AN INVESTIGATION OF THE IMPACT OF CUSTOMER BASELINE (CBL) CALCULATION ON PEAK TIME REBATE PROGRAM

by

Saeed Mohajeryami

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Economics

Charlotte

2016

Approved by:

____
Dr. Peter Schwarz

____
Dr. Badrul Chowdhury

____
Dr. Benjamin W. B. Passty

ABSTRACT

SAEED MOHAJERYAMI. An investigation of the impact of Customer Baseline (CBL) calculation on peak time rebate program. (Under the direction of DR. PETER SCHWARZ)

In this thesis, the impact of customer baseline (CBL) accuracy on the Peak Time Rebate (PTR) program is investigated. In a hypothetical case, PTR is offered to the residential customers and its economic performance is evaluated with respect to the CBL accuracy performance. Since this program relies on CBL for payment settlement, its performance hinges on the accuracy of such calculations. The accuracy of CBL calculations are studied for residential customers. Moreover, for the purpose of this investigation, popular CBL methods of High5of10 (NYISO), Low4of5, Mid4of6, exponential moving average (ISONE) and regression methods and their adjusted forms are selected for CBL calculation. Then, this calculated CBL is utilized to examine the performance of a case of PTR program. The case consists of 262 residential customers. According to the results, in this case study, utility pays at least 50 percent of its revenue as a rebate just because of the inaccuracy of CBL methods. This loss increases if the aforementioned CBL methods get adjusted for their morning consumption. At the end, it is discussed that PTR can cause a significant loss to the customers and cause unfair redistribution of the utility’s revenue. It is shown that aforementioned inefficiencies are because of the failure of CBL calculation methods to predict the customer’s load profile on the event day.

Index Terms— Customer Baseline (CBL), Demand Response (DR), Peak Time Rebate (PTR), percent accuracy metrics, percent bias metrics.

DEDICATION

This thesis is dedicated to the memory of my father, Moslem, and my mother Pari, and my kind and generous wife for their enduring love and support.

ACKNOWLEDGEMENTS

I would like to thank Dr. Peter Schwarz for his mentoring and sharing of insight in electricity markets concerns, as well as his contributions in preparing this thesis. I would like to thank Dr. Badrul Chowdhury for his support, expertise and graciously giving of his time, over the last several years in academia. I would like to thank Dr. Benjamin Passty for his interest in the topic and his input. Additionally, I would like to thank Electric Ireland and Sustainable Energy Authority of Ireland (SEAI) that made the full anonymized data sets from its recent Smart Metering trials for electricity customers available and created this opportunity for the researchers to work on the real data of residential customers.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Introduction

The electricity markets are developing steadily over many years of restructuring and competition. However, there are still some areas in this industry are kept isolated from the market's advancement, one of which is demand-side. Indeed, this area is underdeveloped due to the detachment from market price fluctuations as the regulatory bodies attempted to give immunity to retail customers vulnerable to such fluctuation. For instance, during 2000 and 2001, California experienced a major power crisis under the restructured wholesale market. Although numerous factors could be listed as a reason to create this crisis, most people agree that the lack of demand response exacerbated the situation [1]. Moreover, recent studies demonstrate that demand response (DR) programs can provide an environment that the customers could actively be engaged in the optimization process and change their consumption pattern in response to the wholesale market price signals.

These programs can potentially create a number of possibilities for the system operators and utilities to improve both economic and technical indices of their system. As a matter of fact, power system operators can compensate for the lack of their supply during the peak time by using DR resources [2]. The expensive peak time resources deployed by the supply-side to meet the demand could impose a huge cost burden to the customers. Additionally, it is estimated that the capacity to meet demand during the top 100 peak hours (1.1% of year) accounts for 10-20% of electricity cost annually [3]. On the other hand, utilities can also benefit from DR by taking advantages of lower prices offered by such resources compared to electricity spot market.

Although DR programs look very promising in theory, in practice a host of problems make them difficult to implement. These problems are rooted in diversity of customers, loads and heterogeneity in types of DR programs. Because of all these complications, policy makers are concerned about the way load aggregators compensate customers. Many DR programs rely on customer baseline (CBL) to compensate customers financially for their load reduction. It is worth mentioning that the load reduction in this context refers to a customers' response to a financial incentive and it is basically a change to the customers' normal consumption pattern. Indeed, to detect this shift from the normal pattern, first, it is necessary to find the normal pattern. CBL is a counterfactual consumption level, i.e. the amount of electricity that customers would have consumed in the absence of a DR event. DR event refers to a day that utility believes that the consumption is higher than the level they can meet. In days preceding a DR event, utilities inform customers to lower their consumption on the DR event. CBL is also a basis to measure the performance of DR programs. Moreover, a well-designed baseline could benefit all stakeholders by aligning their incentives, actions and interests. However, baselines are a challenging aspect of DR programs because they represent a "counterfactual", something that is not observable. In this thesis, different methods employed by the industry and the way they have dealt with baseline uncertainties will be reviewed.

In this thesis, the CBL for residential customers are studied whereas previous works in this area merely focus on industrial and commercial customers. Industrial customers as opposed to residential customers have a high degree of predictability due to their pre-scheduled loads [4]. Therefore, this author believes that the findings for industrial customers could not be generalized to residential customers. This thesis goes beyond analyzing accuracy and bias metrics of CBLs and explains how these metrics translate into financial losses for utility and customers. In order to achieve this, an investigation of the economic performance of a case of PTR for residential customers is undertaken. For this investigation, real data of residential customers are employed. In the future sections, the data and the implementation will be explained in detail as the results will be presented.

CHAPTER 2: LITERATURE ON CUSTOMER BASELINE

In this chapter, the different studies about CBL measurement and verification are reviewed and their advantages and disadvantages are highlighted. This review is necessary to shed light on the shortcomings of the existing approaches and find possible avenues for improvement. It is necessary to mention that these findings are based on industrial customers and they must be reexamined for residential customers. However, these findings provide a good starting point. Moreover, CBL challenges are studied in this section. Afterward, CBL from both the customers and utility's points of view are discussed.

2.1 Measurement and Verification

Several recent studies have reviewed and analyzed different methods for calculating demand response customer baseline [5-13]. In this section, many of these methodologies are introduced and compared in both the measurement and verification process. CBL, as previously discussed, is an estimate of the amount of energy the customer would have consumed in the absence of a DR event. Figure 1 illustrates the concepts of DR baseline, actual load and estimated load reduction [5]. Indeed, these three concepts are repeatedly employed hereafter to explain the different methodologies and their measurement and verification processes.

The most extensive review of CBL methods is provided in [6]. This paper examines empirically numerous methods used by utilities and ISOs within the US. For carrying out such a task, it employs the real data from California State. Moreover, in order to evaluate the performance of these methods, the accuracy and bias metrics are utilized. These metrics will be comprehensively elaborated in the future sections of this thesis. This paper concludes from its results that the same day additive or multiplicative adjustment has superior performance to an unadjusted CBL or a CBL using the weather sensitive adjustment. However, the choice of multiplicative or additive adjustment does not change the outcome significantly. Furthermore, this work shows that X of Y methods such as CALISO, PJM economic, mid 4 of 6 and regression approaches, with a same day additive adjustment, produce similar satisfactory results. Y in this context refers to the number of non-DR days before the DR event and X refers to the number of days selected out of these Y days that have certain consumption characteristics. For instance, in HighXofY methods, X are the number of days with highest consumption out of Y days. However, these methods performed poorly for predicting load for variable load customers. Moreover, this paper shows that the regression approaches have higher administrative costs and associated complexity compared to X of Y methods. Therefore, it is recommended not to pursue such approaches to calculate the baseline. One of the striking findings of this work is that explicitly weather–dependent models did not generally outperform models that did not include weather. But given the weather stability of California State, this finding is recommended to be examined in the other States in order to verify its credibility. Incidentally, for readers unfamiliar with the aforementioned methods, they are comprehensively explained in the future sections of this thesis.

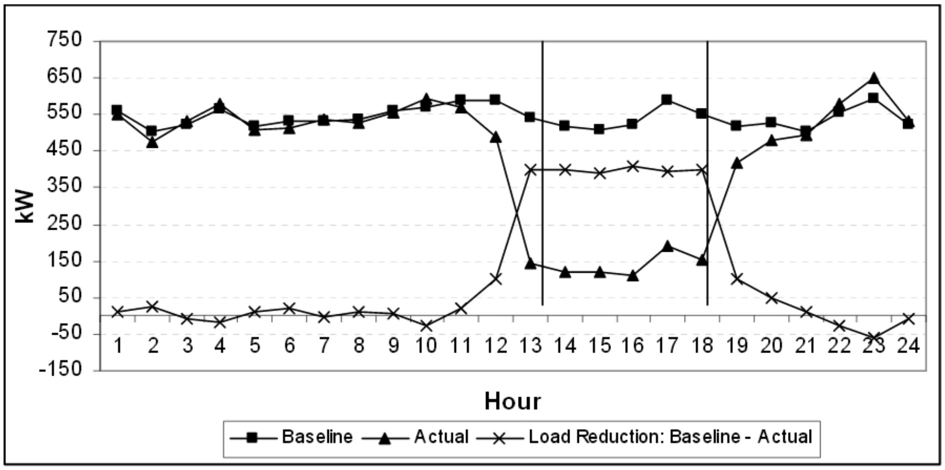

FIGURE 1: Illustration of CBL, actual load and estimated load reduction [5]

Another study from California State is also an analysis of methods to estimate customer baselines [5]. This study is part of the broader evaluation of California's 2004 demand response programs which targeted the industrial and commercial customers. The methods examined in this study are 3-day, 10-day and prior-day baseline methods. In this study, the impacts of load size (small, medium, large or extra-large), business type (commercial, industrial and institutional), event day type (high demand, low demand or consecutive high demand) and weather parameters are investigated on the performance of the aforementioned baseline methods. According to the results, 10-Day Baseline with Same Day Adjustment is the most accurate of the methods. Moreover, 3-day baseline methods is ranked second. The prior-day method shows the poorest performance among the aforesaid methods in terms of variability and predictive accuracy. However, this method has the lowest bias. The obtained results are further confirmed in their next complementary works [7-8].

In another attempt to evaluate CBL methods, [10-11] evaluates CBL methods' performance on non-residential buildings in California. This work is done in Lawrence

Berkeley National Lab (LBNL) on sample data from 32 sites in California. The methods investigated in this study overlap with [5, 7-8]. However, the approach used for weather effect adjustment is different. This work employs a statistical analysis to evaluate the performance of different CBL models for non-residential buildings participating in the DR program with emphasis on the importance of weather effects. According to the results, applying the morning adjustment could significantly reduce the bias and improve the accuracy of all CBL models. Moreover, this work suggests that the characterization of building loads by variability and weather sensitivity could serve as a useful screening indicator that could be used to predict which types of CBL methods suit better to each particular load. However, it shows that none of the examined CBL methods produce satisfactory results for highly variable loads. These types of customers are difficult to characterize with standard CBL models that rely on weather and historical data. Based on the results, this work also suggests that different CBL models should be utilized for different customers. For example, many commercial and institution buildings are weather-sensitive, while industrial customers are not sensitive to the weather. Therefore, the baseline for each load should be based on a method specifically tailored to its characteristics. On the contrary to [6], this work finds that incorporating temperature (e.g. explicit weather models) improves accuracy of the estimated CBL and in cases where it does not improve the accuracy, it has relatively little impact.

Moreover, in a similar attempt, [12] has employed a statistical approach to examine three CBL methods in South Korea. These CBL methods include 10 prior weekdays averaging, 5 highest of 10 prior weekdays averaging and 8 days of 10 prior weekday averaging excluding maximum and minimum. In this paper, the data from 10 industrial customers at 10 weekdays in Nov. 2008 are used to select the best CBL method in terms of minimum total error. It is found that the 10 prior weekdays averaging method works slightly better than the others. However, when these CBL methods are adjusted, all of them have produced very close, good results.

The choice of X relative to Y is one of the key questions in X of Y customer baseline methods. [14] has tried to investigate this question by examining 306 sites from several ISO and utilities' customers. 540 baselines have been produced for each site (i.e. combination of several scenarios) to examine all these baselines for 3 proxy days (non-event days similar to event days) in 2008 and 2009. According to the results, there appears to be a range of X/Y values which minimizes the bias. For all ranges of examined adjustments and all values of Y, error is minimized by $0.4 \leq X/Y \leq 0.8$. Thus, High 3 of 10 is less accurate than High 5 of 10 and on the other hand, High 7 of 10 is more accurate than High 10 of 10. Moreover, this study has shown that the choice of X relative to Y is less critical when the baseline is adjusted. Furthermore, this study has investigated the effect of "cap" on the accuracy. According to the results, uncapped adjustments have slightly higher accuracy than capped adjustments.

