**Data Envelopment Analysis and Software Packages for Measuring Building Energy Efficiency**

Zinovy Radovilsky\*

*California State University, East Bay, Hayward, CA, USA*

Pallavi Taneja

*California State University, East Bay, Hayward, CA, USA*

Payal Sahay

*California State University, East Bay, Hayward, CA, USA*

**\* Corresponding Author; Email: zinovy.radovilsky@csueastbay.edu**

**ABSTRACT**

Data Envelopment Analysis (DEA) is a popular approach in identifying efficiency of similar units. Its utilization in measuring and comparing buildings’ energy efficiency is still an evolving subject in the academic research. In this paper, we addressed and resolved several gaps in the existing DEA research and methodologies for energy efficiency analysis. We introduced energy efficiency indices as a part of DEA models’ outputs and applied transformation output variables to solve these DEA models. We also compared results of applying main DEA models and analytics software packages to identify the most consistent/reliable model and software package for measuring energy efficiency with DEA. Overall, this research will have an important impact in utilizing most appropriate DEA models and software packages for energy efficiency management in buildings.

1. **INTRODUCTION**

The consumption of energy in the building sector has been an important part of the overall energy usage in the United States and on the global scale. According to the U.S. Energy Information Administration, residential and commercial buildings account for about 40% of the total U.S. energy consumption (EIA, 2019). Buildings also emit over one-third of U.S. greenhouse gas emissions, which is more than any other sector of the economy, and, thus, making energy efficiency in buildings incredibly important (ASE, 2019). Energy efficient buildings offer opportunities to save money and reduce greenhouse gas emissions. In addition to overall environmental benefits that arise from a more energy efficient building, there are also personal benefits of saving money through lower energy consumption.

Identification and comparison of energy efficiency in buildings is based on utilizing a variety of metrics. For example, a very commonly used metric in the U.S. and globally for building energy performance is the energy use intensity (EUI), which is also applied for Energy Star Certification. This certification shows that benchmarking energy, e.g., comparing the energy use of buildings with others nationwide, helps effectively identify the opportunities for potential savings and best practices that can be replicated (Energy Star Score, 2018). EUI is a straightforward and easy to understand measure, but it only represents the amount of energy consumption, not energy performance of a building. Another common index for measuring energy efficiency is the energy efficiency index (EEI), also known as the building energy index. It is defined as the ratio of energy input (kWh) to a factor related to energy used, typically, building floor area (Kavousian et al., 2015). A variety of alternative indices for building energy performance has been also discussed in a paper by Goldstein and Eley (2014).

Measuring building energy performance usually takes into consideration a wide range of different factors, including energy usage, floor area, number of occupants, climate condition, energy efficiency of the equipment used, outdoor and indoor temperatures and others (Lee and Lee, 2009). There is also a number of methods applied to identify energy efficiency and compare energy efficiency between buildings. These methods can be summarized into four groups: simulation, regression analysis, artificial neural networks and data envelopment analysis (Ashuri et al., 2019; Wang et al., 2015; Lee and Lee, 2009; Lee, 2008). Each of these methods contains some benefits over other approaches and, at the same time, a number of limitations and drawbacks, a detailed description of which have been discussed in literature sources, e.g., Ashuri et al. (2019). In our paper, we will concentrate on applying Data Envelopment Analysis (DEA) for measuring and comparing building energy efficiency. While the other three methods (simulation, regression, and neural networks) are being extensively used for all kinds of data prediction and data analytics, the DEA method is the only one that is directly utilized, and very successfully, to identify and compare efficiency of similar units, e.g., buildings, for an organization or between organizations. DEA is also a highly popular method for analyzing and comparing efficiency of similar units in various industries including (but not limited to): banking (bank branches), healthcare (hospitals), hospitality (hotels), higher education (universities), etc.

The DEA application in measuring building energy efficiency is still a developing area of DEA research, and, therefore, is a subject of an ongoing discussion in academic literature. Molinos-Senante et al. (2016) consider several issues for DEA approaches – one that focuses on architectural factors (compactness, building shape, etc.), and another – on energy management issues such as equipment efficiency and operations strategy. Another identified DEA issue is on applying various DEA methods in performing energy efficiency comparison for buildings (Lee and Lee, 2009). In addition, the usage of DEA packages from popular analytics software like R or Python, may also affect the DEA results.

This research is motivated by the need to identify an effective DEA model and associated analytics software for comparing buildings’ energy efficiency. The related research objectives are as following:

* Utilize the academic literature review to understand existing DEA methodologies for developing and comparing energy efficiency in buildings.
* Identify appropriate DEA models, input and output variables, and analytics software for measuring buildings’ energy efficiency.
* Compare and contrast internal consistency of various DEA models and DEA-related software packages in order to identify the most reliable DEA model for energy efficiency comparison in buildings.

