AN IMPROVEMENT OF THE LOAD REDUCTION EVALUATION METHODOLOGIES EMPLOYED IN DEMAND RESPONSE (DR) PROGRAMS OFFERED TO RESIDENTIAL CUSTOMERS

by

Saeed Mohajeryami

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Electrical Engineering

Charlotte

2017

Approved by:

____

Dr. Valentina Cecchi

____

Dr. Peter Schwarz

____

Dr. Zia Salami

____

Dr. Yogendra Kakad

____

Dr. Churlzu Lim

ABSTRACT

SAEED MOHAJERYAMI. An Improvement of The Load Reduction Evaluation Methodologies Employed in Demand Response (DR) Programs Offered to Residential Customers. (Under the direction of DR. VALENTINA CECCHI)

In current electricity market, demand side buys power indirectly through electric utilities, and it makes the demand side inelastic to market's price volatility and significantly affects the efficiency of the market. To fix the problem, it is essential to involve the demand side directly. One major step taken towards this goal is demand response programs. These programs offer many benefits and can provide solutions to many issues in the market. However, some challenges are facing their implementation, chief among them is the estimation of load reduction. An accurate measurement of the load reduction needs an accurate estimate of Customer Baseline Load (CBL). In this dissertation, it is observed that the CBL methods developed for large industrial and commercial customers are not satisfactorily accurate when applied to residential customers. Residential customers have a variable load. Moreover, with increasing penetration of distributed generation and storage devices, large industrial and commercial loads are also becoming variable. Therefore, it is an imperative to explore new methods to estimate the CBL for variable loads accurately. In this dissertation, the challenges associated with the presence of CBL are studied carefully, and a k-means clustering method based on the average load and a predictability index is proposed to improve CBL estimation. The advantages of the proposed method are shown both in theory and through an experiment. It is shown this proposed method can improve the error performance of CBL estimation considerably.

ACKNOWLEDGMENTS

I am very pleased that I finally completed this circuitous path towards a notable contribution. I would like to thank, in large part, to exceptionally gracious people who supported me along the way. I like to thank my adviser, Dr. Valentina Cecchi, for all her help and guidance that she has given me over the past two years. I am also extremely fortunate to have committee members Dr. Peter Schwarz, Dr. Zia Salami, Dr. Yogendra Kakad, and Dr. Churlzu Lim. I thank them for their valuable ideas and feedback, always aimed at moving me forward. I would especially like to thank my former advisers, Dr. Zia Salami and Dr. Peter Schwarz. You have set examples of excellence as a researcher, mentor, instructor, and role model, and forever I shall remain indebted to you.

Additionally, I want to thank my colleague and friend, Milad Doostan, who helped me in the simulation, analysis, and writing of a large part of this dissertation. Much of my work would not have been completed without the assistance of some of my fellow graduate students, Tumininu Lawanson, Iman Naziri Moghaddam, and Roozbeh Karandeh.

I would also like to thank my amazing wife, Mahboubeh, for her unremitting love and support. She has been a source of encouragement and inspiration to me since I met her.

Finally, I would like to thank and dedicate this dissertation to my mother, Pari (means "angel" in Persian), for loving me unconditionally, which without her continuous support and encouragement I never would have been able to achieve my goals.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Overview

Traditionally, in the United States, electric utilities were vertically integrated with transmission systems. In other words, one utility was responsible for the entire process of generation, transmission, and distribution of electricity. However, in the 90s, with deregulation in other industries, the electric industry came to the conclusion that deregulation was inevitable. Accordingly, Independent System Operators (ISOs) and Regional Transmission Organizations (RTOs) were established in the light of deregulation of the electric industry. These organizations were introduced to deal with the expanding number of transactions in the electricity market, which is one of the products of deregulation in the electric industry. Around a dozen states chose to deregulate their electricity market. Yet after the California energy crisis of 2000-2001, some pulled back.

Opening up the electricity market was a major step toward deregulation and free markets in the electric industry. However, some parts of the market were not mature enough to participate freely. For instance, the infrastructure that allowed the participation of each individual electricity consumer was non-existent at the time. As a result, smaller electric utilities that used to be only responsible for the distribution of electricity, began to purchase electricity on behalf of their industrial, commercial, and residential customers[64].

In the current state of the electricity market, in a day-ahead market, the power plants offer their electricity generation as a supply and utilities bid for it. The demand in the day-ahead market is not an actual demand; it is a speculative amount based on a demand forecast, which is run by the utilities. The price will be determined by the predicted supply and demand curve. In case something out of order happens, e.g. excess demand or supply shortage, the imbalance will be settled in the spot market. After purchasing electricity for customers, the utilities charge them based on a fixed tariff. In this system, the customers are shielded from the price fluctuations of the wholesale market.

The fact that customers cannot feel the price fluctuations of the market can significantly affect the efficiency of the market. In other words, the efficiency of the free market depends on the elasticity of the supply and demand[9, 50, 10, 11]. In the current wholesale market, the demand is somewhat inelastic to the sudden changes in the supply side. For instance, many economists believe that this isolation is one of the major contributing factors of the California energy crisis[26]. The presence of this issue can lead to problems, including unstable wholesale prices and market power of large electricity suppliers[8, 55]. In order to fix this problem, it is necessary to build the infrastructure for demand-side participation. Part of this infrastructure is technology, and the other part is policy and customer education.

The technological infrastructure of the demand-side participation including advanced smart meters, big data analytics, etc is in development. For example, according to a Federal Energy Regulatory Commission (FERC) survey in 2010, the penetration of smart meters in the US reached a level of 8.7 percent[46]. Furthermore, many utilities are running pilot projects to test these infrastructures and the feasibility of demand participation programs[27, 28, 61, 33, 69].

Another necessary part of the infrastructure is policy. In order to address the inefficiency in the wholesale market, FERC is promoting the adoption of Demand Response (DR) programs. In FERC No. 745, the permission to participate in the wholesale market is granted to the DR owners[29]. In other words, in the wholesale market, ISOs are commanded to treat the load reduction as a supply. DR programs are designed to persuade customers to reduce their consumption in response to some incentives temporarily. According to many researchers in the literature, DR programs can temporarily bring elasticity to the demand side.

1.2 Motivation

Although DR programs can provide workable solutions to many issues in the market, there are many challenges facing their implementation, chief among them is Evaluation, Measurement, and Verification (EM&V) of the load reduction. Measurement of the customers' load reduction is necessary for the purpose of payment settlement, which is a critical part of the most of DR programs. An accurate measurement of the load reduction depends upon the accurate estimation of Customer Baseline Load (CBL). The CBL is the amount of electricity that a customer would have consumed if there was not a DR event. Therefore, it is a "counterfactual" load. Besides, it is called "baseline," because it provides a baseline for computing the load reduction. The difference between the actual load and the CBL is the "curtailment" that DR programs are designed for[43].

The central topic of this dissertation is CBL. The EM&V of CBL and the ways to enhance the accuracy of CBLs were extensively investigated in the literature in recent years. Moreover, FERC 745 order commanded ISOs to develop a CBL for their own DR programs[29]. Some of these methods are tabulated in Table 1. The full description of each method is available in [51]. It is worth noting that the methods in Table 1 were originally created for large Industrial and Commercial (I&C) customers.

Table 1: CBL calculation methods employed by different ISOs

Independent System Operator (ISO)	**CBL Calculation Method**
PJM	Averaging (High4of5)
NYISO	Averaging (High5of10)
CAISO	Averaging (Last10days)
Ontario,Canada	Averaging (High15of20)
ISONE	Exponential Moving Average
ERCOT	Regression Models

In recent years, the high penetration of smart meters in the residential sector, which provide granular hourly consumption data, has created unprecedented opportunities for load aggregators to enter into a contract with residential customers. Load aggregators can bundle the potential residential customers' load reduction and offer it as a supply source in the electricity market. Therefore, it is fair to say that the challenges of CBL estimation are not exclusive to market operators anymore; they have also become a load aggregator problem.

In contrast to ISOs' DR programs that work with large industrial and commercial customers, the load aggregators mainly deal with residential customers. This requires them to establish CBL for each individual customer. The development of

CBL calculation methods for residential customers faces more challenges as the load curves of these customers have much more random characteristics compared to large industrial and commercial customers. This randomness is driven by multitudes of non-correlated personal and household activities. By taking the fluctuations that exist in residential customers' data into consideration, CBL calculation methods ought to be improved in a way that the effect of such volatility is addressed. The authors in [52, 54, 53, 72, 63] show that CBL methods developed for large industrial and commercial customers are not satisfactorily accurate when applied to residential customers. To date, developing CBL methods for residential customers have rarely been seriously scrutinized in the literature.

Moreover, the nature of large industrial and commercial loads is also changing in the current power industry. As more and more customers use Distributed Generation (DG) and energy storage devices, the load shape of these customers is becoming increasingly random and unstable. Therefore, it is possible to speculate that as the nature of these customers changes over time, the old CBL estimation methods are becoming obsolete.

All the aforementioned developments in the power industry make old CBL estimation methods unreliable. Therefore, it is necessary to work on new ways by using state-of-the-art machine learning techniques and sophisticated data analytics to create more reliable and accurate CBL estimation methods.

1.3 Objective

In order to improve the CBL accuracy for residential customers, it is critical to study the nature of residential customers' consumption. One way to examine the nature of residential customers' consumption is to treat it as a time-domain signal and use the signal processing. In this dissertation, it is attempted to decompose the residential customers' hourly consumption data into its underlying frequency-domain signal counterpart and separate the signal into two high and low-frequency components. Then, these two high- and low-frequency components will be reconstructed into their time-domain counterparts. Based on this operation, an index is proposed to be utilized as a means for finding similarity among the customers. The index is called the predictability index, and it reflects the share of low-frequency components of the consumption signal, which is assumed to be predictable.

Another concern and challenge in using CBL as a measure for load reduction is the gaming opportunities. Customers can game the system with inflating their baseline to benefit on a speculated event day. This issue is extensively analyzed in a chapter dedicated to the social welfare analysis. Social welfare is a microeconomics term that explains how much a particular interference can change the dynamics of supply and demand and cause inefficiencies. This analysis is necessary for developing a comprehensive and improved CBL calculation method.

In this dissertation, the author has proposed two approaches to improve the error performance of CBL calculation methods for residential customers. First, according to an economic analysis (social welfare loss analysis,) it is proposed that the load aggregation is necessary to improve the gaming challenges presented by the presence of the CBL as a tool for load reduction evaluation. Also, it is argued that load aggregation has a significant impact on the improvement of error performance. Second, it is proposed to use machine learning techniques such as k-means clustering to separate customers based on the similarity (predictability index) and improve the error performance of CBL calculation methods for residential customers. The data employed in this study belongs to Australian Energy Market Operation (AEMO,) and is the hourly consumption of 189 customers for the time span of a year (2012). In the end, it is shown how the proposed approaches improve the estimation and error performance of the CBLs.

1.4 Summary of contribution

In this section, the contributions of this dissertation are highlighted.

The key contributions are listed as follows:

- The error performance of the well-established CBL estimation methods are evaluated for residential customers.

- The challenges of CBL estimation methods are analyzed in theory by using customers' utility function.

- The impact of FERC 745 order on the current DR programs is evaluated in theory. Moreover, the impact of the order is shown on the social welfare.

- The stochasticity of three categories of customers, i.e. large industrial, commercial, and residential customers is evaluated.

- A novel method is proposed to cluster customers in order to improve the error performance of current CBL estimation methods.
- The impact of the proposed method is analyzed on the social welfare.

1.5 Organization of the dissertation

The rest of this dissertation is organized as follows. Chapter 2 reviews the literature of Evaluation, Measurement, and Verification (EM&V) of demand response programs including the research works on CBL estimation methods for all categories of customers. Moreover, this Chapter describes the challenges of CBL as a means of load reduction evaluation. Furthermore, the Chapter explains how CBL is perceived by customers and utilities. In Chapter 3, the issues with the current FERC 745 order and CBL estimation methods are described. It is shown how this current approach affects the social welfare. Moreover, it is shown that the residential customers have, by far, the highest variability among the customers. In Chapter 4, the proposed methodology of this dissertation is introduced, and the flowchart of its implementation is provided. In this Chapter, it is shown how the proposed method affects the social welfare. In Chapter 5, some of the CBL calculation methods are explained. Moreover, three error metrics are introduced to examine the error performance of the CBL estimations, and also the dataset employed in this dissertation is described. The results for the error performance of regular CBL, CBL with random load clustering, and CBL with k-means clustering and predictability index are provided and discussed. The conclusion along with recommendations for future works are provided in Chapter 6.

CHAPTER 2: CUSTOMER BASELINE LOAD (CBL)

2.1 Overview

In this chapter, the concept of CBL is introduced along with its benefits and also the challenges that it poses to a DR program. Moreover, different methods of Evaluation, Measurement, and Verification (EM&V) used by various programs are investigated. In this investigation, advantages and disadvantages of each method are identified. Besides, to understand how to approach the problem of CBL estimation, it is important to understand how CBL affects utilities and customers. For this reason, this chapter also includes an explanation of how utilities and customers perceive CBL and how it affects them.

2.2 Evaluation, Measurement and Verification (EM&V)

EM&V of DR programs is very important because, without accurate EM&V, it is impossible to justify these programs. Accurate EM&V is also needed for the purpose of payment settlement and customer selection. Moreover, since DR programs are in their early stages of development, they need accurate EM&V to be able to recommend proper policies to policymakers. As mentioned earlier, one of the main issues with EM&V of DR programs is estimating the CBLs. Several studies have investigated CBL methodologies extensively. In this chapter, these studies will be carefully reviewed. CBL is the hypothetical power demand of a customer in the absence of a DR event. CBL is utilized by a utility to determine the level of load curtailment for the customer and subsequently will be used to compensate the customer. Figure 1 illustrates the concepts of DR baseline, actual load, and estimated load reduction. Indeed, these three concepts are repeatedly employed hereafter to explain the different methodologies and their measurement and verification processes. In this figure, the upper black line represents the estimated baseline, and the red line represents the actual consumption. The difference between these two lines during the event hours is the amount of load reduction, which is shaded in the blue color. The difference between the black and red lines is shown by bar charts.

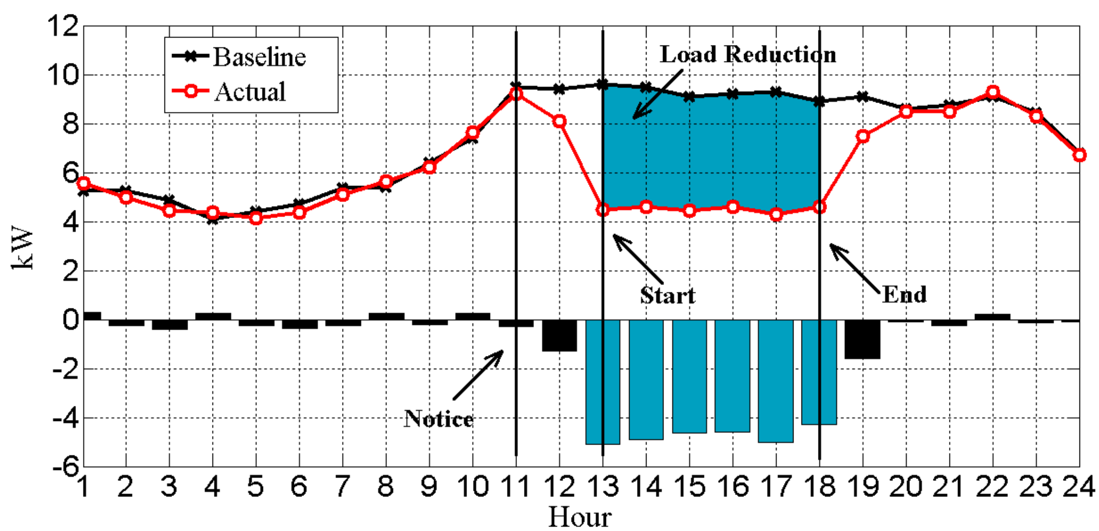

Figure 1: Illustration of CBL, actual load and estimated load reduction

Reference [31] has provided the most comprehensive review of CBL methods. It describes all CBL estimation methodologies used by utilities and ISOs in the U.S. for their large industrial and commercial customers. Here, the algorithms used by the different ISOs are applied to a dataset (collected from California) and their error performance, i.e. accuracy and bias, are compared. Specifics of the error performance metrics are explained in Chapter 5 of this dissertation. The paper concludes from the analysis that:

1. Same-day "additive" or "multiplicative" adjustments have a superior performance compared to an unadjusted CBL or a CBL using weather sensitive adjustment. Additive and multiplicative adjustments are based on the difference and percentage difference between estimated CBL and the actual consumption in the morning of event days, respectively.

2. The choice of "multiplicative" or "additive" adjustment does not change the outcome noticeably.

3. This work shows that X of Y methods such as CAISO (used by California ISO), PJM economic, mid 4 of 6, and regression approaches, with a same day additive adjustment, produce similar satisfactory results.

4. The methods in the above item performed poorly for predicting the CBL for variable load customers.

5. Regression approaches have higher administrative costs and associated complexity compared to X of Y methods.

6. Weather-dependent models did not generally outperform the other models that did not include weather. However, it is probable that this observation is due to the weather stability of California State.

Y and X, in this context, refer to the number of non-DR days before the DR event and the number of days selected out of these Y days, respectively. For instance, in HighXofY methods, X is the number of days with the highest consumption out of Y days. For readers unfamiliar with the aforementioned methods, they are fully described in chapter 5 of this dissertation.

Another study from California is [2]. It is part of the broader evaluation of California's 2004 DR programs. In this study, the impacts of load size (small, medium, large, and extra-large), business type (commercial, industrial and institutional), event day type (high demand, low demand, and consecutive high demand) and weather parameters are investigated on the performance of 3 CBL estimation methods of 3-day, 10-day and prior-day. In this paper, 10-day baseline with the same-day adjustment is reported as the most accurate method. This finding is reaffirmed in [3, 13]. It is worth highlighting that the analysis is performed on data from industrial and commercial customers. In a similar attempt, [74] came to a similar conclusion by performing a statistical analysis on the data from 10 industrial customers in South Korea.

Another study on non-residential buildings in California has been conducted by the Lawrence Berkeley National Lab (LBNL) on sample data from 32 sites in California [21, 20]. The methods investigated in this study overlap with [2, 3, 13]. According to this paper, morning adjustment could significantly reduce the bias and improve the accuracy of all CBL models. Moreover, the results of this paper suggest 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 better suit each particular load. This paper also confirms one of the findings of [31] that none of the examined CBL methods produce satisfactory results for highly variable loads. On the contrary to [31], this work finds that incorporating temperature (e.g. explicit weather models) improves the accuracy of the estimated CBL, and in cases where it does not improve the accuracy, it has relatively little impact.

Another study about X of Y methods suggests that there is an optimal X for any Y values[4]. According to this study, there is a range of X/Y values which minimizes the bias. This study suggests that the error is minimized by $0.4 \leq \frac{X}{Y} \leq 0.8$. However, the study is empirical, and it does not provide any theoretical basis for the results. Therefore, unless another independent study corroborates this finding, it must be regarded with caution.

In a recent attempt to modify regression-based CBL estimation, the authors in [66] have proposed a recursive Bayesian linear regression approach for baseline load analysis. In the paper, the simulation shows that this modified regression works slightly better than the regular regression. Moreover, the modified method can be used as an on-line learning algorithm. In [65], the author has proposed a Kernel adaptive filter for on-line learning. In essence, the author treated the consumption load as a time signal. An adaptive filter refers to a model that adapts the parameters of its transfer function based on the property of incoming signals over time by minimizing a loss (e.g. error) function, which is a deviation of filter output from a typical behavior (the model output).

Besides the right estimation model for CBL, it is very important to understand the sensitivity of the models to implementation choices. The sensitivity analyses can help to improve the performance of CBL methods in an effective manner. The authors in [6, 7] have studied the different implementation choices and their impact on baseline estimation. The study employed data from 38 large commercial buildings and industrial facilities in the Pacific Gas and Electric Company (PG&E). These studies conclude that:

1. The load reduction estimate is strongly sensitive to the source of the outdoor air temperature data.

2. The choice of power outage filter strongly affects the load reduction estimate.

3. The load reduction estimate is less sensitive to data alignment choice. In other words, if there is a time mismatch between temperature and load data, it would not affect the estimate.

4. The load reduction estimate is almost insensitive to the choice of data intervals less than an hour. In other words, three cases of 15-min, 30-min, and 60-min data intervals lead to almost the same estimate.

In another study, the authors in [75] employ the similar data to [6, 7] to study what explanatory variables are better to include in the regression models. They find that outside air temperature and load patterns are good candidates for the regression model. The load patterns could be found from a cluster analysis. Given the uncertainty of business types, a cluster analysis could serve as a useful tool to categorize the data into different types. In addition to the aforementioned attempts to determine the proper implementation choices, there are several papers in the literature that tried to modify the existing CBL methods in order to improve their error performance. The authors in [71] have proposed an exponential smoothing model to calculate CBL. In a sense, this model is like averaging methods with the difference that it weighs past observations with exponentially decreasing weights. This model allows the recent changes to be better reflected in the estimated baseline. The proposed method shows superiority over High5of10 and regression methods for both adjusted and unadjusted cases. In another study about the regression modification, the authors in [62] indicate that the Quantile Regression (QR) algorithm is more effective in reducing the economic cost of the forecasting errors.

Moreover, the authors in [58, 59] have proposed a CBL calculation framework employing data mining techniques. These papers employ the real data from a large industrial building complex in Korea. Their proposed method utilizes two data mining techniques of the Kohonen networks model (self-organizing map) and the unsupervised learning algorithm (k-means clustering) to find the days with the most similar load patterns. According to the results of these papers, the root mean square error is reduced by 15-22% on average compared to the regular averaging methods, and the mean absolute percentage error is reduced by 15-20% on average as well.

Almost all the research works mentioned so far are targeted to industrial and commercial customers. However, in recent years, the technological infrastructure on the distribution system side has become mature enough to allow DR programs to offer their services to residential customers. For example, high penetration of smart meters at the residential level has provided high-quality, high-resolution consumption data [36, 37, 38]. The availability of the residential data has created unprecedented opportunities for load aggregators to offer DR programs to customers. Moreover, it has provided an opportunity for researchers to examine the strength of existing CBL methods for customers.

