A Data-driven Framework to Estimate Saving Potential of Buildings in Demand Response Events

X. Yu[a] and S. Ergan[a]

[a]Department of Civil and Urban Engineering, New York University, United States
E-mail: celiayu@nyu.edu, semiha@nyu.edu

Abstract –

In the U.S., the increasing electricity demand gives pressure on the power grids because of its limited capacity to serve demand. Instead of building new power plants to meet the increasing demand, Demand Response (DR) programs incentivize end-consumers to reduce certain electricity demand during certain periods (e.g., peak demand and emergency times). In the current practice, saving potential of buildings, i.e., the amount of electricity that end-consumers can save during an event, is usually determined using the technical specifications of equipment installed, which is unrealistic and leads to over or underestimation of the expected saving potential. In this study, the authors developed a data-driven framework to quantify the electricity saving potential in buildings. The framework was applied to nineteen campus buildings. Several prediction algorithms were used to fit models to the integrated datasets of these buildings, and models were evaluated using four criteria to avoid over-fitting and under-fitting. The best performance of the models resulting in 0.86 of R^2, which represents high capability to quantify the electricity saving potential. The contribution of this study is the proposed data-driven framework, which provides facility operators with reliable tools to accurately quantify saving potential of buildings. The conducted case study using the framework on 19 test buildings showed that facility operators could avoid unnecessary penalties by eliminating them to sign up for unrealistic targets, and help them to gain the most value out of the DR programs by knowing the true potential of their buildings.

Keywords –

Facility Management; Demand Response; Electricity Saving Potential; Data-driven; Decision Trees; Ada Boost; Random Forest; Energy Efficiency

1 Introduction

The electricity demand in the U.S. is in incline [1], which increases the pressure on power grids; hence the chance of electricity blackouts. In New York and California, electricity blackouts caused billions of dollars of loss to businesses and individuals [2][3]. Demand response (DR), one of the demand side management (DSM) techniques, is able to provide the necessary flexibility to the grid by incentivizing end-consumers to reduce their electricity demand during certain periods such as peak and emergency situations [4]. Meanwhile, building sector accounted for 74% of the total electricity consumption in 2016 [5], making buildings significant candidates for DR programs. In New York, buildings that were enrolled in DR programs managed to provide more than 31K MW/year of load curtailment and millions of benefits in recent years [6][7]. Hence, maximizing energy saving potential in buildings for DR is essential for peak reduction and energy savings.

The saving potential of a building (i.e., DR enrollment) refers to the amount of electricity that the building can save during a DR event. The problem in the current practice is that this potential is usually calculated based on simplified information such as design specifications of equipment in the building or historical metering data, resulting in the loss of opportunities to know the true energy-saving capacity of buildings. In the current practice, end-consumers usually work with third aggregators to determine the saving potential during DR events and customize their DR protocols. DR protocols are instructional statements for building operators to follow to operate major equipment in that building. For buildings which participated in DR program earlier, a yearly assessment is conducted by DR engineers and they will simply increase or decrease the DR enrollment by comparing the average performance on savings for events that happened in that year with the previous DR enrollments [30].

Buildings and the operation procedures are inherently much more complex than the simplified calculation due to their interconnections among their diverse systems [10]. Therefore, using the simplified calculation to estimate the energy saving potential can result in over or underestimation of the DR enrollment of buildings. Such performance issues were observed in the case-buildings analyzed for this study, as shown in Figure 1. The graph on the left indicates that *Building A* did not meet the DR enrollment (determined using the design information), which results in penalties for the end-consumer for exceeding the consumption beyond the enrollment value. The figure on the right illustrates that *Building B* saved more electricity during the event than the DR enrollment value, which indicates the potential money left on the table for the facility for not enrolling in more. Partly the problem was due to the lack of consideration of the context of the building during the event time such as different baseline values (i.e., the amount of the electricity that a building usually consumes under the same condition without DR operation), weather condition, and event time. These and similar examples from the literature [30] show that calculating saving potentials in DR programs simply using average benchmarking values from generic design specifications and not taking into account the context around event times result in a loss for building owners in either way. Hence, facility operators and third aggregators are in need of tools to accurately estimate the true potential of buildings by relying on integrated information including protocols, historical energy performance data, and contextual data around the event times (e.g., baseline values corresponding to the event times, weather data, etc.).