Although finding the right model for CBL is of the utmost importance for a variety of purposes such as measurement and verification, improving DR program design and operation and also financial settlement of DR participation rewards, there are some details in implementation which could plague the efficiency of CBL models and render them ineffective. [15] has tried to compare the different implementation choices to see their impact on DR performance results and the baseline estimation. This study has employed data from 38 large commercial buildings and industrial facilities in the Pacific

Gas and Electric Company (PG&E). The data are a collection of 15-min interval whole-building electric demand and they belong to an automated critical peak pricing (CPP) program between 2007 and 2009. In what follows, the findings of this work are summarized briefly. One of the CBL implementation choices is the selection of weather station, typically based on the proximity of weather station to the target site. In this work, the data from two weather stations are utilized and their impact on the load reduction estimate is examined. According to the results, the load reduction estimate is strongly sensitive to the source of the outdoor air temperature data. Moreover, it is found that the choice of power outage filter strongly affects the load reduction estimate. Power outage filter refers to the process in which some minimum power consumption days are filtered out. These days are those with consumption less than x percent of the average minimum daily power consumption. Three filters of x=0 (no filter), x=50 and x=75 are examined in this work. As previously mentioned, the load reduction estimate is strongly sensitive to the choice of filter. On the other hand, the load reduction estimate is less sensitive to data alignment choice. Temperature and demand data are compared for two cases: a case with 15-min offset and the other case that aligns temperature and data by time stamp (no offset). Moreover, it is found that the load reduction estimate is almost insensitive to the choice of data interval. In this study three cases of 15-min, 30-min and 60-min data intervals are examined. It suggests that it is acceptable to use 60-min interval rather than 15-min interval and it could simplify metering needs and reduce the computational burden. This study is further extended into [16] that only elaborated the details of the aforementioned findings.

Furthermore, implementation choices can affect the performance of regression models. Therefore, it is important to examine the explanatory variables' choices to see how they could be best employed in the regression model. [17] employs the similar data of [15-16] to provide a new regression model. In order to find the proper explanatory variables, it examines load sensitivity to outside air temperature and representative load pattern derived from cluster analysis of CBL. According to the results, the load is sensitive to outside air temperature. Therefore, it could be a suitable candidate as an explanatory variable. Moreover, it is found that the cluster analysis and its algorithms are effective tools to estimate CBL. Given the uncertainty of business types, cluster analysis could serve as a useful tool to categorize the data into the different types. This categorization could be utilized as an explanatory variable in the regression model. Indeed, combination of load sensitivity and cluster analysis improves the performance of the regression models. However, the goodness of fit of regression models, which is expressed by the coefficient of determination, adjusted R^2, is still not ideal.

In addition to the proper implementation choices, several papers have introduced some modifications in order to improve the accuracy and bias of the established CBL methods. [18] has proposed an exponential smoothing model to calculate CBL. This model weighs past observations with exponentially decreasing weights. In this model the recent changes are better reflected in the estimated baseline. The proposed method shows superiority over High5of10 and regression methods whether they are adjusted or not.

Moreover, authors in [19-20] proposed a CBL calculation framework employing data mining techniques. This paper employs the real data from a large industrial building complex in Korea. The number of deployed smart meters in this complex is approximately 2,500. These real data are used to analyze the customers' electricity consumption behaviors for DR. This method utilizes two data mining techniques of Kohonen networks model (self-organizing map) and the unsupervised learning algorithm (k-means clustering). It starts with data preprocessing to remove the outliers and data inconsistency. Then it classifies the load in terms of the seasons of the year and type of weekday to reduce the data size. Self-organizing map (SOP) then is employed to map a multi-dimensional input space onto an output space with greatly reduced dimension. Afterward the SOP output is fed into k-means clustering which partitions the data set into k clusters. The data of each cluster can be used to determine CBL. This method has two key differences from the aforementioned CBL methods. First, unlike the other methods, it finds the most similar day given partial data. Second, it considers other parameters such as average temperature, the gradient of electricity consumption and occupied/unoccupied status. The aforesaid parameters are extracted by learning process through SOM and k-means clustering. According to the results, compared to the averaging methods, the root mean square error is reduced by 15-22% on average and the mean absolute percentage error is reduced by 15-20% on average as well.

Although CBL is best known in the context of DR programs, it may serve other functions. [21] proposed a methodology that employs CBL to identify non-technical losses of the system such as theft of electricity. This methodology utilizes historical demand of a certain customer to estimate the future consumption. Then it compares the estimate with the actual load in order to identify a suspicious reduction.

In a patent application, authors of [22] provided a day-ahead load reduction system based on CBL for inducing a user to efficiently manage her energy consumption.

This system applies an incentive to achieve the desired load reduction and load decentralization. The authors also filed another patent [23] to present a load forecasting analysis system for calculating CBL based on the day-ahead reduction system mentioned before. The patent introduces a system which is composed of multiple components such as the CBL forecaster, a period selector for selecting conditions for forecasting, and eventually a CBL determiner to calculate the error value.

2.2 CBL Challenges

As previously discussed, CBL is the amount of electricity that customers would have consumed in the absence of a DR event. So far, several works related to the measurement and verification of CBL and their challenges are reviewed and discussed. However, there are some other challenges in this area that must be addressed. Many of these challenges stem from the design of DR programs and the nature of the people participating in these programs. In other words, the methods introduced so far neglected many of the real world challenges. But, these practical challenges are as important as the theoretical challenges and a CBL methodology must be able to address such challenges. In what follows some of these issues are reviewed and discussed.

Since participation in DR programs is voluntary, customers have advantage over the utilities in terms of consumption information. The customers know more about their baseline than the utilities. This asymmetric information imposes two possible challenges for utilities and consequently for CBL calculation. In another words, they can be used to game the system. They are an adverse selection and moral hazard problem. The adverse selection problem arises when customers with lower consumption anticipation have more incentive to participate in the program. Therefore, the participation is more likely to be disproportionate. The moral hazard problem arises when customers engage in activities to change their CBL. This means customers might change their normal consumption pattern to affect the future CBL [24]. In what follows, some of the examples of these challenges are described.

a) Baseline manipulation

The improper methodology to determine CBL can encourage the customers to inflate their baselines in order to gain a higher payment from the program. As discussed earlier, the moral hazard problem is a practical challenge and in the absence of an effective mechanism to handle such issues, the customers have incentive to change their CBL. Such manipulations are observed and reported by ISO New England [25]

b) Load shifting behind multiple meters

The customers who have large consumption and several meters can game the system by changing the consumption behind each meter in a way to create illusory demand reduction. The concept is elaborated with a numerical example in [24]

c) Generation relocation and inefficient price formation

This problem is better illustrated by an example. Assuming that the utility offers the flat rate of $70/MWh. It is also possible for the customer to sell her demand reduction as an energy supply in the wholesale market. Suppose the customer bids $80/MWh for demand reduction and the next cheapest generating unit offers $100/MWh in the wholesale energy market. In this case, the customer's bid will be cleared in the market. Therefore, the customer has incentive to use an on-site backup generator that is able to produce $150/MWh or lower. In this case, each MWh consumed by the customer from the backup generator costs $150 or less, but she earns $80/MWh for the demand reduction; therefore, every MWh costs her $70 or less. In this example, everything looks the same for the customer and she consumes energy still with $70/MWh. However, from society's point of view, it's a big loss because the electricity cost increased to $150/MWh rather than $100/MWh. This setup encourages an inefficient investment since energy can often be produced more efficiently in the wholesale market [26].

Furthermore, there are some problems in practice that can challenge customers' decisions. For instance, the poor accuracy performance of CBL also can undermine the efficiency of DR programs that rely upon CBLs for payment settlement. This issue, for some CBL methods, is explained and investigated in detail in [27].

Practical challenges like what have been described so far could plague the effectiveness of CBL methods. Indeed, the performance of some of DR programs hinges on their CBL performance and if CBL could not deliver what is expected of it, it will deteriorate DR performance significantly. From an economic perspective, a properly established customer baseline should meet the following two conditions. First, the customer should be punished for extra consumption above her customer baseline. Second, the customer should be rewarded for its load reduction. In essence, a properly designed customer baseline is two-sided so that demand reduction and demand increase are treated symmetrically.

If the first of the two conditions fails, the customer baseline becomes one-sided; this means that the consumer would still be able to consume electricity above the baseline as permitted in the retail tariff. If both conditions fail, then the customer doesn't have any incentive to partake in the program effectively. Therefore, it could create distorted incentives and gaming opportunities.

2.3 CBL from Customers' Point of View

CBL calculation and the consequent payment settlement have very significant effects on the customer's decision. To start with, the fairness of a program, to some extent, hinges on the accuracy of the CBL method employed by that program. Therefore, if a CBL calculation methodology does not produce an accurate CBL, it can damage the efficiency of the DR programs employing such methodology. Although fairness by itself does not guarantee any positive response, lack of fairness seriously harms the customer's response.

Another fact that can affect the fairness of a program is the potential gaming opportunities of a CBL method. Although this fact might excite a minority of customers, for some behavioral reasons, it is very discouraging for the rest of customers. These potential gaming opportunities create an incentive for some customers to speculate about the possible event days and inflate their CBL in anticipation of higher gain in those event days.

As a customer, load reduction is a means to gain financial rewards, either in terms of rebate or lower payment. Therefore, customers see load reduction in light of its financial impact. For this reason, any CBL shortcoming which affects the perceived financial reward of customers can affect the performance of the DR program employing that CBL. Authors in [28] explain how the financial offering of DR programs is an ultimate determinant of customers' decisions. In this thesis, the main focus is on the impact of accuracy and bias of CBLs on the financial performance of DR programs.

2.4 CBL from Utilities' Point of View

Utilities are interested in CBL for different reasons. CBL is a tool for some DR programs in order to calculate their payments to customers. However, their main interest is the load reduction. Due to the obligation of utilities to serve, they must make sure that they are capable of serving customers in any situation. DR programs can help them out in emergency situations. One of these situations is peak time of some special days that electricity in the wholesale market is either very expensive or unavailable. DR programs can induce customers to lower their peak consumption.

DR also plays a role in delaying investment for new infrastructures. In certain geographies, peak demand has grown significantly while overall energy consumption has not grown proportionally. This growing peak demand prompts the utilities to take an action and invest in new infrastructure which can drive the electricity rate higher. DR can provide an alternative solution to maintain reliability without investing in unnecessary infrastructure. This solution can keep rates low.

In a competitive market, even a single event of violation of obligation to serve can have irreversible negative consequences. For that reason, utilities are more interested in the load reduction aspect of CBL and less sensitive to the financial aspect of it. Another reason for why utilities are less sensitive to the financial aspect of CBL is that they reflect their cost-of-service into their retail rates. Therefore, ultimately the customers are the ones who feel most of the possible financial losses.

However, utilities are aware that CBL accuracy and fair payment settlement can affect the customer's load reduction. For this reason, utilities try to design CBL calculation method in a way to hinder any discouraging effect.

CHAPTER 3: CUSTOMER BASELINE MODELS

In this section, different CBL methods and the details of their implementation such as their associated terminologies and standard mathematical presentation of the models are presented. In what follows, first, some important terminologies used in association with CBL are explained. Then the popular and established CBL methods are described mathematically and the baseline adjustment is explained. The other less popular CBL methods are reviewed and explained afterwards. Moreover, CBL performance metrics such as accuracy and bias are explained in detail. Eventually this section concludes with a thorough discussion of the industries that adopted the aforementioned CBL methods.

3.1 Glossary

a) Weather sensitivity

Weather sensitivity measures to what degree loads are sensitive to the local weather. Temperature and humidity are two variables which could be employed for this purpose. However, in most climates, temperature is the only utilized variable and humidity is regarded as ineffective.