The calculation of CBL for residential customers has its unique challenges. The load curves of residential loads is dependent upon multiple non-correlated personal and household activities. Most of these activities do not follow any schedule. Therefore, load curves of residential customers have much more random characteristics than industrial and commercial customers. Addressing the randomness in loads of residential customers requires a completely different approach. Therefore, it is very critical to review the findings of researchers in addressing the unique challenges of these customers. The authors in [52, 51, 54, 53, 72, 70, 63] show that the conventional CBL methods, which as previously mentioned are developed for large industrial and commercial customers, make considerable amounts of error for residential customers. The error in CBL estimation can lead to a poor distribution of rewards, poor reliability management, and poor resource allocation. With inaccurate CBLs, some customers will be over-compensated, whereas many others will be under-compensated. Insufficient demand reduction leads to emergency load shedding. Moreover, spinning reserve allocation is inaccurate if CBLs over-estimate the capability of loads[76]. Therefore, it is imperative to develop new methods for these customers.

The authors in [43] have reviewed the baseline methodologies for small scale residential DR programs to highlight the challenges of CBL calculation for these customers. The paper has tested five CBL methodologies of High X of Y, Last Y days, Regression, Neural network, and polynomial extrapolation on the actual smart meter data of 66 residential customers. According to the results, the authors found that machine learning technique of neural network and polynomial extrapolation outperform the other methodologies. This research work comes from an electric utility in Australia as they are trying to offer DR programs to residential customers. This work can be regarded as evidence that interest in CBL for residential customers is not merely in academia, but it is one of the current power industry needs.

Another method that is suggested for baseline estimation of residential customers is the Randomized Controlled Trial (RCT) method. This method, unlike the other available CBL calculation methods, is observed only to be utilized for residential customers. This method is recommended by Lawrence Berkeley National Lab (LBNL) as one of the methods that can be used to assess the effects of time-based rates, enabling technologies, and various other treatments on customers' consumption levels and patterns of usage [14]. The RCT method assigns the households into two groups of treatment and control randomly. These two groups are exposed to similar conditions, and the difference between them could be attributed to the treatment effect. RCTs rely on minimal assumptions about the nature of customers; therefore, they can produce unbiased estimates of treatment effects [14]. Additionally, the RCT method has a lower administrative cost compared to the other methods as it requires no historical data for CBL calculation. Therefore, under equal conditions, the RCT is a much better alternative, both regarding lower cost and lower complexity. The authors in [68], have recommended using RCT for evaluating the energy efficiency (i.e. load reduction) in behavior-based efficiency programs. They assert that the RCT method, in comparison with the alternative methods, is more robust and unbiased. Moreover, they acknowledge that due to the counterfactual nature of the real load reduction, it is impossible to measure it, and it can only be estimated; therefore, the RCT estimates would contain inherent randomness. Many other energy saving programs have utilized RCT for the estimation of load reduction. The most popular ones are Statewide Pricing pilot[41], Anaheim CPP program[73], Olympic Peninsula Project[39], Smart metering project, the Integral Energy trial, PowerCentsDC program, Kyushu Electricity Pricing Experiment, and Energy Demand Research Project.

The authors in [40] propose a method to create a suitable control group to measure the treatment effect. They argue that the electric utilities, with the introduction of smart meters, have access to a large amount of quality data about residential customers. This data can be used to select a very effective control group. They propose a new method of control group selection according to individual load curves. This selection aims to minimize the distance between the selected control group's load curve and the load curve of DR participants. The results show a significant improvement compared to conventional methods. Moreover, they find that the greater the size of the control group, the greater the accuracy. In other words, with increasing the control group size, it is argued that the greater and more diversified sample of individual load curves can improve the quality of the control group. The research work is part of a pilot project in France. This DR program is offered to residential customers in Britony, France.

In order to create a better control group, in general, it is necessary to select customers with the same observable characteristics of the treatment group. Although the RCT methods might seem easy and straightforward, they require many individual characteristics to create the best control group, which typically utilities do not collect. Moreover, many such characteristics only show a static view of electricity usage. They are unable to show a dynamic view of the ever-changing behavior of customers. As a matter of fact, there are many major unobservable characteristics that determine the propensity of customers to join DR programs, which are not available.

A clustering method has been proposed in [76]. In the paper, the groups are selected by k-means clustering method. The proposed method clusters the customers based on their power consumption profiles. It uses two criteria of 1) usage level, and 2) usage pattern, to cluster customers into the same group. Then the baseline is estimated by a simple average among the customers of each group. One of the advantages of this method is that it uses the actual data of the event day. Additionally, since it does not rely on historical data, it eliminates some of the moral hazards of the conventional CBL methods such as baseline inflation. However, there are some disadvantages associated with this method, especially when it is not possible to find customers with similar characteristics in one group. For a subset of customers with unique behavior patterns, finding a cluster to accurately replicate their loads is very challenging. Moreover, it is hard to justify why the consumption of others is used for creating a CBL for another customer. According to the results of the paper, among all the methods considered in the article, including several conventional CBL methods, the proposed inner-class-average method according to the k-means clusters consistently show satisfactory results.

In recent years, it is shown that clustering can improve the accuracy of the short-term load forecasting[18, 22], which estimation of CBL can be regarded as one of these load forecasting endeavors. Furthermore, an increasing number of clustering methods have been applied to residential data to identify the right candidates for different programs and planning[30, 47]. In other words, clustering these customers based on their behavioral attributes collected from smart meter data gives a better insight about how to target them in different programs. For example, customers with regular high demand in evenings are not good candidates for DR programs; in contrast, they are ideal candidates for peak shaving by storage devices. One of the challenges of such endeavors is defining the right attributes to cluster the customers. What makes this job more challenging is the high stochasticity and irregularities of residential customers[35]. The authors in [34] have analyzed a case of residential consumption data to better understand the peak demand and also identify major sources of variability in residential customers' behavior. In order to perform this analysis, the authors have used Finite Mixture Model (FMM), which they argue to be a better method compared to k-means clustering. The authors have checked the reliability of the FMM outputs, and according to their results, the final clustering is highly reliable. In other words, with the high degree of certainty, the customers belonged to their appointed cluster.

Although CBL is best known in the context of DR programs, it may serve other functions. [25] proposed a methodology that employs CBL to identify non-technical losses of the system such as theft of electricity. This method utilizes historical demand of a certain customer to estimate the future consumption. Then it compares the estimate with the actual load to identify a suspicious reduction. Also, in a patent application, authors of [44] provided a day-ahead load reduction system based on CBL for inducing a user to manage his/her energy consumption efficiently. This system applies an incentive to achieve the desired load reduction and load decentralization.

The authors also filed another patent [45] to present a load forecasting analysis system for calculating CBL based on the day-ahead reduction system mentioned before.

2.3 CBL challenges

Besides challenges related to the estimation of CBLs, there are some other challenges related to the application of CBL as a means of load reduction estimation. These challenges originate from the rate design of DR programs and the consequent economic incentive of such designs. The studies introduced so far neglected many of these challenges. In what follows some of these issues are reviewed and discussed.

Since participation in DR programs is voluntary, customers have an information advantage over the utilities regarding consumption. This asymmetric information imposes two possible challenges to utilities and consequently to CBL calculation. These two challenges 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, their 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 [15]. In what follows, some of the examples of these challenges are described.

1. **Baseline manipulation**: The improper methodology to determine CBL can encourage customers to inflate their baselines 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 an issue,

the customers have the incentive to change their CBL. Such manipulations are observed and reported by ISO New England in [23].

2. **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 [15].

3. **Generation relocation and inefficient price formation**: This problem is better illustrated by an example. Assume 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 the incentive to use an on-site backup generator that can 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 points 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 [17].

Furthermore, there are some problems in practice that can challenge customers' decisions. For instance, poor accuracy performance of CBL can also 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 [52]. 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 excess 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 does not have any incentive to partake in the program effectively. Therefore, it could create distorted incentives and gaming opportunities.

2.4 CBL from a customer's perspective

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 the event days.

As a customer, load reduction is a means to gain financial rewards, either regarding 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 [56] explain how the financial offering of DR programs is an ultimate determinant of customers' decisions. In this dissertation, the main focus is on the impact of accuracy and bias of CBLs on the financial performance of DR programs.

2.5 CBL from a utility's perspective

Utilities are interested in CBL for different reasons. CBL is a tool for some DR programs to calculate their payments to customers. However, their primary 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 assist in emergencies. 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 provide an incentive for customers to lower their peak consumption.

DR also plays a role in delaying investment for new infrastructures. In some geographical regions, peak demand has grown significantly while overall energy consumption has not grown proportionally. This growing peak demand prompts the utilities to take action and invest in new infrastructures, which can drive the electricity rate higher. DR can provide an alternative solution to maintaining reliability without investing in additional infrastructure. This solution can keep rates low. In a competitive market, even a single event of violation of the obligation to serve can have irreversible negative consequences. For that reason, utilities are more interested in the load reduction aspect of CBL and less interested in the financial side of it. Another reason behind the fact that utilities are less interested in the financial side of CBL is that utilities 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 methods in a way to hinder any discouraging effect.

CHAPTER 3: PROBLEM STATEMENT

3.1 Overview

There are two major problems with CBL and available CBL estimation methods. First, the application of CBL as a means for measurement of load reduction introduces an economic incentive for gaming. Therefore, it is critical to investigate the issue and to find a complementary way to eliminate the gaming incentive. Second, although it is known that residential customers show stochasticity, it is not sure to what degree. In this chapter, the first problem is explored theoretically. In order to address the second problem, residential customers are compared to large industrial and commercial customers to show the degree of difference between them. This study will support the previously discussed argument that CBL estimations for residential customers, due to their nature, needs separate analytical tools.

3.2 Social welfare loss analysis

As discussed earlier, another aspect of employing CBL is that it provides an opportunity for gaming and baseline inflation. Therefore, to improve the performance of CBL calculation of residential customers, it is necessary to examine how CBL impacts the social welfare.

In this chapter, by utilizing the customer's utility function, social welfare loss is investigated. Also, a model of consumer utility maximization is proposed to consider the possibility of gaming and inefficient consumption. The authors in [16] have performed an analysis of the gaming possibilities of FERC Order 745. In this dissertation, the model of [16] is extended to include the effect of inherent inaccuracy of CBLs. Moreover, later, the same analysis is employed to examine the impact of the proposed clustering method on net benefit.

3.2.1 Customer's utility function

In this section, it is shown how customers react to the probability of an event day. For this analysis, a DR program with the compensation rate recommended by FERC 745 is employed. The event day refers to the day that DR program calls the customers to curtail their consumption. Customers try to maximize their utility function at all times. The term "utility" in the economic literature refers to all benefits, tangible or intangible, received by customers. The utility function is a tool to measure the well-being or the level of "satisfaction" or "happiness" that a customer gains from the consumption of various goods.

During a random period t, assuming λ_t for the probability of a DR event, the expected utility of a customer is described as follows:

$$\underset{c_t^e, c_t^n}{Max}(1-\lambda_t)[U_t(c_t^n) - p^r.c_t^n] + \lambda_t[U_t(c_t^e) - p^r.c_t^e + p^\pi(c_t^{BL} - c_t^e)] \tag{1}$$

where:

c_t^n Consumption during non-event period

c_t^e Consumption during event period

U_t Utility function at random period t

p^r Retail rate

p^π Demand reduction credit

c_t^{BL} Customer baseline

CBL is regarded as a forecasting and no one knows the exact value of CBL before its occurrence; therefore, it is subject to forecasting error, and this error ought to be considered in every one of the mathematical statements. After this, CBL will be utilized in the form described below.

$$c_t^{BL} = \alpha . c_t^n \tag{2}$$

where α is the accuracy coefficient for CBL calculation and is defined as:

$$\begin{cases} \alpha = 1 & \text{if CBL is accurate} \\ \alpha < 1 & \text{if CBL is inaccurate, and with negative bias} \\ \alpha > 1 & \text{if CBL is inaccurate, and with positive bias} \end{cases}$$

In order to find the optimal electricity consumption during the event and non-event periods, the first derivative of (1) should be taken with regards to c_t^e and c_t^n, respectively, as in (3) and (4).

$$U_t{}'(c_t^e) = p^r + p^\pi \tag{3}$$

$$U_t{}'(c_t^n) = p^r - \alpha(\frac{\lambda_t}{1-\lambda_t})p^\pi \tag{4}$$

On the event day, the customer's electricity price is p^r; it is therefore expected that consumption would be driven by this price. However, the reward paid by the DR program changes the utility function of the customer. Consequently, the right hand side of (3), i.e. the marginal utility price, changes to the higher value, i.e. $p^r + p^\pi$. The higher marginal utility price, according to the supply and demand curve, drives down the consumption, c_t^e, to a lower level. In other words, consumers consume less power.

Similarly, the consumption on non-event days, c_t^n, would be affected by the DR event because the rewards that are offered on event day are assessed by using the non-event days' consumption. Equation (4) represents the aforesaid argument mathematically.

In the absence of DR program (i.e. $\lambda_t = 0$) the marginal utility price is equal to the retail rate (p^r); nevertheless, when $\lambda_t \neq 0$, the consumption level rises as the marginal utility price drops below the retail rate by a fraction (i.e. $\alpha(\frac{\lambda_t}{1-\lambda_t})$) of the load reduction compensation rate. In order to perform the "welfare loss" analysis, it is essential to find the aforementioned optimal points (marginal utility prices) on the supply and demand curve, and to examine their relation to the equilibrium point. For simplicity, it is assumed that the demand and supply curves are linear as follows:

$$\text{Demand} \rightarrow p = a_1 - a_2.c \tag{5}$$

$$\text{Supply} \rightarrow p = a_3 + a_4.c \tag{6}$$

As mentioned earlier, the ideal consumption in this hypothetical market is where the two curves intersect, i.e. c^*. This point is calculated as (7) and illustrated in Figure 2. This point is referred to in the next section to perform the social welfare analysis of FERC 745 order.

$$c^* = \frac{a_3 - a_1}{a_2 + a_4} \tag{7}$$

3.2.2 FERC 745 introduction

The policy proposed by FERC 745 states that the load reduction must be compensated with a wholesale rate (i.e. LMP). This policy is unable to address incentives for inefficient consumption (non-optimal points) during event period and baseline inflation during non-event period. However, it is argued by FERC that this rate takes many externalities and practical considerations into account.

As stated, this policy compensates the load reduction during the event period with the wholesale market clearing price of the same period as shown in (8).

$$p^{\pi} = p^{w}_{te} \tag{8}$$

This policy leads to formulation of the following marginal utility functions expressed in (9) and (10).

$$U_t{}'(c^e_t) = p^r + p^w_{te} \tag{9}$$

$$U_t{}'(c^n_t) = p^r - \alpha(\frac{\lambda_t}{1-\lambda_t})p^w_{te} \tag{10}$$

As discussed earlier, according to (9), the marginal utility price is higher than the retail price. If marginal utility price is higher than actual payment of customers, there would be an extra incentive payment, which is widely known as "double payment" incentive [15].

Furthermore, (10) indicates that customers, during non-event days, are encouraged to consume more electricity than the optimal consumption level (associated with p^r) to inflate their baseline ($c^n > c^r$).

3.2.3 Social welfare analysis of FERC 745

This section performs an analysis of the FERC 745 policy aimed at defining how the policy impacts social welfare.

During regular hours, in the absence of any DR program, demand side is isolated from the wholesale market, and it only responds to fixed retail price (p^r). As it is shown in (11), demand is independent from the supply in the wholesale market ($p^r \neq p_n^w$).

$$p^r = a_1 - a_2.c^r \tag{11}$$

$$p_n^w = a_3 + a_4.c^r \tag{12}$$

The isolation of demand from market dynamics introduces a social welfare loss known as dead-weight loss. In this subsection, the dead-weight loss is investigated.

According to (10), the marginal utility price on non-event days is smaller than the retail price; therefore, the consumption would increase from c^r to c^n, which causes a dead-weight loss. This is delineated in Figure 2 with the hatched triangles. Based on Figure 2, the right hatched triangle (Δ_{right}) represents the dead-weight loss associated with the difference between non-event day marginal utility price (i.e. the price that customers respond to during non-event days) and equilibrium price. The area of Δ_{right} is calculated as:

$$\Delta_{right} = \frac{1}{2}(c^n - c^*)(p_n^w - p^n) \tag{13}$$

where

$$p^n = p^r - \alpha(\frac{\lambda_t}{1-\lambda_t})p_{te}^w \tag{14}$$

To find the area of Δ_{right} in terms of prices, consumption term $(c^n - c^*)$ should be calculated in terms of prices. For this purpose, (11) and (12) are utilized. In these two equation, according to the FERC 745 order, the marginal utility price of non-event days, changes from p^r to p^n , and the consumption changes from c^r to c^n. These two curves are intersected in order to find $c^n - c^*$ in terms of price as in (15) and (16).

$$p_n^w - p^n = a_3 + a_4.c^n - a_1 + a_2.c^n = a_3 - a_1 + (a_2 + a_4)c^n \tag{15}$$

dividing (15) by $a_2 + a_4$ and replacing $(\frac{a_3 - a_1}{a_2 + a_4})$ term by c^* yields (16)

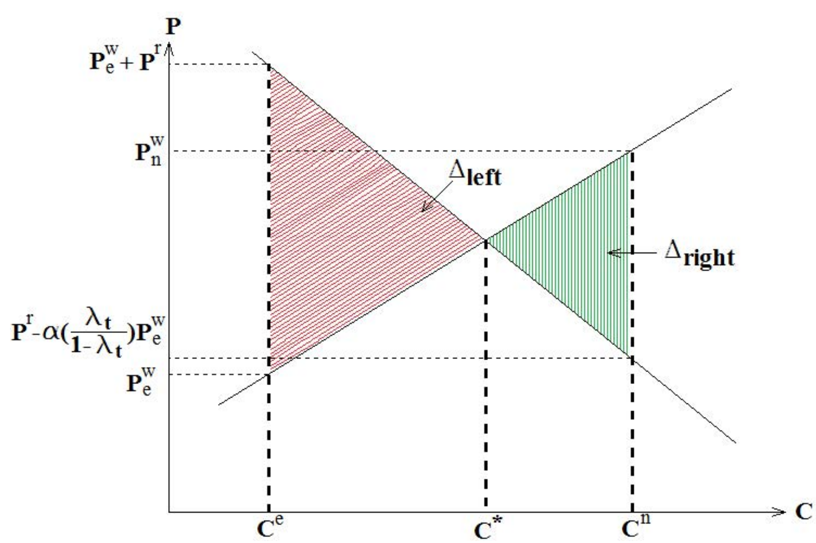

Figure 2: The dynamics of supply and demand during event and non-event periods

$$c^n - c^* = \frac{p^w - p^n}{a_2 + a_4} \tag{16}$$

Using (13) and (16), the dead-weight loss associated with the difference between the non-event day marginal utility price and equilibrium price is shown in (17).

$$\Delta_{right} = \frac{1}{2}\frac{(p^w - p^n)^2}{a_2 + a_4} \tag{17}$$

During an event day, the consumption is determined by the load reduction incentive payment. The consumption during the event period (c_t^e) is illustrated in Figure 2.

As it is demonstrated in Figure 2, the left hatched triangle (Δ_{left}) represents the dead-weight loss associated with the difference between the event day marginal utility price and the equilibrium price. The area of the left triangle is calculated as in (18).

$$\Delta_{left} = \frac{1}{2}(c^e - c^*)(p^r) \tag{18}$$

Similar to the procedure followed to derive (16), $c^e - c^*$ for event day is defined in terms of price by using (11) and (12).

$$c^e - c^* = \frac{p^r}{a_2 + a_4} \tag{19}$$

Using (18) and (19), the dead-weight loss associated with the difference between event day marginal utility price and equilibrium price is as in (20).

$$\Delta_{left} = \frac{1}{2}\frac{(p^r)^2}{a_2 + a_4} \tag{20}$$

The net dead-weight loss is defined as in (21).

$$\Delta_{net} = \lambda\Delta_{left} + (1-\lambda)\Delta_{right} \tag{21}$$

Replacing Δ_{left} and Δ_{right} by equations (20) and (17), respectively, and plugging them into (21) yields (22).

$$\Delta_{net} = \frac{\lambda}{2}.\frac{(p^r)^2}{a_2+a_4} + \frac{(1-\lambda)}{2}.\frac{(p^w-p^n)^2}{a_2+a_4} \tag{22}$$

According to (22), in order to eliminate the dead-weight loss ($\Delta_{net} = 0$), both terms must be zero. The first term points out to the "double payment" incentive. The second term indicates that ideally, the marginal utility price during non-event days must be equal to the same period wholesale price (LMP).

If wholesale rate minus retail rate ($p^w_{te} - p^r$) is adopted as a reward for each load reduction unit, the left-hand side dead-weight loss (Δ_{left}) could be eliminated. Besides, to reduce the value of Δ_{right}, policies and implementation tools must be adopted to minimize the second term. In an ideal case, this second term must approach to zero as shown in (23).

$$p^w - p^n \rightarrow 0 \tag{23}$$

plugging (14) into (23) yields

$$p^w_{tn} - p^r + \alpha(\frac{\lambda_t}{1-\lambda_t})p^w_{te} \rightarrow 0 \tag{24}$$

According to [16] , $p_{tn}^{w} - p^{r}$ is positive; hence, for (24) to hold:

$$\lambda_t \to 0 \quad or \quad \alpha \to 0 \tag{25}$$

It is worth mentioning that $\lambda_t = [0, 1]$ and $\alpha > 0$. The probability of having an event-day (λ_t) depends on the weather, emergencies in the power system, and other unexpected events; thus, this probability cannot be controlled by any exogenous parameter. Also, α represents the accuracy, and it is preferable to have $\alpha = 1$. Nevertheless, if this level of accuracy is not obtainable, according to (25), methods with a negative bias ($\alpha < 1$) are more desirable than methods with positive bias ($\alpha > 1$). However, it is worth mentioning that this conclusion is a solution for a short-run saving as this solution could affect the customers' remuneration adversely in the long run [72].

The discussion on cutting the aforementioned social welfare loss is beyond the scope of this dissertation. Indeed, should CBL calculation methods fail to deliver a satisfactory error performance, further discussion about the gaming incentives of FERC 745 order would be irrelevant. Therefore, it is critical to study the error performance to display the complete picture of the challenges FERC 745 order faces. In chapter 5, the error performance is empirically investigated by examining some well-established CBL calculation methods applied to the residential customers' data.

3.3 Stochasticity of different sectors

As mentioned earlier, CBL methods that worked successfully for large industrial and commercial customers fail to function satisfactorily for residential customers.