This paper’s structure includes five main sections. After the Introduction in section 1, we present in section 2 a literature review of the existing DEA research on measuring and comparing building energy efficiency. In section 3, we describe our research methodology that identifies building data set, input and output variables, DEA models, and DEA-related software packages. Based on this research methodology, we provide in section 4 a detailed Data Envelopment Analysis of measuring and comparing energy efficiency in buildings and interpret its results. We finish the paper with the conclusions in section 5.

1. **LITERRATURE REVIEW**

Data Envelopment Analysis (DEA) is an efficiency evaluation approach, which is used for comparing the performances of similar units in an organization or between organizations. The similar units for DEA are typically referred to as Decision Making Units (DMUs). The efficiency of a DMU is measured relative to all other DMUs. Comparison of peer groups enables the stakeholders (e.g. policy makers, building owners and investors) to derive actionable information from the building energy data that help enhance energy efficiency of buildings (Mathew et al., 2015). Using DEA for evaluating building energy efficiency can help in assessing how the allocation of resources among the group of buildings should be. It also helps to compare the less efficient building with the most efficient one in order to determine buildings that have room for improvement and how to make changes so that the less efficient ones can emulate the most efficient ones (Grösche, 2009; Halkos and Polemis, 2018; Kavousian et al., 2015; Mathew et al., 2015; Molinos-Senante et al., 2016).

One of the most commonly discussed in the literature and widely used in practice DEA approach is the *Charnes-Copper-Rhodes (CCR) model* (Charnes et al., 1978). It can be presented as a linear programming model in the following way (Cooper et al., 2011):

*max z =*

subject to

*j = 1, 2, …, n,* (1)

*=*1  
 *ur ,vi ≥ 0; r = 1, 2, …, s, i = 1, 2, …, m,*

In the CCR model formulation (1), the assumption is that there are *n* DMUs (*n* buildings in case of our research) to be evaluated. Each DMU consumes varying amount of *m* different inputs to produce *s* different outputs. Specifically, *DMUj* consumes amount of input *i* and produces amount of output *r*. Parameters and are the variable weights to be determined by solving the CCR model in formulation (1). The CCR model is built on the notion of efficiency as a ratio of outputs to the relative inputs. A DMU’s efficiency obtained in the model is never absolute as it is always measured relative to other DMUs (Cooper et al., 2011). If the solution for DMU *O* is equal to 1, then this building is considered 100% efficient relative to other buildings in the CCR model.

The CCR model also assumes *constant returns to scale (CRS)* between inputs and outputs (Banker et al., 2004). This implies that an increase in a DMU’s inputs leads to a proportionate increase (or decrease) in its outputs i.e. there is a one-to-one, linear relationship between inputs and outputs. For example, if a 10% increase in inputs yields a 10% increase in outputs, the DMU is operating at constant returns to scale. This means that no matter what scale the DMU operates at, its efficiency will, assuming its current operating practices, remain unchanged.

Another DEA approach frequently discussed in the literature sources is the *Banker, Charnes and Cooper (BCC)* model (Banker et al., 1984). It can be also presented as a linear programming model (Banker et al., 2004) in the following way:

*max z = - uo*

subject to

*j = 1, 2, …, n,* (2)

*=*1  
 *ur ,vi ≥ 0; uo is free of sign; r = 1, 2, …, s, i = 1, 2, …, m,*

Being very similar to the CCR model in (1), the BCC model formulation (2) contains one extra weight variable, *uo*, which makes it possible to effect returns-to-scale evaluations (Cooper et al., 2011 and Banker et al., 2004)). This means that a DMU’s input increase or decrease does not necessarily produce a proportional (constant) change in its outputs, and, thus, the BCC model is considered a model with *variable returns to scale (VRS)*. VRS implies that as a DMU changes its scale of operations, its efficiency will either increase, decrease, or remain the same.

Analysis of research literature on utilizing DEA in measuring building energy efficiency shows that authors utilize both CCR/CRS and BCC/VRS models with various options and extensions in developing and comparing energy efficiency in buildings. Molinos-Senante et al. (2016) applies the CCR model to compare energy performance of office buildings in Chili. Grösche (2009) measures energy efficiency improvements of U.S. single-family homes using a two-stage procedure, in the first stage of which an indicator of energy efficiency is derived by means of employing the CCR model. Wang et al. (2015) also utilize the CCR model to compare energy consumption of single-family buildings.

At the same time, Ashuri et al. (2019) employ a combination of CCR and BCC models to compare energy efficiency of multifamily properties in various U.S. states. Lee and Lee (2009) utilize, in addition to regression analysis, the BCC model to identify efficiency of building energy management in government buildings in Taiwan. Despite the fact that some of these research papers describe the usage of both CCR and BCC models, we did not identify in those papers a discussion or comparison of the merits of each type of these DEA models in measuring building energy efficiency. We consider this as a gap in the existing DEA research on developing and comparing energy efficiency in buildings.