Practically, weather dependence is often represented by using linear regression models. These models try to explain hourly load by utilizing explanatory variables such as hourly temperature. These models, in their complicated form, include lagged variables or more complex functions of temperature.

b) Admissible days

Event days normally happen on weekdays. Therefore, it makes sense to use exclusively normal working days as an input for calculating baseline. Admissible days refers to the days that are used for baseline calculation process. The standard procedure for selecting the admissible days is eliminating weekends, holidays and past curtailment events. However, some ISOs employ further exclusion systems to improve their baseline calculation. For example, PJM utilizes threshold of 25% to exclude the days on which consumption is below the threshold. Moreover, [11] recommended that scheduling information related to shutdowns and large swings in energy be employed by ISOs to improve the process of exclusion. However, the inclusion of threshold and scheduling information in the process of determining the admissible days could increase the complexity of baseline calculation.

c) Proxy event days

Using proxy event days is a valuable means to examine the different baseline calculation methods. The advantage of proxy event days over actual event days in determining the baseline is the availability of the actual loads. Thus, the baseline could be compared with the actual load and the accuracy and bias of the method could be measured. The proxy event days are a subset of the admissible days and are to be selected so they are as similar as possible to the actual event days.

Typically DR events are called on the hottest days. Therefore, the temperature plays a critical role in selecting the proxy event days.

3.2 CBL methods

Several methods are introduced in the literature to calculate the CBL. In this section, these methods are explained mathematically. For the purpose of brevity and clarity, the terminology and nomenclature of [29] are used. Before starting to introduce the methods, some terms should be defined to facilitate the future mathematical presentation. They are as follows:

C	A set of customers
$T = \left\{t_0, \ldots, t_{\lvert T \rvert}\right\}$	Timeslot division within a day
$l_i(d,t)$	Actual load of customer $i \in C$ on day *d* at timeslot $t \in T$
$b_i(d,t)$	Predicted baseline of customer $i \in C$ on day *d* at timeslot $t \in T$
$b_i^*(d,t)$	True baseline of customer $i \in C$ on day *d* at timeslot $t \in T$

DR days are days when DR events are announced. In the absence of a DR event, the other days are called non-DR days hereafter. Two day types are used in this work: weekdays (Monday to Friday), and weekends (Saturday and Sunday).

$D(Y,d)$	A set of Y non-DR days most recently preceding the day *d* having the same day type as *d*.
$l_i(d) = \sum_{t \in T} l_i(d,t)$	Total load of customer $i \in C$ on day *d*

3.2.1 HighXofY Method

This method employs Y non-DR days before the DR event. In order to calculate the baseline, X highest consumption days will be selected from the aforementioned Y days. The baseline is an average load of these X days. If HighXofY is defined as

$High(X,Y,d) \subseteq D(Y,d)$, then the HighXofY baseline of customer $i \in C$ for timeslot $t \in T$ on day d is as follows:

$$b_i(d,t) = \frac{1}{X} \sum_{d \in High(X,Y,d)} l_i(d,t) \tag{1}$$

New York ISO uses High5of10 method and this method is employed in this thesis. The algorithm of NYISO method in its Day-Ahead Demand Response Program for approved demand response providers will be explained shortly [13]. It attempts to select 5 days with the highest consumption level (i.e. average daily kWh usage) out of a pool of 10 days chosen out of days prior to event day that meet certain criteria. These criteria will be defined in Steps 1-4.

Step 1: Demand Side Resource (DSR) signs a contract with customers to participate in the Day-Ahead Demand Response Program.

Step 2: Create a “baseline calculation window.” In other words, create a storage file for 10 selected days. Start from two days prior to the event day and calculate an “average daily event period kWh usage” value for that day. This value is the sum of the 24 hourly demands. This day must not be a weekend, holiday, event day or curtailment day. If so, go backward and pick another day.

Step 3: Go backward and pick the next previous day before the day selected in Step 2. If it is a weekend, holiday, event day or curtailment day, then go backward and pick another day. Calculate an “average daily event period kWh usage” value for that day.

Step 4: Compare the values of "average daily event period kWh usage" in Step 2 and Step 3. If the value of Step 3 is greater than 25% of the value of Step 2, then add it to the "baseline calculation window." Otherwise, discard it.

Step 5: Repeat Step 3 and 4 until 10 days are stored in the "baseline calculation window."

3.2.2 LowXofY Method

This method is the opposite of HighXofY. This method is included because [29] found that this method shows a better bias performance. In addition, excluding the highest use day, it actually might lead to a better outcome for the utility, since highest use could be an anomalous day or could be due to gaming. In order to calculate the baseline, X lowest consumption days will be selected out of Y days. The baseline is an average load of these X days. If LowXofY is defined as $Low(X,Y,d) \subseteq D(Y,d)$, then the LowXofY baseline of customer $i \in C$ for timeslot $t \in T$ on day d is as follows:

$$b_i(d,t) = \frac{1}{X} \sum_{d \in Low(X,Y,d)} l_i(d,t) \tag{2}$$

3.2.3 MidXofY Method

In this method, some of the lowest and highest consumption days will be dropped and the retaining X middle consumption days will be used to calculate the baseline. Let $X, Y \in \mathbb{N}$, $X \leq Y$, and $(Y - X) \bmod 2 = 0$. Moreover, let $Z = (Y - X)/2$. By dropping Z-lowest and Z-highest consumption days, the rest will be X days used in this baseline. If MidXofY is defined as $Mid(X,Y,d) \subseteq D(Y,d)$, then the MidXofY baseline of customer $i \in C$ for timeslot $t \in T$ on day d is as follows.

$$b_i(d,t) = \frac{1}{X} \sum_{d \in Mid(X,Y,d)} l_i(d,t) \tag{3}$$

3.2.4 Exponential Moving Average Method

This method is a weighted average of a customer's historical data from the beginning of her subscription. The weight of each day decreases exponentially with time.

Let $D(\infty, d) = \{d_1, ..., d_k\}$. Moreover, let $1 \le \tau \le k$ be a constant. This constant is the number of days that are used to determine the initial average. $s_i(d_\tau, t)$ is the initial average load of customer $i \in C$ for timeslot $t \in T$.

$$s_i(d_\tau, t) = \frac{1}{\tau} \sum_{j=1}^{\tau} l_i(d_j, t) \tag{4}$$

The exponential moving average for $\tau \le j \le k$ is as follows:

$$s_i(d_j, t) = \left(\lambda . s_i(d_{j-1}, t)\right) + \left((1-\lambda) . l_i(d_j, t)\right) \tag{5}$$

where $\lambda \in [0,1]$. The baseline for customer $i \in C$ on day d for timeslot $t \in T$ is as follows:

$$b_i(d,t) = s_i(d_k, t) \tag{6}$$

In this method, the baseline for days earlier than $d_{\tau+1}$ is undefined.

New England ISO (ISONE) employs this method. ISONE also is employed in this thesis as one of CBL methods. This baseline creation methodology consists of two calculations based on when the customer joined the program [13]. If a customer is a newly joined participant of the program (i.e. no previous record of consumption history), the calculation of the baseline starts with the hourly average of the electricity consumption of the first five business days, Monday through Friday, excluding holidays and other event days. The outcome value of this step is called "Customer Baseline 6." The "6" refers to the day following the first five business days or the sixth day. This step can be formulized as follows:

$$CB_6 = (\text{Sum Meter kW value for the hour})/5 \qquad (7)$$

Once CB_6 is calculated for the customer, the customer can be considered as the current customer and different rules will be applied to calculate her baseline. When the customer has CB_6, the baseline is calculated using (5) with $\lambda = 0.9$.

Every day excluding weekends, holidays and event days, a new baseline is calculated for the customer. The referred calculation equation put a weighting factor of 90% for the previous day and a weighting factor of 10% for the current day. The rationale behind this setup is that by putting more weight on the previous day, the opportunity for customers to "game" the program will be reduced. The topic of "gaming" the baseline is not the concern of this thesis, but will be discussed briefly in the future sections of this thesis.

3.2.5 Regression Method

This method uses linear regression as a basis to calculate the baseline. The baseline of customer $i \in C$ on day d for timeslot $t \in T$ is as follows:

$$b_i(d,t) = \left(\theta_{i,t}\right)^T x_{i,t} + \varepsilon_{i,t} \qquad (8)$$

Where $x_{i,t}$ is the feature vector, $\theta_{i,t}$ is the vector of regression coefficients and $\varepsilon_{i,t}$ is the error term. The feature vector consists of explanatory variables like historical load, temperature, humidity or sunrise/sunset time. This extra information as well as some information about each household habits and their economic standing can improve the results significantly. But in practice, getting access to such information is a cumbersome task. Table 1 provides the key summary of methods introduced so far.

TABLE 1: Key summary of CBL methods

Methods	Type	Data	Proxy day selection Criteria	ISOs using methods	Admissible days
HighXofY	Averaging	Consumption	Days with highest consumption	PJM, CAISO, NYISO	Non-event Working days
MidXofY	Averaging	Consumption	Days excluding highest/lowest consumption	-	Non-event Working days
LowXofY	Averaging	Consumption	Days with lowest consumption	-	Non-event Working days
Exponenti al Moving Average	Rolling averaging	Consumption	All admissible days	ISONE	Non-event Working days
Regressio n	Regression	Consumption/ Calendar/Tem perature/etc	All admissible days	ERCOT	All days (excluding event and holidays)

3.3 Baseline Adjustment

As discussed earlier, the subset of X days are selected to consist of days similar to the event day. Nevertheless, the conditions on the event day are often different from the selected prior days. For this reason, X of Y baseline methods could be adjusted by the event day data. According to North American Energy Standards Board (NAESB) [30], an adjustment to a HighXofY baseline is necessary to more accurately reflect load conditions of the event day. The adjustment is defined by the time frame of adjustment, multiplicative or additive, capped or uncapped and symmetric or asymmetric. In what follows the aforementioned choice of adjustments will be elaborated.

Time frame of adjustment is normally 2-4 hours before the start of the event. The time frame must have two properties. It must be at least one hour earlier for the event not to overlap with people who start the load reduction sooner. Also, it should not be too far away from the event. The inappropriate choice of time frame could penalize customers for early curtailment and inadvertently reward some others for temporary increase of their loads.

In order to adjust the baseline, the difference between the actual load and the estimated baseline in the adjustment time frame could be employed in two ways. Multiplicative adjustment uses the percentage change and applies it to the estimated baseline. For example, if the actual load is 30% higher on average than the estimated baseline during the time frame, the estimated baseline will be adjusted to 130% for the whole duration of the event. Additive adjustment, on the other hand, uses the absolute change. For example, if the actual load is 30kW higher on average than the estimated baseline during the time frame, 30kW will be added to the estimated baseline for the whole duration of the event. Figure 2 illustrates the concept of baseline adjustment.

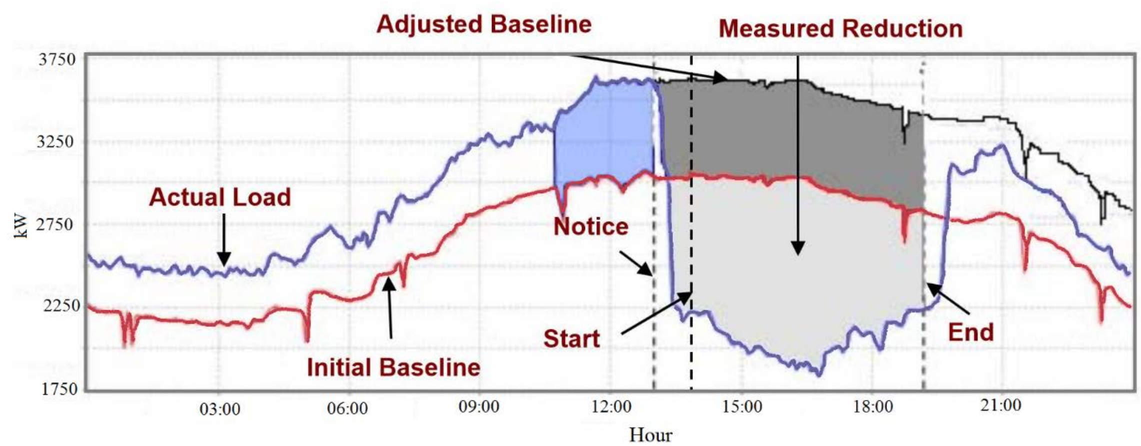

FIGURE 2: Example baseline adjustment [14]

Some programs limit the magnitude of the adjustment by utilizing a cap. For example, if a program uses a 20% cap, in case of 30kW absolute change between the actual load and the estimated baseline, just 6kW (i.e. 20% of 30kW) is allowed to be used for the adjustment purpose.

Some programs limit the adjustment to upward adjustment (i.e. asymmetrical). This practice overestimates the load reduction in cases where the actual loads are lower than the estimated baseline in the adjustment time frame. However, it eschews some unintended consequences of symmetric adjustment such as when a customer decides to start the load reduction very soon in anticipation of the upcoming event [14].