This observation suggests that there must be a fundamental difference in the nature of electricity consumption in customers of different sectors. Therefore, it is critical to study these underlying differences. As a matter of fact, without understanding the nature of these loads, it is impossible to determine the relationship between the strength of the methods and their error performance. Similar studies on the nature of loads have been performed for different purposes. For example, to determine how much Commercial and Industrial (C&I) customers react to dynamic pricing, the authors in [42] use the customers' business type and load pattern characteristics (e.g. Monday load ramp-up behavior, etc.). An extensive review of such attempts is provided in [19]. Almost all of these studies are performed in the time-domain.

In this section, the heterogeneity of customer loads of different sectors, i.e. industrial, commercial, and residential, are investigated in the frequency-domain. To carry out this, Discrete Fourier Transform (DFT) is employed to decompose customer loads into their underlying components. The DFT is a tool that transforms a time-domain signal into a frequency-domain signal. By using DFT, it is possible to sort loads based on the frequency of the repetition of its underlying components [57]. Each frequency in the resultant frequency-domain signal could be regarded as a proxy for certain activities. For example, low-frequency components represent activities that repeat on a daily, weekly, or monthly basis. In order to compare the loads, an index based on the underlying components of loads in frequency-domain is proposed and used. The index is called "predictability index" and can demonstrate the share of high-frequency and random parts of the signal.

3.3.1 Datasets

In this section, four datasets are introduced for the purpose of analysis, i.e. 1) Industrial, 2) Commercial, 3) Residential, and 4) Aggregated residential datasets. The aggregated residential category is made out of residential customers' data. The information of these datasets are provided in detail as follows.

Industrial customers

The industrial data are collected from 11 industrial customers in a multiple transmission zones in the PJM interconnection.

Commercial customers

For commercial customers, the simulated hourly load datasets are used. These datasets are published by OpenEI, which is a data center for energy information and data. The simulation is performed by EnergyPlus software, which is one of the products of the U.S. Department of Energy (DOE). In collaboration with three national laboratories, DOE developed some models for commercial reference buildings. These models are used as a benchmark for pertinent studies conducted by DOE about commercial buildings. These models play a significant role in providing complete descriptions of commercial buildings for the purpose of energy analysis. There are 16 building types that represent approximately 70% of the commercial buildings in the U.S. In this study, as mentioned earlier, the simulated hourly load datasets for these buildings are employed. These data are available for the entire United States.

The 16 building types are as follows: 1) large office, 2) medium office, 3) small office,

4) warehouse, 5) stand-alone retail, 6) strip mall, 7) primary school, 8) secondary school, 9) supermarket, 10) quick service restaurant, 11) full-service restaurant, 12) hospital, 13) outpatient health care, 14) small hotel, 15) large hotel, 16) midrise apartment.

Residential customers

For residential customers, a dataset collected by Australian Energy Market Operation (AEMO) for 200 residential customers has been employed [1]. The dataset is for the leap year of 2012 (366 days). In this dataset, the information for 11 customers was incomplete; therefore, they are excluded from the study, and the data for the remaining 189 customers are used in this dissertation.

3.3.2 Frequency response analysis

In this section, the elements required for the frequency-domain analysis are elaborated. The steps required for decomposing a signal into its underlying components are illustrated in Figure 3.

3.3.2.1 DFT introduction

The discrete Fourier transform is a tool to transform a finite sequence of data (N samples separated by the sampling time) into coefficients of sinusoids, which are ordered by the complex-valued function of frequencies. In other words, the DFT could be regarded as a frequency domain representative of the original time-domain input sequence. After decomposing a time-domain signal into its underlying components, then it is possible to search for any recurring pattern or periodicity in the original time-domain data. Moreover, DFT determines the magnitude of each periodic component,

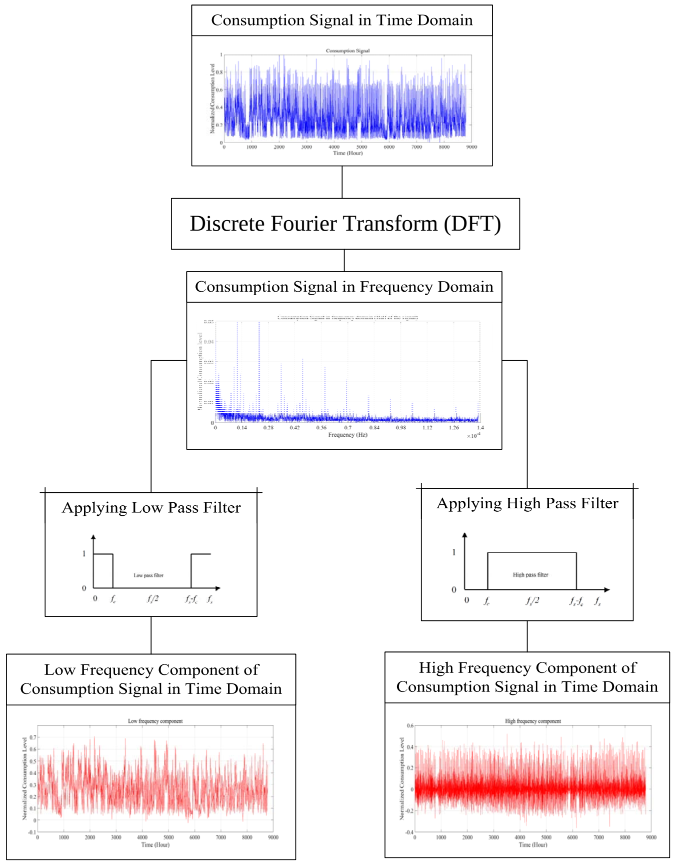

Figure 3: Implementation of DFT and separating time-domain high- and low-frequency components

thereby showing the relative strength of the components. In this section, hourly electricity consumption of each customer for the course of one year is treated as a signal, and by utilizing the DFT and two filters, they are separated into high- and low-frequency signals. The equations for DFT and inverse DFT are shown in (26) and (27).

$$X_k = \sum_{t=0}^{N-1} s[t].e^{\frac{-i2\pi kt}{N}}, k = 0, 1, ..., N-1 \tag{26}$$

$$X[t] = \frac{1}{N}\sum_{t=0}^{N-1} X_k.e^{\frac{+i2\pi kt}{N}}, t = 0, 1, ..., N-1 \tag{27}$$

where N is the number of samples in the signal. The outcome of DFT is another signal with N components in which each component has a different frequency, and these frequencies are listed in monotonically increasing order. An important note about the outcome of DFT is that the real parts in the outcome signal are mirrored over half of the data points. Therefore, only the information of half the signal is relevant; the other half is a repetition of the first. The previous sentence is mathematically expressed in (28).

$$X_k = X^*_{N-k} \quad 1 < k < N-1 \tag{28}$$

where operator (*) refers to the conjugation operator. In the DFT outcome signal, the frequency resolution can be calculated as (29).

$$f_r = \frac{1}{NT_s} \tag{29}$$

Where f_r refers to the frequency resolution and T_s refers to the time resolution in seconds. In this dissertation, the time resolution is 3600 seconds (1 hour) and the frequency resolution, given the 8784 sample points, is 31.6 nHz.

3.3.2.2 Filters

In this dissertation, for the purpose of decomposing the consumption signals into their underlying components, two filters of high and low pass frequency are created around a cut-off frequency (f_c). These two filters separate a frequency-domain signal into high and low-frequency components. The cut-off frequency is the frequency that is assumed to separate the predictable and unpredictable components of the consumption signal. Figure 4 illustrates the filters. f_c is the cut-off frequency, and f_s is the sampling frequency, which is 277.7 microHz. Since the frequency domain signal mirrors itself in a symmetrical way, low- and high-frequency filters must imitate this characteristic too. As a result, both filters look like bandpass filters.

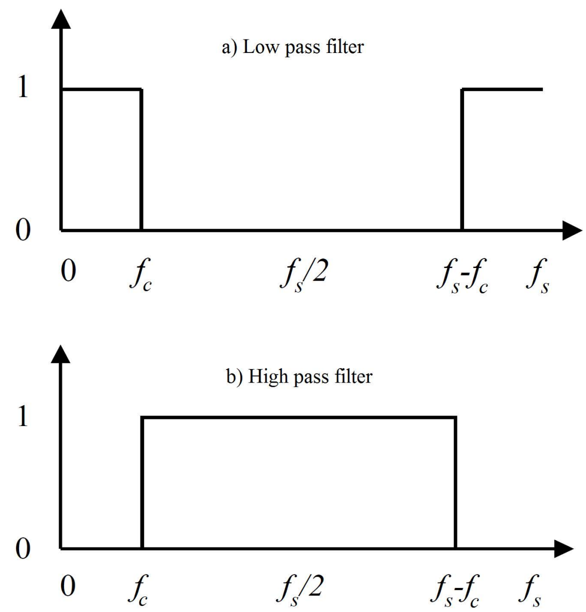

Figure 4: (a) Low pass filter (b) High pass filter

The cut-off frequency differs for different sectors; however, in this dissertation, two cut-off frequencies of 23.1 microHz, equivalent of 12 hours in the time domain, and 11.5 microHz, equivalent of 24 hours in the time domain, are selected. The rationale behind choosing 12 hours is that the frequency of almost all the spontaneous day-time activities is under 12 hours. On the other hand, since industrial activities are different from residential and commercial activities, 24 hours is selected to separate the unscheduled industrial activities that last under 24 hours.

3.3.2.3 Reconstruction

By applying the filters to the frequency-domain consumption signal, two high- and low-frequency signals could be obtained. Then, an inverse DFT could be applied to these frequency-domain signals to reconstruct their high and low-frequency counterparts in the time-domain.

3.3.2.4 Implementation and results

In order to show the frequency components of the consumption signals, one customer is randomly selected from each sector. After normalizing the data, the DFT is applied to them. Figure 5 shows the results. In this figure, green and red vertical lines represent two cut-off points of hours 12 and 24, respectively. According to this figure, it is understood that the residential customers have the highest amount of high-frequency components compared to the customers from other sectors. It is shown that load aggregation removed the high-frequency components of residential customers significantly. Moreover, it is shown that the industrial customers have a relatively small high-frequency contents compared to the other customers. In a case of commercial customers, a customer from the "medium office" building category is selected. In the following sections, all commercial buildings will be analyzed in detail.

In Figure 6, the same frequency-domain signals are illustrated. There are two main differences from figure 5: the horizontal axis is in logarithmic scale, and the results are separated for two cut-off points. The figures on the left side are for the cut-off point of hour 12, and the figures on the right side are for the cut-off point of hour 24. In this figure, the high and low-frequency points are shown with red and blue color,

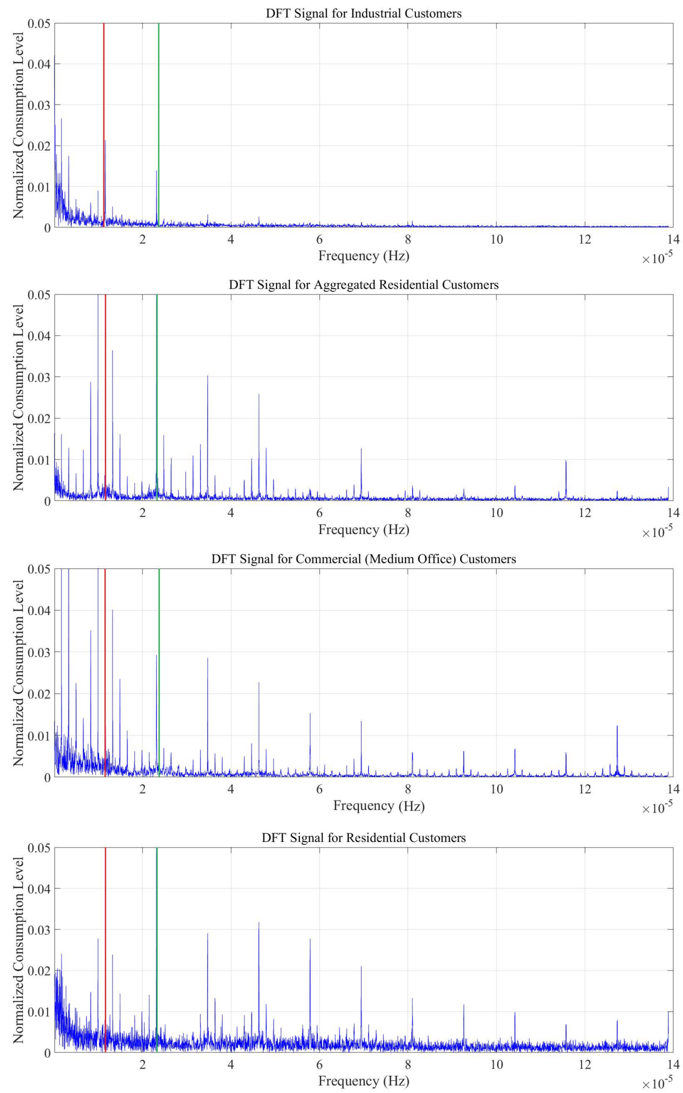

Figure 5: Frequency-domain signals of annual electricity consumption for customers of different sectors

respectively. All plots are on the same horizontal and vertical axises, allowing for easy visual comparison. Again, it is seen that the industrial and residential customers have the lowest and highest contents of high-frequency points.

The filters are applied to the frequency domain signals, and the low and high-frequency components for two cut-off points of hours 12 and 24 are reconstructed. In Figure 7, the low and high-frequency components of the industrial customer in

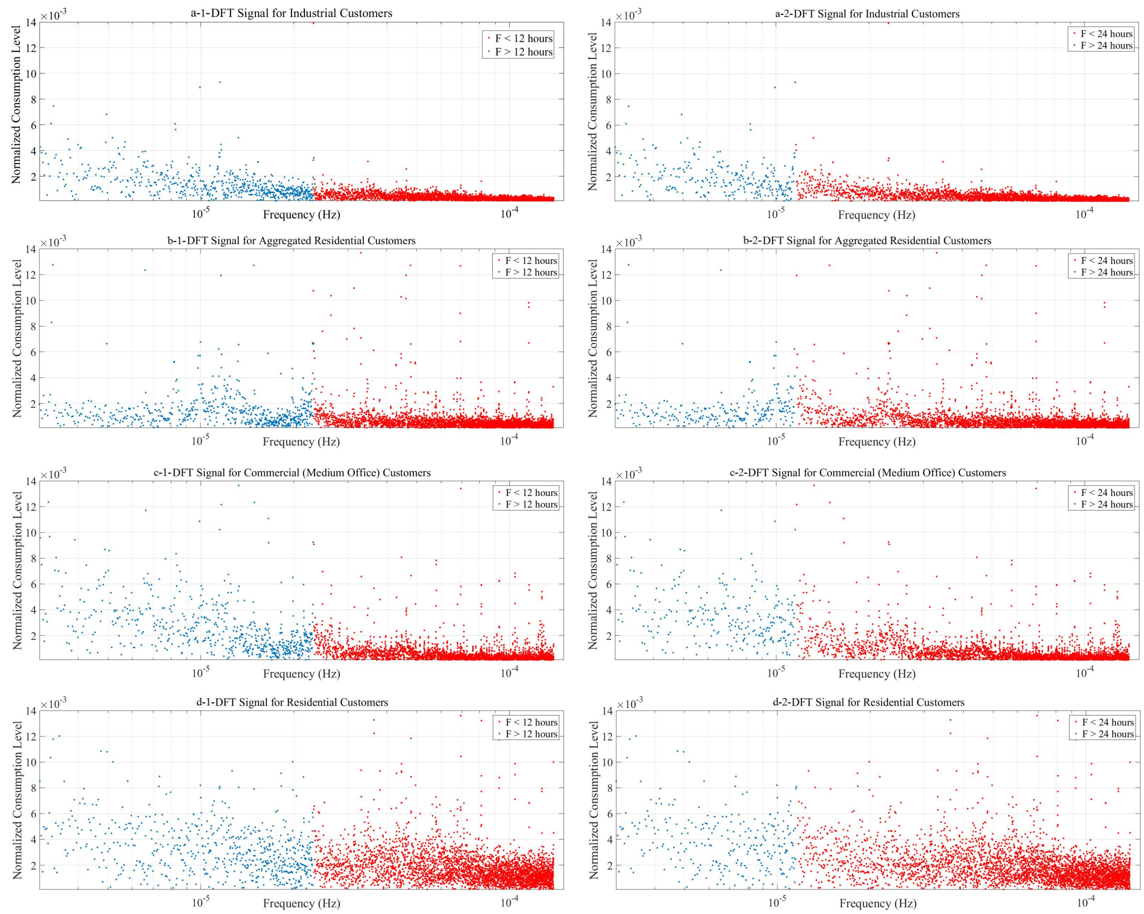

Figure 6: Frequency-domain signals of annual electricity consumption for customers of different sectors (Figures in left are for cut-off point of hour 12 and figures in right are for cut-off point of hour 24. The horizontal axis in each figure has logarithmic scale)

time-domain are illustrated. In this figure, the low and high-frequency signals are shown with blue and red colors, respectively. Industrial customers follow a schedule in their daily operation, and it is preferable not to deviate from the schedule. Since industrial schedules follow a certain pattern, analysis of these signals is expected to result in low-frequency components and few high-frequency components. As shown in this figure, the high-frequency signal is in fact very small and insignificant compared to the low-frequency signal. Another important observation in this figure is that the high-frequency components for these customers did not change for the cut-off hour of 24 compared to the cut-off hour of 12. This means that the amount of electricity consumed by activities with the frequency of 13-24 hours is almost zero. In other words, the spontaneous activities for this industrial customer have a frequency of fewer than 12 hours.

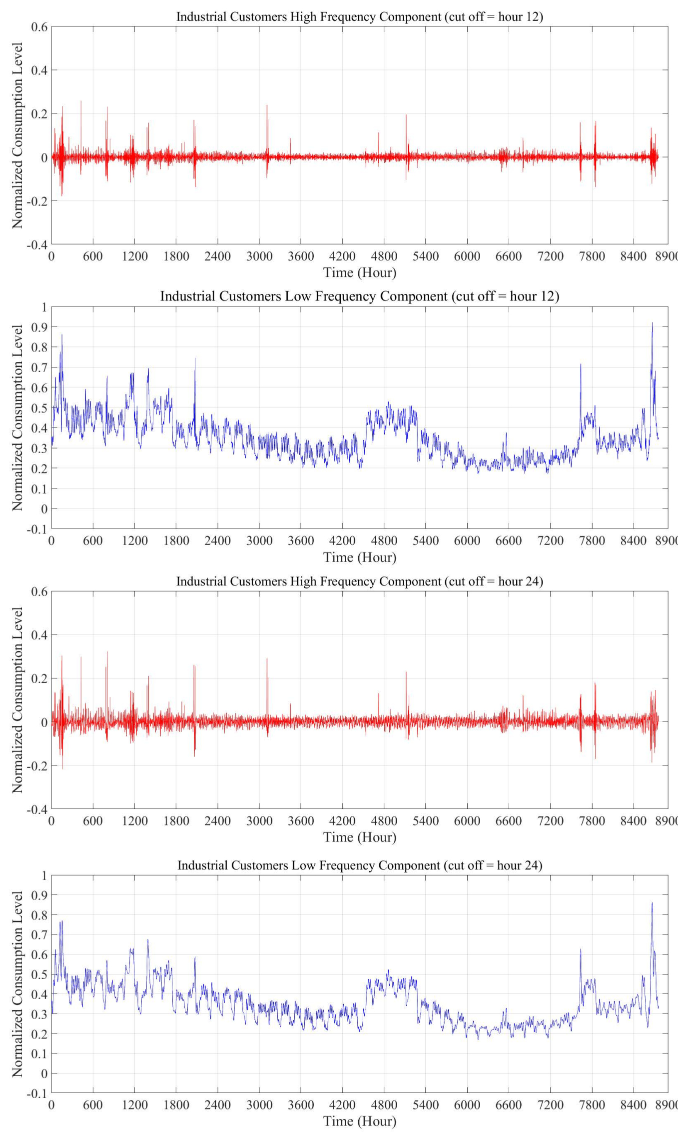

Figure 7: Time-domain electricity consumption signals for industrial customers

Figure 8 shows the time-domain signals of the aggregated residential customers. In comparison to industrial customers, this type of consumption signal has higher high-frequency components. Moreover, the low-frequency components show more fluctuation, which can be attributed to the nature of residential customers' electricity consumption. Due to the diversity and spontaneity of residential households' activities, the consumption signal has more fluctuations. Unlike for the industrial customers, the high-frequency components for aggregated residential customers almost double for the cut-off hour of 24 compared to the cut-off hour of 12. This means that the amount of electricity consumed by activities with the frequency of 1-12 hours is approximately equal to the amount of electricity consumed by activities with the frequency of 13-24 hours.

The time-domain signals of commercial customers are shown in figure 9. As is shown, the share of high-frequency components is much smaller than the low-frequency components. There are some hours in which the low-frequency components are close to zero. Those hours belong to weekends. In weekends, almost all medium office buildings are closed. With ignoring the weekend-related fluctuations, as expected, it is seen that commercial customers show less fluctuation than the aggregated residential. Commercial customers, similar to industrial customers, follow certain schedules; however, the schedule of commercials customers are not as strict as industrial customers. The strictness of schedules varies for different commercial buildings. Similar to aggregated residential customers, the high-frequency components for commercial customers become almost double for the cut-off hour of 24 compared to the cut-off hour of 12. Again, this means that the amount of electricity consumed by activities with the frequency of 1-12 hours is almost equal to the amount of electricity consumed by activities with the frequency of 13-24 hours.