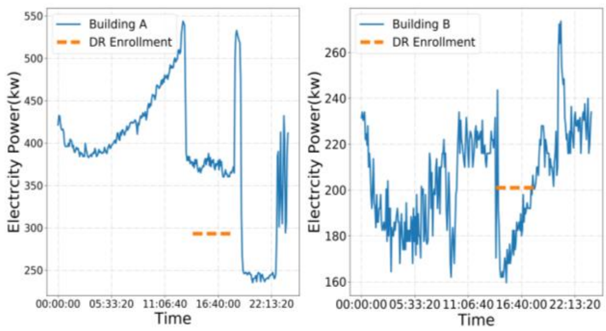

Figure 1. Performance of two buildings (Building A on the left, Building B on the right) during a peak event as compared to the DR enrollment (also marks the time range that an event occurred on the horizontal scale)

The objective of this paper is to quantify the true saving potential of buildings during DR events by integrated analysis of historical electricity consumption data, weather condition, DR protocol statements, and DR profiles (which provide information about DR event times, electricity consumption baselines, and DR enrolment values for buildings). Information from DR protocols can provide insights about buildings and equipment, and the electricity meter provides historical electricity consumption of buildings operated as the DR protocol instructions. In addition, the authors included weather condition information as it plays a crucial role in DR events and electricity consumption of buildings. A framework was developed by the authors for the stated objective, which is composed of two major modules: DR-dataset pre-processing and saving potential estimation. The details of the framework are presented in Section 3. The framework has been evaluated using data and protocols from nineteen buildings and six DR events that occurred in 2016.

The paper is structured as follows. A comprehensive review of the previous efforts on quantifying the saving potential of buildings in the literature has been provided in Section 2. In Section 3, the authors overview the framework and present the criteria to evaluate it. The implementation of the framework on nineteen buildings along with the results is discussed in Section 4. Section 4 also provides discussions on the challenges and future work. Conclusions are presented in Section 5.

2 Literature Review

There are three general types of models that have been used in previous studies to quantify electricity consumption and DR electricity saving potential: physical, statistical and simulation, and hybrid models [10]. Studies that developed physical models use equations and physics laws to predict saving potentials [8][9][18][19]. A large group of previous research studies concentrates on heating, ventilating, and air conditioning (HVAC) systems due to the fact that they account more than half of the energy consumption in buildings. Some of these studies developed simplified equivalent thermal parameters model to simulate saving potential for HVAC systems by adjusting the set-points [19-21]. While other studies implemented physical-statistical models to simulate the impact of adjusting parameters of HVAC systems on electricity consumption, which can further inform about the saving potential of HVAC systems during DR events [10][20][22][29]. Instead of using only HVAC system models, researchers also implemented physical models that include other systems that utilize electricity such as lighting, refrigerating, and appliances, and then quantified the saving potential by simulations of the testbeds [23][25]. These studies provide valuable insights of physical mechanism and potential knowledge. However, they usually are limited to the simplification of model equations and lack of consideration of the stochastic behaviors that happen in buildings- resulting in poor performance [10].

Statistical models, on the contrary, are developed based on experimental data. Large national level datasets of building energy use have been studied to identify the energy saving potential by comparing energy usage and data-driven saving analysis [26][27], yet the results from these studies are often in a very coarse-resolution, which cannot help DR engineers when determining the saving potential of a building for upcoming DR events. Researchers also compared the performance of both physical and simple statistical models for forecasting energy consumption in residential buildings as hybrid models and concluded that they both work with the slightly better performance than the statistical model (artificial neural network) [28]. Other than studies that emphasize the use of overall electricity consumption, load profiles of major household appliances were also utilized to identify their saving potential [24]. Despite the achievement of good results these models had in predictive modeling, they are often limited in the scope of certain purpose they aimed at in a restricted number of buildings/systems due to their inherent computational complexity [10].

This study is motivated by the need for accurately estimating the saving potential of different buildings during DR events. The framework developed in this study aimed at estimating the generic building-level saving potential instead of system-level saving potential by integrating the whole building energy consumption data, DR protocols, and contextual data around event times. This study differs from the studies in literature by providing a way to take into account contextual datasets that relate to DR assessment and analyze them in an integrated way using state of the art data-driven approaches.