3.4 Other Baseline Methods

There are a few other proposed CBL methods that are less popular. They are reviewed in this subsection briefly. “Comparable day” is a CBL method that allows an aggregator to find a day similar to the event day and use its load profile as a baseline. This method unlike average methods uses only data from one day. However, finding a suitable objective criteria to select the target day is very difficult task [14].

Maximum Base Load (MBL) is a method that uses an entirely different approach compared to all aforementioned methods so far. It selects some of the peak hours from the previous year and determines the maximum energy usage expected of each customer. Then it specifies a level to which the customer should drop her consumption during the event. This level is a maximum level minus the committed capacity of the customer. The baseline shape is static in this method and it relies upon previous year’s historical data. Two well-known examples of employing this method are special case resources (SCR) programs in NYISO and emergency load response program (ELRP) in PJM [14]. Although MBL baseline offers simplicity, this benefit is outweighed by its poor accuracy. [14] has performed a comparative study and according to the results, HighXofY methods outperform MBL methods significantly.

Another method is presented by [31] to produce a CBL to be used by the first DR program in Colombia. The method is an adaptation of the forecasting with decomposition approach. It utilizes a multiplicative decomposition to represent the daily consumption of the users. The method is computationally burdensome, but it is argued that it best fits the unique nature of the Colombian electricity industry.

Moreover, [32] proposes an engineering algorithm. Under this approach, the overall client facility should be modeled. It includes all energy consuming elements including variable loads as well as interruptible load equipment. Also, a comprehensive thermodynamic model for building is necessary to accurately determine electrical loads. The measurement of loads can then be employed by the customer during a DR event to determine how much load is being reduced or eliminated by shutting down certain equipments. This approach is mostly beneficial to the customers to reduce or power off the best combination of loads during a DR event. Another engineering algorithm is proposed in [33], which builds its methodology on a norm behavior convention. Norm behavior convention, in this paper, refers to the normal consumption of the flexible devices. This norm behavior can be derived from the status and configuration data of these flexible devices. The flexible device in this paper is defined as either postponable smart devices (e.g. electrical vehicles, smart washing machines, etc.) or (thermally) buffered devices (e.g. air conditioner). In this paper, models for postponable and buffered devices are introduced. Then a cluster of non-controllable and controllable devices is defined. For each of the devices in the cluster a baseline is defined according to its model and then all the baselines are aggregated to compose the CBL. This methodology requires a complete knowledge of the loads deployed by the customers.

3.5 Bias and Accuracy Metrics

Different proposed methods can be compared by their bias and accuracy. For this purpose, two metrics for bias and accuracy are defined in this research. The hourly accuracy and bias of each baseline is defined as follows:

Let *C* be the set of all the customers, *D* be the set of all days in the data set, and *T* be the set of hourly timeslots in a day. Mean Absolute Error (MAE) is utilized for measuring baseline accuracy as shown in (9).

$$\alpha = \frac{\sum_{i \in C} \sum_{d \in D} \sum_{t \in T} \left| b_i(d,t) - l_i(d,t) \right|}{|C|.|D|.|T|} \tag{9}$$

The lower the MAE, the higher the accuracy. Baseline bias is defined as shown in (10).

$$\beta = \frac{\sum_{i \in C} \sum_{d \in D} \sum_{t \in T} \left(b_i(d,t) - l_i(d,t) \right)}{|C|.|D|.|T|} \tag{10}$$

According to (10), baseline methods with positive bias overestimate the customers' actual consumption and vice versa.

3.6. Industry Application of CBL Methods

In this subsection, the methods employed by different ISOs in North America are reviewed and discussed. Different ISOs use different methods primarily based on the nature of their offering programs. For example, within a summer emergency DR program, an event is expected to be driven by extreme weather conditions. Therefore, HighXofY methods suit better to these programs while for some other programs, operators may choose to use midXofY or LowXofY baseline methods.

PJM Economic Load Response Program (ELRP) employs High4of5 for a weekday and High2of3 for a weekend DR event. Beginning in 2012, ELRP upgraded its

CBL calculation. It now includes a symmetric additive adjustment with a three hour adjustment window [34]. In NYISO customers can select between two CBL formulas of average day CBL and weather adjusted CBL. For average day CBL, NYISO utilizes High5of10 for a weekday and High2of3 for a weekend. Moreover, for weather adjusted CBL, the CBL would be adjusted upward or downward based on the actual usage in the two hours prior to the event notification [35-36]. CAISO uses High10of10 for a weekday and High4of4 for a weekend [37]. Ontario, Canada uses High15of20 method [29]. LowXofY and MidXofY have not been employed by industry yet, but as [29] shows they have their own merits.

ISONE uses exponential moving average to calculate the baseline. It uses $\tau = 5$ and $\lambda = 0.9$ for this purpose. As it is discussed before, the baseline is undefined for a customer who joined the program for less than five days.

Regression models are employed by ERCOT and are also developed and introduced in [38].

CHAPTER 4: ACCURACY AND BIAS

In this section, the data employed for the analysis of “classic” metrics of accuracy and bias are introduced and the metrics are defined. Afterwards, these metrics are applied to CBL calculations and the results are presented.

4.1 Setup

In this paper, the Irish CER smart metering trial dataset [39] has been employed. This dataset contains measurements of around 5000 customers over one and a half years (most of the smart meters used in this trial are selected randomly throughout Ireland, more information about the meters used in the field trial is available in [40]) , which is available to the public. The customers consist of a residential sector and small and medium-sized enterprises and the data interval is 30 minutes. The measurements started in July 2009 and ended in December 2010. The purpose of the trials was to assess the impact on consumer’s electricity consumption. The intention of the study is performing the cost-benefit analysis for a national rollout in Ireland. Customers who participated in the trials had an electricity smart meter installed in their homes/premises. They have agreed to take part in research to help to understand how smart metering can be helpful to shape energy usage behaviors across a variety of demographics, lifestyles and home sizes. The trial has two phases of pre-trial and trial. In the dataset, the data from the start of the experiment till Dec. 31st of 2009 are pre-trial and benchmark data. The benchmark data are customers’ consumption in the traditional fixed rate tariff environment. These data were supposed to be utilized later to study the impact of multiple DR programs utilized in this pilot project (e.g. TOU, PTR, …) . In this paper, the benchmark data of 262 customers from Oct. 24th to Dec. 31st (total of 69 days) in 2009 are used for the analysis. The reasons 262 customers were selected are twofold; first, the customers' original data has multiple files and are not sorted well, therefore, it took a lot of time and programming to clean the data, and second, this author must make sure that he has access to all the consumption data (sometimes, probably because of communication failures, the record of some consumption is missing). Figure 3 shows the total consumption of all customers in this dataset. This information is utilized to select the closest day to an event day later.

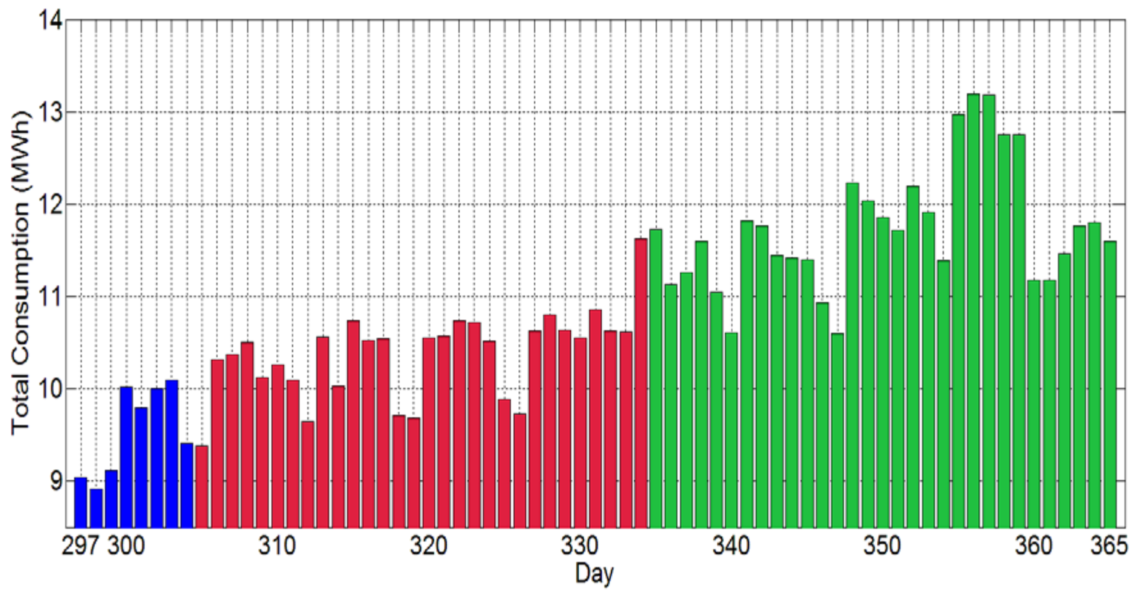

FIGURE 3: Total consumption of the data set (Color-coded for different months, blue for October, red for November, green for December)

4.2 Analysis

In this part, CBL for all customers is calculated. In what follows, the results of calculating CBL for all five methods, mentioned in the previous chapter, are presented. However, as it is shown later, because adjustment did not improve the results, the illustrative results of adjustment of each CBL are not included in this thesis. At the end of this section, all these CBL methods and their adjustments are compared and a discussion about the results of accuracy and bias metrics of the CBLs are provided.

4.2.1 High5of10 (NYISO) Method

Figure 4 shows the results of NYISO CBL and actual loads in an aggregated manner. This picture is helpful to provide a sense about the overall accuracy and bias. Moreover, this aggregated picture shows why many utilities think CBL is an appropriate tool for load reduction, because in the aggregated manner, many random parameters and their effects will even out. In other words, it is because each individual's daily consumption dictates many random variables, while on the aggregate level, all this randomness will even out and the overall daily consumption shows a very predictable behavior.

As it is shown, this method has a positive bias during event hours. But for the whole day, the positive bias during an event will even out with the negative bias during non-event hours, which is not positive. This is because if one uses an entire day index which some reports do, it masks many details about the bias. It is worth mentioning that all the claims about the direction of bias from now on are based on the observation. This author has not done any rigorous calculation to prove the claims.

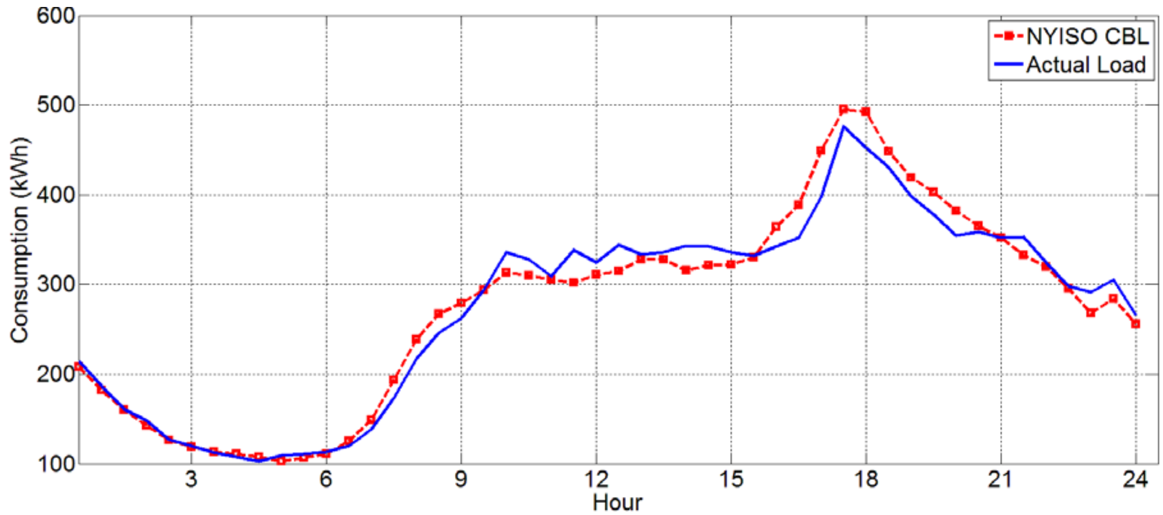

FIGURE 4: Actual data vs. NYISO CBL for all the customers

4.2.2 LowXofY Method

Figure 5 shows the results of Low4of5 CBL calculation method and actual loads in an aggregated manner. As it is shown, this method has a negative bias during event hours and non-event hours.