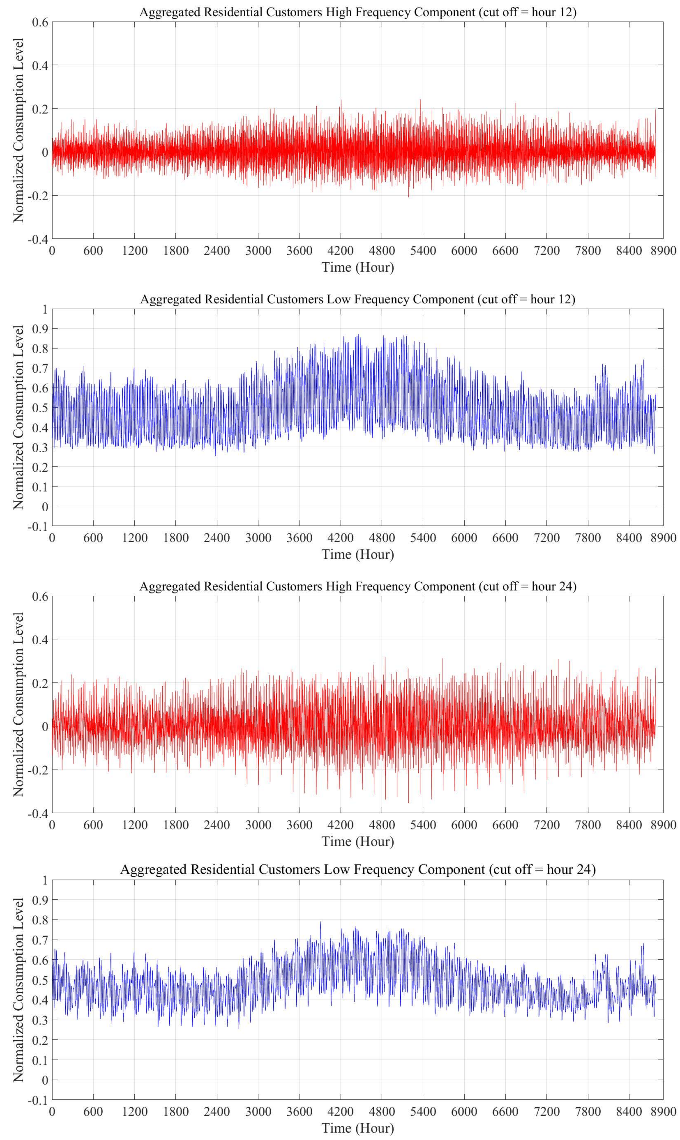

Figure 8: Time-domain electricity consumption signals for aggregated residential customers

Residential customers show different characteristics, as can be seen in figure 10. The share of high-frequency components is much higher than that of low-frequency components. As it is discussed earlier, residential customers, due to diversity and spontaneity of their household activities have highly fluctuating consumption signal. Besides, it can be understood from the signals that the weekends do not have a significant impact on the customers' routines. This observation is critical because, typically,

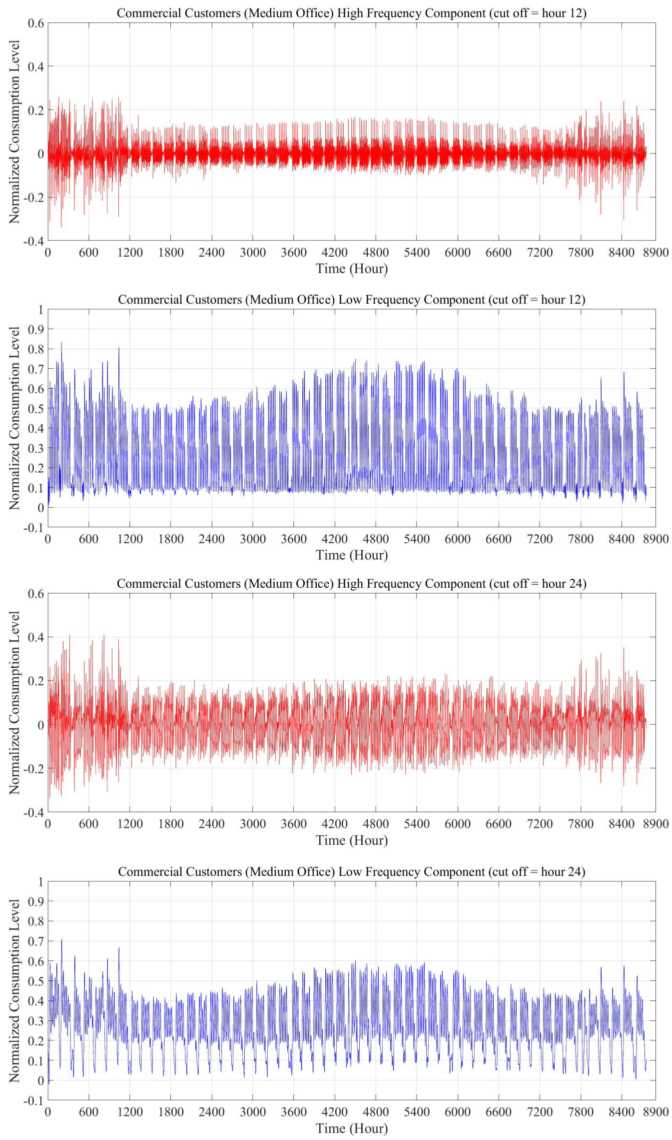

Figure 9: Time-domain electricity consumption signals for commercial customers

many load reduction calculation methods exclude weekends from their process. As is shown, the exclusion of weekends may not be necessary for residential customers. However, further investigation is necessary to examine how the inclusion or exclusion of weekends affect the CBL estimation.

For residential customers, the high-frequency components for the cut-off hour of 24 have a small increase compared to the cut-off hour of 12, which shows that most of the high-frequency activities of residential customers lie in the range of 1-12 hours. Moreover, comparing the low and high-frequency of the customer for the cut-off of hour 24, it is seen that the low-frequency component is very small compared to the high-frequency one. It can be concluded that most of the residential activities have frequency ranges of less than 24 hours. It is because residential customers do not follow a specific schedule for their daily activities, and their consumption is mostly

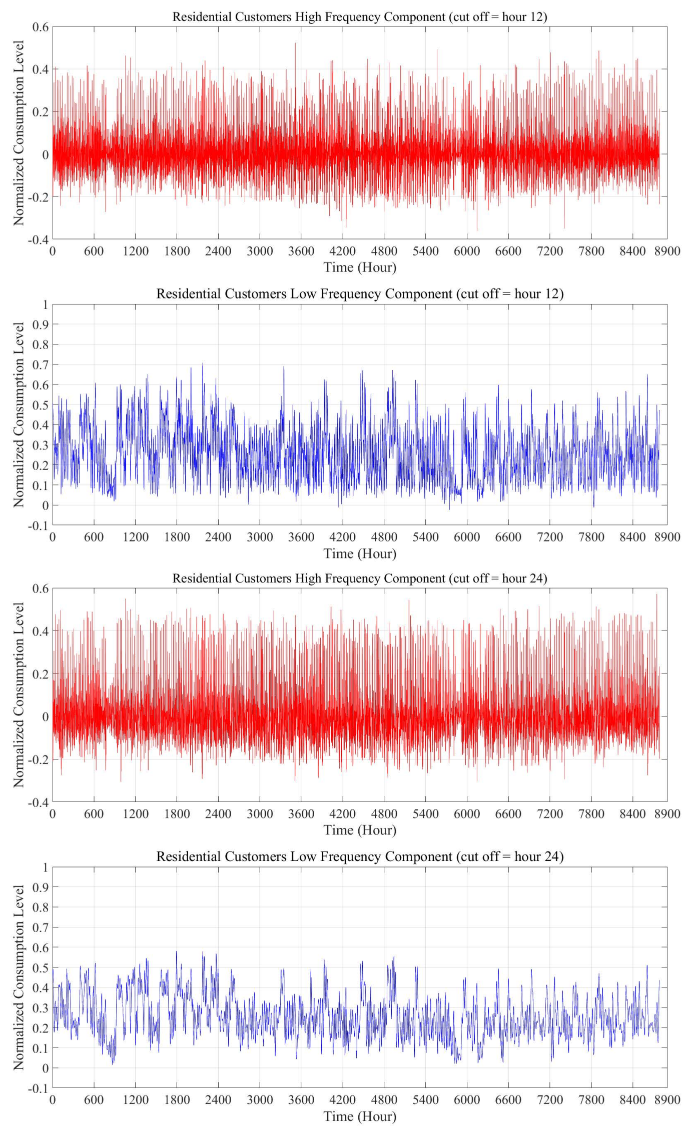

Figure 10: Time-domain electricity consumption signals for residential customers

governed by spontaneity.

3.3.3 Predictability analysis

In this section, first, an index is defined for evaluating the level of predictability. Then, this index is calculated for all the different types of customers, and the results are presented and discussed.

3.3.3.1 Predictability index

In order to determine how predictable a signal is, an index is defined as the predictability index. The rationale behind this index is that if the high-frequency portion of the signal is assumed to be random and hard to predict, then by subtracting that portion from the original signal, the rest of the signal is predictable. In other words, this index demonstrates the share of low-frequency components of the original signal. In order to carry this out, the index sums the share of high-frequency signal and then subtracts this value from one, as shown in (30).

$$P_{index} = 1 - \frac{\sum_{i=1}^{N} abs(c_i^{hf})}{\sum_{i=1}^{N} c_i} \tag{30}$$

3.3.3.2 Commercial buildings

The values of predictability index for commercial buildings are shown in Table 2. The results are for both cut-off hours of 12 and 24. The results are sorted from small value to large value, and the results for both cut-off hours follow almost the same order. As is shown, there is a huge gap between the results of the least predictable building (Strip mall) and the highest predictable building (Hospital). For cut-off hours of 24, the gap is 0.39 points. In other words, commercial buildings are not

Table 2: P_{index} based on annual data for different commercial buildings

	Cut off	
Building type	Hour 12	Hour 24
Strip mall	0.74	0.55
Stand-alone retail	0.76	0.56
Secondary school	0.78	0.61
Quick-service restaurant	0.83	0.73
Full-service restaurant	0.83	0.73
Primary school	0.83	0.68
Warehouse	0.84	0.70
Small office	0.85	0.71
Super market	0.85	0.69
Large hotel	0.85	0.74
Large office	0.86	0.71
Medium office	0.86	0.71
Midrise apartment	0.89	0.79
Small hotel	0.90	0.82
Outpatient health-care	0.92	0.84
Hospital	0.97	0.94

similar, and each one has its characteristics. According to the results, hospitals are almost close to perfect predictability. On the other hand, strip malls have the least predictability among these buildings.

The results of Table 2 are illustrated in figure 11. The difference between the results of cut-off hours of 12 and 24 is more clear in this figure. These results could be used by utilities in determining the effective EM&V methods for DR programs offered to customers of each building.

3.3.3.3 Comparison of different sectors

The predictability index for customers of different sectors is presented in Tables 3 and 4. Table 3 shows the predictability index of different sectors based on annual data, and Table 4 shows the similar results for monthly data. Comparing these two

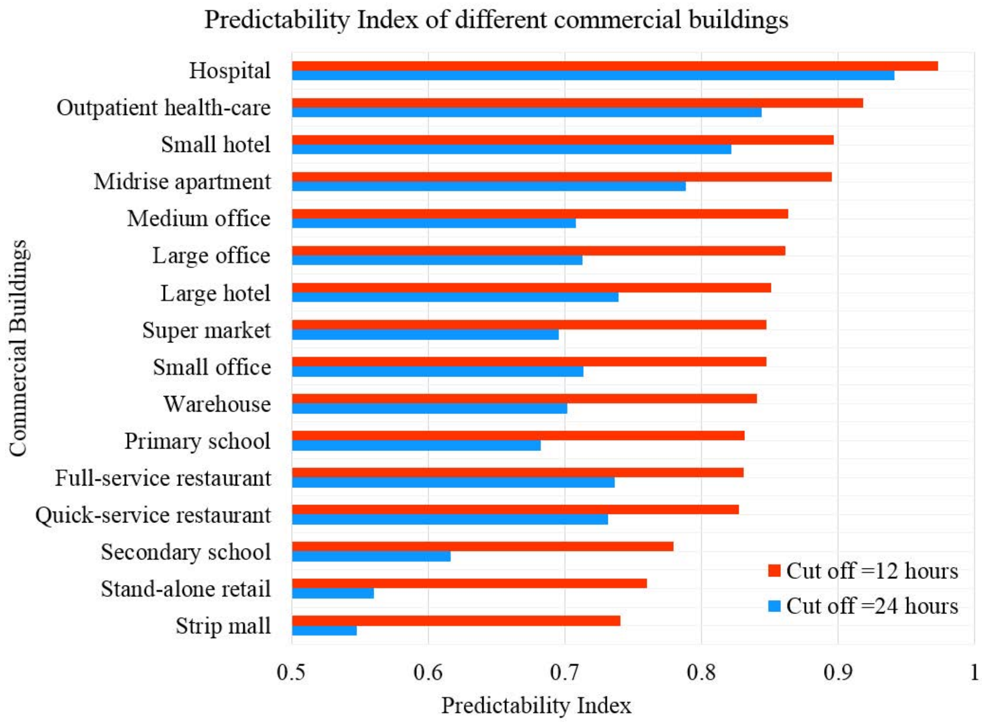

Figure 11: Predictability index (P index) based on annual data for different commercial buildings

tables, it is understood that the results are very close. In other words, only with having monthly data, it is possible to determine the predictability of each sector and group. The results of Table 4 are illustrated in Figure 12. As it is shown, there is a huge gap between the results of large industrial and commercial customers and residential customers. On the other hand, the aggregated residential customers show a lot of similarity with large industrial and commercial customers. According to these results, it is possible to claim that residential customers have a significant randomness element. Moreover, it is not comparable to the other sectors. As a matter of fact, since the large portion of the residential customers' load signal is random, it is impossible to estimate CBL accurately from historical data of the customers.

Table 3: P_{index} based on annual data for different sectors

	Customers' Sector			
Cut off	Industrial	Aggregated Residential	Commercial	Residential
Hour 12	0.95	0.90	0.84	0.44
Hour 24	0.90	0.85	0.72	0.36

Table 4: P_{index} based on monthly data for different sectors

	Customers' Sector			
Cut off	Industrial	Aggregated Residential	Commercial	Residential
Hour 12	0.95	0.90	0.85	0.44
Hour 24	0.90	0.85	0.72	0.36

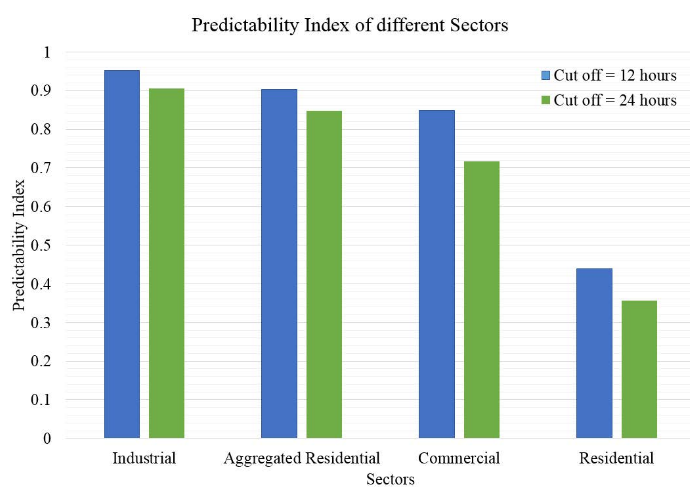

Figure 12: Predictability index values for different sectors

CHAPTER 4: PROPOSED METHODOLOGY

4.1 Overview

As discussed in the previous chapter, residential customers' household activities are highly diverse compared to large industrial and commercial customers as many random parameters influence their electricity consumption. Since randomness in load consumption is a primary source of error in CBL calculation, it could negatively affect the DR performance. This issue, was described in the previous chapter in detail. One way to improve the DR performance is to reduce such randomness. In this dissertation, load aggregation, i.e. grouping customers into different clusters, is proposed to eliminate the randomness.

The underlying mechanism of reducing randomness by aggregating customers into different clusters can be explained by the Central Limit Theorem. This theorem maintains that the aggregation of a large number of mutually independent random variables has a distribution function that can be well-approximated by a normal density function [60]. Residential loads are mutually independent and a large portion of them are random; therefore, it could be claimed that the resultant aggregated dataset exhibits more predictable traits and probabilities. Moreover, as shown earlier, the idea of clustering customers to eliminate their randomness is experimentally confirmed in multiple publications [18, 22, 30, 47].

In the dataset used in this dissertation, the customers' consumption is fairly random and mutually independent; therefore, the proposed approach can be applied to address the randomness problem. In other words, by grouping the customers into different clusters, it is possible to make the consumption loads more predictable. Moreover, it will be shown in this chapter that grouping customers can eliminate the incentive for gaming, which is another critical aspect in the argument for aggregating loads.

If there is no information available on the customers' behaviors, the customers should be grouped randomly; however, if there is available information, one proposed approach is to use k-means clustering for grouping customers [76]. In order to carry this idea out, the P_index, which was proposed and introduced in the previous chapter, will be used in the clustering process. In this chapter, it will be demonstrated that this index has a correlation with the CBL error performance. This correlation is especially important for two reasons: 1. P_index can be used as a means to show the upper limits of the error performance of the CBL calculation for each individual customer, 2. P_index can be utilized as a feature for grouping customers based on their error performance.

4.2 k-means clustering

k-means clustering has originated from signal processing and has been extensively employed for creating different clusters in data analytics. It partitions the raw data into k clusters. All observations are assigned to a specific cluster based on their proximity to the cluster mean. In fact, k-means clustering is a tool for finding similar groups in a dataset.

In order to run a k-means algorithm, a few initial points must be randomly assigned. These points are called cluster centroids. The number of these points are determined based on the preferred number of clusters. k-means is an iterative algorithm, and its two major functions are: 1) assigning a cluster, and 2) moving centroid to minimize an objective function.

Given a set of observations $(x_1, x_2, ..., x_n)$, where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into k $(\preceq n)$ clusters $\mathbf{S} = \{s_1, s_2, ..., s_k\}$ so as to minimize the within-cluster sum of squares (WCSS) (sum of distance functions of each point in the cluster to the K center). In other words, its objective is to solve the optimization relationship explained by the following equation (31).

$$\arg\min_{\mathbf{S}} \sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2 \tag{31}$$

where μ_i is the mean of points in S_i.

For the purpose of illustration, a simple k-means clustering is shown in Figure 13. In this figure, the dataset is grouped into 2 clusters, red and blue stars. For each cluster, the centroid is depicted with a crossed circle. Based on the algorithm of k-means clustering, after making a decision about the number of clusters, in this case, two clusters (k=2), these centroids move within the dataset to partition the dataset into 2 clusters in a way to minimize the equation (31).

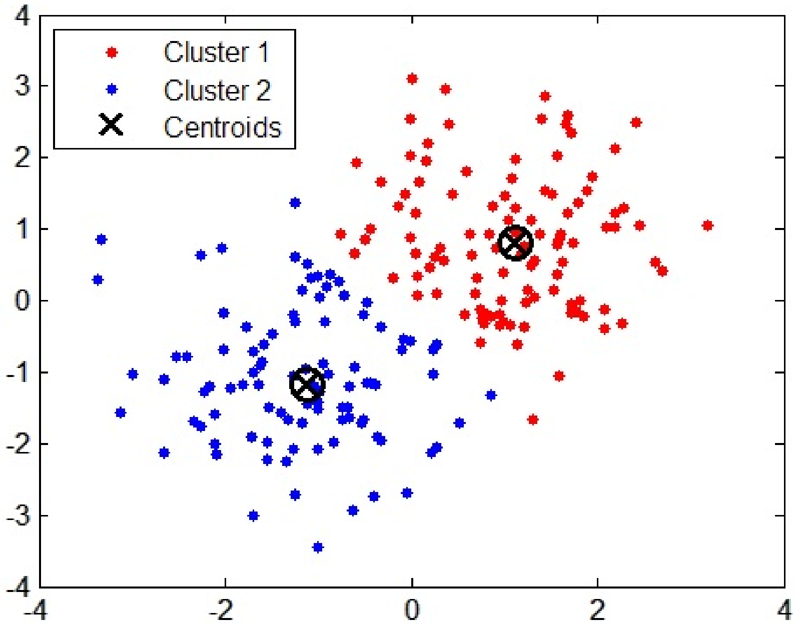

Figure 13: An example of k-means clustering

4.3 Correlation of predictability index and error performance

In this section, the customers are clustered based on 1) P_index values, and 2) event day average hourly consumption. In order to group the customers, the k-means algorithm is utilized. In this dissertation, by using "the elbow criterion" shown in Fig. 14, five cluster bins (k=5) are selected. However, it is worth mentioning that this "elbow" cannot always be unambiguously identified. For more information about k-means clustering, refer to [67]. There are many other ways to cluster the customers [12]; however, k-means clustering is proven to be very efficient. The results of the clustering are shown in Fig. 15. The customers of each bin are shown with different symbols. Customers in bin 1 to 5 are shown with purple stars, green circles, red squares, blue arrows, and black crosses, respectively.

The predictability index is calculated for all customers. Then, the customers are

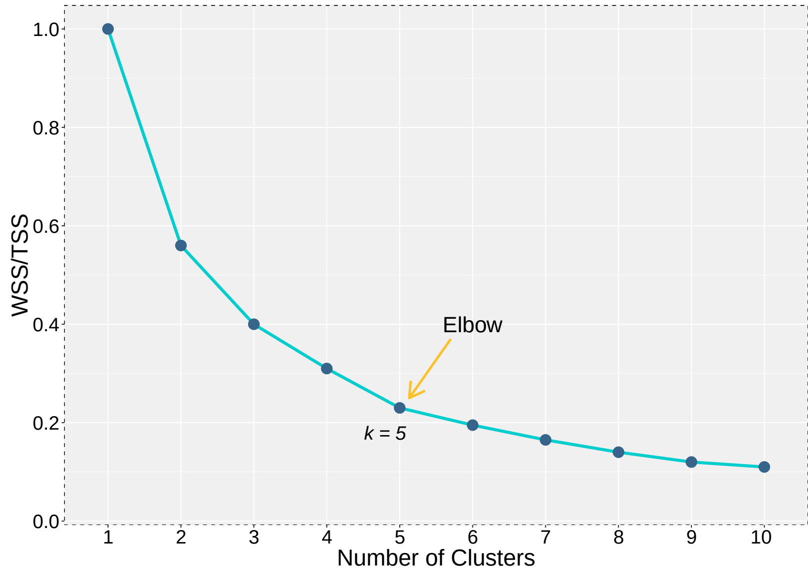

Figure 14: Elbow criterion analysis

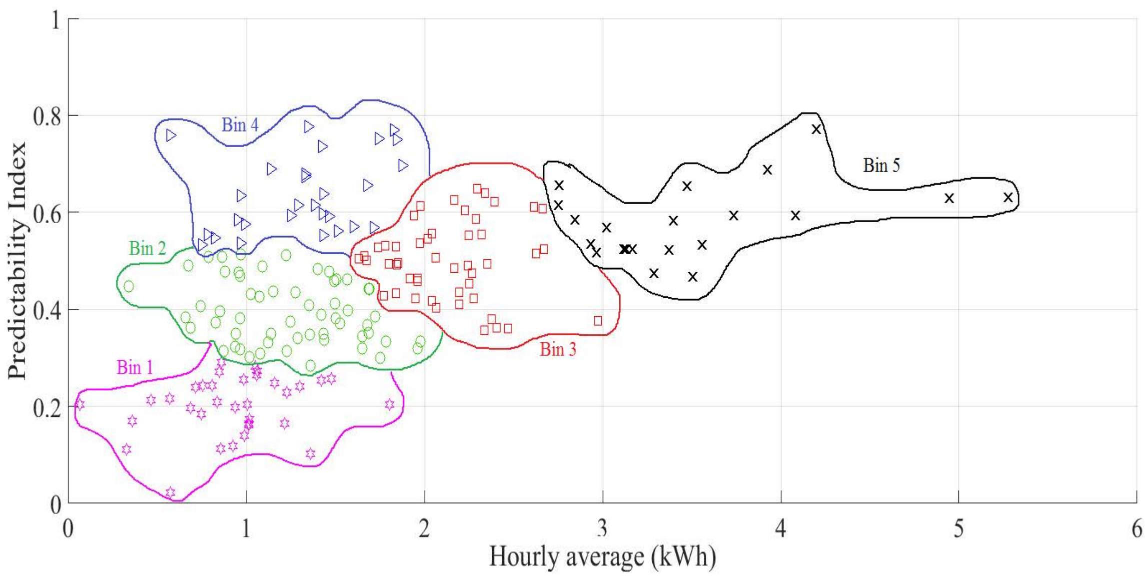

Figure 15: The relationship between average consumption level and predictability index (different clustering bins are coded with different colors)

clustered by k-means clustering into five bins. Later, the CBL for all customers in each bin is calculated, and the average Mean Absolute Error (MAE) for each bin is computed. The information of each bin including the number of customers, the average P_index, the average of CBL calculation's MAE, and the average hourly consumption are provided in columns 1-5 in Tables 5 and 6. Moreover, in order to show the relationship between P_index and accuracy, the error is normalized by the event-day average hourly consumption of each bin. As it is demonstrated in column 5, the average hourly consumption is different for the customers in each bin; hence, merely comparing the average MAE values of each bin (column 4) without considering the difference in the average hourly consumption would be misleading. As a result, for each bin, the value of MAE in column 4 is divided by the value of the event-day average hourly consumption in column 5, and the normalized value for each bin is calculated and enlisted in the 6^{th} column in the aforementioned Tables.