The reviewed literature sources describe a variety of input and output variables used in DEA models. Among input variables in formulations (1) and (2), energy consumption (in kWh or Btu) is the most common (Ashuri et al., 2019; Grösche, 2009; Lee and Lee, 2009). Other input variables applied in the DEA models may include: energy use intensity (EUI) (Lee, 2008), weather normalized EUI (Wang et al., 2015), and outdoor temperature (Ashuri et al., 2019). The output variables in formulations (1) and (2) are much wider in range and scope, and typically include: building floor space, number of occupants in buildings, occupancy intensity, and average outdoor temperature (Grösche, 2009; Lee, 2008; Lee and Lee, 2009; Molinos-Senante et al., 2016). Additional output variables, specifically for residential buildings, may also incorporate: number of apartments, number of bedrooms, number of parking spaces, and number of washing machines (Ashuri et al., 2019; Wang et al., 2015).

Despite the usage of these output measures in DEA models for measuring building energy efficiency, we would like to argue that some of them, e.g., building floor space and number of occupants, are *not output results* of energy consumption, but rather merely represent the inputs to the energy efficiency measurements. Moreover, the actual utilization in practice of a variety of building energy efficiency indices like EUI, normalized energy consumption per number of occupants, and some others (Fairey and Goldstein, 2016; Goldstein and Eley, 2014; Orvis et al., 2016) show that building floor space and number of occupants are, in fact, applied as inputs to energy efficiency indices. From this standpoint, the DEA outputs for building energy efficiency may incorporate some of these indices or normalized energy parameters like energy consumption per space (square footage) or per building’s number of occupants.

Reviewing the literature sources on DEA for building energy efficiency, we identified a surprisingly small number of papers that discuss or briefly mention the usage of DEA-related computer packages. Yoon and Park (2017) describe using Matlab software in comparing building energy performance. Wang (2017) applies Excel Solver to run a DEA model for measuring and comparing building energy efficiency in the multifamily industry. The general analysis of DEA software packages is presented in several papers, for example, in Iliyasu et al. (2015) and Severinsen and Sorensen-Holst (2017). However, we did not discover in the literature any discussion on applicability and efficiency of modern and commonly used data analytics software like Python or R, and their DEA-related packages, in measuring and comparing energy efficiency in buildings. We also consider this as a gap in the existing DEA research on building energy efficiency.

1. **METHODOLOGY**

Based on the literature review and identified gaps in DEA literature sources, our primary goal in this research is to analyze and compare energy efficiency in buildings by applying different DEA models and analytics software in order to identify the best possible model and software package for measuring buildings’ energy efficiency. The research methodology includes the following steps:

1. Identify a data set containing buildings’ energy efficiency parameters.
2. Recognize DEA input/output variables.
3. Select DEA models to be used for developing and comparing building energy efficiency.
4. Choose software packages for applying the selected DEA models.
5. Identify energy efficiency of buildings in the data set using the selected DEA models and software packages; compare their results including model and software consistency/reliability.
6. Interpret the energy efficiency results and identify the most appropriate DEA model and software package for energy efficiency comparison in buildings.

In this research, we employ a data set that contains energy-related characteristics of 17 buildings on a campus of a U.S. organization’s headquarters. The organization’s security and privacy policies require to keep the overall data and location of the campus confidential. For each building on this campus, the data set contains a number of monthly general building characteristics including: energy consumption, monthly energy cost, number of occupants, and space/floor area. In addition, the data set includes some energy-related ratio data, e.g., effective cost rate of energy consumption ($/kWh), energy consumption per occupant, and some others.

Based on the literature review and our comments with respect to input/output variables for DEA models (see section 2), we selected the following four input variables for each building:

* Energy consumption per month, kWh
* Number of occupants, units
* Space area, sq. ft.
* Energy cost per month, $

We also identified the following three output variables:

* Effective rate of energy consumption, $/kWh
* Energy consumption per occupant, kWh/population
* Energy consumption per square foot of space, kWh/sq. ft.

The introduced set of output variables truly reflect the nature of building energy efficiency outputs expressed by energy indices, whereas the set of inputs variables characterize the important building parameters that energy efficiency depends on. The identified variables also necessitate the usage of an output-oriented DEA model, as the goal will be to address the output efficiency indices vs. input building parameters. We also need to address several potential concerns with respect to the selected variables and how we handle them in this research.

The first concern is that the external temperature is not included as an input variable. This is due to the fact that all buildings are situated in the same location, and, thus, external temperature will have a minimal effect in comparing energy efficiency of those buildings. However, it may potentially be added to the input pull of variables if those buildings are located in different areas of the country. The second concern is raised by some DEA authors that the usage of ratio variables, e.g., effective rate or energy consumption per square foot, may violate one of the main assumptions of an efficiency measure in DEA – the convexity axiom (Ashuri et al., 2019; Emrouznejad, and Amin, 2009; Wang, 2017). However, this issue can be resolved by developing a modified CCR or BCC DEA model that can incorporate ratio output or input variables (Olesen et al., 2015; Olesen et al., 2017). Moreover, modern DEA software packages can incorporate ratio variables in energy efficiency analysis without losing integrity of DEA results.