3 Research Approach

3.1 Overview

The framework contains two main modules, which identify and extract the attributes to train the models, integrate datasets, and then fit models to estimate the electricity saving potential of buildings. The input of the DR dataset pre-processing is the DR protocol statements along with the building type, and a set of attributes for each building are extracted by the authors manually. Then, the extracted DR attributes dataset is integrated with other datasets from different sources, such as weather condition dataset, electricity meter dataset, and DR profiles (event time, electricity consumption baseline, DR enrollment). The second module, saving potential estimation, takes the integrated attributes dataset as input to fit machine learning models to estimate the electricity saving potential of buildings and provides the best performing model as a decision-making solution for facility operators for determining the true saving potential of buildings.

3.2 DR Dataset Pre-processing

In this study, the authors examined the DR protocols for a group of buildings to extract related attribute-pairs along with a building attribute: Building_type. The authors first categorized actions based on the equipment types and recorded the quantity of the impacted areas and equipment types. Details of this study are provided in a recent publication [11]. As stated in [11], DR protocols for buildings include five types of equipment: HVAC units (e.g., Air handling, fan coil, fan power units), fans, lights, elevators, and appliances. In this study, appliances are excluded because of its high dependence on occupancy data and lack of access to data on occupancy in spaces. For HVAC units, fans, and lights, the quantity of each equipment along with affected areas were extracted. Therefore, the data on attribute-pairs included: equipment_action and equipment_quantity for each equipment type.

The next step was to link the extracted data to the rest of the datasets. The weather information during the DR events is acquired from weather underground API [12], and the weather attributes included weather condition, temperature, humidity, and wind speed. The data on the DR profile of buildings included the electricity consumption baselines, the enrollment values and event times. More details of the datasets and the merging keys are provided in Section 4.2.

3.3 Saving Potential Estimation

3.3.1 Saving Potential Estimation Methodologies

In this module, the authors fitted several machine learning models to estimate the electricity saving potential of buildings. Because of the relatively small sample size and discrete categorical attributes in the studied problem, the authors chose decision tree regression model along with several boosting methods. There are mainly three types of decision trees: classification and regression trees (CART), C4.5, and C5.0 [13]. Among these decision trees, the CART is very similar to C4.5, yet it supports both categorical and continuous variables. Therefore CART is chosen to fit the dataset and estimate the electricity saving potential in this study. Furthermore, the authors used boosting algorithms such as Ada Boost and Random Forest to improve the performance of decision tree regression models. Both of the algorithms build multiple decision trees through iterations and take the average of the predicted value. Ada Boost is short for adaptive boost, which iterates the training process to build multiple decision trees and modifies the training data during each iteration and gives higher weight to the poorly modeled part [14]. Random forest algorithm, in addition to randomly selecting segmentations of the training data with replacement using the bootstrap method, also randomly selects the attributes when fitting the model.

3.3.2 Model Selection Criteria

One of the major advantages of the decision tree algorithm is that it is easier to interpret the result and can provide logic statements of the model [13][15]. However, it is also very easy to get over-fitted. In this study, the authors used four criteria, bias^2, variance, mean square error, and R^2, together to prune the model to avoid both over-fitting and under-fitting and meanwhile, aiming for adequately good prediction performance. Denoting N as the sample size of the training data, $\{f(1), f(2), \ldots, f(N)\}$ are the predicted values over the training data. The expected predicted value of the fitted model $f(x)$ is shown with $\overline{f(x)}$ (see Equation 1). Bias^2 captures the systematic error of the model (see Equation 2) [16]. When a model has a big bias^2 value, it indicates that the model is under-fitted, whereas big variance indicates the model is over fitted (see Equation 3).

$$\overline{f(x)} = 1/N \sum_{x=1}^{N} f(x) \tag{1}$$

$$Bias^2\left(f(x)\right) = \left(\overline{f(x)} - f(x)\right)^2 \tag{2}$$

$$Variance\left(f(x)\right) = E\left[\left(\overline{f(x)} - f(x)\right)^2\right] \tag{3}$$