As shown in Table 5, the normalized MAE (last column) is decreasing as the average predictability index is increasing. In other words, a higher predictability index means that the CBL can deliver a better error performance. Moreover, it is shown in this Table that the correlation between P_index and normalized MAE (i.e. MAE divided by the event-day average hourly consumption) is -0.98. This value indicates that there is a strong correlation between P_index value and the MAE of the CAISO's CBL calculation method.

The same analysis is performed for the RCT method. a a relatively high correlation of -0.88 is found between P_index and normalized accuracy MAE in this case.

Almost all studies on the EM&V of CBL calculation methods rely upon two metrics of accuracy and bias to evaluate different CBL calculation performance. However, the strong correlation between P_index and the metric of accuracy suggests that P_index could be utilized as an alternative or complementary metric. Furthermore, P_index can be used as a feature to demonstrate the limitation of the CBL calculation

Table 5: The accuracy results of the CAISO CBL calculation method

Bin#	#Cust.	Average P_index	Accuracy MAE	Event-day Average Load	MAE/Normalized Accuracy
Bin 1	35	0.19	0.53	0.80	0.66
Bin 2	57	0.39	0.61	1.16	0.52
Bin 3	48	0.50	0.91	1.94	0.47
Bin 4	28	0.64	0.64	1.51	0.42
Bin 5	21	0.58	1.30	3.26	0.40
Correlation between P_index and MAE/Average					-0.98

Table 6: The accuracy results of the RCT CBL calculation method

Bin#	#Cust.	Average P_index	Accuracy MAE	Event-day Average Load	MAE/Normalized Accuracy
Bin 1	35	0.19	0.75	0.80	0.94
Bin 2	57	0.39	0.87	1.16	0.74
Bin 3	48	0.50	1.16	1.94	0.60
Bin 4	28	0.64	1.00	1.51	0.66
Bin 5	21	0.58	2.18	3.26	0.67
Correlation between P_index and MAE/Average					-0.88

methods. If P_index of a customer is low, it would suggest that no CBL calculation method can work properly to estimate the CBL. On the other hand, if the P_index is high and the CBL method does not deliver a satisfactory performance, it is highly likely that the employed CBL method is a problem. Hence, change of CBL calculation method may prove to be an effective way to improve the error performance.

4.4 Flowchart of the proposed CBL estimation method

In what follows, the structure of the proposed method will be delineated. The technical flow chart of the proposed CBL estimation method is provided in Figure 16. The flow chart starts with the data. The data refers to time series of load consumption of customers, which is collected by residential smart meters (with time increments equal to or less than an hour). Data cleansing (eliminating the incomplete data, etc) is the first step of the process. After detaining an acceptable data set, the data for the target year is extracted, and the signal processing section of the flowchart begins. In order to create a frequency-domain signal, Discrete Fourier Transform (DFT), as described in chapter 3, is applied to the time-domain signal. Two filters of low- and high-pass are then applied to the signal to extract low- and high-frequency components of the signal. Based on the contents of the signal low- and high-frequency components, the predictability index (P_index) is calculated. The process and the formula for calculating predictability index were discussed in chapter 3. After collecting all indexes, data analytics is applied. P_index and average hourly consumption load are used to group customers by k-means clustering as a data analytics tool. After clustering is performed, members of each group are used to create a common CBL for the group. This proposed CBL estimation method will be implemented in the next chapter for evaluation and demonstration purposes.

4.5 Social welfare analysis of the proposed clustering method

One way to reduce the non-event day deadweight loss is to reduce baseline inflation. One method introduced by [54] decouples the customers' individual consumption from CBL. It proposes that similar customers be grouped together by employing matching techniques to establish a common CBL. Utilizing this approach to calculate a common CBL for all individuals in the group decouples the customers' individual activities from the calculated common CBL. Likewise, the proposed CBL method in this dissertation establishes a common CBL. In what follows, it is explained how this

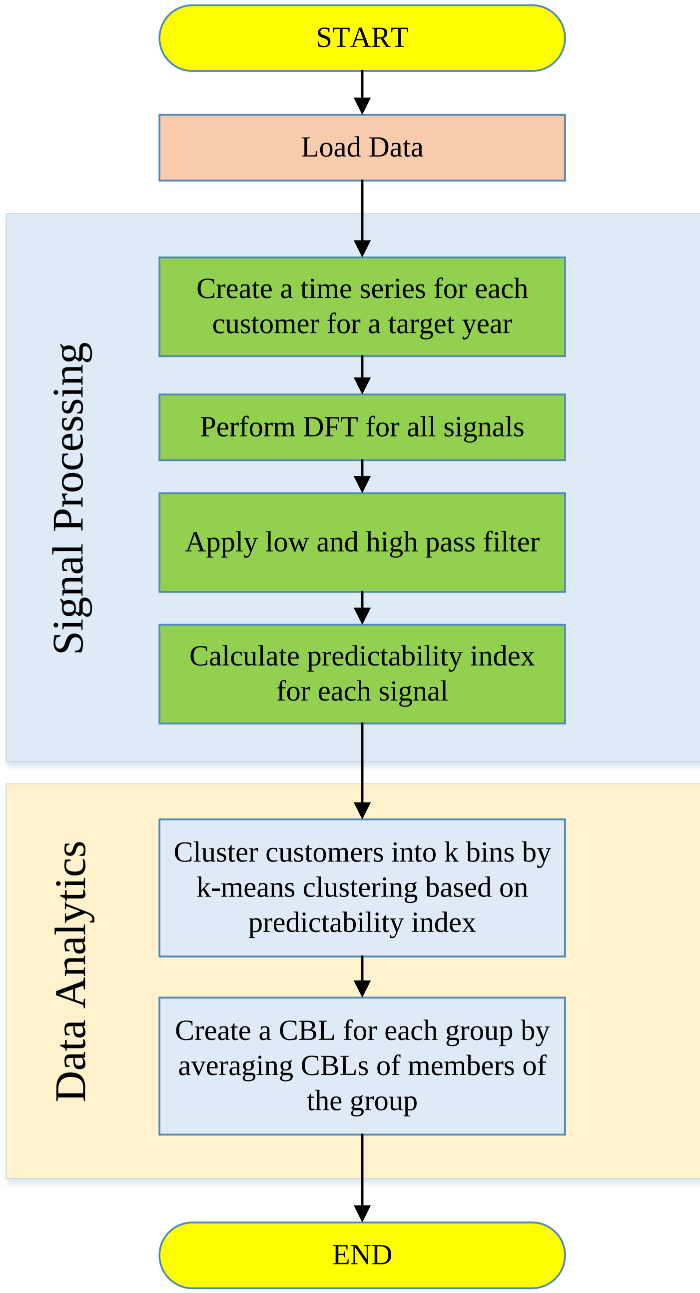

Figure 16: Flowchart of the proposed CBL estimation method

method can remove the economic incentive for inflating the baseline. In addition to proposing a baseline estimation method, a pricing structure is recommended for DR programs. In order to eliminate the deadweight loss associated with the event period, the wholesale rate minus the retail rate is used as a reward. This rate is recommended by almost all the opponents of FERC 745 [16].

$$p^{\pi} = p^{w}_{te} - p^{r} \tag{32}$$

In what follows, the impacts of employing the proposed method on CBL calculation and new event period rate are investigated.

The calculation of a common CBL for a cluster of customers is as follows:

$$c^{BL}_{t} = \frac{1}{N}\sum_{i=1}^{N} c^{BL}_{t,i} \tag{33}$$

where:

$c^{BL}_{t,i}$ CBL for individual i before clustering

c^{BL}_{t} Common CBL of the cluster after clustering

N Number of individuals in the cluster

It is assumed that the number of individuals in each cluster is large enough such that the impact of each individual's CBL on the common CBL is negligible.

With the aforementioned modifications in the FERC 745 policy, the problem of maximization of the expected utility of the customer can be expressed as:

$$\underset{c^{e}_{t},c^{n}_{t}}{Max}(1-\lambda_t)[U_t(c^{n}_{t}) - p^{r}.c^{n}_{t}] + \lambda_t[U_t(c^{e}_{t}) - p^{r}.c^{e}_{t} + p^{\pi}(c^{BL}_{t} - c^{e}_{t})] \tag{34}$$

FOCs for c^{n}_{t} and c^{e}_{t} for (34) are:

$$U_t{}'(c^{e}_{t}) = p^{w}_{te} \tag{35}$$

$$U_t{}'(c^{n}_{t}) = p^{r} \tag{36}$$

Now the marginal utility prices during the event and non-event periods are equal to the wholesale and retail rates, respectively. With these marginal prices, the customers have no incentive to under-consume during event periods or to inflate the baseline during non-event periods. As a result, based on this theoretical analysis, it is understood that the proposed CBL estimation method is capable of eliminating the economic incentive for gaming, which can be regarded as another benefit of the proposed method. In the next chapter of this dissertation, it will be shown that the proposed method is capable of considerably improving the error performance .

CHAPTER 5: IMPLEMENTATION AND RESULTS

5.1 Overview

In this chapter, the CBL estimation methods and metrics used for error analysis are explained. First, the established CBL calculation methods, popular in the industry, are introduced. Then, the metrics necessary for the CBL calculation error analysis are presented. The dataset used in this dissertation is described afterward. Finally, the proposed approach is applied to the dataset, and the results are provided.

5.2 CBL calculation methods

In this section, popular methods for CBL calculation, i.e. HighXofY, Exponential Moving Average, Regression, Randomized Controlled Trial (RCT) are explained in detail. These methods will be used in this dissertation for comparison purposes and to evaluate the performance of the proposed method. As discussed earlier, CBL is the amount of load that is estimated to be consumed by customers in the absence of a DR curtailment signal.

5.2.1 HighXofY method

The HighXofY method involves several steps. First, it selects Y non-DR days. In the absence of a DR event, the days are called non-DR days. Also, weekends are excluded from these non-DR days. Multiple types of day are used in this dissertation, weekdays (Monday to Friday), weekends (Saturday and Sunday), and holidays. Second, X days are chosen from the aforesaid Y days based on the level of consumption. Finally, the baseline is defined as the average load of these X days. If *HighXofY* is defined as $High\left(X,Y,d\right) \subseteq D\left(Y,d\right)$, then the *HighXofY* baseline of customer $i \subseteq C$ for timeslot $t \subseteq T$ on day d is as follows:

$$b_i\left(d,t\right) = \frac{1}{X} \times \sum_{d \in High(X,Y,d)} l_i\left(d,t\right) \tag{37}$$

where:

C is a set of customers;

$T = \left\{t_0, ..., t_{|T|}\right\}$ is timeslot division within a day;

$l_i\left(d,t\right)$ is actual load for customer $i \subseteq C$ on day d at timeslot $t \subseteq T$;

$b_i\left(d,t\right)$ is a predicted baseline for customer $i \subseteq C$ on day d at timeslot $t \subseteq T$;

$D\left(Y,d\right)$ is a set of Y non-DR days preceding the day d having the same day type as d;

$l_i\left(d\right) = \sum_{t \in T} l_i\left(d,t\right)$ is total load for customer $i \subseteq C$ on day d.

New York ISO uses this method with X=5 and Y=10 which is employed in this dissertation as well. The algorithm of NYISO is adequately described in [32]. Moreover, another method from this family is the California ISO (CAISO) method that employs X=10 and Y=10. In the literature, the CAISO method is also known as Last10days. Nevertheless, it is considered as a member of the HighXofY family. In this work, the CAISO is used with one modification. In this study, the weekends, also, are included in the process of the CBL calculation. Another method used in this dissertation is the PJM method with X=4 and Y=5 [49, 48].

5.2.2 Exponential moving average method

This method is a weighted average of customers' historical data from the beginning of their subscription. This method begins with computing an initial average load of the customer. Then, it continues with calculating an exponential moving average using the initial average load. The baseline for the customer is achieved at the end.

Let $D(\infty, d) = \{d_1, ..., d_k\}$; also, $1 \leq \tau \leq k$ be a constant. This constant is the number of days used to determine $s_i(d_\tau, t)$ which is the initial average load for customer $i \subseteq C$ for timeslot $t \subseteq T$.

$$s_i(d_\tau, t) = \frac{1}{\tau}\sum_{j=1}^{\tau} l_i(d_j, t) \tag{38}$$

The exponential moving average for $\tau \leq j \leq k$ is

$$s_i(d_j, t) = (\lambda.s_i(d_{j-1}, t)) + ((1-\lambda).l_i(d_j, t)) \tag{39}$$

where $\lambda \in [0, 1]$. It is understood that the weight of each day decreases exponentially with time.

Finally, the baseline for customer $i \subseteq C$ on day d for timeslot $t \subseteq T$ could be calculated as follows:

$$b_i(d, t) = s_i(d_k, t) \tag{40}$$

In this method, the baseline for days earlier than $d_{\tau+1}$ is undefined. Indeed, if a DR event happens during this short interval, the CBL for this customer cannot be calculated by this method. DR programs using this method do not include these customers in the program unless they access enough days to build the initial average load.

New England ISO (ISONE) employs this methodology, which is used in this dissertation as well. The algorithm of ISONE is as follows. If a customer is a new participant in a DR program, the calculation of the baseline is the hourly average of the previous five business days, Monday through Friday, excluding holidays and other event days. This average is known as "Customer Baseline 6". The "6" refers to the day following the previous five business days or the sixth day. The equation for this baseline calculation is defined by (41)

$$\Phi_6 = \frac{\sum_{i=1}^{5} kWh_{i,h}}{5} \tag{41}$$

Once the "Customer Baseline 6" is calculated, that new customer is now considered as a current customer and her next baseline could be calculated using (39) with $\lambda = 0.9$.

This baseline is calculated every day except for weekends, holidays and event days. Furthermore, the baseline has a weighting factor of 90% for the previous day's CBL and a weighting factor of 10% for the current day consumption [32].

5.2.3 Regression method

This method uses multiple linear regression to calculate the baseline. According to this method, the baseline for customer $i \subseteq C$ on day d for timeslot $t \subseteq T$ is

$$b_i(d,t) = (\theta_{i,t})^T x_{i,t} + \varepsilon_{i,t} \tag{42}$$

where:

$x_{i,t}$ is the feature vector;

$\theta_{i,t}$ is the vector of regression coefficients;

$\varepsilon_{i,t}$ is the error term.

The feature vector could consist of explanatory variables including historical load, temperature, humidity or sunrise/sunset time. In this dissertation, this vector is comprised of the consumption data of different days of the week. The regression method algorithm is fully described in [7].

5.2.4 Randomized Controlled Trial (RCT)

RCTs are very popular and trustworthy as evaluation methods to the extent that, as mentioned earlier, many regard them as the "gold standard" of evaluation methods. RCTs eliminate the selection effects by randomly assigning the households into two groups of treatment and control. Since control and treatment groups are exposed to similar conditions, other possible alternative explanations will be eliminated, and the difference between these two groups could be attributed solely to the treatment. This way, the internal validity would be assured by the construction of the design [24].

The process starts by random assignment of households of the dataset into two groups. One group would serve as a basis for the calculation of CBL for the other group. It is essential that the customers do not exert any control over the assignment process (i.e. conscription) to ensure the internal and external validity of the results [14].

RCTs could be enhanced by some procedures such as using extra information about customers; however, the enhancement is prone to introduction of some factors affecting the performance of such forms of RCTs adversely. RCTs rely on minimal assumptions about the nature of customers; therefore, they can produce unbiased estimates of treatment effects. On the other hand, if these methods are enhanced with matching techniques like propensity score matching, nearest neighbor matching, etc., which require relying on some strong assumptions about the nature of customers. If these assumptions are violated, RCT could produce biased results [14].

The RCT method has lower administrative cost compared to the other methods as it requires no historical data for CBL calculation. Therefore, under equal conditions, RCT is a much better alternative both in terms of lower cost and complexity.

5.2.5 Adjustment of CBL estimation methods

As explained in *HighXofY* method, a subset of X days is selected to collect the days similar to the event day. Nevertheless, the conditions on the event day are often different from the selected prior days. Therefore, X of Y baseline methods could be adjusted by the event day data. In fact, according to North American Energy Standards Board (NAESB) recommendation for large customers, an adjustment to a high X of Y baseline is necessary to reflect load conditions of the event day more accurately [5]. Figure 17 illustrates the concept of baseline adjustment. In this figure, the actual load is the red line. As shown in this figure, the demand decreases on event hours because of the curtailment call. However, on regular hours (non-event hours), the demand is not affected. The green line is the initial baseline estimated by a CBL estimation method. However, as it is illustrated, there is a discrepancy between the actual load and initial baseline on non-event hours. This discrepancy can be adjusted to create the black line, i.e. adjusted baseline.

The adjustment is defined by the time frame. This time frame is normally 2-

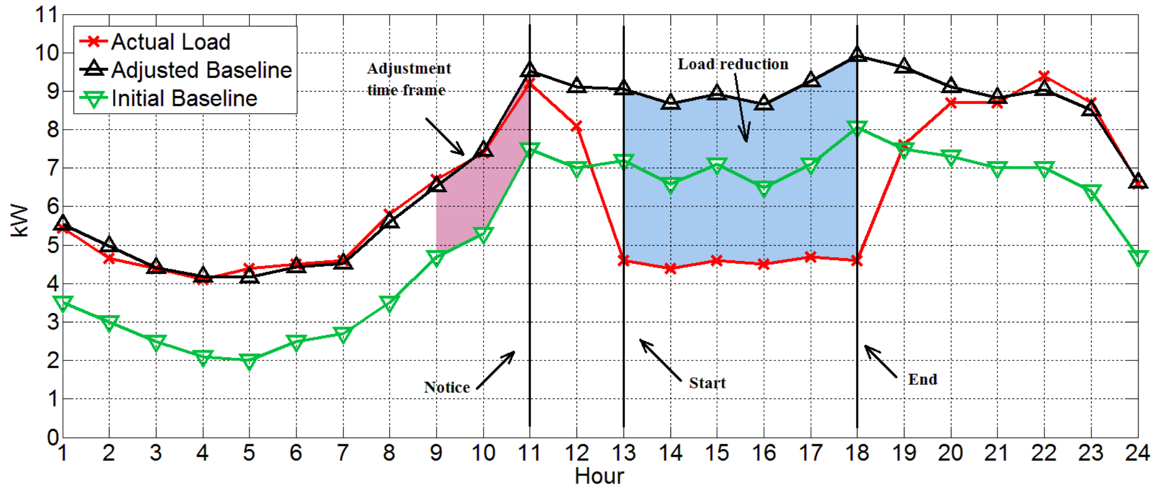

Figure 17: Example baseline adjustment

4 hours before the start of the event. The difference between actual load and the estimated baseline in the adjustment time frame can be employed for adjustment purposes in two ways, multiplicative or additive. Multiplicative adjustment uses the percentage change and applies it to the estimated baseline. Additive adjustment utilizes the absolute change. It is reported that the choice of multiplicative or additive adjustment does not change the outcome considerably [31]. In this dissertation, the additive adjustment is used for adjusting CBL methods. This adjustment can be used for all CBL estimation methods.

5.3 Error performance analysis

In this section, three error metrics of accuracy, bias, and Overall Performance Index (OPI), which are utilized for the purpose of the error analysis, are introduced and elaborated.

5.3.1 Accuracy

The hourly accuracy represents the hourly difference between the estimation and the real consumption. Let C be the set of all 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) for measuring baseline accuracy is defined as shown in (43), and is shown with α symbol. As shown, the lower the MAE, the higher the accuracy.

$$\alpha = \frac{\sum_{i \in C} \sum_{d \in D} \sum_{t \in T} |b_i(d,t) - l_i(d,t)|}{|C| \, . \, |D| \, . \, |T|} \tag{43}$$

5.3.2 Bias

Baseline bias is defined as shown in (44), and is shown with β symbol. The definition of bias is close to accuracy; however, it gives different information about the performance of CBL.

$$\beta = \frac{\sum_{i \in C} \sum_{d \in D} \sum_{t \in T} (b_i(d,t) - l_i(d,t))}{|C| \, . \, |D| \, . \, |T|} \tag{44}$$

The difference between accuracy and bias, as expressed in (43) and (44) is the value of the difference between CBL and the actual consumption, where MAE uses the absolute value of the difference, while bias uses the real value. According to (44), baseline methods with positive bias overestimate the customers' actual consumption and vice versa.