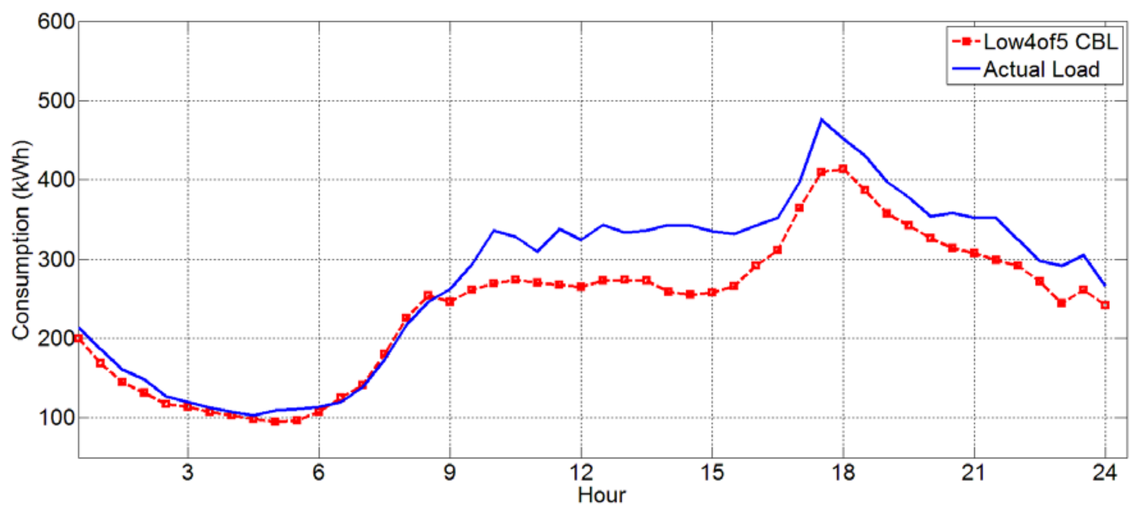

FIGURE 5: Actual data vs. Low4of5 CBL for all the customers

4.2.3 MidXofY Method

Figure 6 shows the results of Mid4of6 CBL calculation method and actual loads in an aggregated manner. As it is shown, this method has a negative bias during event hours and non-event hours.

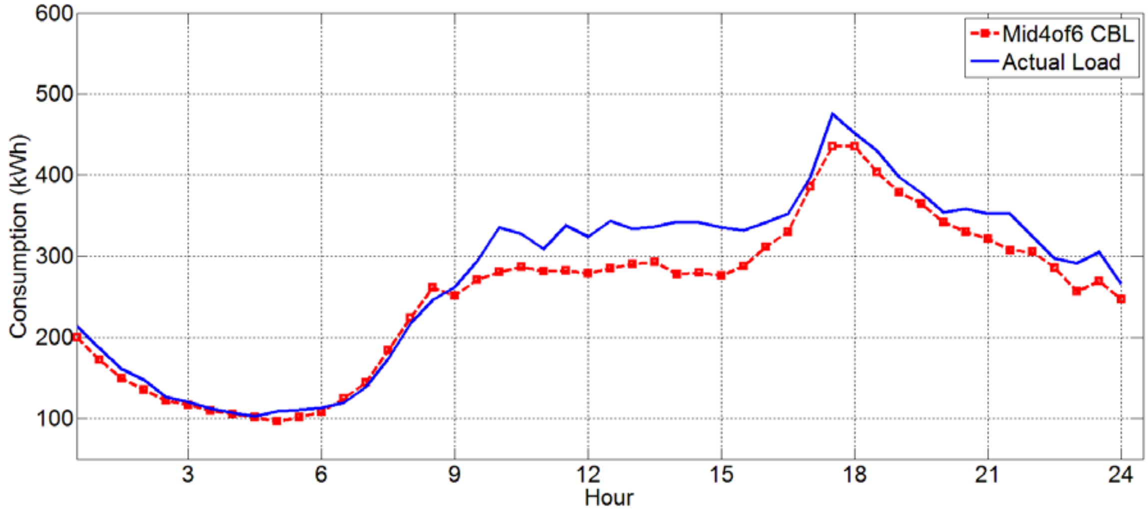

FIGURE 6: Actual data vs. Mid4of6 CBL for all the customers

4.2.4 Exponential Moving Average (ISONE) Method

Figure 7 shows the results of ISONE CBL calculation method and actual loads in an aggregated manner. As it is shown, this method has a negative bias during event hours and non-event hours.

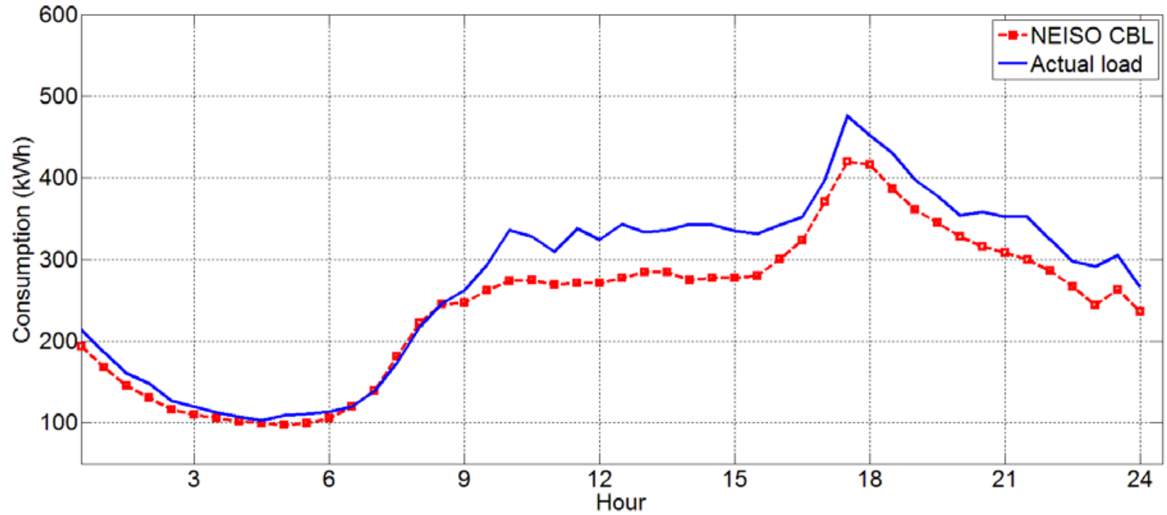

FIGURE 7: Actual data vs. ISONE CBL for all the customers

4.2.5 Regression Method

Figure 8 shows the results of Regression CBL calculation method and actual loads in an aggregated manner. As it is shown, this method has a negative bias during event hours and non-event hours. In the case of regression, care must be taken not to jump to conclusions. Unlike other methods, this method has this potential to be enhanced significantly by adding more explanatory parameters in order to capture the effect of some behavioral patterns of residential customers.

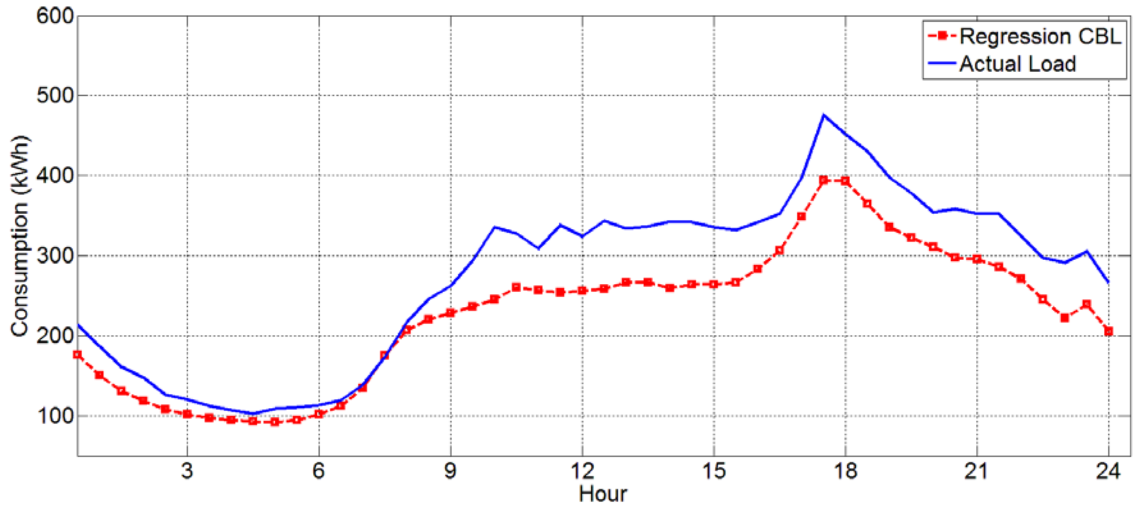

FIGURE 8: Actual data vs. Regression CBL for all the customers

4.2.6 Comparative Study

In this section, the accuracy MAE and bias value of all CBLs employed in this thesis are presented. Table 2 presents an accuracy and bias for event hours and Table 3 shows the same accuracy and bias for the entire event day. Since the payment settlement is based on the load reduction during event hours, the metrics for the event hours are more indicative of the power of CBL than the same metrics for the entire event day.

TABLE 2: Accuracy "MAE" and bias for event hours

	Accuracy MAE (kWh/hr)	Bias (kWh/hr)
NYISO	1.440	+0.1266
Adjusted NYISO	1.6855	+0.7232
Mid4of6	1.2823	-0.2267
Adjusted Mid4of6	1.7816	+0.9333
Low4of5	1.2789	-0.3793
Adjusted Low4of5	2.0470	+1.2957
ISONE	1.1654	-0.3213
Adjusted ISONE	1.4531	+0.4816
Regression	1.4057	-0.4642
Adjusted Regression	1.4534	-0.1466

TABLE 3: Accuracy "MAE" and bias for entire event day

	Accuracy MAE (kWh/hr)	Bias (kWh/hr)
NYISO	0.944	+0.0092
Adjusted NYISO	1.2403	+0.6058
Mid4of6	0.8668	-0.1750
Adjusted Mid4of6	1.5121	0.9850
Low4of5	0.8645	-0.2601
Adjusted Low4of5	1.8651	+1.4149
ISONE	0.7832	-0.2411
Adjusted ISONE	1.1772	+0.5617
Regression	0.9686	-0.3626
Adjusted Regression	1.509	-0.0450

Figures 9 and 10 illustrate the results of Tables 2 and 3. The horizontal axis in these figures is accuracy MAE (kWh/hour) and the vertical axis is bias (kWh/hour). As it is shown, ISONE has the best accuracy among the employed CBL methods. Figure 11 compares the CBL results of all the methods with the actual load data.