The third and most important concern with respect to the variables in our DEA model is related to *undesirable outputs*. In a traditional output-oriented DEA model, we have to maximize outputs relevant to their inputs. However, the output variables for energy efficiency we proposed for DEA modeling, e.g., effective rate ($/kWh) or consumption per person (kWh/population), are those that actually need to be lowered for making more energy efficient buildings, and, therefore, represent “*undesirable outputs*” in terms of the output-oriented DEA model. Literature sources provide a number of approaches of how to deal with undesirable outputs that can be summarized into several major groups of methods: (a) ignoring undesirable outputs (Halkos and Polemis, 2018; Yang and Pollitt, 2009); (b) treating undesirable outputs as inputs ((Jahanshahloo et al., 2005; Seiford and Zhu, 2002) and (c) transforming undesirable outputs into a desirable form (Fare and Grosskopf, 2004: Homayounfar et al., 2014; Scheel, 2001). In our research, we are going to utilize, for comparison purposes, two relatively simple ways of undesirable variable transformation: *inverse transformation* and *min-max normalization*.

The inverse transformation converts a potential undesirable output in formulation (1) or (2) into a desirable form by using the following formula:

*= 1/* (3)

Applying these transformed variables in an output-based DEA model, will maximize outputs (, and, simultaneously, minimize the real output variables, which would be required for efficient energy performance.

The min-max normalization allows to normalize output variables on a uniform scale using the minimum and maximum values of each output variable, *min*() and *max*(

*= [ – max()]/[min() – max()]* (4)

Based on formula (4), the value of outputs is normalized between 0 and 1 (0% and 100%), The smaller the value of , the higher the value of *.* Thus, we can apply these normalized variables in a standard CCR or BCC model to identify proper efficiency measures for each DMU.

The next step in this research is to select the appropriate DEA model(s). In our case, we will consider both CCR and BCC models. As discussed in the literature review (see section 3), the BCC model is used for variable returns to scale (VRS) relationships between inputs and outputs, which is partially the case in our data. For example, the relationships between building’s effective rate as output and square footage or number of occupants as inputs are not proportional and thus are of VRS type. At the same time, some input and output variables may have proportional relationships, e.g., effective rate (output) and energy consumption or energy cost. This implies that we can also consider the CCR model with constant returns to scale (CRS) for our energy efficiency analysis.

An important step in this research is the utilization of various analytics software to identify and compare building energy efficiency. We will be employing four software applications: Microsoft Excel, R, Python, and ~~Minitab~~(MATLAB). All of them are popular software in modern data analytics and also contain specific packages or add-ins (in case of Excel) to run DEA-based models. Using the data set of 17 buildings with their respective values of input and output variables, we will run on this software both CCR and BCC models with inverse transformation and min-max normalization of the output variables.

1. **ENERGY EFFICIENCY ANALYSIS AND INTERPRETATION OF RESULTS**

As discussed in the previous section, the data set for energy efficiency analysis consists of four input and three output variables measured in 17 buildings. The summary statistics of these variables is presented in Table 1.

**Table 1. Summary Statistics of Input and Output Variables for Energy Efficiency DEA.**



The data in the table also incorporates summary statistics of inverse and min-max normalized output variables, which are actually applied in the DEA models. Based on formulations (1) and (2), we introduce CCR and BCC models separately and run each of them for all 17 DMUs by applying different software packages in Microsoft Excel, R, Python, and Matlab (see Table 2).

**Table 2. Software Application and DEA Methods.**



In Excel, we utilize the Excel Solver add-in, which is applied for solving linear and non-linear optimization models. In R, we employ the built-in rDEA package that contains the *dea* function used for both CCR and BCC models. We also apply the Python’s PyDEA package, specifically the *pydea.DEAProblem* function, to run the discussed DEA models. In Matlab, we utilize the build-in package – Data Envelopment Analysis Toolbox (DEA Toolbox) – to perform the energy efficiency analysis with the same models and output variables.

As can be seen from Table 2, for the four described software packages, we consider running CCR and BCC models using either inverse or min-max normalized outputs. For each of these DEA options, we received a set 17 efficiency scores describing energy efficiency of 17 respective buildings (DMUs). The score of 1.0 represents a building’s optimal energy efficiency of 100%, and the score of less than 1.0 implies that the energy efficiency in a building is not optimal. Based on the four software packages, two DEA models and two types of outputs, we overall receive 16 sets of DEA results that are further used to analyze and compare consistency/reliability of the utilized DEA models, variables, and software packages.

The summary of the DEA results for each DMU are presented in Table 3. In this table, for example, the mean of 0.937 for DMU #1 is the average of 16 efficiency scores for this DMU, and variance of 0.013 is the variance of 16 efficiency scores for the same DMU.