The mean squared error (MSE) is composed of Bias^2 and Variance (see Equation 4), and measures the average squares of the errors, where Y denotes the observed data. Meanwhile, R^2 captures the capability of the model in explaining the observed data (see Equation 5).

$$MSE = 1/N \sum_{i=1}^{N} \left(f(i) - Y_i\right)^2 \tag{4}$$

$$R^2 = 1 - \left(\sum_{i=1}^{N} \left(f(i) - Y_i\right)^2 / \sum_{i=1}^{N} \left(Y_i - \bar{Y}\right)^2\right) \tag{5}$$

In Figure 2, the simplified relationship between variance, bias^2, and mean square error is demonstrated. As the complexity of the model increases, the Bias^2 decreases while the variance increases. There is an optimal point in the middle, where the variance meets with the Bias^2, and the mean square error is the lowest.

In this study, the authors visualized all four criteria to prune the models. For decision tree models, the authors iterated the depth of the tree from 1 to 10 to fit the models. For AdaBoost and Random Forest models, two parameters were included to iterate, the depth of the tree and the number of the tree estimators, both from 1 to 10. The optimal choice of the parameters was determined by the authors by visualizing the criteria, and the parameters with a lower bias, variance, MSE and higher R^2 were chosen. At last, the optimal models for each model were fitted and evaluated based on the average R^2 value among the cross-validation of the complete dataset.

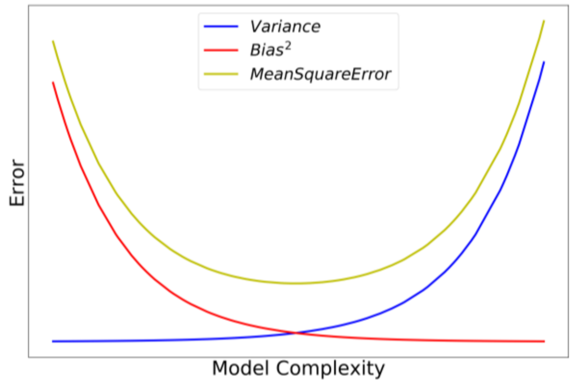

Figure 2. Simplified relationship between Bias^2, Variance, and MSE

4 Implementation of the Framework on a Case Study

4.1 Overview

The framework was tested on a case study of 19 campus buildings (including offices, academic buildings, and dorms) that participated in DR programs in 2016. There were six events that happened in 2016. During the events, data such as weather and electricity meter data was collected from different sources to quantify the saving potential of the buildings.

4.2 DR Dataset Pre-processing

The authors examined the 116 DR protocol statements for nineteen buildings manually to extract the DR attribute-pairs to fit the models. Table 1 listed three examples of the extracted attribute-pairs. As shown in Table 1, there are eight attributes identified from the protocol statements. However, not every protocol

Table 1. Three Examples of the Attribute-Pairs Extracted from the DR Protocols

Building ID	Building_ type	HVAC_ action	HVAC_ quantity	Light_ action	Light_ quantity	Fan_ action	Fan_ quantity	Elevator_ quantity
Building1	Office	Shut Off	4	Shut Off	5	Reduce Power	40%	1
Building2	Dorm	Shut Off	1	Shut Off	27	Shut Off	5	1
Building3	Office	Reduce Power	40%	Shut Off	4	None	-1	2

includes all the attributes, and such cases were reflected as none in the dataset. For example, the DR protocol for Building 3 in Table 1 does not contain any instruction for Fans. Therefore, the authors put ‘None’ as the Fan_action and ‘-1’ as the Fan_quantity.

In addition to the attribute-pairs extracted from the DR protocol statements, the authors also collected the weather condition data, DR profile data, and electricity meter data during the event (based on event times from DR profiles). The description of all the datasets is shown in Table 2. The datasets are integrated with each other by merging keys. For example, the DR attribute-pairs were merged to the DR profiles by linking the dataset based on the building ID. The integrated dataset was pre-processed using forward filling method, which means that the missing value is filled with the nearest previous data.