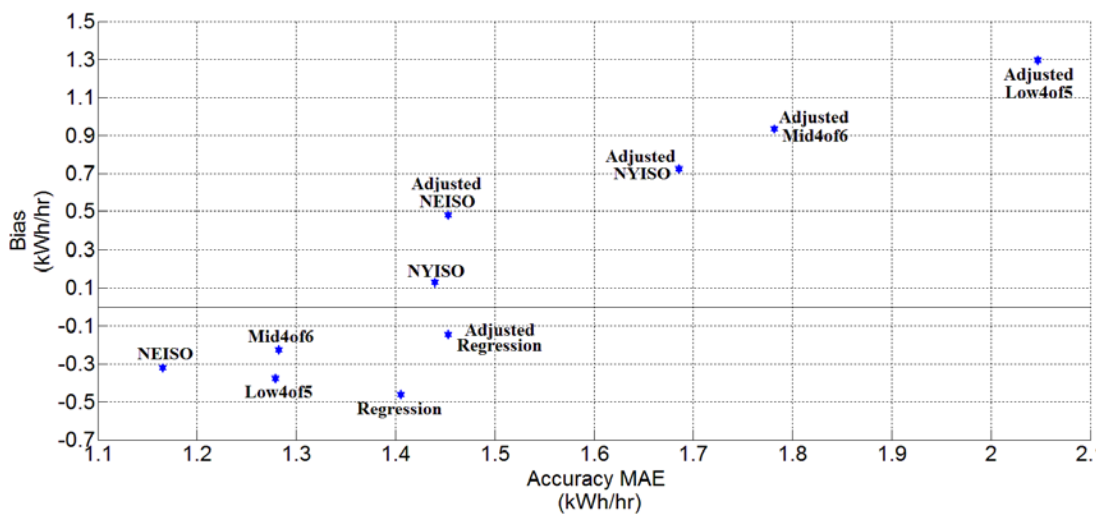

FIGURE 9: Accuracy "MAE" and bias for event hours

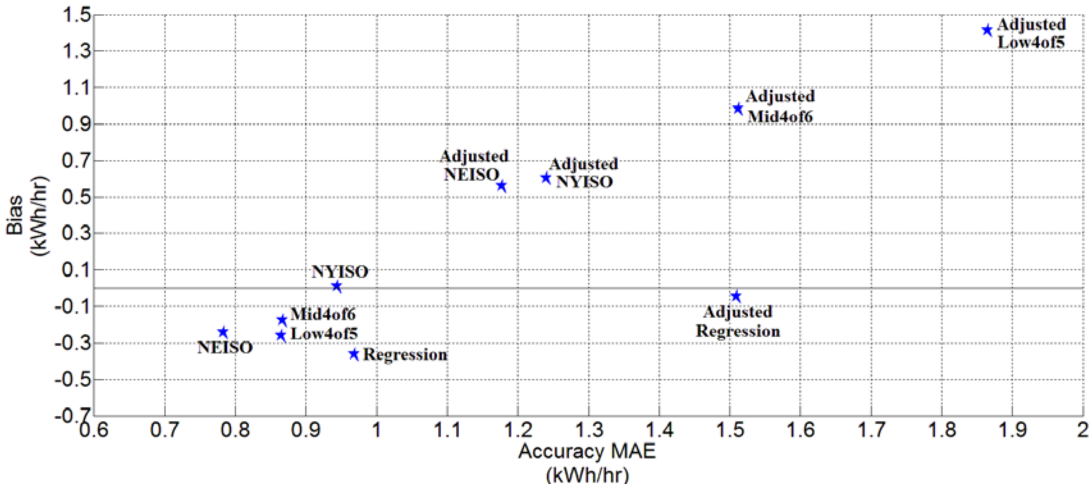

FIGURE 10: Accuracy “MAE” and bias for entire event day

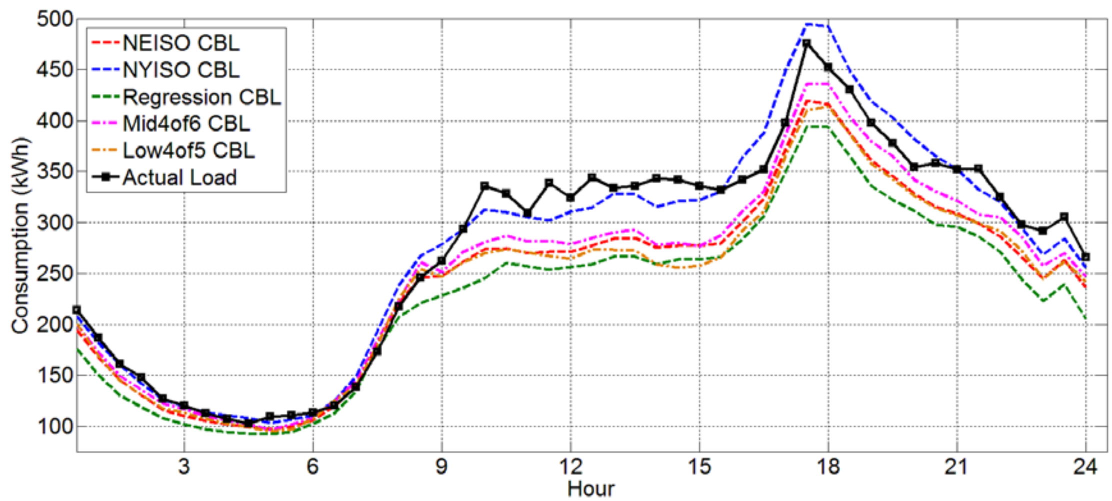

FIGURE 11: Actual data vs. different CBL results for all the customers

CHAPTER 5: CASE STUDY

For economic analysis of PTR program, the accuracy of CBLs employed for payment settlement must be taken into consideration. In this section, a case of PTR program is introduced and its economic performance is analyzed.

PTR is one of the popular DR programs in electricity industry. This program is frequently employed by utilities for their industrial customers. The performance of this program strongly depends on the performance of CBL. Although it is capable of capturing baseline for industrial customers quite accurately, its performance for residential customers is untested.

Moreover, this program is extremely appealing from the policy point of view as it requires a minimal revision to status quo and could provide a huge positive impact if it works correctly. However, it is vulnerable to many implementation deficiencies. Chao in [24] reviews some of these practical issues including opportunities for gaming and problems with CBL methods. Moreover, authors in [28] studied behavioral aspects of customers' involvement in PTR. It is shown that the reward mechanism which PTR employs to incent the customers for load reduction is another source of inefficiency in this DR program. In this thesis, PTR is studied from the different angle.

In PTR program, one challenge is to select a reasonable rebate rate. In order to have a practical rebate and fixed rate values, in this thesis, the relevant values of Anaheim Public Utility (APU) pilot project in residential sector [41] are employed. APU project paid \$0.35/kWh incentive and an average base price of \$0.097/kWh as fixed tariff. It is worth emphasizing that these two values, incentive payment and fixed tariff, are selected from APU project but they are applied to Ireland data.

In this case study, a day with maximum consumption from Figure 3 (i.e. Dec. 22nd) is chosen as an event day for PTR. The event starts from 3:00 p.m. and ends at 9:00 p.m. This time interval is selected for the event, because almost all the peak consumptions statistically are observed to be happening during this interval.

CHAPTER 6: RESULTS AND DISCUSSION

During the event, it is expected that the customers respond to the incentive and decrease their consumption. If one assumes that CBL is able to predict the customers' baseline 100% accurate, then, all the difference between CBL and actual load is because of PTR incentive effect. However, if CBL is inaccurate, the difference between CBL and actual load would have two components; one component is in response to PTR program and the other is because of CBL inaccuracy. In this paper, the focus is on the latter. Since there is no event in the real data, the first component is zero and the difference between CBL and actual data is all because of the second component (i.e. CBL inaccuracy).

The revenue of this hypothetical utility out of these 262 customers on event day is $1279.80 for selling 13,194kWh. Table 4 lists the false load reductions under different CBL methods. It also shows how much rebate this utility must pay to these customers on event day. Moreover, Figure 12 illustrates the rebate as a percent of utility revenue (%) for different CBL methods. As discussed earlier, all the rebate money is incurred because of CBL inaccuracy. According to the results, in this PTR program for residential customers, the inaccuracy of CBLs costs this hypothetical utility at least half of its revenue for the event day.

As discussed earlier, according to the findings of multitudes of successful PTR programs offered to industrial customers, adjustment improves the results of CBL methods significantly, but in this case, adjustment deteriorates the outcome of these methods.

This author believes that the acceptable performance of CBL in industrial sector stems from the predictability of such loads, whereas for the residential customers, the presence of many non-correlated activities makes the loads highly unpredictable. In the absence of such predictability, CBL calculation methods, which worked successfully in the industrial case, perform very poorly as shown in this section.

Due to the obligation of utilities to serve, they must make sure that they have enough electricity to serve in any situation. DR programs can help them out in emergency situations. One of these situations is peak time of some special days that electricity in the wholesale market is either very expensive or unavailable. DR programs can help to relieve part of this pressure. Utilities, in response to such pressures, might accept any available program that induces customers to lower their peak consumption regardless of its damage to their revenue.

However, the utilities reflect their cost-of-service into their retail rates. Therefore, ultimately the customers are the ones who feel most of the aforementioned financial loss. Moreover, this loss of revenue redistributes among the customers randomly. In other words, this program rewards and punishes the wrong customers. This random redistribution of the loss casts a shadow on the fairness of the program.

TABLE 4: Load reduction and "PTR" payment settlement on event day

	False load reduction (kWh)	False load reduction as a percent of consumption on event day (%)	Rebate value ($)	Rebate as a percent of utility revenue (%)
NYISO	2,998	22.7	1049.3	81.9
Adjusted NYISO	5,804	43.9	2,031.3	158.7
Mid4of6	2,175	16.4	761.2	59.4
Adjusted Mid4of6	7,850.9	59.5	2,747.8	214.7
Low4of5	1,900	14.4	665.04	51.9
Adjusted Low4of5	10,312	78.1	3,609.3	282
ISONE	1,704	12.9	596.5	46.6
Adjusted ISONE	5,467	41.4	1,913.4	149.5
Regression	1,905	14.4	666.86	52.1
Adjusted Regression	3,162	23.9	1106.9	86.5

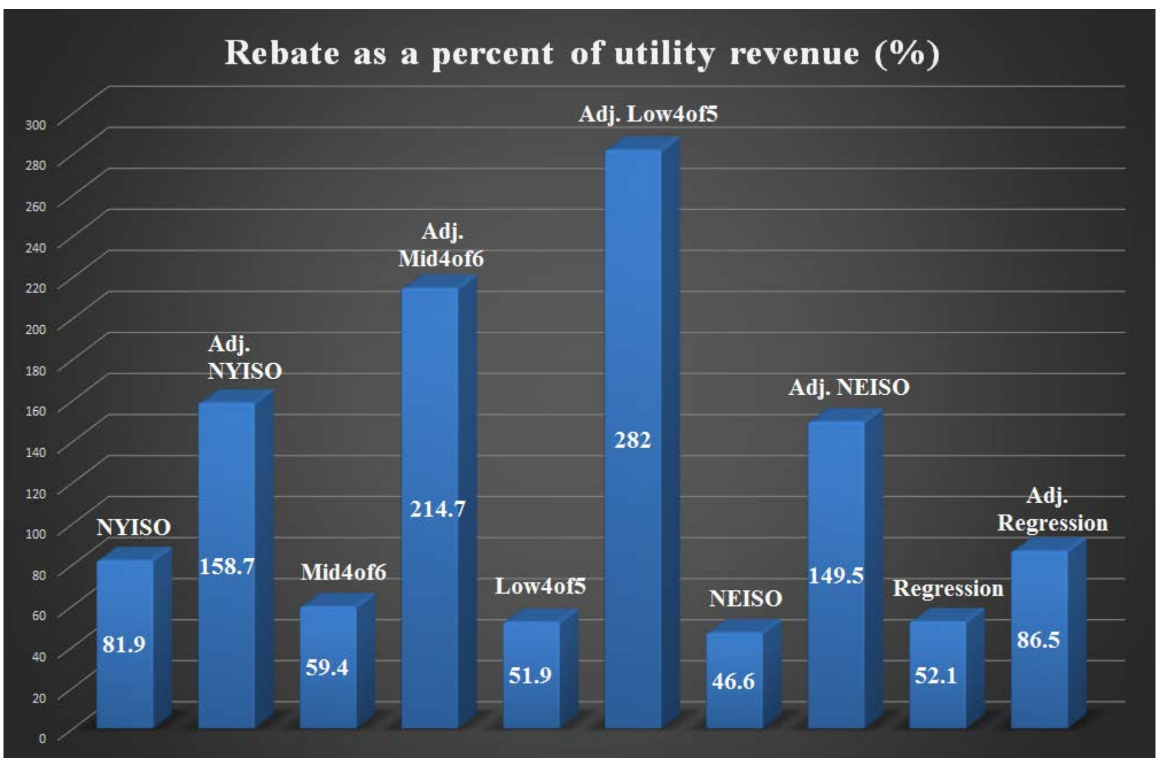

FIGURE 12: Rebate as a percent of utility revenue (%) for different CBL methods on event day

CHAPTER 7: CONCLUSION

In this thesis, the impact of accuracy of CBL on a PTR program offered to the residential customers was investigated. Previous works in this area merely focus on industrial and commercial customers. Residential customers as opposed to industrial customers show a high degree of unpredictability due to multitudes of non-correlated personal and household activities. For the purpose of analysis, High5of10 (NYISO), Low4of5, Mid4of6, exponential moving average (ISONE) and regression methods and their adjusted forms are selected to calculate CBL. The calculated baselines are utilized later to examine the economic performance of the PTR program. According to the results, in the case studied in this paper, for these 262 customers and just for an event day, the hypothetical utility of the case study, pays at least half of its revenue on event day as a rebate just because of the inaccuracy of these CBL calculation methods. Moreover, if these methods were adjusted based on their morning consumption, the results would worsen.

At the end, it is discussed that PTR can cause a significant loss to the customers and cause unfair redistribution of the utility’s revenue. Based on these results, it could be concluded that PTR programs are very inefficient for the residential customers. As discussed previously, these inefficiencies originate from the failure of CBL calculation methods to predict accurately the residential customers’ load profile on event day.