**Table 3. Summary of DEA Energy Efficiency Scores for DMUs.**

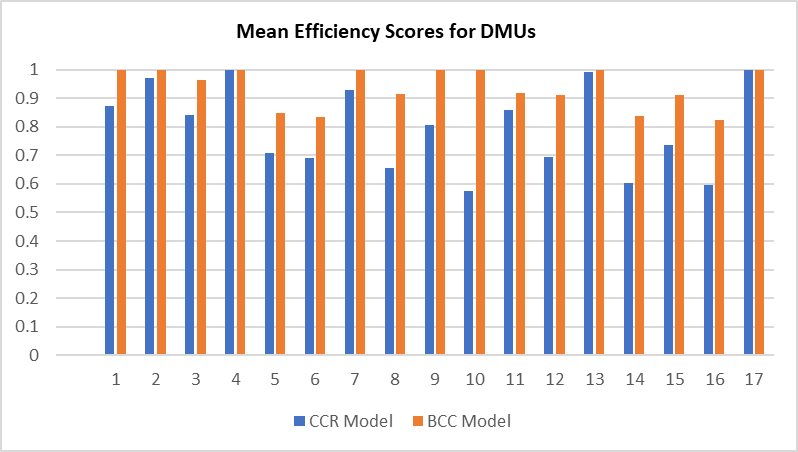


As can be seen from this table, there are only 2 DMUs, #4 and #17, for which the energy efficiency score of 1 (100%) is the same with zero variance and standard deviation for all software packages, DEA methods, and variable types we applied in this research. The rest of DMUs have some non-zero variance of the efficiency scores with means below 100% and standard deviation varying from 0.015 (DMU #13) to 0.226 (DMU #10). If we consider coefficient of variation (CV), a ratio of score’s standard deviation to its respective mean, it would range from 1.6% to 28.7%. Overall, this indicates that the DEA results in terms of DMU’s efficiency scores are mostly variable across different models and software packages applied in our research.

To further examine this variability (inconsistency) of the DEA results for each DMU, we analyzed differences between average efficiency scores for the CCR and BCC models that were developed using all four software packages (see Figure 1).

**Figure 1. Mean Efficiency Scores Derived by CCR and BCC Models with Four**

**Software Packages.**



The chart in Figure 1 shows that, for each DMU, the mean efficiency scores derived by CCR and BCC models are typically different. Moreover, the mean efficiency scores identified by the CCR models is mostly smaller than the respective scores produced by the BCC models. In few cases, specifically for DMUs #4 and #17, these scores are equal.

In addition to the chart in Figure 1, we also apply the two-sample T-test with unequal variances to compare the CCR and BCC models’ mean efficiency scores for each DMU (Table 4). In this test, we hypothesized that a difference between the respective means is considered to be 0 (the population means are equal), with significance α of 0.05.

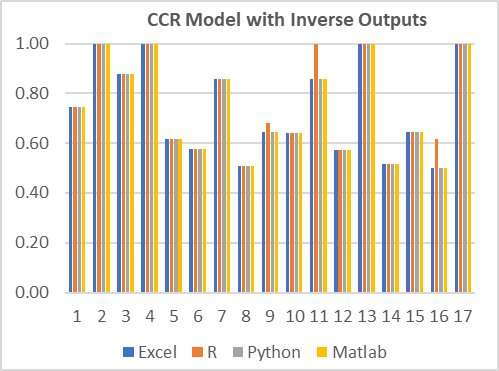
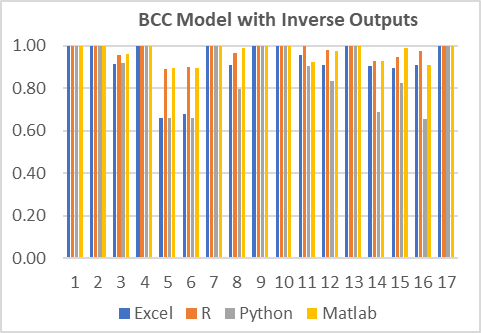
**Table 4. T-Test for Mean Efficiency Scores Using CCR and BCC Models.**

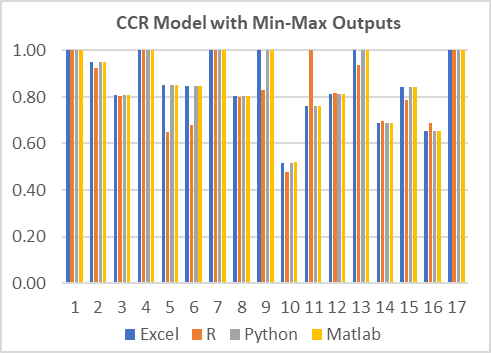
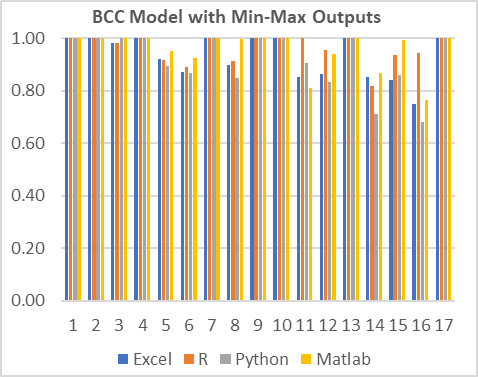