Table 2. Description of the Datasets

Datasets	Variables	Merging Key
DR Attributes Pairs	Building_type, HVAC_action, HVAC_attribute, Light_action, Light_attribute, Fan_action, Fan_attribute, Elevator_attribute	Building ID
Weather Data	Weather Condition, Temperature, Humidity, Wind Speed	Date-time
DR Profiles	Baseline, Enrollment	Building ID
Electricity Meter Data	Electricity Consumption	Date-time; Building ID

4.3 Quantifying Saving Potential in Buildings

4.3.1 Decision Tree

With the integrated dataset from the previous section, the authors fitted the decision tree models and pruned it by iterating the depth of the trees from 1 to 10. Figure 4 demonstrates the results of the MSE, R^2, Variance, and $Bias^2$ of the decision trees with different depths. As shown in Figure 3, MSE and $Bias^2$ decrease when the depth of the tree increases, whereas R^2 and Variance increase when the depth of the tree increases. When the depth of the tree is four, Bias drops drastically with relatively low variance and MSE, and with a fairly good R^2. Therefore, the depth of the decision tree is determined as four, meaning that in this case, the model can provide sufficient capability in estimating the saving potential and not suffering over-fitting and under-fitting issues.

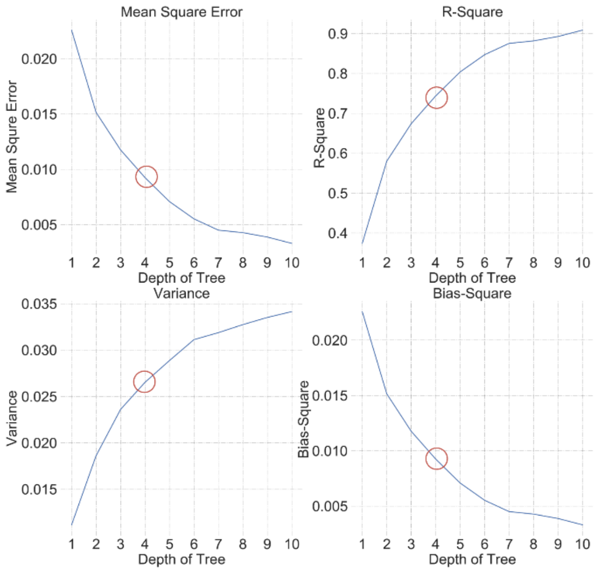

Figure 3. Comparison of MSE, R^2, Variance, and $Bias^2$ of the Decision Trees with Different Depth of the Trees

4.3.2 Ada Boosting

To improve the performance, the authors implemented Ada Boosting decision trees and pruned the parameters by iterating both the depth of the trees and number of the tree estimators from 1 to 10. The heat map of the four criteria is shown in Figure 4 with blue indicating a small value and red indicating a large value. As shown in Figure 2 and Figure 4, variance and $bias^2$ are negatively correlated. Hence the boxes with the color in middle range of the sidebar are selected as candidates. Furthermore, among the candidate boxes, the authors chose the parameters based on the R^2 and Occam’s razor law, which means that the square with a higher R^2 and less depth of tree and number of estimators will be chosen.

Figure 4 shows that when the depth of the tree is equal to five and the number of tree estimators is four, the MSE, variance, and $bias^2$ are smaller than the surrounding cells (meaning that the model is better fitted than the surrounding models), along with a relatively high R^2 (meaning that the model is capable of estimating the saving potential of the buildings). Therefore, the pruning process of Ada Boosting results in five as the depth of the tree and four as the number of estimators.

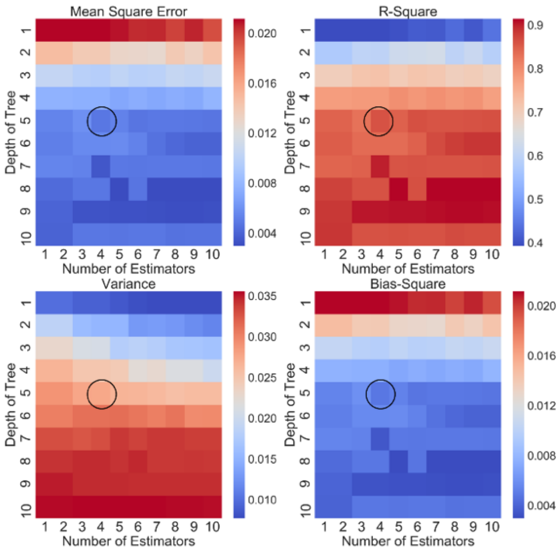

Figure 4. Heat map of MSE, R^2, Variance, and $Bias^2$ of the Ada Boost of Decision Trees with Different Depth of the Trees and Number of Estimator Trees