As a future study, this author plans to confirm these findings with data from smart meters of residential customers within the USA. At this point, the availability of such data is very limited. Moreover, with having broad time span in data, it is possible to select multiple event days. With using multiple event days, some characteristics of a single event day will balance and it is possible to claim that the results are independent from the characteristics of the event day. Furthermore, with having the proper data from a pilot project, the individual characteristics of each household including income, education, house size, weather and etc. can be included in the models.

Moreover, this author plans to expand the economic analysis of this thesis to study the effect of financial performance of CBL-dependent DR programs on economic social welfare.

REFERENCES

[1] A. Faruqui, S. George. "Quantifying customer response to dynamic pricing." *The Electricity Journal*, Vol. 18, no. 4, pp. 53-63, 2005.

[2] S. Mohajeryami , I.N. Moghaddam, M. Doostan, B. Vatani, P. Schwarz, "A Novel Economic Model for Demand Response to TOU Pricing," *Electric Power System Research* , 2016, DOI 10.1016/j.epsr.2016.03.026

[3] J. Wells, D. Haas. "*Electricity markets: consumers could benefit from demand programs, but challenges remain*". DIANE Publishing, 2004.

[4] S. Mohajeryami, M. Doostan, P. Schwarz, "The Impact of Customer Baseline (CBL) Calculation Methods on Peak Time Rebate Program Offered to Residential Customers", *Electric Power System Research,* 2016

[5] Working Group 2 Demand Response Program Evaluation – Program Year 2004 Final Report. Prepared for the Working Group 2 Measurement and Evaluation Committee, by Quantum Consulting Inc. and Summit Blue Consulting, LLC, 2004.

[6] M.L. Goldberg, G.K. Agnew, "Protocol Development for Demand-Response calculations: Findings and Recommendations", Prepared for the California Energy Commission by KEMA-Xenergy. 2003.

[7] Evaluation of 2005 Statewide Large Nonresidential Day-ahead and Reliability Demand Response Programs. Prepared for Southern California Edison and the Working Group 2 Measurement and Evaluation Committee, by Quantum Consulting Inc. and Summit Blue Consulting, LLC, 2006.

[8] A. Buege, M. Rufo, M. Ozog, D. Violette, S. McNicoll, "Prepare for Impact: Measuring Large C/I Customer Response to DR Programs," ACEEE Summer Study on Energy Efficiency in Buildings, Monterey, CA, August 2006.

[9] A. Heshmati, "Demand, Customer Base-Line and Demand Response in the Electricity Market: A Survey", *Journal of Economic Surveys*, Vol. 28, pp.862–888. 2014.

[10] Coughlin, Katie, Mary Ann Piette, Charles Goldman, and Sila Kiliccote. "Statistical analysis of baseline load models for non-residential buildings." Energy and Buildings 41, no. 4 (2009): 374-381.

[11] Coughlin, Katie, Mary Ann Piette, Charles Goldman, and Sila Kiliccote. "Estimating Demand Response Load Impacts: evaluation of baseline load models for non-residential buildings in California." Lawrence Berkeley National Laboratory (2008).

[12] J.R. Won, Y.B. Yoon, J.H. Choi, K.K. Yi, "Case study of customer baseline (CBL) load determination in Korea based on statistical method," in IEEE Transmission & Distribution Conference & Exposition: Asia and Pacific, pp.1-4, 26-30 Oct. 2009.

[13] C. Grimm, ”Evaluating baselines for demand response programs”, In 2008 AEIC Load Research Workshop. San Antonio, TX.

[14] “The Demand Response Baseline,” White Paper, EnerNOC, 2011.

[15] N. Addy, S. Kiliccote, J. Mathieu, and D. S. Callaway, “Understanding the effect of baseline modeling implementation choices on analysis of demand response performance,” ASME Paper No. IMECE2012-86973.

[16] N.J. Addy, S. Kiliccote, D. S. Callaway, and J.L. Mathieu. "How Baseline Model Implementation Choices Affect Demand Response Assessments." *Journal of Solar Energy Engineering*, vol.137, no. 2, 2015.

[17] N. Yamaguchi, J. Han, and G. Ghatikar, “Regression models for demand reduction based on cluster analysis of load profiles,” in *Proc. Sustain. Alternative Energy*, 2010, pp. 1–7

[18] Y.M. Wi, J.H. Kim, S.K. Joo, J.B. Park, J.C. Oh, “Customer baseline load (CBL) calculation using exponential smoothing model with weather adjustment,” in Proceedings of Asian and Pacific Transmission and Distribution Conference and Expo., pp. 1–4., 2009.

[19] S. Park, S. Ryu, Y. Choi, H. Kim, "A framework for baseline load estimation in demand response: Data mining approach," in IEEE International Conference on Smart Grid Communications (SmartGridComm), pp.638-643, 3-6 Nov. 2014.

[20] S. Park, S. Ryu, Y. Choi, H. Kim, “Data-Driven Baseline Estimation of Residential Buildings for Demand Response”. *Energies*, Vol. *8*, 10239-10259. 2015.

[21] P. Faria, Z. Vale, P. Antunes, and A. Souza, “Using baseline methods to identify non-technical losses in the context of smart grids,” in IEEE PES Conference on Innovative Smart Grid Technologies Latin America (ISGT LA), pp. 1–7, 2013.

[22] J.M. Ko, N.J. Jung, I.H. Yu, Y. Kim, "Day-ahead load reduction system based on customer baseline load." U.S. Patent Application 12/576,609, filed October 9, 2009.

[23] J.M. Ko, N.J. Jung, Y. Kim, I.H. Yu, "Load forecasting analysis system for calculating customer baseline load." U.S. Patent 8,406,935, issued March 26, 2013.

[24] H. Chao, "Demand response in wholesale electricity markets: the choice of customer baseline." *Journal of Regulatory Economics*, Vol.39, No.1, pp. 68-88, 2011.

[25] ISO New England. (2008). Docket No. ER08-538-000; Filing of Changes to Day-Ahead Load Response Program (February 5).

[26] Hung-po Chao, "Demand Management in Restructured Wholesale Electricity Markets", ISO New England, May 20, 2012, Available online: http://goo.gl/2RGFDU

[27] S. Mohajeryami, A. Asadinejad, M. Doostan, "An Investigation of the Relationship between Accuracy of Customer Baseline Calculation and Efficiency of Peak Time Rebate Program", in Power and Energy Conference at Illinois (PECI), Feb. 2016

[28] S. Mohajeryami, P. Schwarz, P. Teimourzadeh, "Including the Behavioral Aspects of Customers in Demand Response Model: Real Time Pricing Versus Peak Time Rebate", North American Power Symposium (NAPS), Oct 2015

[29] T.K. Wijaya, M. Vasirani, K. Aberer, "When Bias Matters: An Economic Assessment of Demand Response Baselines for Residential Customers," *IEEE Trans. on Smart Grid*, vol.5, no.4, pp.1755-1763, July 2014.

[30] North American Energy Standard Board. "Business Practices for Measurement and Verification of Wholesale Electricity Demand Response". March 16, 2009.

[31] E. Edgar, C. Angela, C.M. Elsa, D. Hernando, "Model to calculate the consumer baseline for a demand response models in the Colombian power market" In proc. 32nd IAEE International Conference, pp.1-10, San Francisco, EUA, June 2009.

[32] "Using an Engineering Model to Determine a Customer's Baseline Load," White Paper, Viridity Energy, Inc, 2010.

[33] K. Vanthournout, W. Foubert, C. Stuckens, B. Robben, G. Premereur, "A norm behavior based deterministic methodology for demand response base lines," in Power Systems Computation Conference (PSCC), pp.1-7, 18-22 Aug. 2014.

[34] Johnson Controls, "Understanding the PJM Default Baseline Calculation for Economic Load Response", 2012

[35] NYISO, "Emergency Demand Response Program Manual", Oct. 2013.

[36] conEdison, "Energy Efficiency and Demand Management Procedure-General: Calculating Customer Baseline Load", Feb. 2013.

[37] CAISO, "Customer Baseline LoadReview and Recommendation", May 2009.

[38] "PJM Empirical Analysis of Demand Response Baseline Methods," White Paper, KEMA, Inc., 2011.

[39] Smart Metering Trial Data Publication, Irish CER, 2012.

[40] “Electricity Smart Metering Technology Trials Findings Report”, Technical Report, Irish CER, 2011

[41] A. Faruqui, R. Hledik, S. Sergici. "Piloting the smart grid." *The Electricity Journal* 22.7, pp.55-69, 2009

APPENDIX A: DESCRIPTION OF DATA

The following are the description of the data.
Holidays and weekends in Ireland in 2009:

- Oct. 24 Saturday
- Oct. 25 Sunday
- Oct. 26 Bank Holiday
- Oct. 31 Saturday
- Nov. 1 Sunday
- Nov. 7 Saturday
- Nov. 8 Sunday
- Nov. 14 Saturday
- Nov. 15 Sunday
- Nov. 21 Saturday
- Nov. 22 Sunday
- Nov. 28 Saturday
- Nov. 29 Sunday
- Dec. 5 Saturday
- Dec. 6 Sunday
- Dec. 12 Saturday
- Dec. 13 Sunday
- Dec. 19 Saturday
- Dec. 20 Sunday
- Dec. 25 Christmas day
- Dec. 26 Saturday
- Dec. 27 Sunday
- Dec. 28 St. Stephen's day

TABLE 5: The description of the days in data

Day 1	(Sat) Oct. 24	Day 24	(Mon) Nov. 16	Day 47	(Wed) Dec. 9
Day 2	(Sun) Oct. 25	Day 25	(Tue) Nov. 17	Day 48	(Thur) Dec. 10
Day 3	(Mon) Oct. 26	Day 26	(Wed) Nov. 18	Day 49	(Fri) Dec. 11
Day 4	(Tue) Oct. 27	Day 27	(Thur) Nov. 19	Day 50	(Sat) Dec. 12
Day 5	(Wed) Oct. 28	Day 28	(Fri) Nov. 20	Day 51	(Sun) Dec. 13
Day 6	(Thur) Oct. 29	Day 29	(Sat) Nov. 21	Day 52	(Mon) Dec. 14
Day 7	(Fri) Oct. 30	Day 30	(Sun) Nov. 22	Day 53	(Tue) Dec. 15
Day 8	(Sat) Oct. 31	Day 31	(Mon) Nov. 23	Day 54	(Wed) Dec. 16
Day 9	(Sun) Nov. 1	Day 32	(Tue) Nov. 24	Day 55	(Thur) Dec. 17
Day 10	(Mon) Nov. 2	Day 33	(Wed) Nov. 25	Day 56	(Fri) Dec. 18
Day 11	(Tue) Nov. 3	Day 34	(Thur) Nov. 26	Day 57	(Sat) Dec. 19

Day 12	(Wed) Nov. 4	Day 35	(Fri) Nov. 27	Day 58	(Sun) Dec. 20
Day 13	(Thur) Nov. 5	Day 36	(Sat) Nov. 28	Day 59	(Mon) Dec. 21
Day 14	(Fri) Nov. 6	Day 37	(Sun) Nov. 29	Day 60	(Tue) Dec. 22
Day 15	(Sat) Nov. 7	Day 38	(Mon) Nov. 30	Day 61	(Wed) Dec. 23
Day 16	(Sun) Nov. 8	Day 39	(Tue) Dec. 1	Day 62	(Thur) Dec. 24
Day 17	(Mon) Nov. 9	Day 40	(Wed) Dec. 2	Day 63	(Fri) Dec. 25
Day 18	(Tue) Nov. 10	Day 41	(Thur) Dec. 3	Day 64	(Sat) Dec. 26
Day 19	(Wed) Nov. 11	Day 42	(Fri) Dec. 4	Day 65	(Sun) Dec. 27
Day 20	(Thur) Nov. 12	Day 43	(Sat) Dec. 5	Day 66	(Mon) Dec. 28
Day 21	(Fri) Nov. 13	Day 44	(Sun) Dec. 6	Day 67	(Tue) Dec. 29
Day 22	(Sat) Nov. 14	Day 45	(Mon) Dec. 7	Day 68	(Wed) Dec. 30
Day 23	(Sun) Nov. 15	Day 46	(Tue) Dec. 8	Day 69	(Thur) Dec. 31