The P-value in Table 4 of below than 0.05 for one-tail and two-tail T-tests show that we have to reject the zero-difference hypothesis, and thus the population means of CCR and BCC models are not statistically equal. This is true for 13 out of 17 DMUs, the bulk portion of DMUs. For two DMUs, #11 and #13, the null hypothesis cannot be rejected because the respective P-values are greater than 0.05, and for two other DMUs, #4 and 17, the T-test was not applicable, because the variances for these DMUs are equal to 0. Overall, the results in Table 4 confirm a notion that a DMU’s efficiency from the BCC model is typically greater than or equal to the same DMU’s efficiency produced by the CCR model (Coelli et al., 2005; Cooper et al., 2011; Grilo and Santos, 2015).

To further analyze the variability of the DEA results, we compare energy efficiency scores of all DMUs for the CCR and BCC models with inverse outputs and also the same models with min-max normalized outputs (see Figure 2).

**Figure 2. Efficiency Scores for DEA Models with Inverse and Min-Max Outputs.**

**** 

** **

Examining the charts in Figure 2, we can observe that the efficiency scores for each DMU are the most consistent with CCR models, and, in particular, the CCR model with inverse output variables. The respective chart shows variability of efficiency scores in only three DMUs, #9, #11 and #16. Contrary to that, the BCC models with both inverse and min-max outputs produce less consistent efficiency scores for a number of DMUs. For example, for the BCC model with inverse outputs, a high variability of efficiency scores can be observed in 9 DMUs, i.e., #2, #5-6, #8, #11-12, and #14-16.

Visualizing the charts in Figure 2, we can also conclude that the BCC models produce more DMUs with the efficiency scores of 1 than that of the CCR models. This is another confirmation of the fact that a DMU’s efficiency from the BCC model is typically greater than or equal to the same DMU’s efficiency produced by the CCR model (Coelli et al., 2005; Cooper et al., 2011; Grilo and Santos, 2015). In addition, both CCR and BCC models with min-max normalized outputs tend to develop higher efficiency scores than those of the same models with inverse outputs.

To quantitatively measure a degree of consistency within DMU efficiency scores, we utilize the *Cronbach’s α* coefficient of internal consistency/reliability. It measures how closely related a set of items are as a group (Nunnally and Bernstein, 1994). In this research, we identify the Cronbach’s α coefficients for groups of DMU efficiency scores produced by the CCR and BCC models with inverse and min-max outputs that were run on Excel, R, Python, and Matlab (Table 5).

**Table 5. Internal Consistency (reliability) of DEA Results.**

|  |  |  |
| --- | --- | --- |
| **Parameters** | **Cronbach's α** | |
| **CCR Model** | **BCC Model** |
| Inverse Outputs | 0.997 | 0.848 |
| Min-Max Outputs | 0.970 | 0.817 |
| Excel Solver | 0.739 | 0.625 |
| R - rDEA | 0.732 | 0.716 |
| Python - PyDEA | 0.908 | 0.707 |
| Matlab - DEA Toolbox | 0.719 | 0.660 |

According to Nunnally and Bernstein (1994), a reliability coefficient of .70or higher is considered “acceptable” in terms of internal consistency. According to this criterion, the Cronbach’s α coefficients from Table 5 show the overall internal consistency of the DEA models applied with four software packages. However, the highest values of those Cronbach’s α reliability coefficients are associated with the CCR models, and, in particular, the CCR model with inverse variable outputs. Therefore, we can conclude that, in our research, the CCR model with inverse outputs provides the most consistent and reliable results with various software applied.

The data in Table 5 also provides information on Cronbach’s α coefficients for the four software packages we utilized. These coefficients show that, for each software, the CCR models have higher reliability coefficients than those for the BCC models. In fact, for Excel and Matlab, the values of Cronbach’s α coefficients for BCC models are below the acceptable level of 0.7. At the same time, the highest reliability coefficient of 0.908 is achieved by the Python-based CCR model, which is, therefore, considered in our case as the most reliable software package for energy efficiency DEA.

1. **CONCLUSIONS**

The main target of this research was to analyze building energy efficiency by applying different DEA models and various data analytics software in order to identify the best possible model(s) and software packages for buildings’ energy efficiency comparison. In this context, we reviewed a variety of literature sources to understand DEA methodologies, models, and input/output variables applied for buildings’ energy efficiency consumptions.

Overall, the analysis of literature sources on utilizing DEA in building energy efficiency demonstrate that authors apply both CCR and BCC models in measuring and comparing energy efficiency in buildings. However, we did not identify any discussion or comparison of these models, which we consider as a gap in the existing DEA research. We also argued that some common DEA output variables described in the literature, e.g., building floor space and number of occupants, should be a part of the DEA input variables. At the same time, frequently used building energy indices like energy consumption per square footage or per number of occupants, may be a part of outputs in DEA models. We also reveal another gap in the literature sources concerning the lack of discussion and consideration for applicability and effectivity of modern and commonly applied data analytics software and their DEA-related packages in comparing building energy efficiency.