4.3.3 Random Forest

In addition to modifying the training sample to improve the performance, random forest fits the trees with different attributes from the integrated training dataset as well. The parameters for the random forest is the same as Ada boosting, the depth of the tree, and the number of tree estimators, and the best parameters are chosen based on the same process as Ada boosting as well. Figure 5 shows the heat map when pruning the random forest models. When the depth of the tree is six, and the number of tree estimators is four, the $bias^2$, variance, and MSE are all fairly small with a high R^2. Therefore, the best parameters for the random forest model, in this case, are six as the depth of tree and four as the number of estimators.

4.3.4 Comparison of the Models

The authors tested all the pruned models on the datasets using 20 folds cross-validation, which will randomly pick 70% of the data as the training sample to fit the model and 30% of the data as the testing sample to evaluate the performance for 20 times. The pruned parameters and average R^2 from the cross-validation are shown in Table 3. As shown in Table 3, random forest with six as the depth of the tree and four as the number of tree estimators has the best performance in explaining the DR dataset and resulting 0.862 as the R^2, which indicates that the framework has better capability of estimating the true saving potential of the buildings in different types and scales.

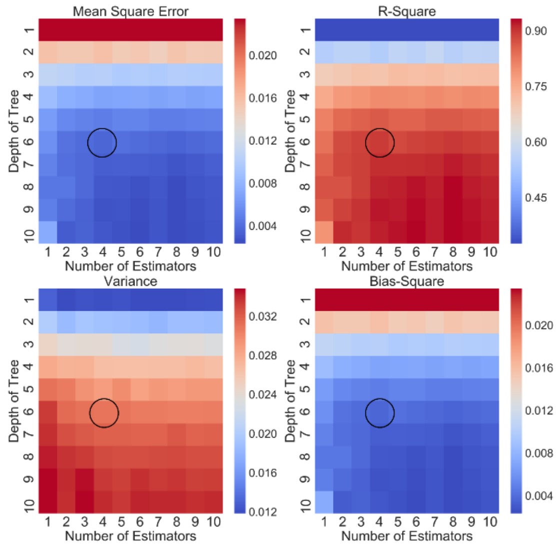

Figure 5. Heat map of MSE, R^2, Variance, and $Bias^2$ of the Random Forest of Decision Trees with Different Depth of the Trees and Number of Estimator Trees

The comparisons between the estimated electricity saving potential from the random forest model, the enrollment value of the buildings, and the actual curtailment during the DR events are shown in Figure 6. In Figure 6, X-axis represents the actual curtailment (percentage of the baseline) of the buildings, and Y-axis represents the estimated electricity saving potential (percentage of the baseline) during the events (predicted by random forest model or the enrollment value from the DR profile).

Table 3. Parameter and Performance of the Models

Model	Depth of Tree	Number of Trees	R^2
Decision Tree	4	-	0.743
Ada Boosting	5	4	0.828
Random Forest	6	4	0.862

The diagonal black line indicates the case when the actual curtailment is equal to the calculated electricity saving potential. The red dots illustrate the enrollment value from the DR profiles versus the actual curtailment, and the blue dots illustrate the estimated saving potential from the random forest model with respect to the actual curtailment. For the red dots, almost half of them indicates that the enrollment value is smaller than the actual curtailment during the events (the red dots that are at the right of the black line), which indicates over-performance of the buildings (i.e., facility operators signed up for a small enrollment value and resulting in saving more electricity) while half of them indicates that the enrollment value is larger than the actual curtailment (the red dots that are at the left of the black line), which indicates under-performance of the buildings (i.e., facility operators signed up for a large enrollment value and resulting in penalties due to less electricity saving). Meanwhile, the blue dots distribute closely around the black line, which demonstrates the capability of the framework for estimating the electricity saving potential for the studied campus buildings. These cases illustrated the potential of the framework in estimating the saving potential of buildings. By enlarging the training set and including more cases, the framework can have a more promising performance over a group of buildings that share common characteristics.

