (2010). *"Evaluating and ranking energy performance of office buildings using fuzzy measure and fuzzy integral."* Energy Conversion and Management 51 (1): 197–203.

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

Lee W, Lee K. (2009) *"Benchmarking the performance of building energy management using data envelopment analysis"*. Applied Thermal Engineering; 29: 3269–73.

Lee, W.-S. and Y.-C. Lin (2011a). *"Evaluating and ranking energy performance of office buildings using Grey relational analysis."* Energy 36 (5): 2551–2556.

Lee, W.-S. and L.-C. Lin (2011b). *"Evaluating and ranking the energy performance of office building using technique for order preference by similarity to ideal solution."* Applied Thermal Engineering 31 (16): 3521–3525.

Lu J, Ashuri B, Shahandashti M. *"A data envelopment analysis model for building energy efficiency benchmarking"*. Construction Research Congress; 2014. p. 1073–82.

Lindberg K.B. and G. Doorman (2013), "*Hourly load modeling of non-residential building stock*", conference proceedings, IEEE PowerTech conference in Grenoble, France, 16–20 June 2013.

Lindberg K.B., J.E. Chacón, G. Doorman, and D. Fischer (2015), "*Hourly Electricity Load Modelling of non-residential Passive Buildings in a Nordic Climate*", IEEE PowerTech conference in Eindhoven, Netherlands, 29 June–2 July 2015.

NVE report nr 7 (2014). "*Evaluering av modeller for klimajustering av energibruk*". Pöyry Management Consulting AS.

Önüt, S. and S. Soner (2006). *"Energy efficiency assessment for the Antalya Region hotels in Turkey."* Energy and Buildings 38 (8): 964–971.

Pedersen, L. (2007) "*Load modeling of buildings in mixed energy distribution systems*". Doctoral thesis, Norwegian University of Science and Technology.

Tu, K-J (2015) *"Establishing the DEA energy management system for individual departments within universities",* Facilities, Vol. 33 Iss: 11/12, pp.716–735.

United Nations Environment Programme. (2007) *Sustainable buildings and construction initiative buildings and climate change: status challenges and opportunities.* Paris: UNEP.

Wang, E., et al. (2015) *"Benchmarking energy performance of residential buildings using two-stage multifactor data envelopment analysis with degree-day based simple-normalization approach."* Energy Conversion and Management, 106: 530–542.

Wang S, Yan C, Xiao F. Quantitative energy performance assessment methods for existing buildings. Energy Build 2012; 55: 873–88.

Wang, K., S. Yu, et al. (2013). *"China's regional energy and environmental efficiency: A DEA window analysis based dynamic evaluation."* Mathematical and Computer Modelling 58 (5–6): 1117–1127.