5.3.3 Overall Performance Index (OPI)

The overall error performance of a method depends on both accuracy and bias. Therefore, in this dissertation, another metrics is defined for measuring the overall performance. It is defined as the weighted sum of the absolute value of accuracy and bias, and it is called Overall Performance Index (OPI) as shown in equation (45).

$$OPI = \lambda\,|\alpha| + (1-\lambda)\,|\beta| \tag{45}$$

A lower OPI means that the CBL is more capable of measuring the customers response to the price incentives in the pertinent DR program.

In this dissertation, the absolute value of accuracy and bias have the same weight (λ= 0.5) as they are equally important for error analysis. However, for some industries, the accuracy value would have more weight than bias as loads are more predictable and bias is more manageable than accuracy.

5.4 Dataset

This section provides a description of the dataset that is utilized to perform error analysis. The data in this dataset is collected by the Australian Energy Market Operation (AEMO) for 199 residential customers, in the leap year of 2012 (366 days). Each electricity distributor in the AEMO market supplies raw data for a sample of 200 customers in each of their supply areas to the market operator to construct load profiles [1]. The customers under study are charged based on a fixed tariff. The data used in this study are broken down into four seasons. Seasons in Australia are as follows:

- Spring: the three transition months September, October, and November.

- Summer: the three hottest months December, January, and February.

- Fall: the transition months March, April, and May.

- Winter: the three coldest months June, July, and August.

In this dissertation, 12 event days (one for each month) are selected for the error analysis, and the information about these days is as follows:

1. Event day = Sept. 8th (252nd day)

2. Event day = Oct. 13th (287th day)

3. Event day = Nov. 25th (330th day)

4. Event day = Dec. 19th (354th day)

5. Event day = Jan. 30th (30th day)

6. Event day = Feb. 22nd (53rd day)

7. Event day = Mar. 18th (78th day)

8. Event day = Apr. 28th (119th day)

9. Event day = May 25th (147th day)

10. Event day = Jun. 23rd (175th day)

11. Event day = Jul. 28th (210th day)

12. Event day = Aug. 11th (224th day)

The total consumption for the four seasons is illustrated in Figure 18. The event is assumed to start from 3:00 p.m. and end at 9:00 p.m.

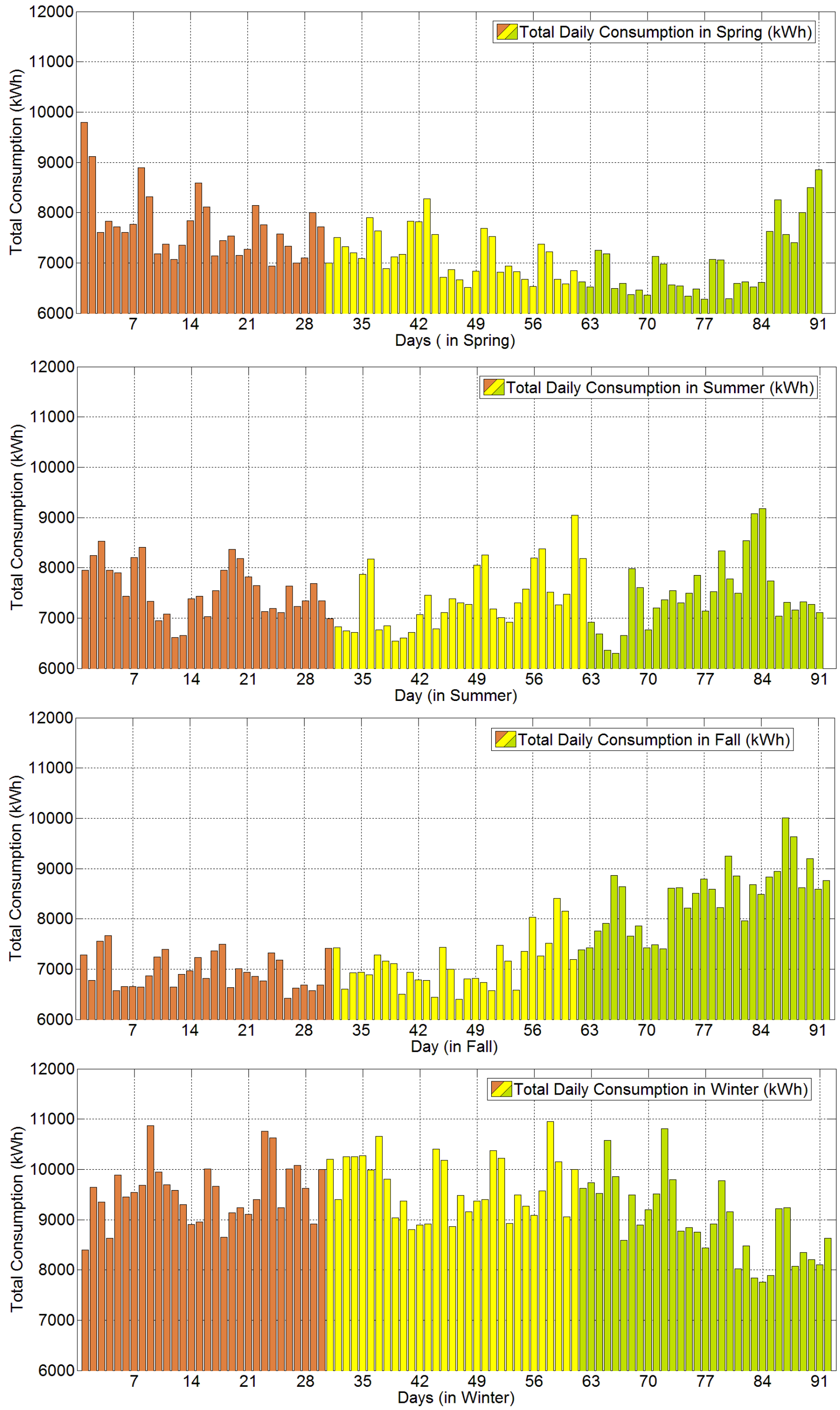

Figure 18: Total daily consumption for the 199 customers in spring, summer, fall, and winter of 2012 (orange for Sept., yellow for Oct., light green for Nov.) (orange for Dec., yellow for Jan., light green for Feb.) (orange for Mar., yellow for Apr., light green for May.) (orange for Jun., yellow for Jul., light green for Aug.)

The scatter and box plots of all customers are illustrated in Figures 19 and 20, respectively. The data is divided into four seasons to see the impact of seasonality. It is shown that different seasons have different load patterns. Nevertheless, the overall patterns are somewhat similar. It is worth mentioning that this figure is obtained from the average of loads. Each individual load might have a completely different shape. However, it is possible to claim that most of the time, individual residential customers follow similar patterns. Furthermore, as shown, at noon, the loads show the highest variability.

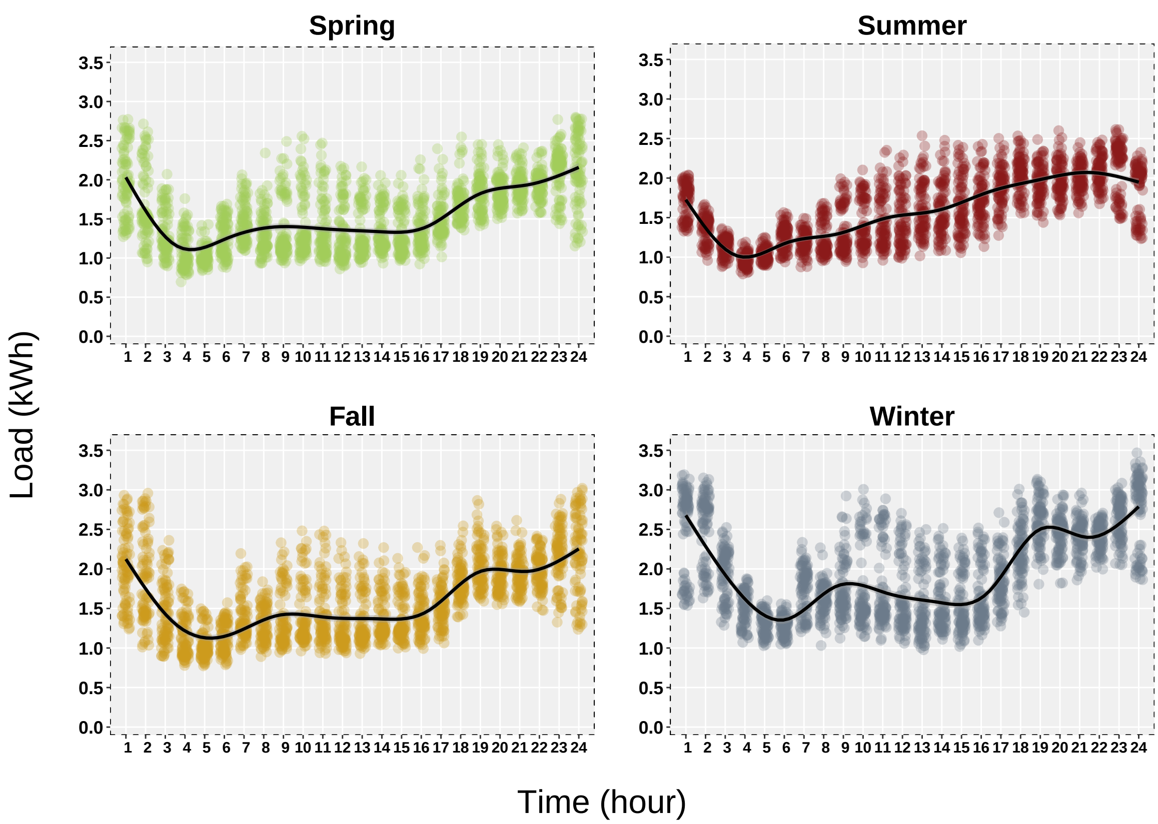

Figure 19: Scatter plot of loads of all customers for 24 hours

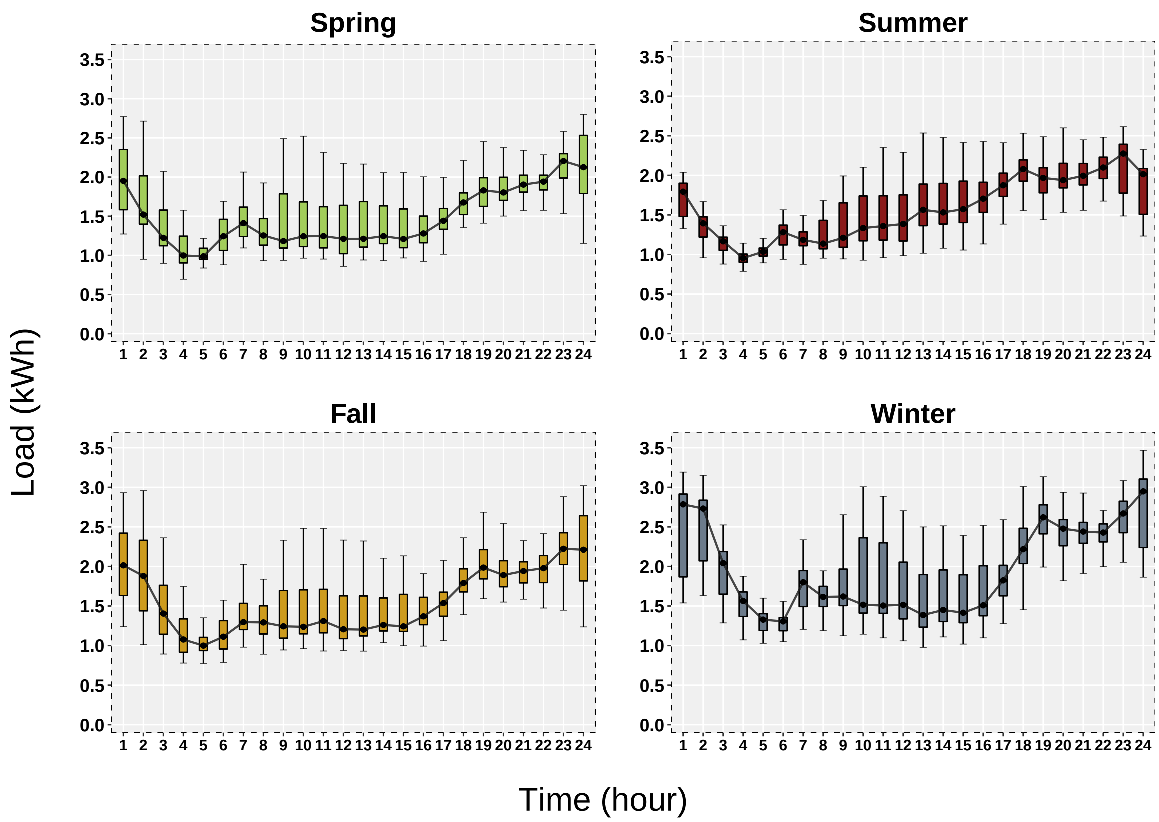

Figure 20: Box plot of loads of all customers for 24 hours

5.5 Results

In this section, the results of the proposed CBL estimation method are provided. Before presenting the final results, setup and steps taken in the analysis are explained in detail.

5.5.1 Setup

In order to examine how the load aggregation can improve the error performance of CBL estimations, a simple and random load aggregation is applied to the dataset. First, the clustering has been done randomly. Then, the clustering will be conducted by using a k-means algorithm based on two features of average consumption and predictability index. For carrying out the analysis for the former, the consumption data for 91 consecutive days of 199 customers are employed. The 91 days are from Oct. 1st, 2012 to Dec. 31st, 2012. In this period, the consumers pay a fixed tariff. Since the maximum consumption day in this partial dataset occurs on Dec. 24th (Day 359), this day is selected as a proxy for an event day.

5.5.2 Error analysis before load aggregation

Table 7 applies the three metrics to the CBL calculations and their adjusted forms. The calculated values are for event hours. According to the results, the morning adjustment exerts a detrimental effect on the accuracy of each CBL method. However, for OPI values, the adjustment improved it significantly for ISONE and regression methods. Unlike the two aforementioned methods, adjustment did not improve the OPI of NYISO.

5.5.3 Error analysis after simple load aggregation

In this section, a simple load aggregation is applied to the sample data to examine how a load aggregation and random clustering can improve the accuracy, bias, and OPI values of CBL calculation methods that were provided in Table 7. Furthermore, a sensitivity analysis on the number of customers in each cluster is performed to show

Table 7: Accuracy MAE, bias, and OPI of classic CBL methods at event hours

CBL Methods	Accuracy MAE (kWh/hr)	Bias (kWh/hr)	OPI (kWh/hr)
NYISO	1.1923	-0.0436	0.6179
Adjusted NYISO	1.4396	+0.3294	0.8845
ISONE	1.0656	-0.4552	0.7604
Adjusted ISONE	1.1814	+0.1846	0.6830
Regression	1.3341	-0.6232	0.9786
Adjusted Regression	1.3591	+0.1961	0.7776

the impact of the number of customers in each cluster on error performance.

As discussed earlier, one of the potential benefits of load aggregation is harnessing the randomness of individual customers. In order to show how aggregation can help improve the predictability of the loads, four aggregated loads with different number of customers in each one are illustrated in Figure 21. In this figure, the average load for five days for four groups with 1, 5, 10, and 25 customers are shown for comparison purposes. Moreover, a 20-degree polynomial is fitted to loads of each group to illustrate that as the number of customers in each group increases, the variability decreases and the group becomes more predictable.

In order to show more details about aggregation, and how it helps improve the predictability, the box plots for the four groups with 1, 5, 10, and 20 customers are illustrated in Figure 22. Moreover, in this figure, the outliers are shown. It is understood that there are many outliers in the groups with one customer (no grouping); however, as more customers are added to a group, the number of outliers decreases significantly.

The 200 customers are grouped into 40 clusters of 5 households. Since the data is only available for 199 customers, one customer's data is repeated in the dataset to create the artificial 200th customer. We treat the aggregated consumption of each 5-household cluster as an individual. Table 8 provides the accuracy, bias, and OPI values of the new CBL calculations.

Figures 23 and 24 illustrate accuracy, bias, and OPI values of all CBL methods before and after applying simple load aggregation and clustering. As shown in Figure 23, simple load aggregation and clustering clearly improve the accuracy while the bias

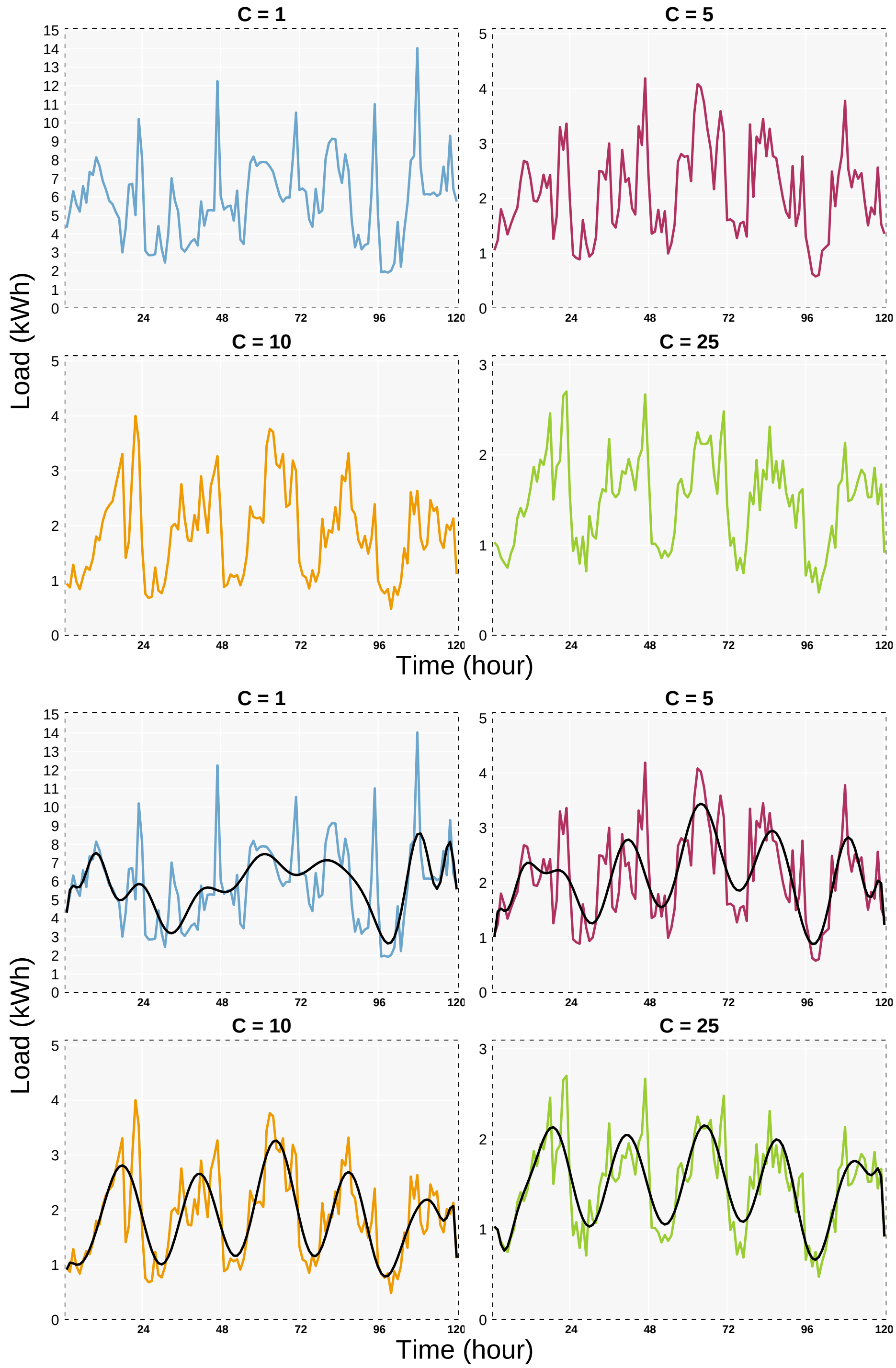

Figure 21: (top) The average load for five days for groups with 1, 5, 10, and 25 customers, (bottom) The average load for five days for groups with 1, 5, 10, and 25 customers with 20 degree polynomial fitted to the graphs

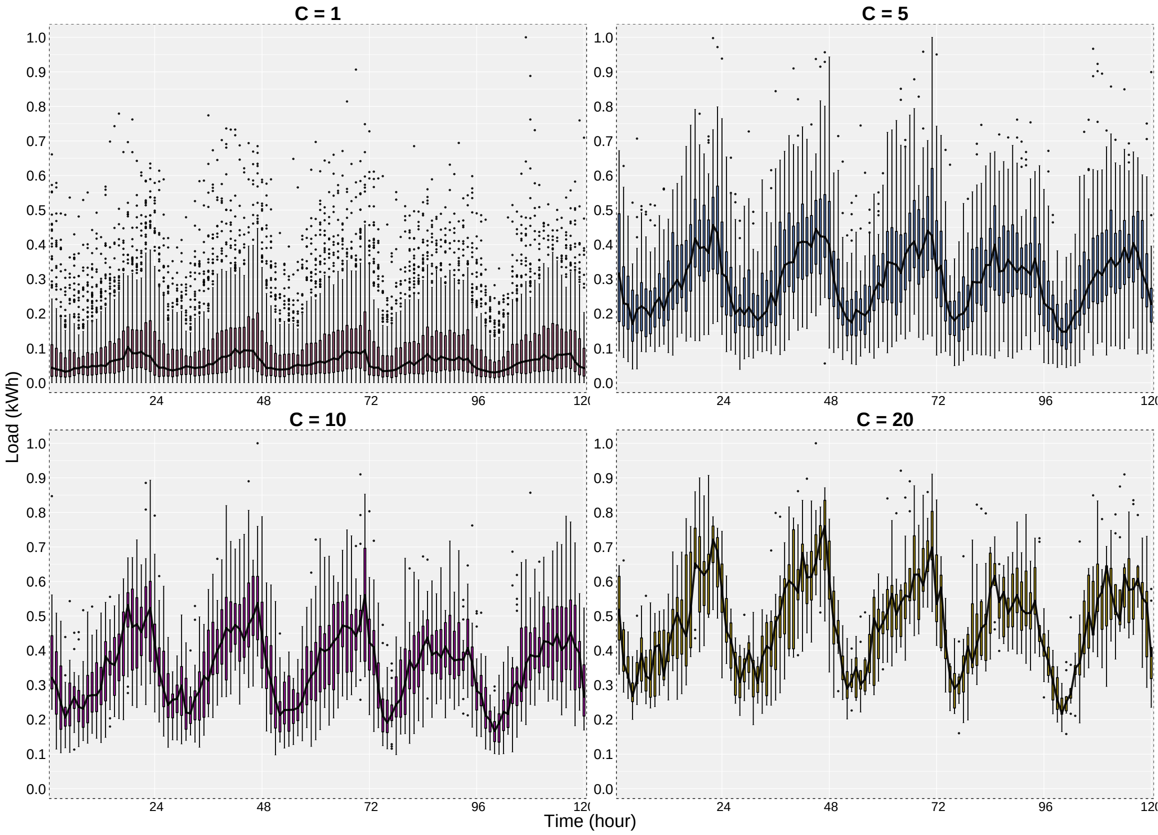

Figure 22: The average load for five days for groups with 1, 5, 10, and 20 customers. The median load of all hours is connected with a line.

remains almost unchanged. Also, Figure 24 demonstrates that the OPI is markedly improved after applying clustering. Unlike in the case before applying simple load aggregation and clustering, the additive event-day adjustment applied to the clusters has a positive effect on the accuracy of clustered methods of ISONE and regression.