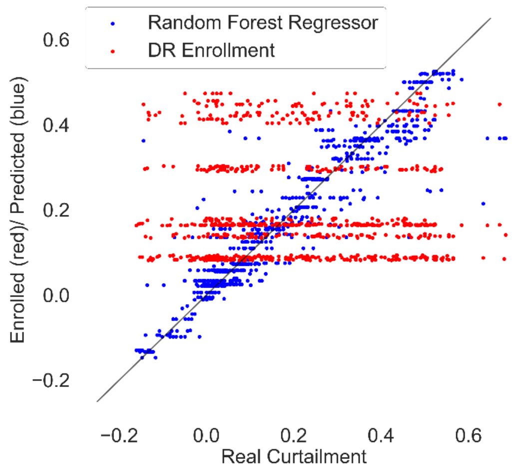

Figure 6. Comparison of MSE, R^2, Variance, and Bias of the Decision Trees with Different Depth of the Tree

4.4 Challenges and Future Work

When implementing the framework on the DR datasets, the authors faced multiple challenges. Firstly, DR protocol statements were not explicitly written, which result in inaccuracy and difficulties when estimating the electricity saving potential of buildings. For example, when a DR protocol statement instructs the facility managers to shut down the AHUs in common areas, it is hard to quantify the saving potential since the quantity of involved equipment remains vague. Secondly, occupancy status in buildings plays an essential role impacting the electricity consumption, which gives inherent complexity in predicting the electricity consumption. Furthermore, occupancy data is rarely available in buildings, and the available occupancy data usually contains a large amount of noise. These situations lead to difficulties when estimating the electricity saving potential considering the uncertainty of the occupancy status.

For future work, the authors plan to include more data to improve the performance of the framework presented in this study. In addition to extracting coarse building and equipment information from DR protocols, the authors intend to include Building Information Models (BIMs) to provide building and equipment configuration information to reduce the vagueness in DR protocol statements. Furthermore, the authors aim to collect DR related equipment sensor readings from building automation systems, which will provide much more details about the equipment behaviors during event times.

5 Conclusion

In this study, the authors presented a data-driven framework to estimate the electricity saving potential of buildings. The framework developed in this study fills the gap in the previous studies by providing a granular building-level data-driven approach to estimate the electricity saving potential for buildings by integrating various data sources containing data for DR events. By implementing the framework on 19 campus buildings that vary in type and scale, the authors demonstrated the capability of it in estimating the saving potential of buildings during DR events. The results from the case study indicate that the building owners and facility operators can benefit from the accurately determined saving potential by avoiding penalties and making the most out of the DR programs. The framework is extensible by integrating more equipment sensor data from BAS and BIM in the future to further improvement of the performance.

References

[1] U. S. Energy Information Administration. *Annual energy outlook 2017*. 2017.

[2] Electric Consumer Research Council (ELCON). The Economic Impacts of the August 2003 Blackout. 2004.

[3] Lawton L., Sullivan M., Van Liere K., Katz A. and Eto J. A framework and review of customer outage costs: Integration and analysis of electric utility outage cost surveys. *Lawrence Berkeley National Laboratory*, 2003.

[4] Qdr Q. Benefits of demand response in electricity markets and recommendations for achieving them. *US Dept. Energy, Washington, DC, USA, Tech. Rep*, 2006.

[5] U. S. Energy Information Administration. *Electric Power Annual 2016*, 2017.