APPENDIX B: MATLAB CODES

Part One: Code for each CBL methods

a) NYISO code

```
%%%%
for f=1:262
    for i=1:69     %%% creat an empty matrix
        r(i,f)=0;
    end
    %%% selecting 10 admissible day
    for i=45:49
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            r(i,f)=data(k,ff)+r(i,f);
        end
    end
    for i=52:56
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            r(i,f)=data(k,ff)+r(i,f);
        end
    end
%%%%%%%%%%%%%% sorting the consumption in an ascending
order %%%%%%%
    rsort=sort(r(:,f));
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%
%%%%%% finding the days with higher consumption level
%%%%%%%%%%%%%
    for i=1:10
        no(i)=0;
    end
    for j=1:10
        for i=1:69
            u=70-j;
            if rsort(u)==r(i,f)
                no(j)=i;
            end
        end
    end
```

```
%%%%% finding customer baseline (5 out of 10 days)
%%%%%%%%%%%%%%%%%%%%%
    for i=1:48
        cbl(i,f)=0;
        act(i,f)=0;
    end
    for e=1:5
        i=no(e);
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            cbl(j,f)=data(k,ff)+cbl(j,f);
        end
    end
    for j=1:48
        cbl(j,f)=cbl(j,f)/5;
    end
%%%%%%%%%%%%%%%% creating a matrix for actual consumption
%%%%%%%%%
    for i=60
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            act(j,f)=data(k,ff)+act(j,f);
        end
    end
end
```

b) Mid4of6 code

```
%%%%
for f=1:262
    for i=1:69  %%% creat an empty matrix
        r(i,f)=0;
    end
%%% selecting 6 admissible day
    for i=52:56
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            r(i,f)=data(k,ff)+r(i,f);
        end
    end
    for i=59
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
```

```
            r(i,f)=data(k,ff)+r(i,f);
        end
    end
%%%%%%%%%%%%%% sorting the consumption in an ascending
order %%%%%%%
    rsort=sort(r(:,f));
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%
%%%%%% finding the days with higher consumption level
%%%%%%%%%%%%%
    for i=1:6
        no(i)=0;
    end
    for j=1:6
        for i=1:69
            u=70-j;
            if rsort(u)==r(i,f)
                no(j)=i;
            end
        end
    end
%%%%% finding customer baseline (4 out of 6 days)
%%%%%%%%%%%%%%%%%%%%%
    for i=1:48
        cbl(i,f)=0;
        act(i,f)=0;
    end
    for e=2:5
        i=no(e);
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            cbl(j,f)=data(k,ff)+cbl(j,f);
        end
    end
    for j=1:48
        cbl(j,f)=cbl(j,f)/4;
    end
%%%%%%%%%%%%%%%% creating a matrix for actual consumption
%%%%%%%%%
    for i=60
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            act(j,f)=data(k,ff)+act(j,f);
        end
```

```
    end
end
```

c) Low4of5 code

```
%%%%
for f=1:262
    for i=1:69   %%% creat an empty matrix
        r(i,f)=0;
    end
%%% selecting 5 admissible day
    for i=53:56
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            r(i,f)=data(k,ff)+r(i,f);
        end
    end
    for i=59
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            r(i,f)=data(k,ff)+r(i,f);
        end
    end
%%%%%%%%%%%%%% sorting the consumption in an ascending
order %%%%%%%
rsort=sort(r(:,f));
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%
%%%%%% finding the days with higher consumption level
%%%%%%%%%%%%%
    for i=1:5
        no(i)=0;
    end
    for j=1:5
        for i=1:69
            u=70-j;
            if rsort(u)==r(i,f)
                no(j)=i;
            end
        end
    end
%%%%% finding customer baseline (4 out of 5 days)
%%%%%%%%%%%%%%%%%%%%%
    for i=1:48
        cbl(i,f)=0;
```

```
        act(i,f)=0;
    end
    for e=2:5
        i=no(e);
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            cbl(j,f)=data(k,ff)+cbl(j,f);
        end
    end
    for j=1:48
        cbl(j,f)=cbl(j,f)/4;
    end
%%%%%%%%%%%%%%%% creating a matrix for actual consumption
%%%%%%%%%
    for i=60
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            act(j,f)=data(k,ff)+act(j,f);
        end
    end
end
```

d) ISONE code

```
%%%%%%%%%%%%% creat empty matrices %%%%%%
for f=1:262
    for i=1:48
        cbl6(i,f)=0;
        cbl(i,f)=0;
        act(i,f)=0;
    end
end
%%%%%%%%%%% for all customers %%%%%%
for f=1:262
    %%%%%%%% creat CBL6 out of first five days %%%%%%%%
    for i=4:7        %%% day 4, 5, 6 and 7
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            cbl6(j,f)=data(k,ff)+cbl6(j,f);
        end
    end
    for i=10      %%%% and day 10
        for j=1:48
            k=48*(i-1)+j;
```

```
            ff=3*f;
            cbl6(j,f)=data(k,ff)+cbl6(j,f);
        end
    end
    for j=1:48
        cbl6(j,f)=cbl6(j,f)/5;
    end
%%%%%%%%%%%%%%% creat CBL %%%%%%
    for i=1:48
        cbl(i,f)=cbl6(i,f);
    end

    for i=11:14
        for j=1:48
            ff=3*f;
            k=48*(i-1)+j;
            cbl(j,f)=0.1*data(k,ff)+0.9*cbl(j,f);
        end
    end
    for i=17:21
        for j=1:48
            ff=3*f;
            k=48*(i-1)+j;
            cbl(j,f)=0.1*data(k,ff)+0.9*cbl(j,f);
        end
    end
    for i=24:28
        for j=1:48
            ff=3*f;
            k=48*(i-1)+j;
            cbl(j,f)=0.1*data(k,ff)+0.9*cbl(j,f);
        end
    end
    for i=31:35
        for j=1:48
            ff=3*f;
            k=48*(i-1)+j;
            cbl(j,f)=0.1*data(k,ff)+0.9*cbl(j,f);
        end
    end
    for i=38:42
        for j=1:48
            ff=3*f;
            k=48*(i-1)+j;
            cbl(j,f)=0.1*data(k,ff)+0.9*cbl(j,f);
        end
```

```
    end
    for i=45:49
        for j=1:48
            ff=3*f;
            k=48*(i-1)+j;
            cbl(j,f)=0.1*data(k,ff)+0.9*cbl(j,f);
        end
    end
    for i=52:56
        for j=1:48
            ff=3*f;
            k=48*(i-1)+j;
            cbl(j,f)=0.1*data(k,ff)+0.9*cbl(j,f);
        end
    end
    for i=60     %%%%% CBL for event day
        for j=1:48
            ff=3*f;
            k=48*(i-1)+j;
            cbl(j,f)=0.1*data(k,ff)+0.9*cbl(j,f);
        end
    end

%%%%%%%%%%%%%%%% creating a matrix for actual consumption
%%%%%%%%%
    for i=60
        for j=1:48
            ff=3*f;
            k=48*(i-1)+j;
            act(j,f)=data(k,ff)+act(j,f);
        end
    end
end
```

e) Regression code

```
%%%%%%%%
for f=1:262
    a=1;
    for i=1:3312
        sort(i,1)=0;
        sort(i,2)=0;
        sort(i,3)=0;
        sort(i,4)=0;
        sort(i,5)=0;
        sort(i,6)=0;
        sort(i,7)=0;
```

```
        sort(i,8)=0;
        sort(i,9)=0;
    end
%%%%%%% all the consumptions of each half an hour will be
sorted after
%%%%%%%% each other in one place. Therefore, the first 69
rows are the
%%%%%%%% first half an hour of all the 69 days
    for i=1:48
        t=1;
        for j=1:69
            k=48*(j-1)+i;
            ff=3*f;
            sort(a,1)=data(k,ff);
            sort(a,2)=t;
            a=a+1;
            t=t+1;
        end
    end
    for i=1:3312
        if mod(sort(i,2),7)==1;
            sort(i,3)=1;
        elseif mod(sort(i,2),7)==2;
            sort(i,4)=1;
        elseif mod(sort(i,2),7)==3;
            sort(i,5)=1;
        elseif mod(sort(i,2),7)==4;
            sort(i,6)=1;
        elseif mod(sort(i,2),7)==5;
            sort(i,7)=1;
        elseif mod(sort(i,2),7)==6;
            sort(i,8)=1;
        else
            sort(i,9)=1;
        end
    end

X=[sort(:,3),sort(:,4),sort(:,5),sort(:,6),sort(:,7),sort(:
,8),sort(:,9)];
    Y=sort(:,1);
%%%%%%%%%%%%%%%% mvregress %%%%%%%%%%%%%%%%%%%%%%%
    for i=1:48
        for j=1:59
            k=(i-1)*69+j;
            XX(j,:)=X(k,:);
        end
```

```
        for j=1:59
            k=(i-1)*69+j;
            YY(j,:)=Y(k,:);
        end
        [beta,sigma,resid]=mvregress(XX,YY);
        cbl(i,f)=beta(4,1);
    end
%%%%%%%%%%%%%%%% creating a matrix for actual consumption
%%%%%%%%%
    for i=60
        for j=1:48
            k=48*(i-1)+j;
            ff=3*f;
            act(j,f)=data(k,ff);
        end
    end
end
```

Part Two: Code for adjustment

```
%%%%%%%%%%%%
for p=1:262    %%%% create empty matrices
        adjXX(p)=0;
        adjYY(p)=0;
        adjaax(p)=0;
        adjyyx(p)=0;
        adjAA(p)=0;
        adjacc(p)=0;
        adjacc1(p)=0;
        adjaccuracy(p)=0;
        adjbias(p)=0;
        adjbias1(p)=0;
        adjbias(p)=0;
    end
for f=1:262
    for j=21:24
        adjXX(f)=act(j,1)+adjXX(f);
    end
    adjaax(f)=adjXX(f)/2;
    for j=21:24
        adjYY(f)=cbl(j,1)+adjYY(f);
    end
    adjyyx(f)=adjYY(f)/2;
    adjAA(f)=adjaax(f)-adjyyx(f);    %%%% adjustment value

    for j=1:48
```

```
        adjcbl(j,f)=cbl(j,f)+adjAA(f);           %%%%
adjusted CBL
    end
end
```

Part Three: Accuracy and Biases

```
%%%%
for f=1:262
%%%% accuracy with additive adjustment %%%%
%%%% for entire day
    for j=1:48
        adjacc(f)=abs(adjcbl(j,f)-act(j,f))+adjacc(f);
    end
    adjaccuracy(f)=(adjacc(f)/24);
%%%% bias with additive adjustment %%%%
%%%% for entire day
       for j=1:48
        adjbias(f)=adjcbl(j,f)-act(j,f)+adjbias(f);
    end
    adjbiasday(f)=(adjbias(f)/24);
%%%% accuracy with additive adjustment %%%%
%%%% for event hours
    for j=29:42
        adjacc1(f)=abs(adjcbl(j,f)-act(j,f))+adjacc1(f);
    end
    adjaccuracyevent(f)=(adjacc1(f)/7);
%%%% bias with additive adjustment %%%%
%%%% for event hours
       for j=29:42
        adjbias1(f)=adjcbl(j,f)-act(j,f)+adjbias1(f);
    end
    adjbiasevent(f)=(adjbias1(f)/7);
end
AAregeventday=mean(adjaccuracy);
ABregeventday=mean(adjbiasday);
AAregeventhours=mean(adjaccuracyevent);
ABregeventhours=mean(adjbiasevent);
```

Part Four: Payment Settlement

```
%%%%%%%%%%%% CBL payment settlement after adjustment
%%%%%%%%%%%%%%%%
for u=1:48
    axis(u)=u;
end
for i=1:48
       aatotal(i,1)=0
```

```
  end
  for i=1:48
       aatotal(i,1)=sum(act(i,:));
  end
  plot(axis,aatotal,'--rs')
utilityrevenue=sum(aatotal)*0.097;
%%%%%%%%%%%%rebate payment%%%%%%%%%%%%%%%%%%%%%%%%
for f=1:262
    for j=1:48
        diff(j,f)=adjcbl(j,f)-act(j,f);
    end
end
for f=1:262
    for j=1:48
        if diff(j,f)>0
            posdiff(j,f)=diff(j,f);
        else
            posdiff(j,f)=0 ;
        end
    end
end
  for i=1:48
       hourlyrebatetotal(i,1)=sum(posdiff(i,:));
  end
falseloadreduction=sum(hourlyrebatetotal);
utilityrebate=sum(hourlyrebatetotal)*0.35;
```