Table 8: Accuracy MAE, bias, and OPI for clustered methods at event hours

CBL Methods	Accuracy MAE (kWh/hr)	Bias (kWh/hr)	OPI (kWh/hr)
Clustered NYISO	0.6808	-0.2256	0.4532
Adjusted Clustered NYISO	0.7439	+0.2988	0.5213
Clustered ISONE	0.6846	-0.4571	0.5708
Adjusted Clustered ISONE	0.6172	0.1935	0.4053
Clustered Regression	0.9074	-0.6611	0.7842
Adjusted Clustered Regression	0.6921	0.1901	0.4411

Moreover, the OPI of these two methods is improved significantly. On the other hand, clustering has little effect on the accuracy and the OPI values of the NYISO method.

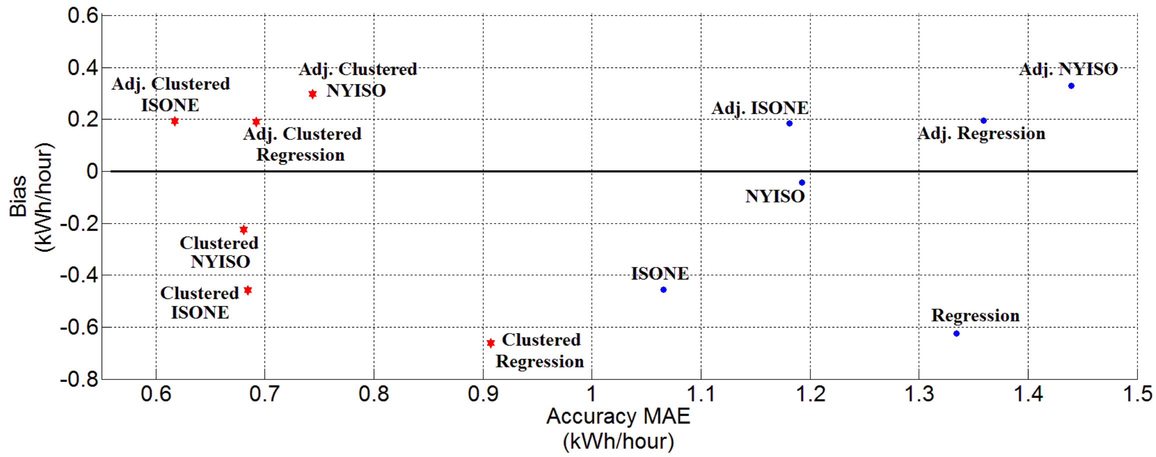

Figure 23: CBL method accuracy "MAE" and bias for event hours

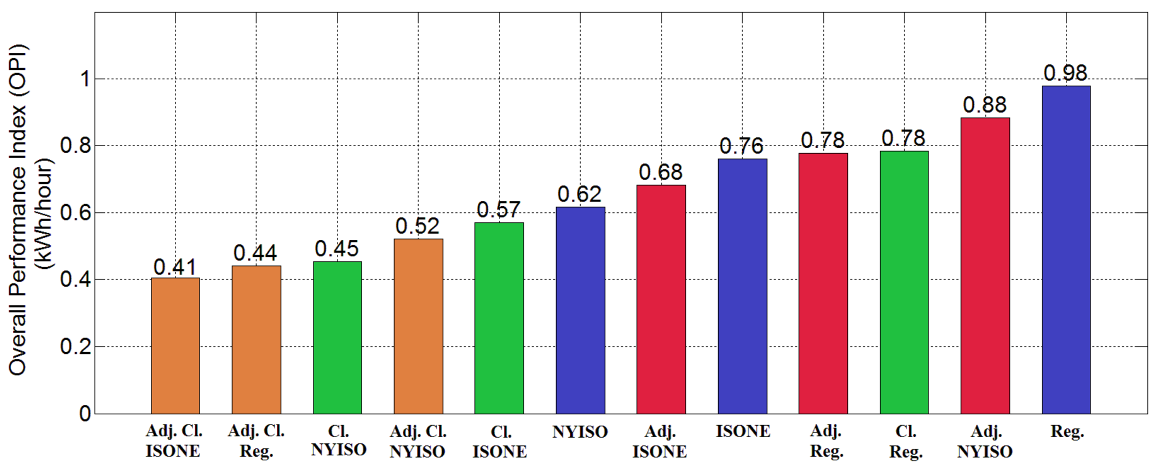

Figure 24: CBL method OPI (Clustered is green, traditional is blue, adjusted clustered is light brown, and adjusted is red)

5.5.3.1 Sensitivity analysis

A sensitivity analysis on the number of customers in each cluster for three methods of NYISO, ISONE, and regression is performed to show how the variation in the number of customers in each cluster would impact the error analysis. The results are provided in Table 9 and Figure 25. The figure shows that as the number of customers in each cluster increases, the OPI improves accordingly for all three methods. According to the results, even grouping two customers in a cluster can improve OPI by 13.99%, 12.73%, and 11.80% for NYISO, ISONE and regression method, respectively. These values increase as the number of customers in each cluster increases, except for the case of two clusters with 100 customers per cluster.

Table 9: Sensitivity analysis on the number of clusters in each cluster's impact on overall performance index

No. of clusters (No. of customers)	Overall performance index (kWh/hour)		
	NYISO	ISONE	Regression
2 (100)	0.3704	0.4593	0.6385
4 (50)	0.3514	0.4609	0.6356
5 (40)	0.3597	0.4722	0.6524
8 (25)	0.3645	0.4854	0.6616
10 (20)	0.384	0.5008	0.6891
20 (10)	0.4140	0.5257	0.7161
25 (8)	0.4088	0.5288	0.7802
40 (5)	0.4532	0.5708	0.7842
50 (4)	0.4779	0.5979	0.8161
100 (2)	0.5314	0.6636	0.8631
No Cluster	0.6179	0.7604	0.9786

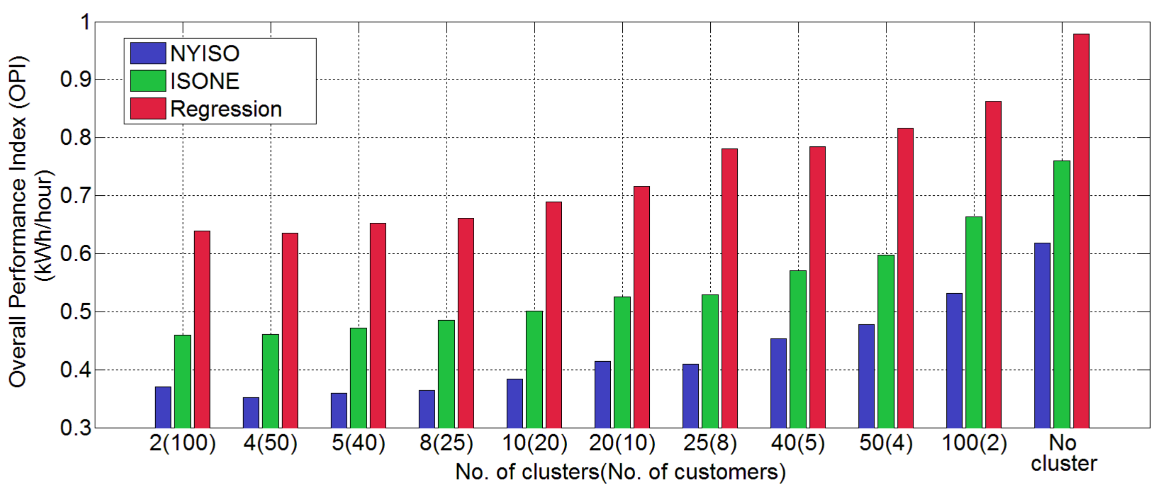

Figure 25: Sensitivity analysis on the number of customers in each cluster's impact on overall performance index

5.5.4 Error analysis for the proposed CBL estimation method

In this section, the proposed CBL estimation method is applied to the dataset. As illustrated in the flowchart in chapter 3, the estimation starts with clustering customers into different bins based on average consumption and predictability index. After obtaining the predictability index of all customers and applying k-means clustering, five bins (k=5) are selected. These bins are shown in figure 15.

The box plot of customers in each bin is illustrated in figure 26. Similar to Figures 19 and 20, the data are divided into four seasons to see the impact of seasonality. As shown in the figure, customers in bin 4 and 5 have the highest variability. On the other hand, bin 1 has the lowest variability among the bins.

The next step is to randomly group the customers of each bin into different clusters and estimate CBL for each cluster, and calculate the error performance. For carrying out this analysis, two CBL methods of CAISO and PJM are selected. The error performance results, i.e. MAE values for CAISO method, are listed in Table 10. Likewise, the similar error results for PJM method are provided in Table 11. These results are the average of MAE for 12 events. In order to claim that the presence of customers in each group is random, five random combinations of customers in each group are selected, and the value of MAE are averaged for each event.

The results of Tables 10 and 11 are illustrated in Figures 27 and 28. The trend for CAISO is that as the number of customers in each group increases, the MAE decreases. The similar trend is observed for PJM method. In CAISO, the highest drop in MAE is observed in bin 1 (54.1%), and the lowest is observed for bin 4

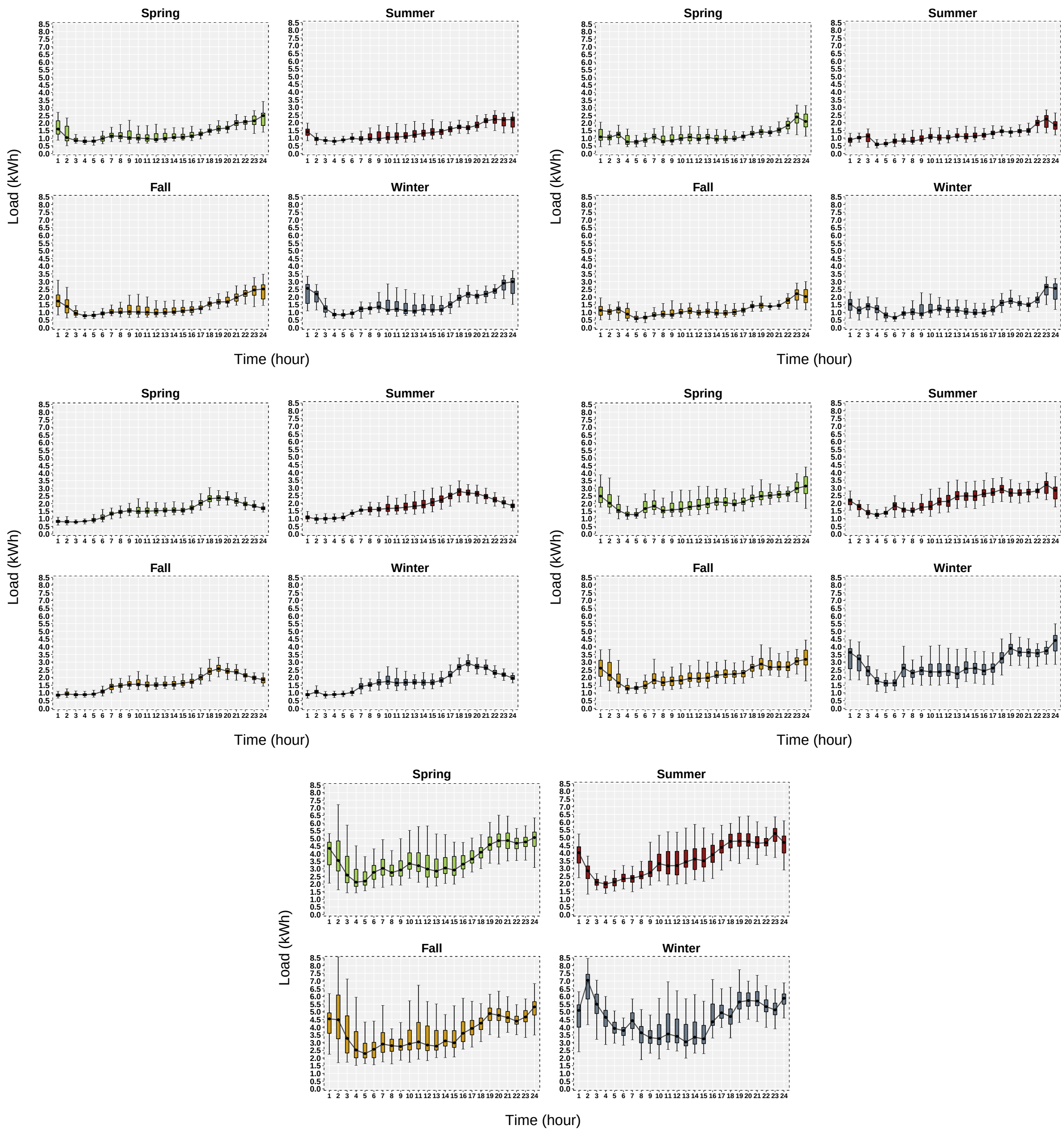

Figure 26: (top left) Box plot of loads of all customers in bin 1 for 24 hours
(top right) Box plot of loads of all customers in bin 2 for 24 hours
(middle left) Box plot of loads of all customers in bin 3 for 24 hours
(middle right) Box plot of loads of all customers in bin 4 for 24 hours
(bottom) Box plot of loads of all customers in bin 5 for 24 hours

Table 10: Mean absolute error for CAISO method

	No cluster	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
Bin 1	1.201	0.904	0.76	0.679	0.685	0.563	0.615	0.48	0.57	0.55
Bin 2	0.621	0.467	0.432	0.338	0.366	0.342	0.248	0.32	0.30	0.28
Bin 3	0.819	0.671	0.558	0.531	0.493	0.463	0.449	0.42	0.41	0.39
Bin 4	0.723	0.607	0.541	0.483	0.396	0.432	0.323	0.39	0.36	0.37
Bin 5	1.727	1.4	1.011	1.111	1.06	0.989	0.605	0.93	0.89	0.86

Table 11: Mean absolute error for PJM method

	No cluster	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
Bin 1	1.17	0.878	0.739	0.654	0.669	0.533	0.584	0.45	0.55	0.53
Bin 2	0.64	0.46	0.462	0.337	0.385	0.354	0.244	0.32	0.31	0.3
Bin 3	0.801	0.661	0.556	0.531	0.501	0.479	0.458	0.42	0.42	0.39
Bin 4	0.712	0.613	0.544	0.49	0.378	0.426	0.313	0.40	0.37	0.38
Bin 5	1.718	1.407	0.992	1.16	1.084	0.971	0.566	0.92	0.92	0.93

(48.4%). Almost similar observation could be made for PJM. The highest drop in MAE is observed in bin 1 (54.1%), and the lowest is observed for bin 5 (45.7%). As a result, it could be asserted that the drop in MAE value is sensitive to the variability. In fact, as the variability increases, the drop in MAE decreases as the number of customers in a group decreases.

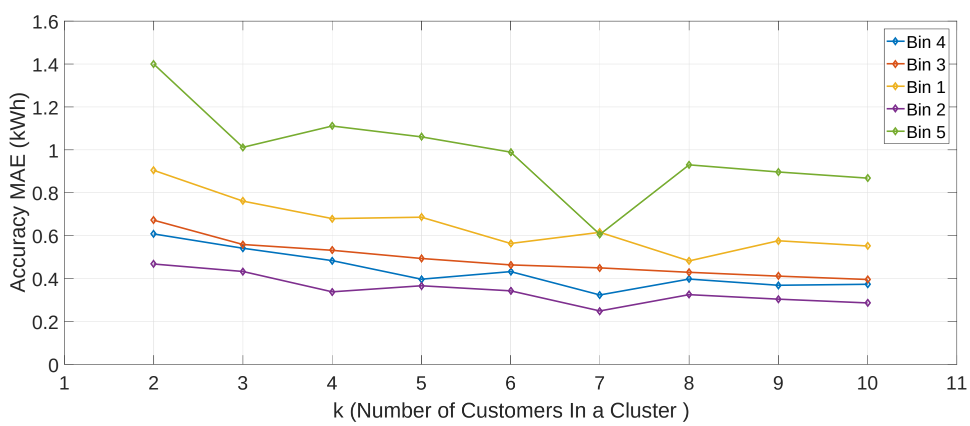

Figure 27: Mean absolute error for customers of all bins in groups with k customers obtained by the CAISO method

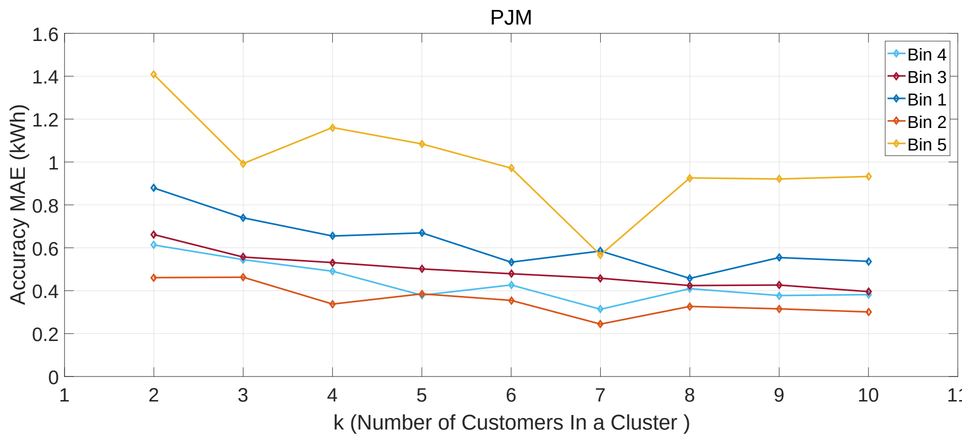

Figure 28: Mean absolute error for customers of all bins in groups with k customers obtained by the PJM method

Another metric for error performance is "bias," which its values for the CAISO method are listed in Table 12. The similar bias results for the PJM method are listed in Table 13. These results are the average of bias for 12 events. In order to claim that the presence of customers in each group is random, five random combinations of customers in each group are selected, and the value of MAE are averaged for each event.

The results of Tables 12 and 13 are illustrated in Figures 29 and 30. In both CAISO and PJM methods, it can be understood that the bias is not sensitive to the number of customers in each cluster. Moreover, it can be seen that there is no relation between variability and the bias value.

Table 12: Bias value for CAISO method

	No cluster	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
Bin 1	-0.255	-0.23	-0.24	-0.23	-0.26	-0.21	-0.26	-0.19	-0.25	-0.25
Bin 2	-0.064	-0.06	-0.06	-0.05	-0.05	-0.05	-0.04	-0.06	-0.06	-0.06
Bin 3	-0.156	-0.15	-0.15	-0.15	-0.15	-0.15	-0.15	-0.15	-0.15	-0.15
Bin 4	-0.103	-0.10	-0.10	-0.10	-0.08	-0.11	-0.07	-0.10	-0.10	-0.10
Bin 5	-0.488	-0.48	-0.4	-0.48	-0.49	-0.48	-0.38	-0.46	-0.48	-0.48

Table 13: Bias value for PJM method

	No cluster	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
Bin 1	-0.256	-0.24	-0.23	-0.21	-0.25	-0.24	-0.25	-0.22	-0.25	-0.25
Bin 2	-0.071	-0.05	-0.07	-0.06	-0.07	-0.07	-0.05	-0.07	-0.07	-0.07
Bin 3	-0.154	-0.15	-0.14	-0.15	-0.15	-0.15	-0.15	-0.15	-0.15	-0.15
Bin 4	-0.105	-0.10	-0.10	-0.10	-0.08	-0.10	-0.07	-0.10	-0.10	-0.10
Bin 5	-0.532	-0.53	-0.44	-0.53	-0.53	-0.53	-0.41	-0.53	-0.53	-0.53

5.5.5 Comparing the proposed method with simple load aggregation

In this section, it is shown how the proposed approach successfully enhances the simple load aggregation. As shown earlier, simple load aggregation improves the

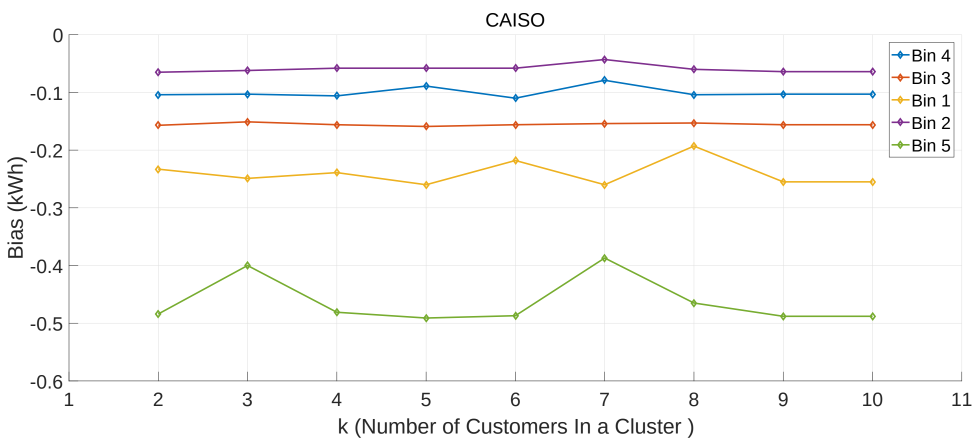

Figure 29: Bias value for customers of all bins in groups with k customers obtained by the CAISO method

Figure 30: Bias value for customers of all bins in groups with k customers obtained by the PJM method

Table 14: The accuracy MAE for simple load aggregation and the proposed method for CAISO and PJM CBL estimation methods

		k (Number of customers in a cluster)									
		1	2	3	4	5	6	7	8	9	10
CAISO	Simple load aggregation	1	0.83	0.68	0.64	0.58	0.56	0.52	0.51	0.52	0.52
	Proposed method	1	0.81	0.66	0.62	0.60	0.55	0.44	0.51	0.51	0.49
PJM	Simple load aggregation	1	0.83	0.73	0.66	0.62	0.58	0.54	0.53	0.52	0.49
	Proposed method	1	0.77	0.64	0.59	0.55	0.47	0.41	0.42	0.43	0.42

error performance of CBL estimation methods significantly. Moreover, the proposed method builds upon that foundation and tries to improve it further. Table 14 lists the results of both simple load aggregation and the proposed method for CAISO and PJM cases. In this Table, k refers to the number of customers in a cluster. It is worth mentioning that k equals one means that there is no cluster, and it is included in the Table to give a reference for the future comparisons.