[6] Neenan B., Pratt D., Cappers P., Doane J., Anderson J., Boisvert R., Goldman C., Sezgen O., Barbose G. and Bharvirkar R. How and Why Customers Respond to Electricity Price Variability: A Study of NYISO and NYSERDA 2002 PRL Program Performance (No. LBNL--52209). *Ernest Orlando Lawrence Berkeley National Laboratory*, Berkeley, CA (US), 2003.
[7] Federal Energy Regulatory Commission. *Assessment of demand response and advanced metering*, 2008.
[8] Brandemuehl M.J. and Braun J.E. The impact of demand-controlled and economizer ventilation strategies on energy use in buildings. *ASHRAE Transactions*, 105:39, 1999.
[9] Yin R., Xu P., Piette M.A. and Kiliccote S. Study on Auto-DR and pre-cooling of commercial buildings with thermal mass in California. *Energy and Buildings*, 42(7):967-975, 2010.
[10] Lü X., Lu T., Kibert C.J. and Viljanen M. Modeling and forecasting energy consumption for heterogeneous buildings using a physical–statistical approach. *Applied Energy*, 144:261-275, 2015.
[11] Yu X. and Ergan S. BIM Coverage in Demand Response Management: A Pilot Study in Campus Buildings. In *Construction Research Congress 2018*, New Orleans, LA, United States, 2018.
[12] Weather Underground API. Online: https://www.wunderground.com/history/, Accessed: 27/12/2017.
[13] Tso G.K. and Yau K.K. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *Energy*, 32(9):1761-1768, 2007.
[14] Freund Y. and Schapir R.E. A desicion-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory*, pages 23-37, Springer, Berlin, Heidelberg, 1995.
[15] Perner P., Zscherpel U. and Jacobsen C. A comparison between neural networks and decision trees based on data from industrial radiographic testing. *Pattern Recognition Letters*, 22(1):47-54, 2001.
[16] Dietterich T.G. and Kong E.B. Machine learning bias, statistical bias, and statistical variance of decision tree algorithms. *Technical report*, Department of Computer Science, Oregon State University, 1995.
[17] Fortmann-Roe S. Understanding the bias-variance tradeoff, 2012.
[18] Ascione F., Bianco N., De Masi R.F., de'Rossi F. and Vanoli G.P. Energy refurbishment of existing buildings through the use of phase change materials: Energy savings and indoor comfort in the cooling season. *Applied Energy*, 113:990-1007, 2014.
[19] Lu N. An evaluation of the HVAC load potential for providing load balancing service. *IEEE Transactions on Smart Grid*, 3(3):1263-1270, 2012.
[20] Ghahramani, A., Castro, G., Becerik-Gerber, B. and Yu, X. Infrared thermography of human face for monitoring thermoregulation performance and estimating personal thermal comfort. *Building and Environment*, 109:1-11, 2016.
[21] Yin R., Kara E.C., Li Y., DeForest N., Wang K., Yong T. and Stadler M. Quantifying flexibility of commercial and residential loads for demand response using setpoint changes. *Applied Energy*, 177:149-164, 2016.
[22] Chassin D.P., Stoustrup J., Agathoklis P. and Djilali N. A new thermostat for real-time price demand response: Cost, comfort and energy impacts of discrete-time control without deadband. *Applied Energy*, 155:816-825, 2015.
[23] Muratori M., Roberts M.C., Sioshansi R., Marano V. and Rizzoni G. A highly resolved modeling technique to simulate residential power demand. *Applied Energy*, 107:465-473, 2013.
[24] Pipattanasomporn M., Kuzlu M., Rahman S. and Teklu Y. Load profiles of selected major household appliances and their demand response opportunities. *IEEE Transactions on Smart Grid*, 5(2):742-750, 2014.
[25] Rotger-Griful S., Jacobsen R.H., Nguyen D. and Sørensen G. Demand response potential of ventilation systems in residential buildings. *Energy and Buildings*, 121:1-10, 2016.
[26] 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*, 140:85-93, 2015.
[27] Yu Y. and Guo J. Identifying electricity-saving potential in rural China: Empirical evidence from a household survey. *Energy Policy*, 94:1-9, 2016.
[28] Neto A.H. and Fiorelli F.A.S. Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. *Energy and buildings*, 40(12):2169-2176, 2008.
[29] Aduda K.O., Labeodan T., Zeiler W., Boxem G. and Zhao Y. Demand side flexibility: Potentials and building performance implications. *Sustainable Cities and Society*, 22:146-163, 2016.
[30] Bârgăuan B., Fati O., Ceclan A., Micu D.D., Șteț D., Czumbil L. and Muresan P. Demand response on blocks of buildings - Romanian pilot site innovation project. In *Modern Power Systems (MPS), 2017 International Conference, IEEE*, pages 1-4, Cluj-Napoca, Romania, 2017.