We developed, based on the literature review and recognized DEA gaps, a research methodology to address these gaps. First, we identified four input and three output variables for DEA that truly reflect the way that energy efficiency comparison is done in buildings. We then proceeded with managing undesirable outputs in this case by utilizing inverse and min-max normalized output variables. We also decided to apply both CCR and BCC models with four DEA-related software packages (add-ins) including Excel Solver, R’s rDEA, Python’s PyDEA, and Matlab’s DEA Toolbox.

Using the four software packages, CCR and BCC models and two types of outputs, we overall received 16 sets of DEA results that were further used to analyze and compare consistency/reliability of the utilized DEA models, variables, and software. Analysis of these results revealed a number of insights into utilization of different DEA models, output variables, and software packages.

In particular, the results indicated that the DMU’s efficiency scores are mostly varying across different models and software packages applied. In particular, the DMUs’ mean efficiency scores identified by the CCR models is mostly smaller than the respective scores produced by the BCC models, which is consistent with the previous research notion that, for the same data set, the BCC model’s efficiency is typically greater than or equal to the efficiency measured in the CCR model.

We also compare energy efficiency scores of all DMUs for the CCR and BCC models with inverse outputs and also the same models with min-max normalized outputs. The comparison of these models by utilizing Cronbach’s α led to a conclusion that the efficiency scores for each DMU are the most consistent with CCR models, and, in particular, the CCR model with inverse output variables. In addition, the Cronbach’s α coefficients for the four applied software packages show that, for each software, the CCR models have higher reliability coefficients than those for the BCC models, with the highest reliability coefficient appeared to be for the Python-based CCR model. In summary, we can conclude that the CCR model with inverse outputs provides the most reliable results regardless of the software application applied, with the Python’s PyDEA package being the most reliable in terms of running the CCR model.

Overall, this paper produced a number of important contributions to the theory and practice of applying DEA models for measuring and comparing energy efficiency in buildings. In particular, we addressed and resolved several gaps in the existing DEA research and methodologies for building energy efficiency analysis, specifically in terms of: (a) comparing various DEA models to identify the best model for energy efficiency; (b) introducing energy efficiency indices as a part of DEA outputs and applying transformation variables to solve DEA models, (c) comparing various analytics software packages for DEA models, and (d) identifying the most reliable model and software package for measuring energy efficiency with DEA. All these results will have an important impact in utilizing most appropriate DEA models and software packages for energy efficiency management in buildings.

In the future, we would like to expand this research by including into consideration buildings from various locations, and also by incorporating uncontrollable variable like outside temperature into measuring and comparing energy efficiency. We can expand our research with applying more software applications, and, in particular, specialized DEA software packages like DEAP and DEAFrontier (Excel add-in). We may also consider more input and output variables to be used in the DEA models and apply additional ways of output variable transformations for the variables with undesirable outputs.

1. **LITERATURE SOURCES**

ASE (Alliance to Save Energy), “Energy Use of Buildings”, *https://www.ase.org/initiatives/buildings*, (accessed December 25, 2019).

Ashuri, B., Wang, J., Shahandashti, M., and Baek, M., “A data envelopment analysis (DEA) model for building energy benchmarking”, *Journal of Engineering Design and Technology,* 17(4), 2019, 747-768.

Banker R.D., Charnes A., and Cooper, W.W., “Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis”, *Management Science*, 30, 1984, 1078–1092.

Banker, R.D., Cooper W.W., Seiford, L.M., Thrall, R.M., and Zhu, J., “Returns to scale in different DEA models”, *European Journal of Operations Research*, 154, 2004, 345-362.

Charnes, A., Cooper, W.W., and Rhodes, E., “Measuring the efficiency of decision making units”, *European Journal of Operational Research*, 2(6), 1978, 429-442.

Coelli T. J., Rao D. S. P., O'Donnell C. J., and Battese G. E., “*An Introduction to Efficiency and Productivity Analysis”*, 2nd Edition, Springer, New York, 2005.

Cooper, W.W., Seiford, L.M., and Zhu, J., *Handbook on Data Envelopment Analysis*, Springer Science+Business Media, 2011.

Energy Star Score, *Energy Start Portfolio Manager: Technical Reference*, August 2018, 1-14.

# EIA (U.S. Energy Information Administration), “How much energy is consumed in U.S. residential and commercial buildings?”, *https://www.eia.gov/tools/faqs/faq.php?id=86&t=1*, (accessed December 25, 2019).

Emrouznejad, A. and Amin, G.R., “DEA Models for Ratio Data: Convexity Consideration”, *Applied Mathematical Modeling*, 33(1), 2009, 486-498.

Fairey, P. and Goldstein, D. B., “Metrics for Energy Efficient Buildings: How Do We measure Efficiency?”, *2016 ACEEE Summer Study on Energy Efficiency in Buildings*, 2017, Retrieved from www.aceee.org, 1-13.