According to the results of Table 14, the average MAE for all k values for simple load aggregation and the proposed method in CAISO case are 0.601 and 0.58, respectively. These results indicate that the proposed method shows 3.43% improvement in MAE value. Moreover, the average MAE for all k values for simple load aggregation and the proposed method in PJM case are 0.616 and 0.527, respectively, which is equivalent to 14.42% improvement. The improvement is more significant in PJM case. However, it is observed that the bulk of the improvement is because of the load aggregation, while the clustering based on predictability index causes a small

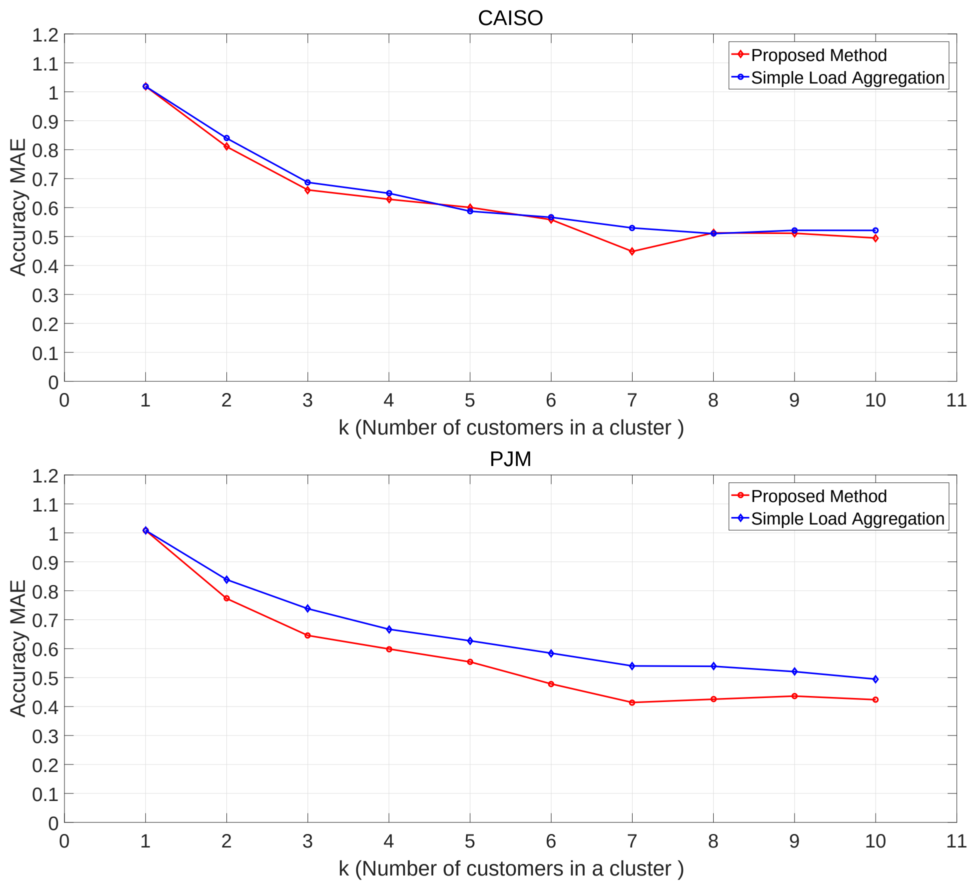

Figure 31: (top) Comparing the accuracy MAE of the proposed method and the simple load aggregation for CAISO case, (bottom) Comparing the accuracy MAE of the proposed method and the simple load aggregation for PJM case

extra improvement. Consequently, if the proposed clustering based on predictability index incurs a large administrative cost, the utilities can use simple load aggregation and have a significant improvement. In large DR programs that even a small improvement in accuracy can translate into a significant financial gain, the proposed clustering method can be extremely beneficial. The results of Table 14 are illustrated in figure 31.

CHAPTER 6: CONCLUSION AND FUTURE WORKS

6.1 Conclusion

In this dissertation, it is argued that demand response programs need reliable Evaluation, Measurement, and Verification (EM&V) methods. One of the major EM&V challenges in DR programs is an accurate estimation of the load reduction, particularly for residential customers that show significant variability. To estimate the load reduction for DR programs offered to residential customers, it is necessary to estimate the Customer Baseline Load (CBL) accurately. The presence of CBL introduces a lot of challenges to DR programs. In this dissertation, the challenges associated with the presence of CBL, e.g. gaming and poor error performance, are described in detail. In order to address these problems, a method is proposed to improve the error performance of the CBL estimation. The merits of this proposed methods are shown in theory. Moreover, by an analysis of residential dataset, it is demonstrated that this proposed method can improve the error performance.

The key conclusions are:

- FERC 745 policy provides an incentive for over-consumption during non-event days;
- FERC 745 policy provides an incentive for under-consumption during event days;

- If perfect accuracy is not obtainable, methods with negative bias deliver better performance than methods with positive bias with regards to social welfare loss;
- The event-day additive morning adjustment has a detrimental effect on the accuracy of each CBL method. However, it improves the OPI for ISONE and regression methods, but not the OPI of NYISO;
- The simple load aggregation improves the accuracy significantly while the bias stays almost unchanged;
- If the morning adjustment is applied to the simple load aggregation, it affects the accuracy and bias values positively;
- OPI is noticeably improved after applying simple load aggregation;
- In both cases of individual and simple load aggregation, the morning adjustment changes the negative sign of the bias to positive; consequently, it increases the utility's revenue loss;
- In the proposed clustering method, the drop in MAE value is sensitive to the variability.
- In the proposed clustering method, as the variability increases, the drop in MAE decreases as the number of customers in a group decreases.
- In the proposed clustering method, it is shown that there is no relation between variability and the bias value.

- The proposed method shows an improvement compared to the simple load aggregation method.

6.2 Summary of contribution

The key contributions of this dissertation can be summarized as follows:

- The error performance of the well-established CBL estimation methods are evaluated for residential customers.
- The challenges of CBL estimation methods are analyzed in theory by using customers' utility function.
- The impact of FERC 745 order on the current DR programs is evaluated in theory. Moreover, the impact of the order is shown on the social welfare.
- The stochasticity of three categories of customers, i.e. large industrial, commercial, and residential customers is evaluated.
- A novel method is proposed to cluster customers in order to improve the error performance of current CBL estimation methods.
- The impact of the proposed method is analyzed on the social welfare.

6.3 Future works

Based on the results of this dissertation, it is possible to suggest a few paths for the future works. The key suggestions are as follows:

- There are other signal processing techniques such wavelet transform (WT) that are better equipped to handle the time-domain data with variable frequency.

It would be an interesting research to produce a different predictability index according to WT outputs.

- Another interesting path is to apply the proposed methodology in RCT methods to create a control group.

- It is very important to re-examine the results of this dissertation with a much larger dataset, preferably on residential customers from multiple locations in the US.

- If personal data from customers are available, it is possible to use them to see the relation between predictability index and individual characteristics of each customer. In other words, to examine how household income, education, and house size can change the predictability index.

- With actual data from a DR program, it is possible to study the impact of the proposed method on revenue stream of utilities.

- With actual data from a DR program, it is also possible to study temperature. Temperature is another important feature that can be used to improve the error performance of CBL estimation methods. In this dissertation, the author could not find a good source of data with temperature information.

- Residential customers who use renewable energies, e.g. rooftop solar panels, show different behaviors, another interesting direction is to examine the proposed method for these customers.

- Residential customers show many outliers. It is worth examining how eliminating such outliers would affect the error performance of CBL estimation methods.

REFERENCES

[1] Australian Energy Market Operation (AEMO) load profile data. http://www.aemo.com.au/Electricity/Data/Metering.

[2] Working group 2 demand response program evaluation. program year 2004 final report. prepared for the working group 2 measurement and evaluation committee. Technical report, by Quantum Consulting Inc. and Summit Blue Consulting, LLC, 2004.

[3] 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. Technical report, by Quantum Consulting Inc. and Summit Blue Consulting, LLC, 2006.

[4] The demand response baseline. Technical report, EnerNOC, 2011.

[5] Business practices for measurement and verification of wholesale electricity demand response. Technical report, North American Energy Standards Board, March 16, 2009.

[6] N. Addy, J. L. Mathieu, S. Kiliccote, and D. S. Callaway. Understanding the effect of baseline modeling implementation choices on analysis of demand response performance. In *ASME 2012 International Mechanical Engineering Congress and Exposition*, pages 133–141. American Society of Mechanical Engineers, 2012.

[7] 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*, 137(2):021008, 2015.

[8] J. Aghaei and M.-I. Alizadeh. Demand response in smart electricity grids equipped with renewable energy sources: A review. *Renewable and Sustainable Energy Reviews*, 18:64–72, 2013.

[9] A. Asadinejad and K. Tomsovic. Optimal use of incentive and price based demand response to reduce costs and price volatility. *Electric Power Systems Research*, 144:215–223, 2017.

[10] A. Asadinejad, K. Tomsovic, and C. f. Chen. Sensitivity of incentive based demand response program to residential customer elasticity. In *2016 North American Power Symposium (NAPS)*, pages 1–6, Sept 2016.

[11] A. Asadinejad, K. Tomsovic, and M. G. Varzaneh. Examination of incentive based demand response in western connection reduced model. In *2015 North American Power Symposium (NAPS)*, pages 1–6, Oct 2015.

[12] A. Asadinejad, M. Varzaneh, K. Tomsovic, C.-F. Chen, and R. Sawhney. Residential customers elasticity estimation and clustering based on their contribution at incentive based demand response. In *IEEE Power and Energy Society (PES) General Meeting*, 2016.

[13] A. Buege, M. Rufo, M. Ozog, D. Violette, and 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.

[14] P. Cappers and et al. Quantifying the impacts of timebased rates, enabling technology, and other treatments in consumer behavior studies: Protocols and guidelines. *Lawrence Berkeley National Lab (LBNL), LBNL-6301E*, July 2013.

[15] H.-p. Chao. Demand response in wholesale electricity markets: the choice of customer baseline. *Journal of Regulatory Economics*, 39(1):68–88, 2011.

[16] H.-p. Chao and M. DePillis. Incentive effects of paying demand response in wholesale electricity markets. *Journal of Regulatory Economics*, 43(3):265–283, 2013.

[17] H.-p. Chao and I. N. England. Demand management in restructured wholesale electricity markets. *ISO New England*, 2010.

[18] M. Chaouch. Clustering-based improvement of nonparametric functional time series forecasting: Application to intra-day household-level load curves. *IEEE Transactions on Smart Grid*, 5(1):411–419, Jan 2014.

[19] G. Chicco. Overview and performance assessment of the clustering methods for electrical load pattern grouping. *Energy*, 42(1):68–80, 2012.

[20] K. Coughlin, M. A. Piette, C. Goldman, and S. Kiliccote. Estimating demand response load impacts: evaluation of baseline load models for non-residential buildings in california. *Lawrence Berkeley National Laboratory*, 2008.

[21] K. Coughlin, M. A. Piette, C. Goldman, and S. Kiliccote. Statistical analysis of baseline load models for non-residential buildings. *Energy and Buildings*, 41(4):374–381, 2009.

[22] P. G. Da Silva, D. Ilic, and S. Karnouskos. The impact of smart grid prosumer grouping on forecasting accuracy and its benefits for local electricity market trading. *IEEE Transactions on Smart Grid*, 5(1):402–410, 2014.

[23] I. N. England. Filing of changes to day-ahead load response program. *Docket No. ER08-538-000*, Feb. 2008.

[24] EPRI. Guidelines for designing effective energy information feedback pilots: Research protocols. Technical report, Electric Power Research Institute, ID1020855, Apr. 2010.

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

[26] A. Faruqui and S. George. Quantifying customer response to dynamic pricing. *The Electricity Journal*, 18(4):53–63, 2005.

[27] A. Faruqui and S. Sergici. Household response to dynamic pricing of electricity: a survey of 15 experiments. *Journal of regulatory Economics*, 38(2):193–225, 2010.

[28] A. Faruqui and S. Sergici. Arcturus: international evidence on dynamic pricing. *The Electricity Journal*, 26(7):55–65, 2013.

[29] FERC. Demand response compensation in organized wholesale energy markets. *Docket No. RM10-17-000, Order No. 745*, March 2011.

[30] C. Flath, D. Nicolay, T. Conte, C. van Dinther, and L. Filipova-Neumann. Cluster analysis of smart metering data. *Business & Information Systems Engineering*, 4(1):31–39, 2012.

[31] M. Goldberg and G. K. Agnew. Protocol development for demand-response calculations: Findings and recommendations, kema-xenergy. Technical report, CEC 400-02-017F, 2003.

[32] C. Grimm. Evaluating baselines for demand response programs. In *AEIC Load Research Workshop*, 2008.

[33] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke. Smart grid and smart homes: Key players and pilot projects. *IEEE Industrial Electronics Magazine*, 6(4):18–34, 2012.

[34] S. Haben, C. Singleton, and P. Grindrod. Analysis and clustering of residential customers energy behavioral demand using smart meter data. *IEEE Transactions on Smart Grid*, 7(1):136–144, 2016.

[35] S. Haben, J. Ward, D. V. Greetham, C. Singleton, and P. Grindrod. A new error measure for forecasts of household-level, high resolution electrical energy consumption. *International Journal of Forecasting*, 30(2):246–256, 2014.

[36] R. J. Hamidi, S. H. Hosseinian, S. H. H. Sadeghi, and Z. Qu. A novel approach to utilize plc to detect corroded and eroded segments of power transmission lines. *IEEE Transactions on Power Delivery*, 30(2):746–754, April 2015.

[37] R. J. Hamidi and H. Livani. Traveling-wave-based fault-location algorithm for hybrid multiterminal circuits. *IEEE Transactions on Power Delivery*, 32(1):135–144, Feb 2017.

[38] R. J. Hamidi, H. Livani, and R. Rezaiesarlak. Traveling-wave detection technique using short-time matrix pencil method. *IEEE Transactions on Power Delivery*, PP(99):1–1, 2017.

[39] D. Hammerstrom, R. Ambrosio, J. Brous, T. Carlon, D. Chassin, J. DeSteese, R. Guttromson, G. Horst, O. Järvegren, R. Kajfasz, et al. Pacific northwest gridwise testbed demonstration projects. *Part I. Olympic Peninsula Project*, 210, 2007.

[40] L. Hatton, P. Charpentier, and E. Matzner-Løber. Statistical estimation of the residential baseline. *IEEE Transactions on Power Systems*, 31(3):1752–1759, 2016.

[41] K. Herter, P. McAuliffe, and A. Rosenfeld. An exploratory analysis of california residential customer response to critical peak pricing of electricity. *Energy*, 32(1):25–34, 2007.

[42] D. Jang, J. Eom, M. G. Kim, and J. J. Rho. Demand responses of korean commercial and industrial businesses to critical peak pricing of electricity. *Journal of Cleaner Production*, 90:275–290, 2015.

[43] J. Jazaeri, T. Alpcan, R. Gordon, M. Brandao, T. Hoban, and C. Seeling. Baseline methodologies for small scale residential demand response. In *Innovative Smart Grid Technologies-Asia (ISGT-Asia), 2016 IEEE*, pages 747–752. IEEE, 2016.

[44] K. Jong-Min, J. Nam-Joon, Y. In-Hyeob, and Y.-I. Kim. Day-ahead load reduction system based on customer baseline load, Oct. 9 2009. US Patent App. 12/576,609.

[45] K. Jong-Min, J. Nam-Joon, Y.-I. Kim, and Y. In-Hyeob. Load forecasting analysis system for calculating customer baseline load, Mar. 26 2013. US Patent 8,406,935.

[46] M. Lee, O. Aslam, B. Foster, D. Kathan, J. Kwok, L. Medearis, R. Palmer, P. Sporborg, and M. Tita. Assessment of demand response and advanced metering. *Federal Energy Regulatory Commission, Tech. Rep*, 2013.

[47] Y. Li, P. Guo, and X. Li. Short-term load forecasting based on the analysis of user electricity behavior. *Algorithms*, 9(4):80, 2016.

[48] S. Moghadasi and S. Kamalasadan. Optimal fast control and scheduling of power distribution system using integrated receding horizon control and convex conic programming. *IEEE Transactions on Industry Applications*, 52(3):2596–2606, 2016.

[49] S. M. Moghadasi, A. Kazemi, M. Fotuhi-Firuzabad, and A. A. Edris. Composite system reliability assessment incorporating an interline power-flow controller. *IEEE Transactions on Power Delivery*, 23(2):1191–1199, 2008.

[50] S. Mohagheghi, F. Yang, and B. Falahati. Impact of demand response on distribution system reliability. In *2011 IEEE Power and Energy Society General Meeting*, pages 1–7, July 2011.

[51] S. Mohajeryami. An investigation of the impact of customer baseline (CBL) calculation on peak time rebate program. Master's thesis, Department of Economics, UNC Charlotte, Charlotte, NC, 2016.

[52] S. Mohajeryami, A. Asadinejad, and M. Doostan. An investigation of the relationship between accuracy of customer baseline calculation and efficiency of peak time rebate program. In *in IEEE Power and Energy Conference at Illinois (PECI)*, Feb 2016.

[53] S. Mohajeryami, A. Asadinejad, M. Doostan, and P. Schwarz. Error analysis of customer baseline load (CBL) calculation methods for residential customers. *IEEE Transactions on Industry Applications*, 2016.

[54] S. Mohajeryami, M. Doostan, and P. Schwarz. The impact of customer baseline load (cbl) calculation methods on peak time rebate program offered to residential customers. *Electric Power Systems Research*, 137:59–65, 2016.

[55] S. Mohajeryami, I. N. Moghaddam, M. Doostan, B. Vatani, and P. Schwarz. A novel economic model for price-based demand response. *Electric Power Systems Research*, 135:1–9, 2016.

[56] S. Mohajeryami, P. Schwarz, and P. T. Baboli. Including the behavioral aspects of customers in demand response model: Real time pricing versus peak time rebate. In *North American Power Symposium (NAPS)*, pages 1–6. IEEE, 2015.

[57] I. Naziri Moghaddam and B. Chowdhury. Optimal sizing of hybrid energy storage systems to mitigate wind power fluctuations. In *IEEE PES General Meeting Conference & Exposition, Boston, MA*, 2016.

[58] S. Park, S. Ryu, Y. Choi, and H. Kim. A framework for baseline load estimation in demand response: Data mining approach. In *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 638–643, 2014.

[59] S. Park, S. Ryu, Y. Choi, J. Kim, and H. Kim. Data-driven baseline estimation of residential buildings for demand response. *Energies*, 8(9):10239–10259, 2015.

[60] J. Rice. *Mathematical statistics and data analysis*. Nelson Education, 2006.

[61] H. Sæle and O. S. Grande. Demand response from household customers: Experiences from a pilot study in norway. *IEEE Transactions on Smart Grid*, 2(1):102–109, 2011.

[62] H. Sangrody and N. Zhou. An initial study on load forecasting considering economic factors. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5, July 2016.

[63] R. Sevlian and R. Rajagopal. Short term electricity load forecasting on varying levels of aggregation. *arXiv preprint arXiv:1404.0058*, 2014.

[64] R. Sharifi, S. Fathi, and V. Vahidinasab. A review on demand-side tools in electricity market. *Renewable and Sustainable Energy Reviews*, 72:565–572, 2017.

[65] N. H. Tehrani. *Demand-side participation & baseline load analysis in electricity markets*. PhD thesis, University of Victoria, 2016.

[66] N. H. Tehrani, U. T. Khan, and C. Crawford. Baseline load forecasting using a bayesian approach. In *2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–4, May 2016.

[67] R. Tibshirani, G. Walther, and T. Hastie. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):411–423, 2001.

[68] A. Todd, E. Stuart, C. Goldman, and S. Schiller. Evaluation, measurement, and verification em&v) for behavior-based energy efficiency programs: issues and recommendations. *Proceedings of the 2012 ACEEE Summer Study on Energy Efficiency in Buildings*, 2012.

[69] J. Torriti, M. G. Hassan, and M. Leach. Demand response experience in europe: Policies, programmes and implementation. *Energy*, 35(4):1575–1583, 2010.

[70] Y. Weng and R. Rajagopal. Probabilistic baseline estimation via gaussian process. In *2015 IEEE Power Energy Society General Meeting*, pages 1–5, July 2015.

[71] Y.-M. Wi, J.-H. Kim, S.-K. Joo, J.-B. , and J.-C. Oh. Customer baseline load (cbl) calculation using exponential smoothing model with weather adjustment. In *IEEE Transmission & Distribution Conference & Exposition: Asia and Pacific*, 2009.

[72] T. K. Wijaya, M. Vasirani, and K. Aberer. When bias matters: An economic assessment of demand response baselines for residential customers. *IEEE Transactions on Smart Grid*, 5(4):1755–1763, 2014.

[73] F. A. Wolak. Residential customer response to real-time pricing: The anaheim critical peak pricing experiment. *Center for the Study of Energy Markets*, 2007.

[74] J.-R. Won, Y.-B. Yoon, J.-H. Choi, and K.-K. Yi. Case study of customer baseline (cbl) load determination in korea based on statistical method. In *2009 Transmission & Distribution Conference & Exposition: Asia and Pacific*, pages 1–4. IEEE, 2009.

[75] N. Yamaguchi, J. Han, G. Ghatikar, S. Kiliccote, M. A. Piette, and H. Asano. Regression models for demand reduction based on cluster analysis of load profiles. In *IEEE PES/IAS Conference on Sustainable Alternative Energy (SAE)*, pages 1–7, 2009.

[76] Y. Zhang, W. Chen, R. Xu, and J. Black. A cluster-based method for calculating baselines for residential loads. *IEEE Transactions on Smart Grid*, 7(5):2368–2377, 2016.