Fare, R. and Grosskopf, S., “Modelling undesirable factors in efficiency evaluation: Comment”, *European Journal of Operational Research*, 157, 2004, 242–245.

Goldstein, D. B. and Eley, C., “A classification of building energy indices”, *Energy Efficiency*, July 2014, 353-375.

Grilo, A. and Santos, J., “Measuring Efficiency and Productivity Growth of New Technology-Based Firms in Business Incubators: The Portuguese Case Study of Madan Parque”, *The* *Scientific World Journal*, 2015, March 2015, 1-22. online publication (accessed October 17, 2019).

Grösche, P., “Measuring residential energy efficiency improvements with DEA”, *Journal of Productivity Analysis*, 31(2), 2009, 87-94.

Halkos, G.E. and Polemis, M.L., “The impact of economic growth on environmental efficiency of the electricity sector: A hybrid window DEA methodology for the USA”, *Journal of Environmental Management*, 211, 2018, 334-346.

Homayounfar, M., Amirtteimoori, A.R., and Toloie-Eshlaghy, A., “Production planning considering undesirable outputs – A DEA based approach”, *International Journal of Applied Operations Research*, 4(3), Summer 2014, 1-11.

Iliyasu, A., Mohamed, Z.A., and Terano, R., “Data Envelopment Analysis Models and Software Packages for Academic Purposes”, *Pertanika Journal of Scholarly Research Reviews*, 1(1), 2015, 27-32.

Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Shoja, N., Tohidi, G., and Razavyan, S., “Undesirable inputs and outputs in DEA Models”, *Applied Mathematics and Computation*, 169, 2005, 917-925.

Kavousian, A., Rajagopal, R., and Fischer, M., “Ranking appliance energy efficiency in households: Utilizing smart meter data and energy efficiency frontiers to estimate and identify the determinants of appliance energy efficiency in residential buildings”, *Energy and Buildings*, 99, 2015, 220-230.

Lee, W-S. and Lee, K-P., “Benchmarking the performance of building energy management using data envelopment analysis”, *Applied Thermal Engineering*, 29(16), 2009, 3269-3273.

Lee, W-S., “Benchmarking the energy efficiency of government buildings with data envelopment analysis”, *Energy and Buildings*, 40, 2008, 891-895.

Mathew, P.A., Dunn, L.N., Sohn, M.D., Mercado, A., Custudio, C. and Walter, T., “Big-data for building energy performance: lessons from assembling a very large national database of building energy use”, *Applied Energy*, 140, 2015, 85-93.

Molinos-Senante, M., Encinas, F., and Ureta, F., “Benchmarking the Energy performance of Office Buildings: A Data Envelopment Analysis Approach,” *Revista Electronica de Comunicaciones y Trabajos de ASEPUMA*, 16, 2016, 179-190.

Nunnally, J. C., and Bernstein, I. C. Psychometric theory, 3rd edition, New York: McGraw-Hill, 1994.

Olesen, O.B., Petersen, N.C., and Podinovski, V.V., “Efficiency analysis with ratio measures”, *European Journal of Operations Research*, 245(2), 2015, 446-462.

Olesen, O.B., Petersen, N.C., and Podinovski, V.V., “Efficiency measures and computational approaches for data envelopment analysis models with ratio inputs and output”, *European Journal of Operations Research*, 261(2), 2017, 640-655.

Orvis, R., Aggarwal, S., and O’Boyle, M., “Metrics for Energy Efficiency: Options and Adjustment Mechanisms”, *www.energyinnovation.org*, April 2016, (accessed October 17, 2019).

Scheel, H., “Undesirable outputs in efficiency valuations”, *European Journal of Operations Research*, 132, 2001, 400-410.

Seiford, L.M. and Zhu, J., “Modeling undesirable factors in efficiency evaluation”, *European Journal of Operations Research*, 142, 2002, 16-20.

Severinsen, A. and Sorensen-Holst, H. M., “Using machine learning and mathematical programming to benchmark energy efficiency in buildings”, *ECEEE Summer Study Proceedings on Consumption, Efficiency & Limits*, August 2017, 1083-1089.

U.S. Environmental Protection Agency. What Is Energy Use Intensity (EUI)? *http://www.eia.gov/emeu/efficiency/measure\_discussion.htm* (accessed on 11 January 2017).

Wang, E., Shen, Z., Alp, N. and Barry, N., “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, 2015, 530-542.

Wang, J., “Benchmarking Building Energy in the Multifamily Industry: A Data Envelopment Analysis (DEA) Model”, *Georgia Institute of Technology*, May 2017, 1-93.

Yang, H. and Pollitt, M., “Incorporating both undesirable outputs and uncontrollable variables into DEA: The performance of Chinese coal-fired power plants”, *European* *Journal of Operational Research*, 197, 2009, 1095–1105.

Yoon, S.-H. and Park, C.-S., “Objective Building Performance Benchmarking Using Data Envelopment Analysis and Monte Carlo Sampling”, *Sustainability*, 9, 2017, 780, 1-12.